

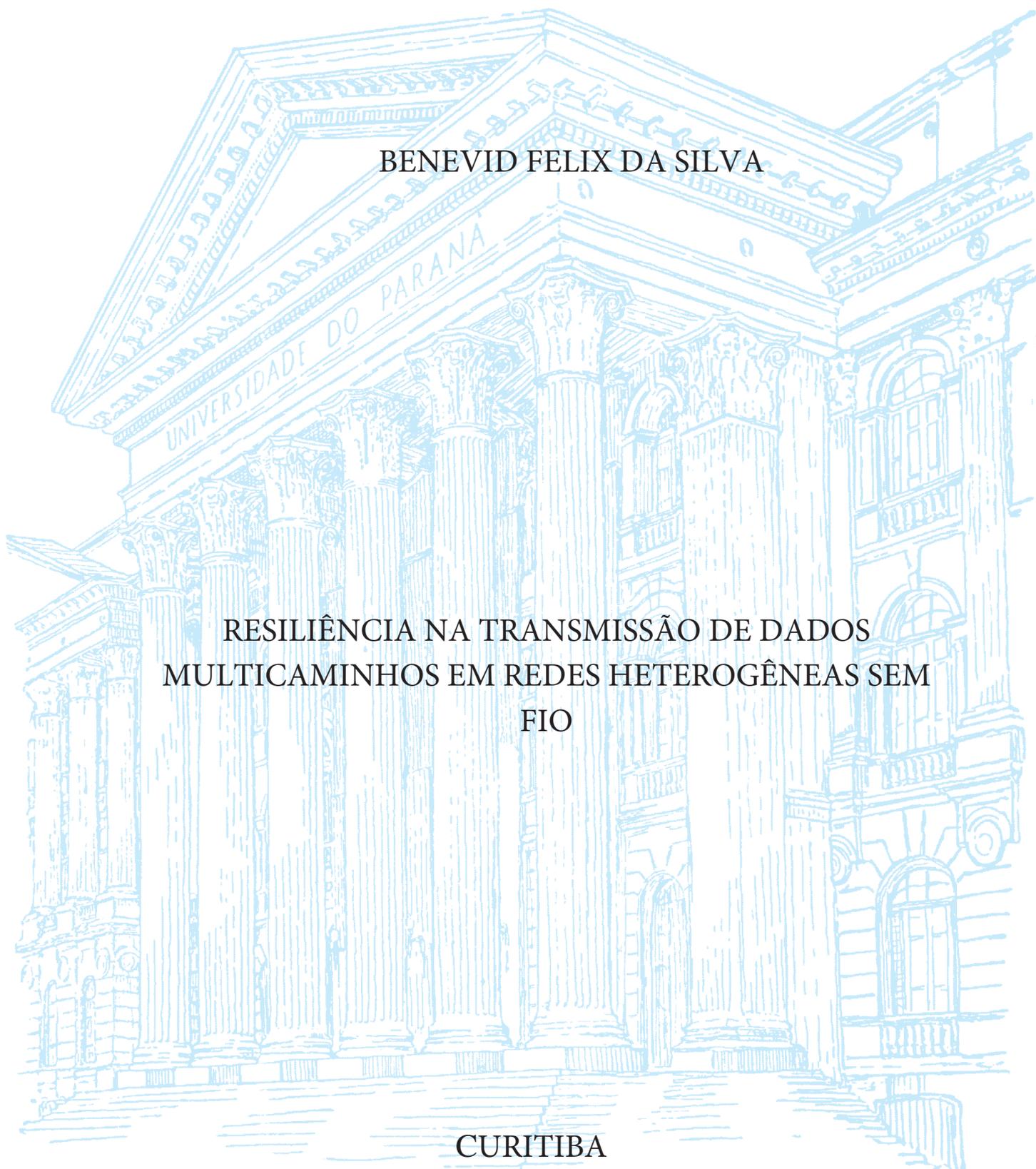
UNIVERSIDADE FEDERAL DO PARANA

BENEVID FELIX DA SILVA

RESILIÊNCIA NA TRANSMISSÃO DE DADOS
MULTICAMINHOS EM REDES HETEROGÊNEAS SEM
FIO

CURITIBA

2018



BENEVID FELIX DA SILVA

RESILIÊNCIA NA TRANSMISSÃO DE DADOS
MULTICAMINHOS EM REDES HETEROGÊNEAS SEM
FIO

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Informática, no Programa de Pós-Graduação em Informática, setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Profa. Dra. Michele Nogueira Lima.

CURITIBA

2018

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR
BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

SI586r

Silva, Benevid Felix da

Resiliência na transmissão de dados multicaminhos em redes heterogêneas sem fio / Benevid Felix da Silva. – Curitiba, 2018.

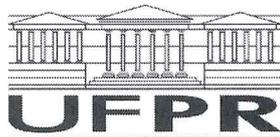
Tese (Doutorado) - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Informática, 2018.

Orientadora: Profa. Dra. Michele Nogueira Lima.

1. Tecnologia da informação. 2. Realidade virtual. 3. Redes de acesso sem fio.
I. Universidade Federal do Paraná. II. Lima, Michele Nogueira. III. Título.

CDD: 004.6

Bibliotecária: Romilda Santos - CRB-9/1214



MINISTÉRIO DA EDUCAÇÃO
SETOR SETOR DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO INFORMÁTICA

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **BENEVID FELIX DA SILVA** intitulada: **Resiliência na Transmissão de Dados Multicaminhos em Redes Heterogêneas sem Fio**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa. A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 22 de Junho de 2018.

MICHELE NOGUEIRA LIMA
Presidente da Banca Examinadora

ALDRI LUIZ DOS SANTOS
Avaliador Interno

JOSE AUGUSTO SURUAGY MONTEIRO
Avaliador Externo

LUIS HENRIQUE MACIEL KOSMALSKI COSTA
Avaliador Externo

MARCELO EDUARDO PELLEENZ
Avaliador Externo



Dedico este trabalho à minha família, esposa, filhos, meus pais, irmãos, sobrinhos, amigos e meus alunos. Espero lhes servir de exemplo para que trilhem os caminhos da ciência em busca de respostas para muitos dos problemas sem solução ou desconhecidos.

Agradecimentos

A conclusão deste trabalho é fruto de um esforço não só pessoal, mas de um grupo de pessoas e de instituições que contribuíram com sua realização.

Inicialmente eu agradeço a minha família, em especial, minha esposa Vanessa e meus filhos Myrella, Mariane e Murilo. Durante estes quatro anos sacrifiquei muito do tempo que a eles deveria ser destinado. Mesmo assim, estiveram sempre presentes, me compreendendo e dando força para manter firme e com foco na realização das inúmeras tarefas. Agradeço à todos os meus irmãos, em especial ao meu irmão/padrinho Jânio, que se esforçou para me oferecer um estudo de melhor qualidade, à minha irmã Nilce, a qual foi meu exemplo de dedicação aos estudos. Em nome dos dois, agradeço aos demais Durce, Nelci, Solene, Adelmo, Bill, Elson, Benício e Osman por sempre estarem presentes. E como não poderia deixar de lembrar, um agradecimento mais que especial, à minha mãe Maria e suas orações em todos os momentos importantes da minha vida, e ao meu pai Bené, que mesmo sem concluir seus estudos nunca mediu esforços para que todos o fizessem.

Agradeço também aos colegas do grupo de pesquisa NR2¹ e CCSC² que compartilharam informações, aflições e realizações nestes anos que se passaram. Desde minha chegada, em 2014, os colegas sempre foram receptivos, e isso sempre me motivou a repassar as gentilezas aos próximos. Foram muitos colegas, entre eles: Danilo Evangelista, Alisson Puska, Robson Gomes, Thiago Druciak, Renan, Rafael Araújo, Otto, Jefferson, Adi, Ricardo Tombesi, Cristian Vegas, Arthur, Mateus Peloso, Carlos Pedroso, Gustavo, Paulo Lenz, Nelson, Bruno Marques, Euclides, Igor Steuck, Danilo Possati, Cainã Passos, Diego, Agnaldo, Lúgia e Fernando.

Deixo meu agradecimento a todos os demais colegas dos Labs vizinhos, em especial, para Renato Melo, Edgar Cabral, Santiago Viertel, Ivan Pires, Alane e Daniel.

Agradeço também aos professores Aldri Santos e a Professora Michele Nogueira. Deixo um agradecimento especial a Profa. Michele, minha orientadora, que considero uma profissional de excelência, em termos de pesquisa, e que tem uma grande sensibilidade em compreender as dificuldades individuais e de motivar em momentos de desespero.

Agradeço a CAPES, UNEMAT, UFPR e IEEE, que fomentaram a realização dos estudos em suas diferentes etapas.

Por fim, agradeço a todos que direta ou indiretamente estiveram presentes durante a realização deste trabalho, em especial meu santo de fé, São Dimas, o protetor dos estudantes.

Meu muito obrigado !

¹Núcleo de Redes Avançadas

²Center for Computational Security sScience

RESUMO

A evolução da Internet tem sido marcada pelo crescente número de dispositivos conectados e pela grande quantidade de dados trafegados. Diante deste crescimento, cabe ressaltar a popularização dos dispositivos móveis inteligentes, como *Smartphones* e *Tablets*, que têm agregado grande poder computacional, capacidade multimídia e múltiplas interfaces de rede sem fio (e.g., WiFi e 3G/4G). Isto tem impulsionado uma crescente demanda por transmissões com Qualidade de Serviço (QoS), como aplicações de realidade virtual, realidade aumentada, videochamadas e outras, que requerem baixa latência e uma alta vazão. As redes de acesso sem fio têm evoluído constantemente para atender a esta demanda. A próxima geração de redes celulares (5G) prevê um aumento na densidade da rede de acesso, incluindo inúmeros pontos de acesso de diferentes tecnologias, a fim de oferecer uma ampla cobertura com alta capacidade de transferência e baixa latência. Neste cenário os dispositivos móveis terão acesso a múltiplas redes sem fio heterogêneas (HetNets) e estarão sujeitos a inúmeras transições entre os pontos de acesso. Isto traz muitos desafios, especialmente para transmissão de dados confiável, feita pelo TCP/IP. Embora seja o protocolo de transporte mais utilizado atualmente, o TCP não suporta a troca de endereços de rede (IP) associados a uma conexão ativa. Neste sentido, a transmissão de dados multicaminhos visa sobrepor as limitações do TCP associando uma conexão a múltiplos endereços. Isto permite alcançar um melhor desempenho, com a agregação da largura de banda de diferentes interfaces sem fio, e aumentar a resiliência das transmissões, através dos múltiplos caminhos entre os dispositivos de origem e destino. No entanto, a heterogeneidade causa o problema do reordenamento que reduz o desempenho da transmissão em termos de latência e vazão e limita os benefícios alcançados com a agregação de banda dos múltiplos caminhos. Atualmente existem várias abordagens que visam lidar com estes desafios. Contudo, uma única solução não é capaz de lidar com os desafios da heterogeneidade e prover a diversidade de caminhos necessária para que a transmissão atenda aos diferentes requisitos de QoS. Neste sentido, este trabalho apresenta a tese de que a diversidade, em termos de soluções, é necessária para lidar com os desafios enfrentados pela transmissão multicaminhos no cenário das HetNets. Assim, propõe-se uma estratégia de resiliência a fim de aumentar o desempenho e melhorar a resiliência das transmissões multicaminhos. A estratégia proposta provê um arcabouço de resiliência cujos módulos são organizados dentro de um ciclo evolutivo e adaptativo, com objetivo de que as medidas de resistência e/ou recuperação se adaptem conforme o desafio e sejam constantemente melhoradas. Para verificar a viabilidade da tese, foram realizados estudos de caso propondo medidas de reconhecimento, resistência e recuperação a serem empregadas pelo arcabouço de resiliência. De um modo geral, os resultados dos estudos de caso demonstram que as medidas empregadas melhoram o desempenho e a resiliência das transmissões multicaminhos, validando a tese definida.

Palavras-chave: transmissão multi-caminhos, resiliência, robustez, confiança.

ABSTRACT

The Internet grown in number of connected devices and amount of data transmitted. In the face of this growth, it is important to highlight the popularization of mobile devices, such as Smartphones and Tablets, which have added great computational power, multimedia capability and multiple wireless network interfaces (e.g. WiFi and 3G/4G). This has driven a increasing demand for Quality of Service (QoS), by delay-sensitive applications, such as virtual reality, augmented reality and video call, which require low latency and high throughput. Wireless access networks constantly evolve to meet this demand. The next generation of cellular networks (5G) foresees an increase in the access network density, including innumerable access points from different technologies to provide broad coverage with high throughput and low latency. In this scenario mobile devices will have access to multiple heterogeneous wireless networks (HetNets) and will be subject to numerous transitions between access points. This brings many challenges, mainly for reliable data transmission by TCP/IP. Although TCP is the most commonly used connection-oriented protocol, it does not support the exchange of network addresses associated with an active connection. In this case, multipath data transmission aims to override the limitations of TCP by associating a connection to multiple addresses. This provides better performance by aggregating the bandwidth of different wireless interfaces and increasing transmission resiliency through multiple paths between the source and destination devices. However, heterogeneity reduces transmission performance in terms of latency and throughput, and it limits the benefits achieved by multipath bandwidth aggregation. There are currently several approaches to address these challenges. However, a single solution is not able to deal with the heterogeneity challenges and provide path diversity necessary for meeting different QoS requirements. In this sense, this work presents the thesis that diversity, in terms of solutions, is needed to address the challenges faced by multipath transmission in HetNets environment. Thus, a resilience strategy is proposed in order to increase the performance and to improve the resilience of multipath transmissions. The strategy provides a resilience framework whose modules are organized within an evolutionary and adaptive cycle, so that resistance and/or recovery measures adapt to the challenge and are constantly improved. To verify the feasibility of this thesis, case studies were carried out proposing recognition, resistance and recovery measures to be employed by the resilience framework. In general, case studies results demonstrate that the measures employed improve the performance and resilience of multipath transmissions, validating the defined thesis.

Keywords: multipath transmission, resilience, robustness, trustworthiness.

Lista de Figuras

1.1	Visão geral dos problemas tratados neste trabalho..	18
1.2	Construção da tese	18
1.3	Relação dos estudos de caso desenvolvidos.	20
2.1	Desafios para as redes de comunicação [Cetinkaya e Sterbenz, 2013]	24
2.2	Detecção de desafios multiestágios [Fry et al., 2010]	25
2.3	Disciplinas de resiliência	26
2.4	Tipos de Transmissão multicaminhos.	31
2.5	Cenários para transmissão multicaminhos	32
2.6	Sequenciamento DSN e SSN no MPTCP.	34
2.7	Estabelecimento de Subfluxos [Ford et al., 2013]	35
2.8	Componentes do MPTCP [Ford et al., 2013].	36
2.9	Associação entre endereços e interfaces de rede	36
2.10	Metodologias <i>Push</i> e <i>Pull</i> [Barré et al., 2011b]	39
2.11	Round-Robin	39
2.12	LRF (menor RTT).	40
3.1	Tipos de desafios e organização das abordagens	43
3.2	Bloqueio HOL	45
3.3	Organização das abordagens	47
3.4	Organização dos Métodos de Reconhecimento.	48
3.5	Organização dos Métodos de Seleção	52
3.6	Número de caminhos vs. número de subfluxos.	55
3.7	Organização das Políticas de Alocação	56
3.8	Processo de envio em ordem	57
3.9	Processo de envio fora de ordem	58
4.1	Requisitos de resiliência	63
4.2	Diversidade de Caminhos.	64
4.3	Componentes da estratégia de resiliência.	66
4.4	Arcabouço de Resiliência.	66
4.5	Funções da base de informações	70
4.6	Posicionamento do gerenciador de contexto	70

5.1	Taxonomia das Abordagens	73
5.2	Cenário de coleta dos dados	77
5.3	Distribuição cumulativa do atraso e da vazão dos caminhos	78
5.4	Distribuição dos valores do coeficiente c	79
5.5	Distribuição dos valores do coeficiente c	80
5.6	Valores de r durante o experimento para o método MPERT	80
5.7	Acurácia dos métodos.	81
6.1	Escalonador RED com MPTCP	85
6.2	Fases do escalonador RED	86
6.3	Esquema de janela deslizante	86
6.4	Coeficiente de Correlação (c_s)	87
6.5	Cenário de Avaliação	89
6.6	Valores de c_s durante um experimento	90
6.7	Acurácia do RED-CM vs. MSCM	91
6.8	CDF do Atraso e Vazão (<i>Throughput</i>)	92
6.9	CNDA vs. CNR	93
6.10	MPTCP Original Scheduler vs. RED: Overhead	93
7.1	Diagrama de Fluxo do CoDel.	96
7.2	Topologia do cenário de simulação	97
7.3	Desempenho da transmissão	100
7.4	Descartes vs. Goodput	100
7.5	Estado da fila com CoDel vs. MPTCP CC	101
7.6	Atraso na fila vs. pilha	101
7.7	Diagrama de Fluxo do LIFO-CoDel	102
7.8	Descartes vs. Goodput	103
7.9	LIFO-CoDEL com LIA.	104
7.10	Estado da fila	104

Lista de Tabelas

3.1	Escalonadores: Métodos, Problemas e Cenários	46
5.1	Vantagens e Desvantagens	76
6.1	Parâmetros de simulação	89
7.1	Parâmetros de Simulação	98

Lista de Acrônimos

3G	<i>Third-generation cellular networks</i> Terceira geração de redes celulares
AAA	<i>Authenticity, Authorisability, Auditability</i> Autenticidade, Autorização e Auditoria
ABW	<i>Available Bandwidth</i> Largura de Banda Disponível
ACK	<i>Acknowledgement</i> Reconhecimento ou Confirmação
AIMD	<i>Additive increase/multiplicative decrease</i> Incremento aditivo/decremento multiplicativo
ANA	<i>Autonomic Network Architecture</i> Arquitetura Autonômica de Rede
API	<i>Application Programming Interface</i> Interface de Programação de Aplicações
AQM	<i>Active Queue Management</i> Gerenciamento Ativo de Filas
BALIA	<i>Balanced Linked Adaptation Congestion Control Algorithm</i> Algoritmo de CC com Adaptação do Balanceamento Acoplado
BER	<i>Bit Error Rate</i> Taxa de Erro de Bit
CMT	<i>Concurrent Multipath Transfer</i> Transferência Multi-caminhos Concorrente
CoDel	<i>Controlled Delay</i> Atraso controlado
CWND	<i>Congestion Window</i> Janela de Congestionamento
D²R²+DR	<i>Defend, Detect, Remediate, Recover more Diagnose, Refine</i> Defender, Detectar, Remediar, Recuperar mais Diagnosticar e Refinar
DDoS	<i>Distributed Denial of Service</i> Negação de Serviço Distribuído
DRR	<i>Deficit Round Robin</i>
DSN	<i>Data Sequence Number</i> Número de Sequência de Dados
DSL	<i>Digital Subscriber Line</i> Linha Digital de Assinante
FIFO	<i>First In, First Out</i> Primeiro que chega, Primeiro que sai
FQ-CoDel	<i>Flow Queue Controlled Delay</i> Atraso Controlado por Fila de Fluxo

HOL	<i>Head-of-line blocking</i> Bloqueio de Cabeça de linha ou Bloqueio da fila de recebimento
IDS	<i>Intrusion Detection System</i> Sistema de Detecção de Intrusão
IETF	<i>Internet Engineering Task Force</i> Força Tarefa ou Grupo de Trabalho para Engenharia da Internet
IGA	<i>Grey Relational Analysis</i> Análise Relacional Cinza
IoT	<i>Internet of Things</i> Internet das Coisas
IP	<i>Internet Protocol</i> Protocolo de Internet
LIA	<i>Linked Increase Algorithm</i> Algoritmo de Incremento Acoplado
LRF	<i>Lowest-RTT-First</i> Primeiro o menor RTT
LTE	<i>Long Term Evolution</i> Evolução de Longo Prazo
MAC	<i>Medium Access Control</i> Controle de Acesso ao meio
MP	<i>Multipath</i> Multi-caminhos
MPLS	<i>Multiprotocol Label Switching</i> Comutação de Rótulos Multiprotocolo
MPTCP	<i>MultiPath TCP</i> TCP Multi-caminhos
NAT	<i>Network Address Translation</i> Tradução de Endereço de Rede
OLIA	<i>Opportunistic Linked Increase Algorithm</i> Algoritmo de Incremento Acoplado Oportunístico
OWD	<i>One-Way Delay</i> Atraso de uma via
PGM	<i>Packet Gap Mode</i> Modo de intervalo de pacotes
PIT	<i>Probe Interval Time</i> Tempo de intervalo de sondagem
PoMo	<i>Postmodern Internet Architecture</i> Arquitetura da Internet pós-moderna
RCVW	<i>Received Window</i> Janela de Recebimento
ResumeNet	<i>Resilience and Survivability for Future Networking</i> Resiliência e Sobrevivência para Redes Futuras
RFC	<i>Request for Comment</i> Requisição para comentário
RFO	Recebimento fora de Ordem
RMT	<i>Resilient Multipath Transfer</i>

	Transmissão Multicaminhos Resiliente
RR	<i>Round-Robin</i>
RSRQ	<i>Reference Signal Received Quality</i> Sinal de Referência da Qualidade Recebida
RTO	<i>Retransmission Timeout</i> Tempo Limite de Retransmissão
RTT	<i>Round Trip Time</i> Tempo de ida e volta
SACK	<i>Selective Acknowledgments</i> Reconhecimento Seletivo
SCTP	<i>Stream Control Transmission Protocol</i> Protocolo de Transmissão de Controle de Fluxo
SDN	<i>Software Defined Networking</i> Rede Definida por Software
SNDQ	<i>Send Queue</i> Fila de Envio
SP	<i>Single Path</i> Caminho Único
SPTCP	<i>Single Path TCP</i> TCP de Caminho Único
SSN	<i>Subflow Sequence Number</i> Número de Sequência de Sub-fluxo
TCP	<i>Transmission Control Protocol</i> Protocolo de Controle de Transmissão
TFRC	<i>TCP-Friendly Rate Control</i> Controle de Taxa amigável com TCP
UDP	<i>User Datagram Protocol</i> Protocolo de Datagrama do usuário
UE	<i>User Equipment</i> Equipamento do Usuário
WiFi	<i>Wireless Fidelity</i> Fidelidade Wireless
wVegas	<i>Delay-based Congestion Control for MPTCP</i> Controle de Congestionamento baseado em atraso para o MPTCP
XCP	<i>Explicit Congestion Control</i> Controle de Congestionamento Explícito

Sumário

1	Introdução	16
1.1	Problema	17
1.2	Objetivos	18
1.3	Contribuições	19
1.4	Organização do texto	20
2	Resiliência e Transmissão Multicaminhos	22
2.1	Termos e Definições	22
2.2	Resiliência em Redes de Comunicação	23
2.2.1	Desafios	24
2.2.2	Disciplinas	25
2.2.3	Estratégias de Resiliência	27
2.2.4	Princípios para Construção de Redes Resilientes	28
2.3	Transmissão Multicaminhos	31
2.4	Protocolo de Transmissão Multicaminhos	33
2.4.1	MPTCP	34
2.4.2	Mecanismos Resilientes do MPTCP	40
2.5	Resumo	41
3	Principais Desafios para a Transmissão Multicaminhos	42
3.1	Contextualização	42
3.2	Desafios	42
3.2.1	Heterogeneidade	43
3.2.2	Falta de Diversidade	43
3.2.3	Equidade	44
3.2.4	Reordenamento	44
3.2.5	Consumo de Energia	45
3.2.6	Segurança	45
3.3	Classificação das Abordagens	46
3.3.1	Métodos de Reconhecimento	47
3.3.2	Métodos de Seleção	52
3.3.3	Políticas de Alocação	55
3.4	Resumo	61

4	Uma Estratégia de Resiliência para a Transmissão Multicaminhos	62
4.1	Contextualização e Trabalhos Relacionados	62
4.2	Requisitos	63
4.2.1	Restrições e Complexidade	63
4.2.2	Diversidade	63
4.2.3	Redundância	64
4.2.4	Conectividade e Continuidade	64
4.2.5	Comportamento	65
4.3	Estratégia de resiliência	65
4.3.1	Arcabouço de Resiliência	65
4.3.2	Base de Informações	69
4.3.3	Gerenciador de Contexto	70
4.4	Resumo	70
5	Análise da Diversidade em Transmissões Multicaminhos	72
5.1	Contextualização	72
5.2	Taxonomia.	73
5.2.1	Intracamada	73
5.2.2	Intercamadas	75
5.2.3	Vantagens e Desvantagens	75
5.3	Avaliação	76
5.4	Importância dos resultados para a tese	82
5.5	Resumo	82
6	Escalonador Redundante com Seleção de Caminhos Diversos	83
6.1	Contextualização	83
6.2	Trabalhos Relacionados.	84
6.3	Escalonador RED	85
6.3.1	Coleta dos Dados	86
6.3.2	Determinando a Correlação dos Caminhos.	87
6.3.3	Alocação dos Pacotes	88
6.4	Avaliação	88
6.5	Resultados.	90
6.6	Importância dos resultados para a tese	94
6.7	Resumo	94
7	Resistindo ao <i>Bufferbloat</i> em Caminhos Compartilhados	95
7.1	Contextualização Trabalhos Relacionados	95
7.2	Cenário de Avaliação	97
7.3	Impacto do <i>bufferbloat</i> na Transmissão multicaminhos	99

7.4	Mitigando o problema em nível de rede	101
7.5	Resultados.	103
7.6	Importância dos resultados para a tese	105
7.7	Resumo	105
8	Conclusões	106
8.1	Problemas tratados e resultados alcançados	106
8.2	Questões em Aberto e Direções Futuras	107
8.3	Publicações	108
	REFERÊNCIAS.	110

1 Introdução

A Internet se popularizou no início da década de 90, expandindo-se para todos os continentes e alcançando 1 milhão de dispositivos e 1 trilhão de Bytes transferidos por mês [Zakon, 1997]. Atualmente, conforme relatório da Cisco [Cisco, 2017a] atualizado em 2017, o tráfego global de dados passou de 11 Exabytes por mês e o número de dispositivos ultrapassou a marca de 8 bilhões. A previsão para 2021, conforme o relatório, é que o tráfego de dados seja de 49 Exabytes por mês e o número de dispositivos superior a 10 bilhões. Diante deste crescimento cabe ressaltar a evolução dos dispositivos móveis inteligentes, como celulares e Tablets, que possuem múltiplas interfaces de rede sem fio (e.g., WiFi e 3G/4G), grande poder computacional e capacidade multimídia. Em 2016, estes dispositivos superaram o número de computadores conectados à Internet e atualmente correspondem a mais de 55% do total de dispositivos, e em 2021 este número passará de 75%.

Em paralelo a este crescente número de dispositivos inteligentes, especialmente os dispositivos com capacidade multimídia, aumenta a demanda na rede por Qualidade de Serviço (QoS), ou seja, por baixa latência e alta vazão. Novas aplicações, como as de realidade virtual (VR, do Inglês *Virtual Reality*) e realidade aumentada (AR, do Inglês, *Augmented Reality*), utilizadas em diferentes setores, como jogos, medicina, educação, turismo, e transmissões de vídeo ao vivo, são exemplos desta demanda [Cisco, 2017b]. Em relação a esta última, a previsão é que em 2021 ela corresponda a 13% do tráfego de vídeo na Internet. Do mais, cabe ressaltar que atualmente o tráfego de vídeo corresponde a 60% do tráfego dos dispositivos móveis e esse valor está projetado para 78% em 2021 [Cisco, 2017a].

Para atender a este grande número de dispositivos, o grande volume de tráfego, e os requisitos das aplicações e serviços, as redes de acesso estão em constante evolução. Ainda segundo o relatório da Cisco, em 2021, a quinta geração de redes celulares (5G) irá superar o tráfego de dados dos dispositivos que ainda usam a segunda geração (2G), sendo respectivamente 1,5% contra 0,2%, enquanto a maioria dos dispositivos usarão as redes da quarta geração. No entanto, a projeção é que a maior parte do tráfego descarregado por estes dispositivos seja através das redes WiFi, em residências e locais públicos, ou pontos de acessos de curto alcance, disponibilizadas pelas operadoras de rede celular. A estimativa é que estas redes sejam responsáveis por descarregar 63% do tráfego.

As redes de quinta geração (5G) irão trazer grandes avanços para as redes móveis, cujos objetivos são prover uma cobertura mais ampla, uma alta taxa de transferência e uma baixa latência, dando suporte ao grande volume de tráfego e a grande quantidade de dispositivos móveis [Song et al., 2017]. Para alcançar estes requisitos, as redes 5G resultam no aumento da densidade da rede, incluindo um grande número de pequenas células [Ge et al., 2016]. Um exemplo de tecnologia associada às pequenas células é a comunicação por ondas milimétricas (do Inglês, *Millimeter-wave Communications*). Essa tecnologia possui restrição de 100 metros para difusão do sinal, o que exige um grande número de pontos de acesso para prover uma ampla cobertura. As pequenas células remetem a uma característica comum das redes 5G que se refere à adoção de múltiplas tecnologias, provendo o acesso por redes sem fio heterogêneas (HetNets).

O cenário das redes 5G oferece grandes desafios em relação à mobilidade dos nós, principalmente para os protocolos de transporte, como o TCP (do Inglês, *Transmission Control Protocol*). Neste cenário, dado o grande número de pontos de acesso, os dispositivos móveis estão sujeitos a inúmeras transições entre redes de diferentes tecnologias e domínios [Ge et al., 2016]. O TCP não suporta a transição entre redes heterogêneas com troca de endereços IP (do Inglês, *Internet Protocol*), como de uma rede WiFi para uma rede celular (3G/4G/LTE). Isto porque cada conexão se associa com apenas um endereço IP do dispositivo. Neste caso, a transição interrompe todas as conexões TCP ativas. Isto prejudica seriamente a transmissão de dados, especialmente para as aplicações sensíveis a atraso (i.e., aplicações de tempo real, transmissões de voz e vídeo ao vivo) que são tendência neste cenário. Mesmo que a rede empregue mecanismos ou protocolos para dar suporte à mobilidade do nó e o processo de transição fique transparente, o TCP ainda tem problemas. A diferença de atraso entre as redes faz com que os pacotes sejam recebidos fora de ordem no destinatário. Isto reduz o desempenho do TCP [Song et al., 2017] porque provoca um aumento na latência e uma redução da vazão em função do atraso no recebimento dos pacotes que preenchem a sequência correta dos dados para sua subsequente entrega ao receptor.

A transmissão de dados multicaminhos, em nível de camada de transporte, associa uma conexão a múltiplos endereços IP. Com isso, a transmissão multicaminhos permite associar uma conexão com múltiplas interfaces de rede (i.e., endereços IP) e, conseqüentemente, com diferentes redes de acesso [Domżał et al., 2015]. O acesso a múltiplas redes permite que os dados sejam intrinsecamente transmitidos por múltiplos caminhos entre os dispositivo de origem e destino. Isto provê uma maior capacidade de transmissão, com a agregação da largura de banda, e uma maior resiliência, com o uso de múltiplos caminhos. Atualmente existem dois protocolos de transporte multicaminhos padronizados pelo IETF (*Internet Engineering Task Force*)¹, o SCTP (*Stream Control Transmission Protocol*) [Stewart, 2007] e o MPTCP (*MultiPath TCP*) [Ford et al., 2011].

Os avanços dos dispositivos móveis e das redes de acesso sem fio têm impulsionado a transmissão de dados multicaminhos. Isto pode ser observado pelos dados apresentados no início deste capítulo e pelo interesse de grandes empresas do setor de telefonia móvel. Cabe ressaltar o caso específico da Apple, que é uma das maiores fabricantes de Smartphones atualmente, a qual emprega a transmissão multicaminhos desde a versão 7 do iOS, seu sistema operacional usado em dispositivos móveis. Recentemente, a empresa disponibilizou aos desenvolvedores de aplicativos a API que provê acesso à transmissão multicaminhos. O objetivo é que eles possam utilizar as múltiplas interfaces de rádio disponíveis nos dispositivos com iOS a fim de melhorar a confiabilidade e o desempenho das transmissões [app, 2018].

1.1 Problema

Esta tese aborda dois grandes problemas enfrentados pelas transmissões multicaminhos no cenário das HetNets, a redução do desempenho e a falta de resiliência, como ilustrado na Figura 1.1. Estes problemas resultam da mobilidade dos nós e da heterogeneidade da rede. A mobilidade causa desconexões dos caminhos, quando ocorre a troca de endereço IP, e variações no atraso, quando a rede se encarrega de fazer a transição sem a troca de endereço IP. Por outro lado, o acesso simultâneo a redes de tecnologias diferentes faz com que os caminhos fim-a-fim entre os dispositivos envolvidos na conexão sejam heterogêneos, apresentando diferentes valores de atraso, taxa de perda, largura de banda, e outros. A mobilidade e a heterogeneidade têm

¹O IETF corresponde a uma força tarefa envolvendo técnicos, fabricantes, pesquisadores, e outros profissionais, com objetivo de identificar e propor padrões para tecnologias e protocolos da Internet - <https://www.ietf.org>.

grande impacto na entrega ordenada dos dados, que é uma das características dos protocolos de transporte confiáveis, como o TCP e o MPTCP. Estes protocolos fazem o sequenciamento dos dados recebidos da camada de aplicação para realizar a transmissão confiável e a entrega ordenada ao destinatário. No entanto, os caminhos heterogêneos e dinâmicos aumentam o recebimento de pacotes fora de ordem. O recebimento fora de ordem é conhecido como problema do reordenamento [Yedugundla et al., 2016]. Este problema ocasiona um aumento da latência e uma redução na vazão das transmissões afetando as aplicações sensíveis ao atraso.



Figura 1.1: Visão geral dos problemas tratados neste trabalho.

Além da mobilidade e da heterogeneidade, outro aspecto que afeta diretamente o desempenho e a resiliência das transmissões multicaminhos são os caminhos compartilhados, que embora não sejam específicos das HetNets, são mais difíceis de serem detectados devido as características dinâmicas neste ambiente. Uma transmissão por caminhos compartilhados está sujeita a falhas correlacionadas e gargalos que limitam os benefícios alcançados com a agregação de banda dos múltiplos caminhos. Na transmissão multicaminhos, os dispositivos não podem tomar decisões sobre um caminho individualmente, isto é função dos protocolos de roteamento. No entanto, eles têm a capacidade de decidir qual caminho utilizar para transmitir os dados. Esta decisão faz parte de um dos principais componentes dos protocolos de transmissão multicaminhos, o escalonador de pacotes, que tem como função principal alocar os pacotes de dados sobre os caminhos disponíveis conforme a política adotada.

1.2 Objetivos

Este trabalho tem como objetivo prover uma nova abordagem para aumentar a resiliência das transmissões de dados no cenário de HetNets. Para alcançar este objetivo, defende-se a hipótese

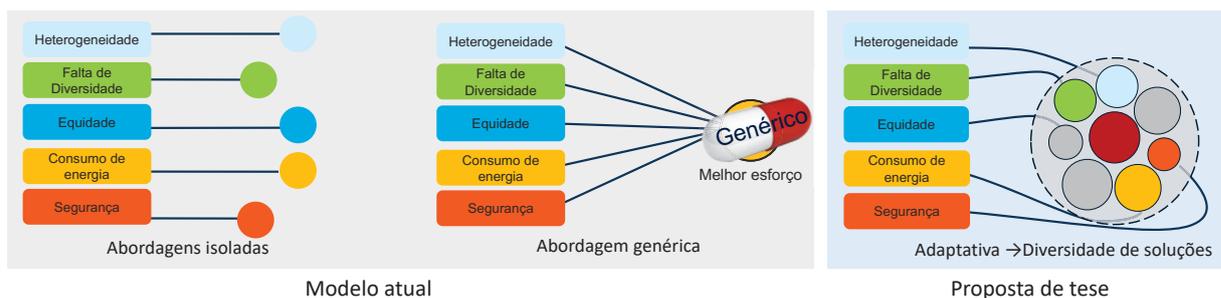


Figura 1.2: Construção da tese

de que *com diversidade é possível prover um melhor desempenho e uma maior resiliência às transmissões multicaminhos*. Esta hipótese tem como base a premissa de que uma única abordagem não é capaz de resolver todos os desafios enfrentados pela transmissão e foi inspirada no trabalho de inteligências múltiplas, do psicólogo Howard Gardner (1996) [Brualdi Timmins, 1996]. O conceito de inteligências múltiplas refuta o conceito de inteligência tradicionalmente definido até aquele momento, com a hipótese de que um único método (e.g., teste de Quociente de Inteligência) não era capaz de descrever a variedade de habilidades cognitivas humanas. Deste modo, a teoria defende que cada indivíduo possui um conjunto diferente de inteligências. Neste sentido, problemas complexos, que dependem de diferentes habilidades (inteligências) e competências, requerem o trabalho coordenado entre diferentes indivíduos. Traduzindo isto em um rápido exemplo do cotidiano, como a construção de um edifício, vê-se que desde o projeto até a finalização da construção, diferentes profissionais, incluindo arquitetos, engenheiros, pedreiros, pintores, eletricitas, decoradores e outros, são envolvidos e trabalham em uma ou mais etapas do processo. Neste exemplo, a diversidade de inteligências está relacionada aos diferentes profissionais que precisam trabalhar coletivamente para finalizar o projeto.

No contexto da transmissão multicaminhos, existem diferentes desafios que afetam o seu desempenho e sua resiliência. Conforme ilustra a Figura 1.2, o modelo atualmente empregado pelos protocolos de transporte aplicam abordagens isoladas para cada desafio conhecido, ou, uma abordagem genérica do tipo melhor esforço, não importando o desafio enfrentado. Esta proposta desta tese defende que é necessário prover uma diversidade de soluções, empregando a abordagem adequada ao desafio enfrentado. A fim de prover esta diversidade propõe-se uma estratégia de resiliência que tem como objetivo agregar múltiplas abordagens a fim de prover medidas de resistência e de recuperação para os diferentes desafios da transmissão multicaminhos no cenário das HetNets. A estratégia envolve um arcabouço de gerenciamento de resiliência, um gerenciador de contexto e uma base de conhecimento. O arcabouço visa garantir o desempenho e a resiliência das transmissões. O gerenciador de contexto tem como função refletir as políticas definidas nos protocolos e serviços existentes nas diferentes camadas. A base de informação é responsável por guardar os modelos de desafios, métricas, políticas empregadas pelo arcabouço de resiliência, histórico das medidas empregadas, entre outras informações. Defende-se que esta estratégia é essencial para prover múltiplas soluções para os desafios de transmissão em HetNet a fim de prover uma maior resiliência e um melhor desempenho do que é alcançado atualmente.

Para verificar a hipótese levantada acima, este trabalho realizou estudos de caso que demonstram como diferentes desafios impactam na resiliência das transmissões. O objetivo dos estudos de casos é demonstrar a viabilidade da estratégia de gerenciamento de resiliência.

1.3 Contribuições

Este trabalho apresenta inicialmente como contribuição uma estratégia de resiliência para a camada de transporte, especialmente designada para o contexto das transmissões de dados multicaminhos. A partir desta estratégia, foram desenvolvidos diferentes estudos de caso que resultaram em diferentes contribuições para melhorar o desempenho e a resiliência das transmissões multicaminhos. As contribuições estão associadas com o arcabouço de gerenciamento de resiliência, que faz parte da estratégia de resiliência proposta. Conforme ilustra a Figura 1.3, as os estudos de caso estão associados aos módulos de reconhecimento e de resistência do arcabouço. Em relação ao módulo de reconhecimento, foi realizada uma análise das abordagens que tratam o problema da falta de diversidade em transmissões multicaminhos. A partir dessa análise, realizou-se um estudo de caso em que se avaliou um método de correlação suscetível às dinâmicas do cenário das HetNets. O método foi implementado em conjunto

com um escalonador redundante, com objetivo de prover uma transmissão redundante e com diversidade de caminhos, como um dos requisitos para aumentar a resiliência da comunicação.

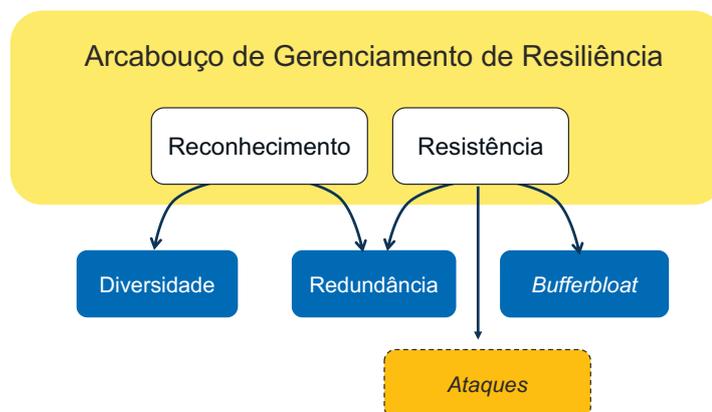


Figura 1.3: Relação dos estudos de caso desenvolvidos

As contribuições relacionadas ao módulo de resistência referem-se às contramedidas que podem ser aplicadas para que a transmissão possa resistir aos desafios impostos pelo cenário das HetNets. Essas contramedidas incluem: um estudo de caso envolvendo a redundância e diversidade de caminhos, o qual refere-se à proposta de um escalonador que tem como objetivo equalizar as diferenças entre os caminhos e prover uma maior resiliência; dois estudos de caso desenvolvidos de forma colaborativa, um sobre o impacto do ataque de negação de serviço LDoS em que foi proposta uma medida de resistência a fim de mitigar os seus efeitos quando afeta caminhos compartilhados, publicado em [Felix et al., 2018b], e outro sobre como a transmissão multicaminhos pode prover comunicação segura sobre a Internet em casos de ataques do homem do meio (do Inglês, *Man-in-the-middle*), publicado em [Nguyen et al., 2017]; um estudo de caso, realizado preliminarmente, sobre o desempenho da transmissão quando os caminhos compartilham o fenômeno do *Bufferbloat*. Este fenômeno refere-se a alta latência em redes com comutação de pacotes causada pelo excesso de enfileiramento de pacotes ². A fim de mitigar os efeitos do *Bufferbloat*, foi proposta uma nova disciplina de controle ativo de filas. Os resultados dos estudos de caso demonstram que as medidas empregadas melhoram o desempenho, em termos de redução da latência e aumento da vazão, e aumentam a resiliência das transmissões multicaminhos. Além das contribuições principais, esta tese apresenta uma revisão da literatura sobre os desafios da transmissão multicaminhos e uma descrição detalhada das abordagens empregadas para lidar com esses desafios.

1.4 Organização do texto

Este manuscrito está organizado da seguinte forma. O Capítulo 2 apresenta os fundamentos relacionados à resiliência para redes de comunicação e à transmissão de dados multicaminhos, que são necessários para a compreensão do problema tratado. Em seguida, o Capítulo 3 apresenta uma revisão do estado da arte sobre as estratégias, em nível de protocolo de transporte, empregadas para lidar com os desafios da transmissão multicaminhos. O Capítulo

²Alguns fabricantes de equipamentos de comunicação colocaram *buffers* excessivamente grandes em alguns de seus produtos de rede. Com isso, o *bufferbloat* ocorre quando um enlace de rede fica congestionado, fazendo com que os pacotes fiquem enfileirados no *buffer* durante muito tempo. Em uma fila FIFO (do Inglês, First In, First Out), os *buffers* excessivamente grandes resultam em filas mais longas, com maior latência. Disponível em <https://pt.wikipedia.org/wiki/Bufferbloat>.

4 descreve a estratégia de resiliência para a transmissão de dados multicaminhos. O Capítulo 5 apresenta um estudo acerca das abordagens empregadas para prover diversidade à transmissão de dados multicaminhos. O Capítulo 6 apresenta um estudo de caso sobre a redundância e diversidade nas transmissões multicaminhos. O Capítulo 7 apresenta um estudo de caso sobre o desempenho das transmissões com caminho que compartilham os efeitos do fenômeno do Bufferbloat e apresenta uma nova disciplina de fila para mitigar o problema. Por fim, o Capítulo 8 conclui o trabalho e descreve as questões em aberto e os trabalhos futuros.

2 Resiliência e Transmissão Multicaminhos

Este capítulo apresenta os fundamentos sobre os princípios de resiliência para redes de comunicação e uma visão geral sobre as transmissões de dados multicaminhos relacionados ao tema da pesquisa. O capítulo está organizado da seguinte forma. A Seção 2.1 apresenta os termos e as definições empregados neste capítulo e no restante do trabalho. A Seção 2.2 apresenta os princípios e as definições de resiliência para redes de comunicação, abordando os desafios, as disciplinas e as estratégias de resiliência com base na literatura. A Seção 2.3 apresenta uma visão geral sobre a transmissão multicaminhos, os tipos de transmissão, os cenários e os principais protocolos. Esta seção ainda descreve com mais detalhes a arquitetura e os modos de operação do protocolo TCP Multicaminhos (MPTCP). A Seção 2.5 resume o capítulo.

2.1 Termos e Definições

As terminologias consideradas neste trabalho estão em conformidade com as RFCs 6182 [Ford et al., 2011], 6356 [Raiciu et al., 2011b], 6824 [Ford et al., 2013] e 6897 [Scharf e Ford, 2013]. Com base nestas RFCs, um **caminho** representa uma sequência de ligações entre um emissor e um receptor, que conforme o protocolo da camada de transporte, pode ser identificado por uma tupla de 4 elementos composta pelos endereços IP e portas de origem e destino. Um **subfluxo** é compreendido como um fluxo de segmentos TCP, operando sobre um caminho e faz parte de uma única conexão. Uma **conexão** multicaminhos é representada por um conjunto de um ou mais subfluxos. O termo *nível de dados* é usado de forma similar ao *nível de conexão*, em contraste com nível de subfluxo, que se refere de forma específica a propriedades de um subfluxo.

A seguir apresenta-se um conjunto de definições formais com objetivo de equalizar as diferenças entre os símbolos e termos matemáticos empregados neste documento e também nos trabalhos relacionados. Os termos e definições têm como base os modelos formais apresentados em [Peng et al., 2014, Becke et al., 2012].

Definição 2.1.1. Uma rede é definida como $\mathcal{R} = (N, L)$, sendo:

- N - o conjunto de nós;
- L - o conjunto de enlaces.

Definição 2.1.2. Um caminho na rede refere-se a um conjunto de enlaces e nós entre a ligação fim-a-fim de um nó de origem N_s e um nó de destino N_d [Rohrer et al., 2014]

Definição 2.1.3. Um *fluxo* é associado aos pacotes de dados transmitidos sequencialmente através de uma única conexão.

Definição 2.1.4. Um *subfluxo* refere-se a uma subdivisão do fluxo de dados de uma conexão multicaminhos. Em uma transmissão multicaminhos, o fluxo de dados da conexão pode ser subdividido em x subfluxos, sendo $1 \leq x \leq |P_{ij}|$.

Definição 2.1.5. Um **ciclo de escalonamento** se refere ao processo de distribuição de pacotes para os caminhos disponíveis. Um ciclo pode ser controlado pelo número de pacotes disponíveis (fila de envio), pela disponibilidade de envio dos caminhos e pela política de escalonamento.

Definição 2.1.6. O RTT (do Inglês, *Round Trip Time*) refere-se ao tempo transcorrido desde o envio do segmento pelo subfluxo até o recebimento de um reconhecimento para o mesmo. Os protocolos baseados no TCP seguem a RFC¹ 6298 [Paxson et al., 2011] como base para calcular o RTT e o tempo limite de retransmissão (RTO, do Inglês *Retransmission Timeout*).

Com base na RFC 6298, o RTT é estimado a cada segmento enviado e pode mudar constantemente conforme características do caminho. Devido a estas flutuações, que são comuns na Internet, a maioria das implementações utilizam um RTT suavizado, denotado $SRTT$, que se trata de uma média ponderada do RTT e do $SRTT$ anterior. No início de cada transmissão o RTT não recebe nenhum valor. Ao receber a primeira medição, o valor do RTT suavizado passa a ser $SRTT = RTT$. Então, cada medição é iterada pela seguinte equação $(1 - \alpha)SRTT + \alpha \times RTT$. O valor de α depende da implementação, mas normalmente é fixado em 1/8. Isto significa que o novo $SRTT$ é calculado com 7/8 do $SRTT$ antigo e 1/8 do RTT atual.

Além do RTT e $SRTT$, muitas implementações calculam o desvio do RTT (RTT_D). O cálculo do RTT_D utiliza os valores anteriores dos RTTs, onde no início da transmissão não se atribui valor. Após a primeira medição, atribui-se $RTT_D = RTT/2$. Em seguida, após cada medição, tem-se $RTT_D = (1 - \beta)RTT_D + \beta \times |SRTT - RTT|$. Similar ao valor de α , o cálculo do valor de β depende da implementação, contudo, comumente utiliza-se 1/4. O valor do RTO é dado pelos valores do $SRTT$ e RTT_D . No entanto, ele possui um valor inicial e comumente emprega-se 3 [Paxson et al., 2011]. A cada medição do RTT, o RTO é atualizado conforme equação: $RTO = SRTT + 4 \times RTT_D$.

2.2 Resiliência em Redes de Comunicação

As falhas nos elementos de rede ou nos enlaces são inevitáveis. Elas podem ocorrer em face a um grande número de **desafios** [Fry et al., 2010]. Um desafio é uma condição ou evento adverso que impacta na operação normal da rede [Cetinkaya e Sterbenz, 2013], como desastres naturais, erros humanos, ataques maliciosos e outros. A resiliência é a habilidade da rede em manter um nível aceitável de serviço na presença de tais desafios. Ela tem se tornado uma propriedade das redes de comunicação, abrangendo diversas áreas temáticas, como segurança de rede e da informação, tolerância a falhas, dependabilidade² de software e sobrevivência de redes. Segundo [Smith et al., 2011], muitos estudos têm sido realizados abrangendo estes diferentes temas, tendo como foco mecanismos específicos para resiliência e um determinado subconjunto de desafios. Isto tem produzido abordagens com uma visão parcial, ou seja, não cobrem aspectos correlacionados com diferentes desafios. Neste sentido, uma estratégia de resiliência deve considerar uma abordagem sistemática, ou seja, abrangente a diferentes desafios.

Antes de apresentar os detalhes de uma abordagem sistemática para prover resiliência de uma rede, são apresentados os principais desafios e possíveis impactos que podem ocasionar um

¹RFC (do Inglês, *Request for Comments*) se refere a um documento de padrões, disponibilizado pelo IETF, responsável pela manutenção dos padrões da Internet.

²Um roteiro para exploração dos conceitos básicos de tolerância a falhas. <http://www.inf.ufrgs.br/taisy/disciplinas/textos/Dependabilidade.pdf>.

desvio na operação normal da rede ou de um serviço. Além disso, são apresentadas as principais disciplinas que servem de base para a resiliência das redes. Conhecer estes desafios e o impacto que eles causam são essenciais para definir os requisitos de um protocolo ou arquitetura de uma rede resiliente [Sterbenz et al., 2010].

2.2.1 Desafios

Um desafio incorre em falhas que podem incorrer em uma interrupção do serviço de transporte e afetar a operação normal da rede [Cetinkaya e Sterbenz, 2013]. A Figura 2.1 ilustra os principais desafios para as redes de comunicação.

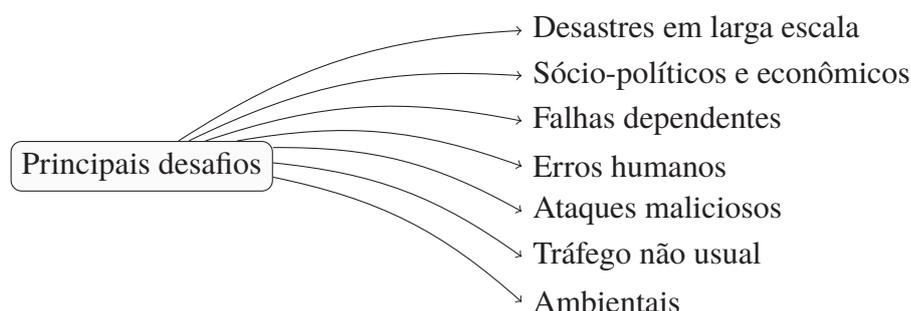


Figura 2.1: Desafios para as redes de comunicação [Cetinkaya e Sterbenz, 2013]

Os **desastres em larga escala** podem ser causados por forças naturais (desastres naturais), como terremotos, tsunamis, furacões, erupções vulcânicas, e outros. Eles podem causar interrupções significantes na infraestrutura física da rede, afetando dispositivos e enlaces de comunicação [Mauthe et al., 2016]. Os desafios **sócio-políticos e econômicos** incluem as consequências de um ato de terrorismo ou decisões políticas e econômicas deliberadas com o intuito de interromper a operação normal da rede. As **falhas dependentes** ou em cascata ocorrem quando um sistema depende de outro, como por exemplo, uma falha na rede de energia pode afetar a infraestrutura da Internet. Os **erros humanos**, em consequência de ações não maliciosas, também são um grande desafio para as redes e podem causar defeitos catastróficos. Outro grupo de desafios se refere aos **ataques maliciosos**, que deliberadamente tentam interromper o funcionamento de um serviço provido por um hardware ou software. O impacto de um ataque malicioso pode alcançar uma extensão global. O ataque DDoS é um exemplo. Ele tem ocorrido com frequência e tem como objetivo gerar um grande volume de tráfego direcionado a um serviço de rede para causar o mau funcionamento ou queda do mesmo³. Além dos ataques, que geram grandes volumes de tráfego na rede, grandes eventos ou eventos que têm um impacto social muito grande, como um jogo de futebol ou um ato terrorista, podem gerar uma grande quantidade de **tráfego não usual**. Se o volume for maior que a capacidade do sistema, isto pode ocasionar a interrupção parcial ou total dos serviços. E por fim, os desafios **ambientais** são aqueles inerentes ao ambiente de comunicação. São exemplos deste tipo de desafio a alta mobilidade dos nós e atrasos inesperados nos caminhos relacionados a eventos de conectividade.

A identificação imediata destes desafios, especialmente quando compartilham sintomas em comum, é uma tarefa difícil. Como por exemplo, o aumento repentino no volume de tráfego na rede pode estar associado a um ataque DDoS ou a um evento social de grandes proporções. Detectar de maneira apropriada o desafio é uma tarefa importante para que se possa determinar as estratégias de resiliência a serem aplicadas. Alguns desafios podem ser detectados

³<http://www.digitalattackmap.com>

utilizando mecanismos inseridos nas redes, como a detecção baseada em assinaturas de ataques conhecidos ou detecção baseada em comportamentos de eventos não usuais. Além disso, existem mecanismos que estão fora da rede, como sistemas que monitoram a previsão do tempo e indicam a possibilidade de tempestades que podem interromper o funcionamento da rede ou de um serviço. Neste caso, estes dados podem ser utilizados para migrar os dados para caminhos alternativos.

Em [Fry et al., 2010], os autores propõem uma abordagem multiestágios para detecção de desafios. Conforme ilustra a Figura 2.2, a abordagem inclui a detecção dos sintomas, a identificação da causa principal e a determinação do seu potencial de impacto. Este último implica em uma abordagem prática de gestão de riscos, que avalia o impacto do desafio contra os custos para remediar.

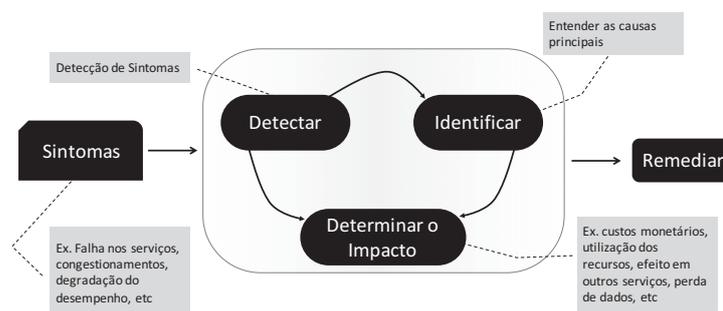


Figura 2.2: Detecção de desafios multiestágios [Fry et al., 2010]

A abordagem de detecção multiestágios pode ser aplicada em um cenário de redes móveis para identificar um determinado desafio e suas causas [Fry et al., 2010]. Neste cenário, uma falha na camada física (ex. perda de sinal) ocasionada em decorrência da mobilidade pode ser identificada pela degradação percebida do sinal, como por exemplo, através do aumento da taxa de erro de *bit* (BER, do Inglês *Bit Error Rate*) ou pela deterioração da qualidade de serviço, indicada pela redução da vazão ou aumento do atraso. Neste caso, após a detecção, é importante localizar o ponto de falha a fim de distribuir as notificações necessárias para remediar os efeitos negativos no desempenho da rede. O retorno para o estado de operação normal pode ser novamente alcançado se as causas principais da falha forem eliminadas.

A próxima seção apresenta uma descrição sobre as estratégias empregadas para remediar o impacto dos desafios, sendo organizadas conforme classificação sistemática apresentada em [Sterbenz et al., 2010].

2.2.2 Disciplinas

As disciplinas de resiliência referem-se a abordagens de projetos de redes para prover continuidade dos serviços. As disciplinas podem ser divididas em duas categorias, ilustradas pela Figura 2.3. A categoria **tolerância de desafios** lida com o projeto de sistemas que continuam a prover seus serviços mesmo em face dos desafios. E a categoria **confiança** descreve as características mensuráveis do sistema de comunicação analisado. A relação entre estas duas categorias se refere à **robustez**, que é um indicador de desempenho de uma rede em condições de ameaça. A categoria tolerância a desafios pode ser decomposta em *sobrevivência*, que inclui *tolerância a falhas*, e está ligada a desafios que afetam a infraestrutura das redes de comunicação; *tolerância à interrupção*, para caminhos resistentes a falhas e desconexões; e *tolerância de tráfego*, para desafios relacionados ao tráfego. A sobrevivência possui diferentes definições na literatura.

Deste modo se considera uma definição simples porém abrangente: uma rede sobrevivente é aquela que tem capacidade de continuar o serviço na presença de desafios [Rak, 2015].

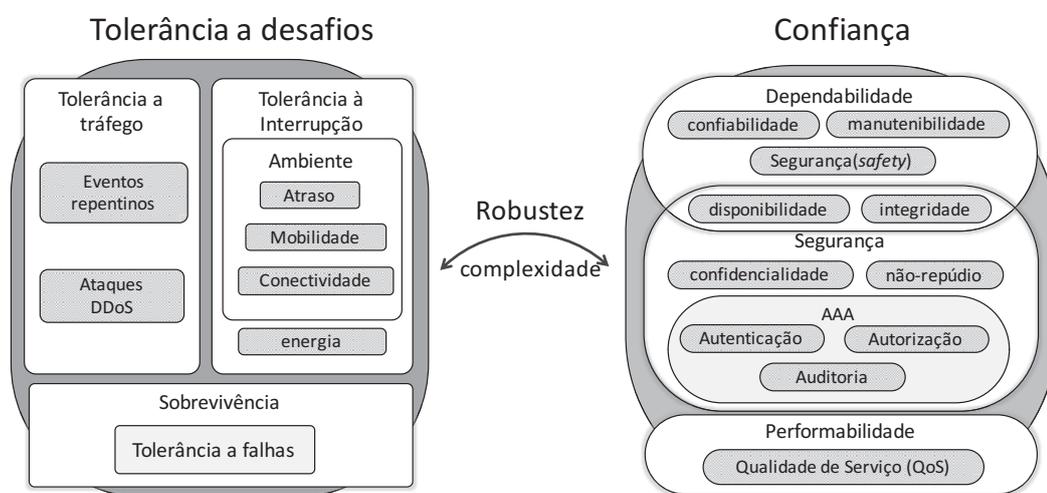


Figura 2.3: Disciplinas de resiliência

A **sobrevivência** tem um escopo mais amplo que a tolerância a falhas. Ela compreende falhas correlacionadas com eventos que estão fora do limite da rede, como aquelas relacionadas a atividades humanas maliciosas ou não maliciosas, mas que afetam parte da infraestrutura das redes de comunicação. Comparando com a tolerância a falhas, além da redundância necessária para a recuperação de um serviço, a sobrevivência também requer diversidade, a fim de que uma mesma falha não afete múltiplos elementos de uma rede de comunicação. A quantificação da sobrevivência de uma rede em relação à tolerância a falhas também é mais complexa. A **tolerância a falhas** pode ser compreendida como a capacidade da rede lidar com falhas de modo a não causar um desvio no funcionamento normal do serviço. A redundância é utilizada para compensar as falhas aleatórias e não correlacionadas. Contudo, a tolerância a falhas não provê meios suficientes para a recuperação da rede após múltiplas falhas correlacionadas, sendo considerada um subconjunto da sobrevivência.

A **tolerância à interrupção** pode ser definida como a habilidade de um sistema tolerar interrupções de conectividade entre seus componentes. As características dos caminhos de comunicação são utilizadas para avaliar a conectividade. Neste caso, os desafios ambientais, como mobilidade, longos atrasos, e limitações de energia podem afetar a conectividade. A **tolerância de tráfego** se refere à habilidade da rede ou serviço tolerar uma carga de tráfego não esperada. Exemplos comuns deste tipo de tráfego não esperado incluem ataques DDoS e atividades relacionadas a eventos sociais de grandes proporções, citados anteriormente.

A **confiança**, conforme ilustra a Figura 2.3, pode ser definida em termos de características mensuráveis acerca da prestação de serviços, como a garantia de que o sistema funcionará como esperado. Ela compreende três disciplinas: a dependabilidade; a segurança; e a performabilidade. A **dependabilidade** é utilizada para quantificar o nível de segurança de um serviço. Ainda fazem parte desta disciplina a **confiabilidade** e a **disponibilidade**. A primeira se refere a uma medida da continuidade do serviço, que pode ser definida como a probabilidade do serviço permanecer operacional em um determinado intervalo de tempo $(0, t)$. A disponibilidade de um serviço ou sistema de comunicação pode ser definida pelo quão disponível ele está, em termos de probabilidade, em um determinado instante de tempo. Ela pode ser estimada através do indicador de disponibilidade $A = MTTF/MTBF = MTTF/(MTTF + MTTR)$, sendo *MTTF* (do Inglês, *Mean Time to Failure*) o período de tempo no qual o serviço não é interrompido, o *MTBF* (do

Inglês, *Mean Time Between Failures*) o tempo médio entre falhas consecutivas e *MTTR* (do Inglês, *Mean Time to Repair*) o tempo médio de reparo. A confiabilidade tem sido um importante fator para serviços de transmissão e entrega confiável dos dados, especialmente para aqueles os quais requerem relativamente longos valores de MTTF. Por outro lado, para aplicações baseadas em fluxos de curta duração, como aplicações de transferência de hipertexto, as falhas frequentes do sistema não têm tanta importância. A **manutenibilidade** se refere à capacidade do sistema de se atualizar e evoluir. A **segurança** (do Inglês *safety*) é a medida de dependabilidade de um sistema sob falhas catastróficas, referindo-se particularmente ao efeito e não à causa de uma falha. E a integridade pode ser compreendida como a ausência de alterações inapropriada ou não autorizadas no sistema ou serviço.

A **segurança** (do Inglês, *Security*) se refere à habilidade de um sistema se proteger de várias atividades não autorizadas. Ela compartilha com a dependabilidade as propriedades anteriores de disponibilidade e integridade. Além disso, inclui as propriedades de **Autenticação**, **Autorização** e **Auditoria** (AAA), **confidencialidade** e a capacidade de **não-repúdio**. Por fim, a **performabilidade** é uma disciplina usada para prover medidas de desempenho de um sistema, tendo como referência os requisitos de QoS de um determinado serviço, como atraso, variação do atraso, vazão, taxa de entrega, etc.

2.2.3 Estratégias de Resiliência

Atualmente existem inúmeras estratégias de resiliência para redes de comunicação na literatura [Mauthe et al., 2016]. Contudo, esta seção descreve apenas a estratégia ResiliNets, a qual é vista como uma revisão e síntese das estratégias anteriores a ela [Sterbenz et al., 2010]. A estratégia ResiliNets tem desenvolvido um arcabouço de resiliência, inicialmente como parte dos projetos ANA (do Inglês, *Autonomic Network Architecture*) e PoMo (do Inglês, *Postmodern Internet Architecture*). Esta iniciativa faz parte do projeto ResumeNet (*Resilience and Survivability for Future Networking: Framework, Mechanisms, and Experimental Evaluation*) [Jabbar, 2010]. O arcabouço proposto pela ResiliNets tem como base os princípios de resiliência descritos na seção anterior.

A estratégia ResiliNets define um esquema de resiliência chamado D^2R^2+DR (Defender, Detectar, Remediar, Recuperar mais Diagnosticar e Refinar). Esse esquema possui seis etapas, divididas em duas fases, apresentadas a seguir.

Primeira fase

A primeira fase, a D^2R^2 , consiste de um ciclo de 4 etapas envolvidas com as operações de rede e serviços simultaneamente, sendo:

- **Defender** - medidas defensivas para proteger a rede contra desafios e ameaças à operação normal. Em primeiro lugar, qualquer sistema resiliente precisa defender-se contra os desafios para que as falhas não resultem em falhas observáveis. Em segundo lugar, a defesa precisa isolar estes eventos que afetam a operação da rede dos serviços prestados por ela. A tolerância a falhas e a capacidade de sobrevivência são consideradas as primeiras linhas de defesa. Elas podem ser alcançadas usando mecanismos passivos, como redundância e diversidade. A segunda linha de defesa inclui medidas para impedir que o serviço seja afetado mesmo na ocorrência de eventos adversos, utilizando mecanismos ativos de autoproteção (e.g., *firewalls*). Um exemplo de medidas defensivas inclui o uso de paradigmas de conectividade que possibilitam a comunicação fim-a-fim mesmo quando os caminhos estão instáveis. Contudo, mesmo com os mecanismos de

defesa alguns desafios são imprevisíveis e podem ocorrer e neste caso o sistema deve detectar essas falhas.

- **Detectar** - a capacidade de detectar em tempo real desafios ou eventos adversos que não foram antecipados. Assim como a defesa, a detecção pode ocorrer em duas fases. Primeiro, o mecanismo de detecção pode detectar as mudanças nas condições operacionais da rede. Isso requer que o sistema compreenda operações normais (Axioma 1) e detecte os possíveis desvios. Em segundo lugar, o sistema pode detectar alterações nos parâmetros do serviço. Para isso, o sistema deve compreender os requisitos primários dos serviços.
- **Remediar** - se as medidas de defesa falharem e isso for detectado, deve-se empregar medidas para corrigir os efeitos do evento adverso antes que a operação da rede seja comprometida, ou seja, minimizar o impacto da falha.
- **Recuperar** - retomar as operações originais e normais da rede ou serviço após a ocorrência do evento adverso. Para isto, a rede deve detectar o fim do evento adverso e o mecanismo de recuperação se encarrega de restaurar a rede para seu estado original.

A segunda fase atende à evolução a longo prazo do sistema, através do diagnóstico das causas do desafio e do refinamento da operação do sistema. Esta fase consiste em duas etapas:

- **Diagnosticar** - diagnóstico da falha que foi a causa principal de um erro ou defeito. As falhas não podem ser observadas ou detectadas diretamente, elas só podem ser detectadas quando elas se manifestam como erros ou defeitos. Neste caso, após a ocorrência de um evento adverso, deve-se diagnosticar qual foi a falha principal que ocasionou o defeito.
- **Refinar** - realizar uma análise dos ciclos D^2R^2 anteriores. As observações do ciclo D^2R^2 na presença de eventos adversos são utilizadas para refinar, aprimorar e evoluir o processo para que a rede seja mais resiliente aos desafios futuros.

2.2.4 Princípios para Construção de Redes Resilientes

As definições da estratégia ResiliNets e as disciplinas de resiliência descritas na seção 2.2.2, determinam um conjunto de diretivas para a construção de redes resilientes [Sterbenz et al., 2014]. Estas diretivas incluem os pré-requisitos, os prós e contras, os possibilitadores e o comportamento esperado para uma rede resiliente [Sterbenz et al., 2010].

Requisitos

A seguir estão os requisitos necessários para criar redes resilientes:

- **Requisitos dos serviços** - os requisitos de serviço das aplicações precisam ser determinados para entender o nível de resiliência que o sistema deve fornecer.
- **Comportamento normal** - pode ser compreendido como uma combinação que envolve a especificação e o projeto da rede, juntamente com o estado enquanto não ocorre um evento adverso, a fim de aprender os parâmetros normais de operação da rede.

- **Modelos de ameaças e desafios** - são essenciais para a compreensão e detecção de potenciais eventos e condições adversas. Sem eles não é possível compreender, definir e implementar os mecanismos de resiliência.
- **Métricas** - quantificam os requisitos de serviço (normal ↔ parcialmente degradado ↔ severamente degradado) e o estado operacional (aceitável ↔ prejudicado ↔ inaceitável). Estas métricas são necessárias para detectar, remediar e quantificar a resiliência para refinar o comportamento futuro. As métricas são fundamentais para quantificar os requisitos anteriores.
- **Mecanismos heterogêneos, confiança e política** - estes três itens são as realidades do mundo atual. Nenhuma tecnologia é apropriada para todos os cenários, e as escolhas mudam com o tempo. Os mecanismos de resiliência devem lidar com tecnologias heterogêneas, endereçamento, encaminhamento, roteamento, sinalização, tráfego e mecanismos de gerenciamento de recursos. Os mecanismos de resiliência também devem admitir explicitamente a existência de conflitos de confiança e políticas.

Para construir sistemas resilientes, a medição, avaliação e compreensão do aspecto de operações (comportamento normal) e de serviço (requisitos) da rede, juntamente com os desafios, são requisitos. Sem uma estrutura para quantificar esses aspectos da rede, não se pode avançar com sucesso para a resiliência [Jabbar, 2010].

Vantagens e Desvantagens

As vantagens e desvantagens representam os ganhos e as perdas que devem ser considerados no desenvolvimento de um sistema resiliente, sendo:

- **Recursos** - determinam a implementação de mecanismos resilientes. Isto envolve um planejamento dos recursos para balancear e otimizar a relação entre resiliência e custo. O custo pode limitar o uso de medidas de resiliência, como redundância e diversidade.
- **Complexidade** - resulta da interação de sistemas em vários níveis de hardware e software e está relacionada à escalabilidade. Enquanto muitos dos princípios e mecanismos de resiliência aumentam essa complexidade, a própria complexidade torna os sistemas difíceis de entender e gerenciar, ameaçando a resiliência. Deve haver um equilíbrio no grau de complexidade.
- **Gerenciamento de estado** - é parte essencial de qualquer grande sistema complexo e está relacionado à resiliência de duas maneiras: primeiro, a escolha do gerenciador de estado afeta a resiliência da rede. Segundo, os próprios mecanismos de resiliência requerem um estado, sendo importante que eles alcancem sua meta de aumentar a resiliência global através do gerenciamento dos estados.

Possibilitadores

A seguir é apresentado um conjunto de possibilitadores para a construção de sistemas resilientes [Sterbenz et al., 2010]. Esses possibilitadores incluem:

- **Autoproteção e a segurança** - são propriedades essenciais das entidades para se defenderem contra os desafios em uma rede resiliente. A autoproteção inclui os mecanismos de autenticação, autorização e responsabilização, bem como os mecanismos de segurança convencionais de confidencialidade, integridade e não-repúdio.

- **Conectividade e a associação** - estas propriedades devem ser mantidas entre entidades comunicantes, quando possível. O fluxo de informações ainda deve ocorrer mesmo quando não existir um caminho fim-a-fim estável.
- **Redundância** - se refere à replicação de entidades na rede, geralmente para prover tolerância a falhas. A redundância no espaço, tempo e na informação aumentam a resiliência contra falhas e alguns desafios caso as defesas sejam penetradas.
- **Diversidade** - no espaço, no tempo, no meio e nos mecanismos aumentam a resiliência contra os desafios. A diversidade consiste em fornecer alternativas para que, mesmo quando os desafios impactam alternativas específicas, outras alternativas evitam a degradação das operações normais.
- **Resiliência multinível** - é necessária em três dimensões ortogonais: camadas de protocolo, em que a resiliência em cada camada fornece uma base para a próxima camada acima; planos de dados, controle e gerenciamento; arquitetura de rede, de dentro para fora, de componentes tolerantes a falhas, através de sub-redes e topologias sobreviventes, até a Internet, incluindo sistemas anexos e aplicações.
- **Conhecimento do contexto** - necessário para que os nós resilientes monitorem o ambiente da rede (condições do canal, estado do enlace, estado operacional dos componentes da rede, etc.) e detectem eventos ou condições adversas. Mecanismos de correção devem levar em consideração o contexto atual de operação do sistema.
- **Translucência** - permite interações entre camadas e planos, sendo fundamental para otimizar a resiliência entre os diferentes níveis. Os sistemas complexos são estruturados em vários níveis para abstrair a complexidade e separar as preocupações. Um exemplo simples de interação entre as camadas seria a passagem dos requisitos definidos pelas aplicações, com base em seus requisitos de serviço, para a rede abaixo. Estes requisitos são utilizados pela rede para otimizar seu comportamento [Sterbenz et al., 2014].

Todos os possibilitadores especificados acima tentam melhorar a resiliência. No entanto, para um determinado cenário, é difícil avaliar qual combinação resulta em uma melhor ou maior resiliência [Jabbar, 2010].

Comportamento

A arquitetura ResiliNets propõe três comportamentos de um sistema resiliente, conforme discutido abaixo:

- **Comportamento auto-organizado e autonômico** - necessário para a resiliência da rede, que é altamente reativa, com mínima intervenção humana. Uma rede resiliente deve inicializar e operar com mínima configuração, gerenciamento e intervenção humana. Idealmente, a intervenção humana deveria ser limitada e apenas quando desejado, com base em políticas operacionais de alto nível. As fases da rede autônoma consistem em inicialização, auto-configuração, auto-organização, autogestão, auto-otimização, autodiagnóstico e autoreparo [Jabbar, 2010].
- **Adaptabilidade** - os componentes de rede resilientes precisam adaptar seu comportamento com base em condições dinâmicas da rede, a fim de se remediar de eventos ou condições adversas, e para se recuperar, retomando as operações normais.

- **Evolutividade** - necessária para refinar o comportamento futuro para melhorar a resposta aos desafios, bem como para a arquitetura de rede e os protocolos responderem às ameaças emergentes e demandas de aplicativos. O refinamento baseia-se no ciclo de defesa, detecção e remediação da estratégia D^2R^2+DR .

2.3 Transmissão Multicaminhos

A transmissão multicaminhos consiste em uma transmissão em que dois ou mais caminhos estão disponíveis para transmitir dados entre dois dispositivos [Becke, 2014]. Em nível de camada de transporte, cada caminho está associado a um subfluxo e o conjunto destes a uma única conexão. O uso dos múltiplos caminhos está associado a dois objetivos principais: (i) aumentar a capacidade de transmissão através da transmissão multicaminhos concorrente (CMT); (ii) prover uma maior resiliência através da transmissão multicaminhos resiliente (RMT) [Ford et al., 2011]. Conforme ilustra a Figura 2.4, a CMT agrega a capacidade de transmissão dos caminhos disponíveis para transmitir em paralelo os dados de um único fluxo de dados. Por outro lado, a RMT utiliza os caminhos de forma redundante [Bonaventure et al., 2015]. Neste último caso, os caminhos primários são utilizados para transportar os dados e os caminhos auxiliares (*backup*) apenas quando estes falharem [Bonaventure et al., 2015]. A RMT não é exclusiva à CMT, ou seja, os dados podem ser transmitidos em paralelo sobre múltiplos caminhos, mas tendo a resiliência como objetivo principal [Domżał et al., 2015].

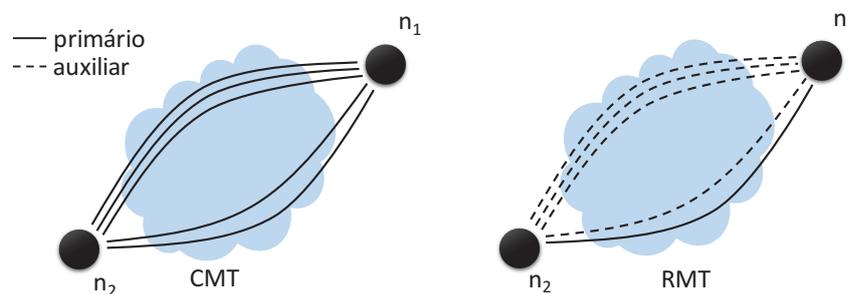


Figura 2.4: Tipos de Transmissão multicaminhos.

Agregar a capacidade de múltiplas interfaces para a transmissão de dados sobre múltiplos caminhos pode aumentar o desempenho e a resiliência do transporte confiável de dados sobre a Internet. No entanto, existem desafios impostos por alguns cenários que dificultam alcançar estes benefícios [Xu et al., 2016]. Este trabalho destaca três diferentes cenários de comunicação multicaminhos: o cenário C1 é homogêneo, uma vez que os caminhos têm pouca diferença entre suas características; o cenário C2 é de baixa heterogeneidade porque os caminhos possuem diferentes características; o cenário C3 é de alta heterogeneidade e dinâmico, dado as características heterogêneas e dinâmicas dos caminhos devido à mobilidade dos usuários. Conforme ilustra a Figura 2.5, estes cenários estão associados a diferentes tipos de redes:

- *Cenário homogêneo (C1)* - comumente representa uma rede de centro de dados. Nesta rede os dispositivos possuem interfaces de alto desempenho e são interligados através de redes de alta velocidade e baixa latência [Wang et al., 2016]. A transmissão multicaminhos neste cenário permite melhorar o desempenho e alcançar alta disponibilidade [Paasch e Bonaventure, 2014]. Para exemplificar, o servidor S_1 possui duas interfaces de rede conectadas a diferentes comutadores (i.e., *switches*). A conexão

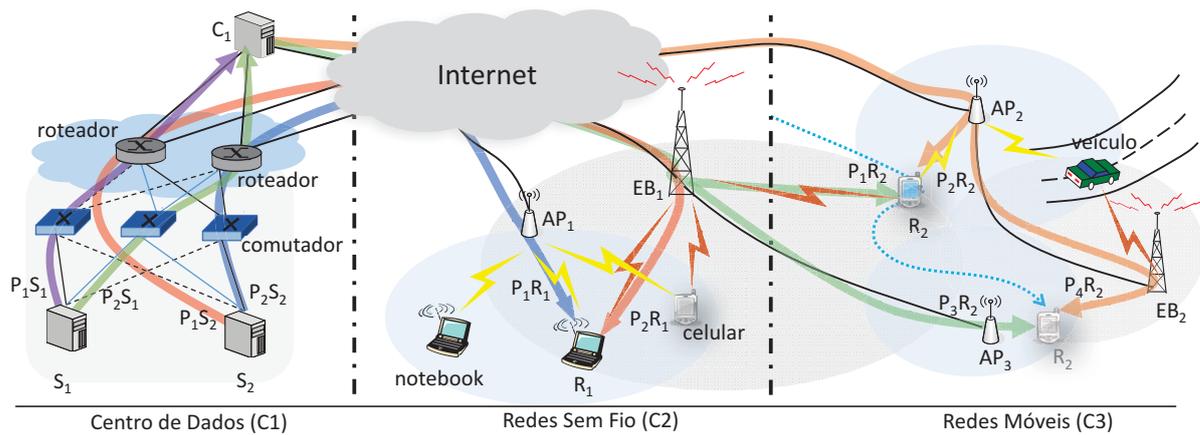


Figura 2.5: Cenários para transmissão multicaminhos

entre C_1 e S_1 estabelece um subfluxo sobre cada interface de rede. Para melhorar o desempenho da transmissão, os dados são transmitidos simultaneamente através dos caminhos P_1S_1 , P_2S_1 . Os caminhos também são utilizados de modo redundante, e caso ocorra uma falha na interface ou em um dos caminhos a transmissão permanece ativa nos caminhos restantes.

- *Cenário de baixa heterogeneidade (C2)* - cenário comumente composto por dispositivos sem fio que têm pouca ou nenhuma mobilidade. Este cenário representa uma rede residencial ou de escritório. Neste exemplo, essas redes são cobertas por duas diferentes redes de acesso sem fio, a WiFi, por meio do ponto de acesso AP_1 , e a rede Celular, através da estação base EB_1 . Essas redes possuem características diferentes e ambas têm variações em função de problemas com a propagação e interferência do sinal de rádio. O dispositivo R_1 estabelece uma conexão com C_1 através das duas redes de acesso. Neste caso a transmissão multicaminhos pode ser CMT ou RMT. A CMT é utilizada para aumentar a taxa de transferência agregando a capacidade dos caminhos. A RMT utiliza uma interface, como a que oferece um menor atraso ou a maior largura de banda, deixando a segunda ociosa. Em caso de falhas ou mudança do estado da rede, os dados são redirecionados para a segunda interface.
- *Cenário de alta heterogeneidade (C3)* - cenário heterogêneo e dinâmico, em que os dispositivos têm mobilidade moderada e alta, como redes celulares e redes veiculares. As transições entre os pontos de acesso e entre as estações base são frequentes, causando perda de conectividade através de determinados caminhos. Devido à mobilidade, os problemas com a propagação do sinal de rádio são ainda maiores. Neste sentido, a disponibilidade do caminho e suas características, como atraso, taxa de perdas e largura de banda, mudam frequentemente. O dispositivo R_2 está em movimento. Em um determinado momento estabelece uma conexão com o servidor C_1 utilizando os caminhos P_1R_2 e P_2R_2 . R_2 se movimenta, perdendo a conexão com o AP_2 em um determinado momento. Ele estabelece conexão com o AP_3 ao entrar em sua área de cobertura. Continuando sua trajetória, ele perde conexão com EB_1 , e adiante, quando entra na área de cobertura do EB_2 estabelece novamente conexão com a rede celular. Em todo este trajeto é desejável que a conexão com C_1 não seja interrompida.

O cenário homogêneo tem sido um dos mais promissores para utilização da transmissão multicaminhos. A homogeneidade das características dos caminhos, como baixa latência e enlaces

de alta velocidade, facilitam a obtenção das estimativas de atraso e largura de banda disponível. Estas informações são úteis para prevenir o recebimento fora de ordem (RFO) de pacotes quando múltiplos caminhos são utilizados em paralelo para transmitir um único fluxo de dados. Neste cenário, o RFO tem um impacto pequeno [Raiciu et al., 2010], e os desafios para a transmissão multicaminhos incluem realizar o balanceamento de carga, compartilhar de forma justa os recursos da rede, e melhorar a robustez [Raiciu et al., 2011a]. O balanceamento da carga deve obedecer ao princípio de *Resource Pooling*, no qual se busca mover o tráfego para caminhos menos congestionados, reduzir o número de caminhos subutilizados ou ociosos e aumentar a resiliência da rede [Wischik et al., 2008]. Ao compartilhar um gargalo, uma conexão multicaminhos deve obedecer ao princípio da amigabilidade do TCP (*TCP-friendliness*) [Baker e Fairhurst, 2015], ou seja, os subfluxos não devem ser mais agressivos que um fluxo TCP padrão. Por fim, uma transmissão multicaminhos robusta deve ser resiliente a falhas. Se ocorrer uma falha em um dos caminhos, a transmissão deve estar apta a redistribuir a carga sobre os demais caminhos restantes.

Nos cenários heterogêneos, os caminhos (subfluxos) normalmente possuem diferentes características de atraso, taxa de perda, largura de banda, etc. Estas diferenças fazem com que os pacotes, sequencialmente distribuídos entre os caminhos, cheguem fora da ordem no destinatário. Uma vez que os pacotes precisam ser entregues em ordem, eles permanecem na fila única de recebimento até o recebimento do pacote faltante. Como a fila geralmente tem um tamanho limitado, o RFO pode ocasionar o enchimento e conseqüentemente o bloqueio da fila, conhecido como bloqueio HOL [Sarwar et al., 2013]. Quanto maior as diferenças entre os caminhos, maior é a possibilidade de ocorrer o bloqueio HOL durante a transmissão. O bloqueio HOL degrada o desempenho da transmissão multicaminhos, limitando principalmente os benefícios da agregação de banda.

2.4 Protocolo de Transmissão Multicaminhos

Embora amplamente utilizado na Internet, o protocolo TCP não tem suporte a transmissão de dados multicaminhos. Ele estabelece uma conexão fim-a-fim com apenas um par de endereços IP de origem e destino. Logo, um fluxo TCP não pode ser associado a um novo endereço, sendo ou não de uma nova interface de rede, sem que a conexão seja reestabelecida [Comer, 2008]. Muitos protocolos foram propostos com objetivo de utilizar a transmissão multicaminhos sobre dispositivos com múltiplas interfaces [Li et al., 2016], porém, os únicos padronizados pelo IETF foram o CMT-SCTP [Iyengar et al., 2006] e o MPTCP [Ford et al., 2013]. Embora existam diferenças entre as arquiteturas implementadas pelo CMT-SCTP e MPTCP [Becke et al., 2013], os desafios da transmissão multicaminhos, como o reordenamento, compartilhamento dos recursos, entre outros, são comuns a ambos os protocolos.

O SCTP surgiu antes do MPTCP, porém existem alguns problemas que podem ter impedido que o mesmo fosse empregado largamente. Ele não é compatível com a maioria dos serviços de rede, como *firewalls* e NATs, que são incapazes de processar os pacotes SCTP, portanto acabam descartando-os. Outra razão se deve ao fato que o SCTP possui uma API diferente da API padrão do TCP, utilizada pelas aplicações. Assim, os serviços de rede não adicionam o suporte ao SCTP porque poucos aplicativos o estão utilizando. Por outro lado, os novos aplicativos não usam o SCTP porque os serviços de rede descartam seus pacotes. Embora existam tentativas de romper esse círculo vicioso, encapsulando SCTP em cima do UDP (*User Datagram Protocol*) e expondo esta API para os aplicativos, seu uso ainda está distante de ser generalizado [Paasch e Bonaventure, 2014].

Se estes problemas têm dificultado a adoção do protocolo SCTP, por outro lado, tem contribuído para tornar o MPTCP uma das principais soluções de transporte multicaminhos.

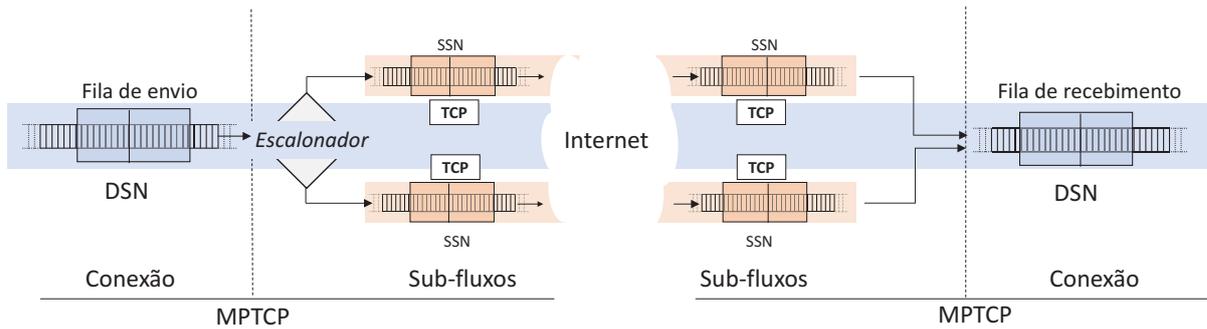


Figura 2.6: Sequenciamento DSN e SSN no MPTCP

Adicionalmente, o MPTCP foi projetado para evitar estes problemas, mantendo compatibilidade com os serviços de rede e aplicações e também não se limita apenas ao tráfego de fluxo de dados de vídeo e voz [Becke et al., 2013].

2.4.1 MPTCP

O protocolo MPTCP [Barré et al., 2011a] tem como base o protocolo TCP e possibilita a transmissão de dados através de múltiplos caminhos. O MPTCP é o esforço mais recente de padronizar um protocolo de transporte multicaminhos confiável para uso genérico na Internet. Responsável pela padronização desde 2009, o grupo de trabalho do MPTCP⁴ do IETF tem como responsabilidade a padronização do protocolo, envolvendo aspectos da arquitetura [Ford et al., 2011], modos de operação [Ford et al., 2013], API [Scharf e Ford, 2013], controle de congestionamento [Raiciu et al., 2011b] e a segurança da transmissão em [Bagnulo, 2011] e [Bagnulo et al., 2015].

Um dos principais objetivos do MPTCP como protocolo de transporte multicaminhos trata-se de obter um desempenho melhor que o oferecido por um único fluxo TCP, em termos de melhorar a vazão e reduzir a latência [Bonaventure et al., 2012]. Por outro lado, o MPTCP também possibilita aumentar a resiliência da comunicação, utilizando os caminhos de forma redundante para persistir a conexão em caso de falhas. Uma comunicação MPTCP provê a troca de dados bidirecional entre dois nós comunicando assim como o TCP padrão, não requerendo qualquer mudança na aplicação. O MPTCP possibilita que os nós utilizem diferentes caminhos e diferentes endereços para transmitir pacotes pertencentes a uma mesma conexão. Para a camada de rede, cada subfluxo é visto como um subfluxo TCP padrão, mas que transportam no cabeçalho TCP um tipo de campo *Option* específico do MPTCP. Todas as operações do MPTCP são sinalizadas através do campo *Option*. Para controlar o envio de pacotes através de diferentes caminhos o MPTCP possui dois níveis de reconhecimento: DSN (do Inglês *Data Sequence Number*), que sequencia os pacotes recebidos da aplicação; SSN (do Inglês *Subflow Sequence Number*), que faz um novo sequenciamento para controle de cada subfluxo. Conforme ilustra a Figura 2.6, O DSN faz o sequenciamento dos dados em nível de conexão, antes dos dados serem alocados pelo escalonador aos subfluxos. Ele tem como função controlar o envio e recebimento ordenado dos dados entre os nós de origem e destino. O SSN tem o mesmo comportamento do sequenciamento utilizado no TCP e faz o controle do envio dos dados em nível de subfluxo.

A comunicação entre dois nós utilizando o MPTCP pode ser resumida nos seguintes passos [Ford et al., 2013]. Uma conexão MPTCP inicia de modo similar a uma conexão TCP, utilizando a troca de pacotes SYN, SYN/ACK e ACK, conhecido como aperto de mão de

⁴<https://tools.ietf.org/wg/mptcp/>

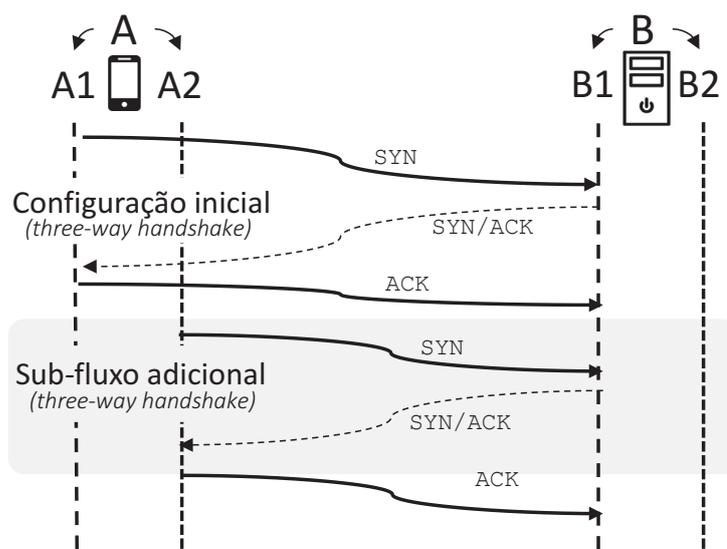


Figura 2.7: Estabelecimento de Subfluxos [Ford et al., 2013]

três vias (do Inglês, *three-way handshake*). A Figura 2.7 ilustra de uma forma simplificada como se estabelece uma conexão MPTCP entre os nós A e B a partir dos endereços A1 e B1, respectivamente. Após a configuração inicial, os nós trocam informações sobre os endereços adicionais e um novo *three-way handshake* é realizado para adicionar um subfluxo com os endereços A-A2 e B-B1. O MPTCP ainda permite que outros subfluxos sejam criados, por exemplo, combinando os endereços A1-B2 e A2-B2.

A estrutura do protocolo MPTCP, descrita na RFC 6182 [Ford et al., 2011], inclui quatro entidades, como ilustra a Figura 2.8, sendo: o gerenciador de caminhos [Ford et al., 2013]; o escalonador; o controle de congestionamento [Raiciu et al., 2011b]; e a interface de subfluxos. Estes componentes se relacionam da seguinte forma: O gerenciador de caminhos se responsabiliza pela descoberta dos caminhos (i.e., subfluxos) entre os dois nós. O escalonador de pacotes, após receber um fluxo de dados da aplicação, realiza suas operações, como segmentar os dados, adicionar um número de sequência em nível de conexão e distribuir os segmentos para os subfluxos conforme política adotada. Um subfluxo, ao receber o segmento, adiciona o seu próprio número de sequência e encaminha-o através da rede. No destinatário, os dados são recebidos em nível de subfluxo e passados a fila de recebimento que, caso necessário, faz o reordenamento em nível de conexão. Por fim, o controle de congestionamento realiza, entre outras tarefas, o controle do fluxo de dados e a prevenção de congestionamentos nos subfluxos.

O MPTCP visa suportar de forma eficiente os objetivos do transporte multicaminhos [Ford et al., 2011]. O MPTCP faz uso das sessões TCP (subfluxos) para prover o transporte por múltiplos caminhos e também manter compatibilidade com a rede atual. As informações específicas do MPTCP são transportadas nos subfluxos TCP. As próximas seções descrevem as operações dos principais componentes do MPTCP, o Gerenciador de Caminhos, o Controle de Congestionamento e o Escalonador.

Gerenciador de Caminhos

O gerenciador de caminhos tem como função adicionar, divulgar e remover os subfluxos com base nos endereços IP dos nós. A adição de caminhos pode ser feita de duas formas. Durante a configuração do primeiro subfluxo, conforme ilustrado previamente na Figura 2.7, o subfluxo adicional é estabelecido utilizando-se a opção `MP_JOIN`. A divulgação ou sinalização

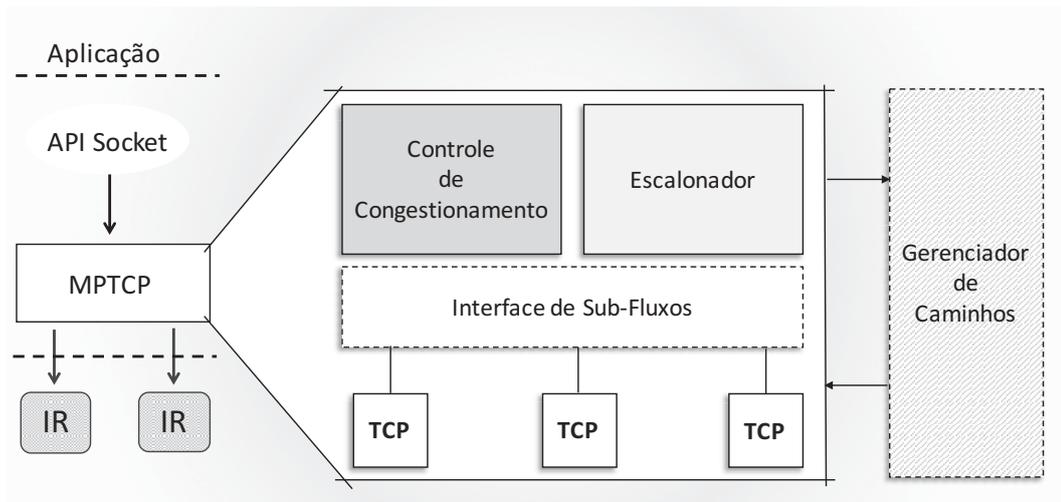


Figura 2.8: Componentes do MPTCP [Ford et al., 2013]

dos endereços adicionais é feita através da opção `ADD_ADDR`. Se durante a conexão um endereço se torna indisponível (falha na interface), o nó comunica a outra ponta utilizando a opção `REMOVE_ADDR`, afim de que sejam removidos os subfluxos relacionados com este endereço. Estes mecanismos permitem a mudança de endereços, ou seja, a adição e remoção de caminhos, durante o tempo de vida da conexão.

O mecanismo de gerenciamento de caminhos do MPTCP ainda provê duas formas de associação de subfluxos. A primeira, conforme ilustra a Figura 2.9.a, associa cada subfluxo a uma interface/endereço IP. Na segunda, conforme ilustra a Figura 2.9.b, o gerenciador de caminhos associa múltiplos subfluxos com uma mesma interface/endereço IP, alterando apenas as portas lógicas. Neste último caso, o MPTCP visa alcançar uma maior diversidade de caminhos, uma vez que cada subfluxo são tratados pela rede como fluxos TCP independentes e deste modo, podem ser encaminhado por caminhos distintos, aumentando a resiliência da transmissão.

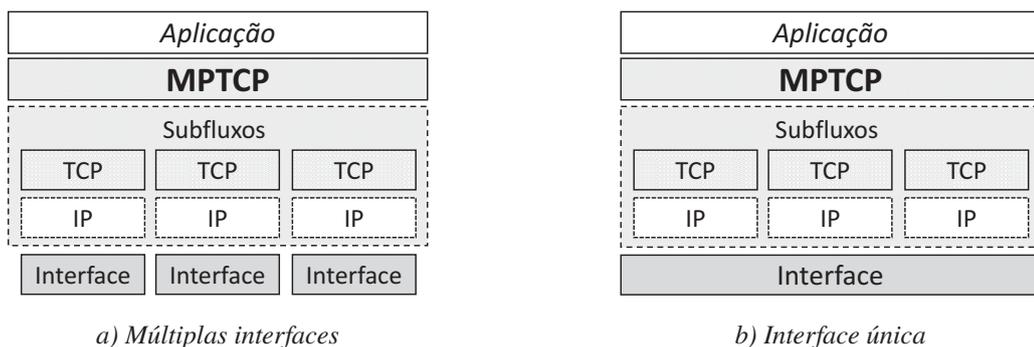


Figura 2.9: Associação entre endereços e interfaces de rede

Conforme a RFC 6824, o gerenciador de caminhos está relacionado com os modos de operação do MPTCP [Ford et al., 2013]. Estes modos consideram fatores relacionados ao interesse dos usuários [Paasch et al., 2012], como desempenho, consumo de energia e custo de transmissão. Isto porque alguns usuários podem optar por uma melhor taxa de transferência, enquanto outros querem maximizar o uso da bateria ou podem preferir priorizar a rede de acesso com menor custo monetário. Os principais modos de operação implementados pelo MPTCP são:

- *default* (padrão) - o nó de origem não anuncia os diferentes endereços IPs e nem inicializa a criação de subfluxos. No entanto, o mesmo aceita a criação passiva de novos subfluxos, quando requisitado pelo outro nó.
- *FullMesh* (Malha completa) - tem como objetivo atender os usuários que querem obter uma melhor taxa de transferência. Este modo cria toda a malha possível de subfluxos, utilizando os pares de endereços IP origem/destino, entre os nós envolvidos na conexão;
- *Ndiffports* (Portas Diferentes)- cria uma determinada quantidade de subfluxos através do mesmo par de endereços IP, modificando a porta de origem. O número de subfluxos pode ser controlado pelo usuário.

Controle de Congestionamento

O mecanismo de controle de congestionamento (CC) é o responsável por controlar a taxa de entrega de dados para conexão em função do congestionamento percebido. Ele faz isto ajustando o tamanho da janela de congestionamento (CWND) em cada subfluxo de uma conexão de forma individual. Além do controle de fluxo, o CC tem como função balancear o tráfego entre os subfluxos, evitando aqueles mais congestionados e garantindo que o MPTCP seja amigável com os demais fluxos TCP (TCP-Friendly) [Barré et al., 2011b]. O CC executado pelo MPTCP tem como base os algoritmos do TCP, como a variante *NewReno*. Todos eles modificam apenas a fase de prevenção de congestionamento (*Congestion Avoidance*), ficando as fases de inicialização lenta (*Slow Start*), retransmissão rápida (*Fast Retransmit*) e recuperação rápida (*Fast Recovery*) inalteradas [Singh et al., 2013]. O CC é executado em nível de subfluxo, utilizando uma janela de congestionamento (CWND) individual e uma janela de recebimento (RCVW) compartilhada para facilitar a entrega ordenada dos dados [Zhou et al., 2015]. Em síntese, seus objetivos principais incluem aumentar a vazão, não prejudicar fluxos concorrentes e balancear o congestionamento.

A maioria dos algoritmos de CC propostos para o MPTCP aplicam diferentes abordagens às fases de incremento aditivo e decremento multiplicativo. O algoritmo LIA (*Linked Increase Algorithm*), adotado como padrão pelo IETF [Ford et al., 2013], especifica apenas como deve ocorrer o incremento da janela de congestionamento ao receber um ACK, mantendo o decremento padrão como no TCP. O crescimento das janelas dos subfluxos são acoplados (Eq. 2.1). O parâmetro α , Eq. 2.2, controla a agressividade dos subfluxos de modo que o incremento da janela não seja superior ao de um fluxo de caminho único com o mesmo tamanho de janela. Nas Eqs. 2.1 e 2.2, w_i e RTT_i referem-se ao tamanho da janela de congestionamento e o tempo de ida e volta (RTT) de um subfluxo i , respectivamente. O parâmetro W representa o tamanho total das janelas de congestionamento de todos os subfluxos.

$$w_i = \begin{cases} \min(\alpha/W, 1/w_i) & \text{incremento} \\ w_i/2 & \text{decremento} \end{cases} \quad (2.1)$$

$$\alpha = W * \frac{\max(w_i/RTT_i^2)}{(\sum_i w_i/RTT_i)^2} \quad (2.2)$$

Um dos objetivos do algoritmo de controle de congestionamento é garantir o compartilhamento justo da largura de banda (i.e., princípio de *fairness*) e o uso eficiente dos caminhos. Deste modo, se espera que a transmissão multicaminhos tenha um desempenho, no mínimo, superior a uma transmissão SPTCP pelo melhor caminho. Contudo, alguns trabalhos têm mostrado que utilizar o MPTCP não traz os benefícios esperados e que os subfluxos podem ser muito agressivos com os demais em um ambiente compartilhado [Khalili et al., 2012].

Escalonador

O escalonador é um dos componentes principais de qualquer arquitetura de agregação de banda [Ramaboli et al., 2012]. Para os protocolos de transmissão de dados multicaminhos, o escalonador é utilizado para distribuir os segmentos através de um ou mais caminhos. O funcionamento do escalonador basicamente consiste de fazer a alocação dos pacotes (segmentos) de dados nos subfluxos disponíveis. O IETF não define uma localização específica para o escalonador, isto é, uma parte explícita dentro da arquitetura do protocolo multicaminhos [Ford et al., 2011]. Isto porque o escalonador não é descrito como uma instância ou módulo, e sim como um processo e suas etapas. Contudo, a implementação do protocolo MPTCP tem se baseado em diferentes estratégias [Barré et al., 2011b]. Inicialmente o protocolo utilizava uma estratégia *Push* e atualmente a estratégia *Pull*. As diferenças entre as estratégias são significativas e determinam o modo de operação do escalonador. Basicamente, uma estratégia de implementação se refere à metodologia utilizada para alocar os dados recebidos da camada de aplicação aos subfluxos. Em [Singh et al., 2012], os autores descrevem as metodologias empregadas:

- *Push* - cada subfluxo tem sua própria fila de envio (SNDQ) onde mantém alocados os segmentos não enviados (N-ENV) e os segmentos enviados mas não reconhecidos (N-ACK), como ilustrado pela Figura 2.10;
- *Pull* - os pacotes de dados provenientes da aplicação são colocados na fila única de envio (SNDQ). Esta fila armazena os pacotes não enviados, enviados (N-ENV) e os não reconhecidos (N-ACK), como ilustrado pela Figura 2.10. Os pacotes de dados são alocados pelo escalonador somente quando algum subfluxo adverte sobre o reconhecimento de um pacote (N-ACK). Isto pode ocorrer com mais de um subfluxo ao mesmo tempo, cabendo ao escalonador decidir como os segmentos serão distribuídos;
- *Híbrida* - combina as duas abordagens. Cada subfluxo mantém uma fila de pacotes não enviados de forma complementar à fila de envio.

As estratégias possuem vantagens e desvantagens. Com a estratégia *Pull*, cada subfluxo pode requisitar os dados diretamente a SNDQ desde que não seja restrito pela política do escalonador. Se um subfluxo falhar, automaticamente ele para de puxar dados da SNDQ. Isto não acarreta prejuízos à transmissão porque o segmento não reconhecido em nível de subfluxo também está pendente na SNDQ, e provavelmente será retransmitido. Além disso, quando o subfluxo se recuperar da falha, ele passa automaticamente a puxar mais segmentos da SNDQ. No entanto este esquema afeta as políticas adaptativas, que precisam alocar os segmentos dinamicamente entre os caminhos. Isto porque, uma vez que os segmentos são construídos e colocados na fila compartilhada, o escalonador somente é executado quando um subfluxo libera espaço em sua CWND. Antes disso, ele não pode alterar a lógica já executada. Isto dificulta a implementação de políticas que consideram o tamanho dos pacotes e as características dos caminhos, dada a complexidade das políticas quando esta ação depende do recebimento de um ACK.

O método *Push* suporta a designação de segmentos com tamanhos pré-definidos aos subfluxos. Isto porque os segmentos são construídos antes de serem designados para a fila dos subfluxos. O escalonador, por sua vez, designa os segmentos assim que são recebidos da aplicação para as filas dos subfluxos e não é mais requisitado durante a transmissão. Contudo, isto aumenta a complexidade de implementar uma política de alocação de dados eficiente. O escalonador deve predizer, com base nas informações dos subfluxos, o tempo (atraso) que um pacote leva para ser transmitido da origem ao destino. Predizer estes valores em HetNets é uma tarefa complexa.

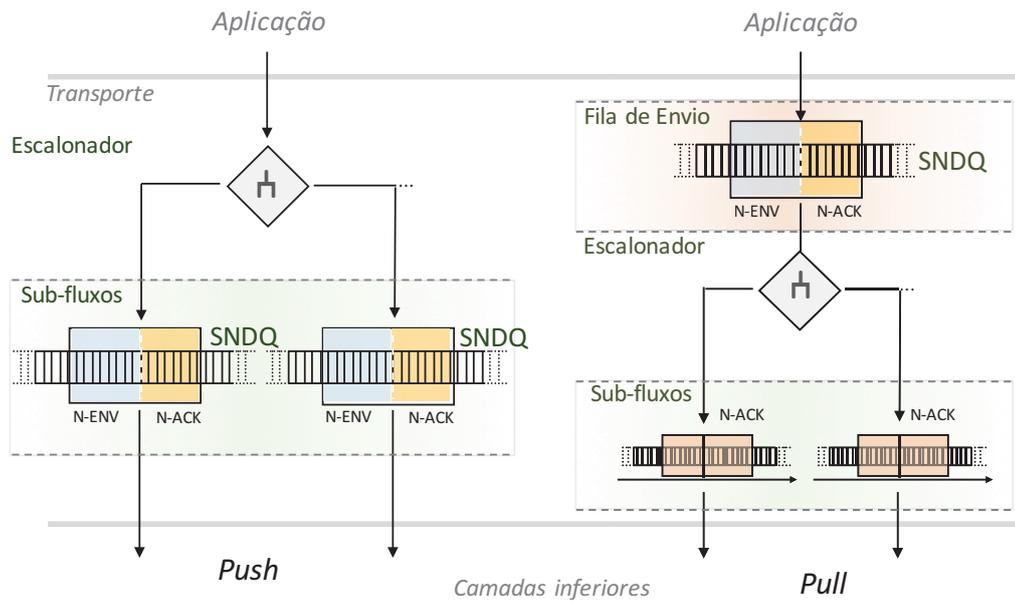


Figura 2.10: Metodologias *Push* e *Pull* [Barré et al., 2011b]

A estratégia *Híbrida* tem como objetivo resolver os problemas existentes nos dois métodos. Neste caso, cada subfluxo utilizaria uma pequena fila, a fim de facilitar a predição e o escalonador teria possibilidade de alocar os dados da fila compartilhada para a fila dos subfluxos, ou realocar, em caso de retransmissões. O problema desta estratégia *Híbrida* está na complexidade envolvida no processo de implementação. No entanto, as estratégias *Push* e *Híbrida* têm sido desconsideradas porque incorrem no problema do reordenamento [Becke, 2014], uma vez que os pacotes alocados aos subfluxos não podem ser reescalonados.

O MPTCP não possui um escalonador padronizado pelo IETF. Contudo, sua implementação [mpt, 2018], disponibilizada para diferentes sistemas operacionais, tem uma estrutura modular, o que permite optar por diferentes escalonadores. Existem três opções de políticas: o escalonador **Round-Robin** (RR), que distribui os dados de modo circular entre os subfluxos disponíveis; o escalonador LRF (**Lowest-RTT-First**), que distribui os dados primeiramente nos caminhos com menor RTT; e por fim, o escalonador **Redundante**, que replica os dados sobre os subfluxos disponíveis.

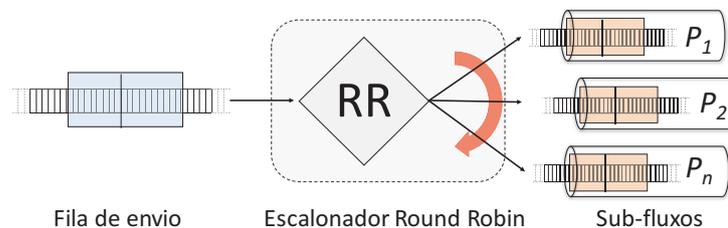


Figura 2.11: Round-Robin

Os escalonadores RR e LRF possuem um funcionamento similar. O RR seleciona os subfluxos de uma maneira circular, como ilustrado pela Figura 2.11. A quantidade de segmentos enviados de forma consecutiva em cada rodada pode ser configurada. Por outro lado, o escalonador pode optar por preencher a CWND de cada subfluxo a cada ciclo. No primeiro caso, o escalonador distribui igualmente os dados entre os subfluxos. No segundo caso, o escalonamento

não é de fato um RR, uma vez que os subfluxos recebem mais dados apenas quando liberam espaço em sua CWND. Este atraso para realizar um novo escalonamento é conhecido como o *tempo de espera pelo reconhecimento (Acknowledgement-Clock)*. O LRF inicialmente ordena os caminhos a partir do menor atraso (RTT), conforme ilustra a Figura 2.12. Posteriormente escalona os pacotes preenchendo a CWND dos subfluxos. Assim como o RR, logo que a CWND é preenchida, o escalonador se torna *Acknowledgement-Clock*. O escalonador redundante visa equalizar as diferenças entre os subfluxos, comuns em cenários heterogêneos, alcançando uma menor latência ao custo de desperdiçar recursos da rede. Ele também provê uma comunicação de alta confiabilidade, por transmitir os segmentos de forma redundante. Além disso, suprime as falhas, quando ocorrem em um dos subfluxos, de forma instantânea. Os escalonadores RR e LRF provêm uma transmissão multicaminhos concorrente (CMT), enquanto o escalonador redundante uma transmissão multicaminhos resiliente (RMT).

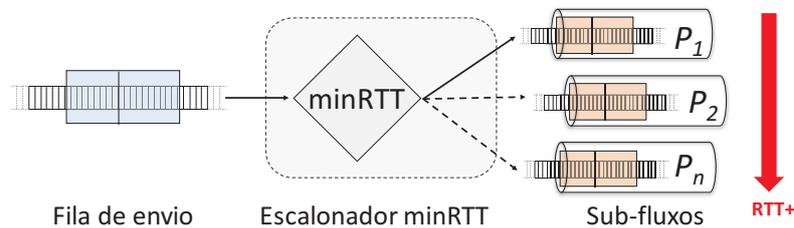


Figura 2.12: LRF (menor RTT)

Para compensar a heterogeneidade dos caminhos e evitar os problemas com o bloqueio HOL, o MPTCP usa o mecanismo de **Retransmissão e Penalização** [Paasch et al., 2014]. Este mecanismo reencaminha o segmento que possa estar causando o bloqueio HOL o mais rapidamente pelo subfluxo disponível (com espaço na CWND). O objetivo é tornar mais rápida a recuperação das situações de bloqueio, compensando a diferença entre os RTT dos caminhos. No entanto estas técnicas afetam principalmente a CMT, uma vez que reduz capacidade de envio de determinados subfluxos e conseqüentemente o desempenho global da transmissão multicaminhos [Ferlin et al., 2016]. O MPTCP também implementa um algoritmo para mitigação de atrasos nos caminhos, especialmente causados pelo fenômeno *bufferbloat*. O objetivo do algoritmo é limitar a quantidade de dados encaminhada no subfluxo que apresenta grande atrasos. O esquema empregado captura o *bufferbloat* monitorando a diferença entre o mínimo $RTT_{S_{min}}$ (RTT suavizado) e o RTT_S atual. Quando estes valores se distanciam, o esquema limita o crescimento da janela de congestionamento conforme Eq. 2.3, onde λ determina a tolerância entre $RTT_{S_{min}}$ e RTT_S .

$$CWND_{limit} = \lambda \times (RTT_{S_{min}} / RTT_S) \times CWND \quad (2.3)$$

2.4.2 Mecanismos Resilientes do MPTCP

O MPTCP emprega os mecanismos do TCP (e.g., ACKs, retransmissões, entrega ordenada, etc.) para prover uma comunicação fim-a-fim confiável [Comer, 2008]. Além disso, possui alguns mecanismos específicos para prevenir e se recuperar de falhas no processo de estabelecimento da conexão, perda de conectividade e presença de congestionamentos nos caminhos [Ford et al., 2011]. Caso ocorra uma falha ao estabelecer uma conexão, o MPTCP retorna para uma conexão TCP regular. Isto porque o nó, com quem ele esteja tentando abrir a conexão, ou algum serviço de rede, como um *firewall* ou NAT, pode não suportar o protocolo. Se a falha ocorrer durante a abertura de um subfluxo adicional, o MPTCP continua tentando

adicionar um novo subfluxo utilizando outros endereços e portas disponíveis. Adicionalmente, estas informações de falhas são guardadas, durante o tempo de vida da conexão, a fim de indicar que uma determinada interface ou endereço de destino falham e devem ser evitados. Este mesmo comportamento tem sido empregado para gerenciar os subfluxos, a fim de não utilizar, mesmo que temporariamente, aqueles que falhem regularmente ou tem um desempenho ruim [Oh e Lee, 2015, Deng et al., 2014, Hu et al., 2016].

O MPTCP implementa um mecanismo para tentar recuperar a transmissão se todos os subfluxos falharem ao mesmo tempo. O MPTCP utiliza segmentos `DATA-FIN`, para encerrar o envio de dados em nível de conexão, e um segmento `FIN`, para encerrar o envio em um subfluxo individual. Se todos os subfluxos são fechados, com a troca de segmentos `FIN`, mas nenhum segmento `DATA-FIN` foi recebido e reconhecido, a conexão somente será encerrada após o estouro do temporizador. Isto tem uma grande importância para sobrevivência da conexão. Em um cenário móvel, em que o nó perde a conectividade temporariamente, devido a mobilidade, este tempo de espera mantém a conexão ativa até que um subfluxo seja reestabelecido. Além deste mecanismo, o controle de congestionamento acoplado do MPTCP move o tráfego para caminhos menos congestionados. Isto permite que a transmissão multicaminhos possa reagir a um aumento não esperado de tráfego causado por eventos repentinos ou ataques de negação de serviço distribuídos (DDoS, do Inglês *Distributed Denial of Service*).

O MPTCP não tem como garantir a diversidade dos caminhos, ou seja, que subfluxos sejam encaminhados por caminhos distintos na rede. No entanto, ao criar múltiplos subfluxos TCP, ele espera que os mesmos ou um subconjunto deles, sejam encaminhados por caminhos distintos. Mesmo sem garantia disto, isto torna o MPTCP mais resiliente que o TCP, que utiliza apenas um caminho.

2.5 Resumo

Este capítulo descreveu os princípios de resiliência em redes de comunicação, destacando os desafios e as principais considerações da literatura para o desenvolvimento de uma estratégia de resiliência. Além disso, apresentou os fundamentos da transmissão multicaminhos, seus objetivos, cenários, desafios e os modos de operação do protocolo MPTCP. Os princípios de resiliência e os fundamentos da transmissão multicaminhos servem de base para proposta de tese apresentada no Capítulo 4. O protocolo MPTCP é utilizado como referência para a revisão bibliográfica, descrita no Capítulo 3, e nos estudos de caso, apresentados a partir do Capítulos 6.

3 Principais Desafios para a Transmissão Multicaminhos

Este capítulo apresenta uma revisão da literatura sobre as abordagens empregadas para tentar resolver os desafios enfrentados pela transmissão multicaminhos. O escopo da revisão tem como foco principal as abordagens propostas para o protocolo MPTCP. Isto porque ele tem um caráter de protocolo de transporte genérico, assim como o TCP, e engloba a maior parte das abordagens existentes. A pesquisa complementa a revisão referente ao controle de congestionamento multicaminhos apresentada em [Xu et al., 2016]. De forma específica, a revisão se direciona às abordagens que visam atender aos desafios e requisitos da transmissão multicaminhos, abordando com detalhe as soluções empregadas. O capítulo está organizado da seguinte forma. A Seção 3.2 apresenta os principais desafios para transmissão de dados multicaminhos. A Seção 3.3 apresenta uma descrição detalhada das abordagens que lidam com desafios apresentados na seção anterior, dividida em métodos de reconhecimento, seleção e políticas de alocação. A Seção 3.4 apresenta um resumo do capítulo.

3.1 Contextualização

A falhas nos elementos de rede são inevitáveis e podem ocorrer em consequência de vários desafios, como visto na Seção 2.2.1. Embora sejam diversos os eventos que podem ocorrer, a transmissão multicaminhos está sujeita à maioria deles. Isto ocorre devido ao caráter fim-a-fim da conexão, que esta sujeita aos desafios que afetam os caminhos na Internet, as redes de acesso e os próprios dispositivos móveis. Conforme ilustra a Figura reffig:cp2-desafios, esses desafios incluem a heterogeneidade e a falta de diversidade dos caminhos, a segurança, equidade, o custo e o consumo de energia. Na literatura existem inúmeras abordagens que tratam estes diferentes desafios, as quais foram organizadas conforme os grupos: métodos de reconhecimento, métodos de seleção e políticas de alocação. As próximas seções descrevem os desafios e apresenta as abordagens existentes conforme a organização citada acima.

3.2 Desafios

A Figura 3.1 apresenta uma relação dos principais desafios para a transmissão de dados multicaminhos. Estes desafios foram identificados através da revisão da literatura, tendo como foco os trabalhos direcionados para o protocolo de transporte MPTCP.

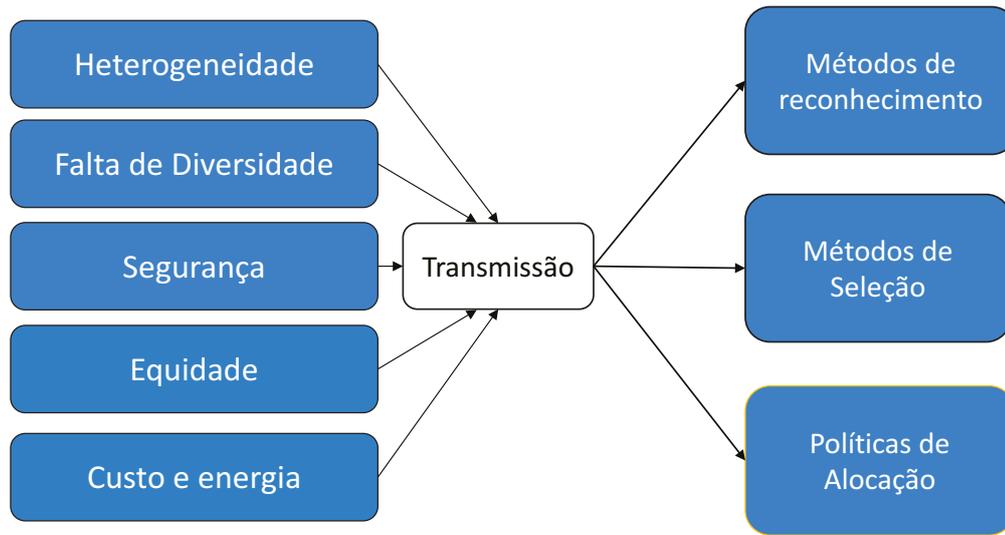


Figura 3.1: Tipos de desafios e organização das abordagens

3.2.1 Heterogeneidade

A heterogeneidade tem sido um dos grandes desafios que afetam o desempenho das transmissões multicaminhos. Isto ocorre porque os caminhos possuem diferentes características, como variação da largura de banda, atraso, taxa de perda, entre outras [Ferlin et al., 2014]. Os pacotes enviados pelos diferentes caminhos acabam chegando fora de ordem no receptor causando o problema do reordenamento e o bloqueio da fila de recebimento, denominado HOL (do Inglês, *Head Of the Line blocking*). O reordenamento degrada o desempenho da transmissão (vazão), o que tem resultado no **problema da adaptação**. Este problema se refere à diminuição da vazão da transmissão multicaminhos devido à diferença de desempenho entre os caminhos, o que impede de se aproximar do melhor desempenho [Liu et al., 2011]. Isto ocorre quando a agregação de banda entre dois ou mais subfluxos não têm um resultado significativo em termos de desempenho. Para exemplificar, dado um caminho com maior capacidade P_f e outro com menor capacidade P_s , onde C_f e C_s refere-se a capacidade de transmissão, respectivamente. A largura de banda agregada total, considerando que sejam disjuntos, é dada por $B_t = C_f + C_s$. No caso ótimo, o desempenho alcançado com a agregação, representado por B_a , seria um valor próximo de B_t . Contudo, se a largura de banda alcançada tende ao que é alcançado com apenas um dos caminhos, tem-se o problema da adaptação. Quando $B_t \rightarrow C_f$, considerando o melhor caminho, o desempenho é similar ao alcançado sem a agregação utilizando o melhor caminho. O problema é pior quando $B_t \rightarrow C_s$, ou seja, inferior ao alcançado sem a agregação e pode ocorrer em face da ausência de estratégias eficientes para mitigá-lo.

3.2.2 Falta de Diversidade

A Internet originalmente foi projetada para suportar mais que um caminho entre dois dispositivos a fim de garantir a confiabilidade e a resiliência da rede [Kurose e Ross, 2013]. No entanto, devido às regras de roteamento no núcleo da Internet e a dependência entre os protocolos TCP/IP, os fluxos de dados são comumente roteados por um mesmo caminho. Isto evita que os pacotes cheguem fora de ordem e ocasionem o problema do reordenamento [He e Rexford, 2008]. Deste modo, os protocolos de roteamento tradicionais não distinguem os fluxos pertencentes a uma conexão multicaminhos, fazendo com que eles sejam roteados por caminhos que compartilham

determinados pontos da rede. Isto resulta na falta de diversidade entre os caminhos e afeta o desempenho das transmissões, uma vez que a largura de banda agregada pode ser limitada por um gargalo compartilhado, e a resiliência afetada por pontos comuns de falhas.

Identificar e medir a diversidade dos caminhos pode ser um grande desafio [Rohrer et al., 2014]. Dado um conjunto de caminhos, a diversidade pode variar entre alta e baixa. O conjunto tem uma diversidade alta se a maioria dos caminhos são disjuntos, ou seja, não compartilham enlaces ou nós através da rede [Hu et al., 2016]. Por outro lado, a diversidade é baixa se a maioria dos caminhos são compartilhados. Os caminhos disjuntos aumentam a robustez da transmissão, provendo uma maior tolerância a falhas e um melhor desempenho, uma vez que não compartilham pontos comuns através da rede. Identificar se os caminhos são disjuntos é um grande desafio, especialmente no *cenário 3*, em que a variação do estado da rede dificulta a obtenção precisa das informações [Liao et al., 2011]. Os caminhos compartilhados têm grande impacto no princípio da eficiência, ou Ótimo de Pareto, discutido em [Khalili et al., 2012]. Este princípio estabelece que, quando existe competição entre fluxos em um gargalo, é impossível aumentar a vazão de uma conexão sem reduzir a vazão de outra, ou aumentar o custo de congestionamento. Para uma transmissão multicaminhos o desafio é ser Pareto-ótima, ou seja, um fluxo transmitido sobre múltiplos caminhos precisa ter um desempenho melhor que se fosse transmitido por um único caminho.

3.2.3 Equidade

O princípio da equidade (do Inglês, *fairness*) corresponde ao compartilhamento justo de um determinado recurso (e.g., gargalo) [Becke et al., 2012]. Um subfluxo pode ser considerado amigável se age como um único fluxo TCP sob condições de congestionamento. O objetivo da equidade é de alcançar um estado em que todos os subfluxos compartilhem a largura de banda igualmente. A transmissão multicaminhos deve respeitar o princípio da equidade *fairness* quando os subfluxos compartilham um gargalo. Neste caso, para que o compartilhamento seja justo, os subfluxos não devem ser mais agressivos que um único fluxo TCP (*TCP-friendliness* [Aydin et al., 2012]). O controle de congestionamento do MPTCP se preocupa com o princípio da equidade, uma vez que acopla os subfluxos para que se comportem como se fossem apenas um subfluxo TCP em situações de congestionamentos. O agrupamento dos recursos (subfluxos) é conhecido como o princípio de *Resource Pooling* (RP) [Wischik et al., 2008]. O princípio de RP se refere à agregação do subfluxos para que se comportem como apenas um fluxo TCP, especialmente quando compartilham gargalos. Alguns estudos mostram que o MPTCP sofre com alguns problemas, como por exemplo, não sendo Pareto-ótimo e podendo ainda ser agressivo contra fluxos TCP [Khalili et al., 2012].

3.2.4 Reordenamento

O problema do reordenamento pode ocorrer em transmissões TCP devido a questões ligadas ao roteamento dos pacotes no núcleo da rede, embora seja pouco expressivo [He e Rexford, 2008]. Contudo, seu impacto se torna mais expressivo em transmissões multicaminhos devido à heterogeneidade dos caminhos. Como os caminhos possuem diferentes características (e.g., atraso), os pacotes acabam chegando fora da ordem. O reordenamento dos pacotes afeta principalmente as aplicações de tempo real, devido ao aumento do atraso ocasionado pela espera dos pacotes faltantes. Além disso, o reordenamento gera problemas com as retransmissões, uma vez que os pacotes fora de ordem são interpretados como uma perda, e ocasiona o bloqueio HOL, devido ao esgotamento da fila de recebimento. Conforme ilustra a

Figura 3.2, o bloqueio HOL ocorre porque os pacotes faltantes impedem que os pacotes recebidos sejam entregues de forma ordenada à camada superior. Neste caso, os pacotes recebidos ficam armazenados até encherem a fila, que possui um tamanho limitado, causando o bloqueio HOL. Alocar os dados sobre os múltiplos caminhos heterogêneos sem causar um reordenamento dos dados é um grande desafio, especialmente quando existem flutuações no estado do caminho durante o processo de transmissão.

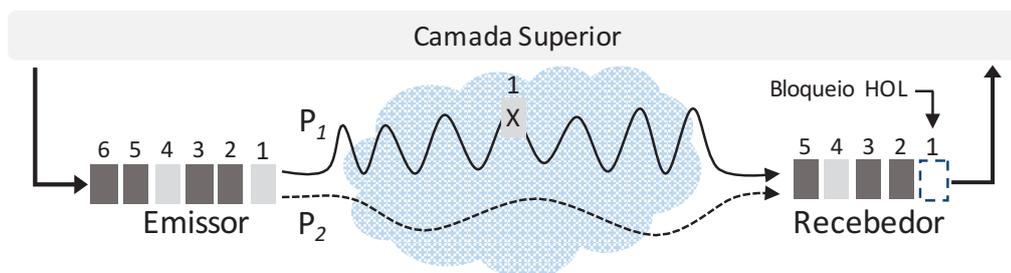


Figura 3.2: Bloqueio HOL

3.2.5 Consumo de Energia

Os dispositivos móveis possuem capacidade limitada de energia e isto impõe novos desafios à transmissão multicaminhos [Morawski e Ignaciuk, 2017b]. As interfaces de rádio de um dispositivo móvel podem divergir significativamente em termos de consumo de energia. Além da interface, o consumo de energia está associado também à capacidade de transferência e o tipo de fluxo de dados (i.e., aplicação) [Pluntke et al., 2011]. E em relação ao fluxo de dados, as aplicações de tempo real e que fazem transferência de arquivos causam impactos diferentes no consumo de energia [Peng et al., 2014]. Definir estratégias eficientes, que otimizem o consumo de energia a fim de prolongar a vida útil do dispositivo, tem sido um grande desafio [Nikraves et al., 2016].

3.2.6 Segurança

A infraestrutura de segurança existente nas redes atuais ainda não está preparada para os novos protocolos de transmissão multicaminhos [Afzal e Lindskog, 2015]. Em uma transmissão multicaminhos o emissor pode controlar a quantidade de caminhos a se utilizar e ainda alterar de forma dinâmica a alocação dos dados sobre eles. Isto permite, por exemplo, que ele faça um ataque de fragmentação, ou seja, fragmenta o fluxo de dados recebido da aplicação por caminhos diferentes a fim de dificultar a identificação por parte de sistemas que inspecionam o tráfego da rede, como NATs, *firewalls* e sistemas de detecção de intrusão. Esses sistemas comumente utilizam técnicas baseadas em assinaturas pré-definidas e que são utilizadas para detecção de anomalias. As assinaturas são baseadas em premissas e observações dos padrões de tráfego ao longo dos anos. A fragmentação e a troca de endereços faz com que a maioria das assinaturas não sejam válidas, resultando na possibilidade de se evadir destes sistemas.

Além disso, embora uma conexão multicaminhos tenha uma maior confiabilidade que as conexões TCP, isto também pode impor desafios para a segurança da rede. Com o TCP, um único caminho comprometido poderia ser usado para manipular o tráfego, mas com o MPTCP, a parte comprometida tem capacidade de manipular inúmeros caminhos. Com o advento da Internet das Coisas (IoT, do Inglês *Internet of Things*), os inúmeros dispositivos podem ser comandados e controlados para efetuar ataques em larga escala utilizando múltiplos caminhos e fazendo uso da

fragmentação dos dados, para dificultar o reconhecimento, e a troca de endereços, para se evadir dos sistemas de detecção [Pearce e Thomas, 2014].

3.3 Classificação das Abordagens

Tabela 3.1: Escalonadores: Métodos, Problemas e Cenários

Algoritmo	Método Utilizado	Problema	Cenário
DRA(2011) [Liu et al., 2011]	N/A (Pull)	Reordenamento	Heterogêneo
MDP(2011) [Pluntke et al., 2011]	N/A (Híbrido)	Consumo de energia	Heterogêneo
MPTCP-LB(2012) [Diop et al., 2012]	N/A (Pull)	Equidade	Heterogêneo
RCPS(2012) [Kim et al., 2012]	N/A (Push)	Reordenamento	Heterogêneo
CWA-MPTCP(2013) [Zhou et al., 2013]	N/A (Push)	Reordenamento	Heterogêneo
EAPC(2013) [Yang et al., 2013a]	Push	Bloqueio HOL	Heterogêneo
eMTCP(2013)[Chen et al., 2013]	Pull	Consumo de energia	Heterogêneo
DAPS(2014) [Kuhn et al., 2014]	N/A (Push)	Reordenamento	Heterogêneo
HDDS(2014) [Bhat e Talmale, 2014]	Híbrido	Reordenamento	Heterogêneo
F ² P-DPS(2014) [Ni et al., 2014]	N/A (Pull)	Reordenamento	Heterogêneo
NTS(2014) [Liu et al., 2014]	Híbrido	Reordenamento	Heterogêneo
OTIAS(2014) [Yang et al., 2014]	Push	Bloqueio HOL	Heterogêneo
RTT-CTS(2014) [Chen et al., 2014]	N/A (Push)	Reordenamento	Heterogêneo
TSC(2014) [Park et al., 2014]	N/A (Push)	Reordenamento	Heterogêneo
AMPTCP(2015) [Li et al., 2015]	N/A (Pull)	Reordenamento	Homogêneo
CPS(2015) [Oh e Lee, 2015]	Híbrido	Reordenamento	Heterogêneo
FDPS(2015) [Le e Bui, 2015]	Pull	Reordenamento	Heterogêneo
FrPS(2015) [Hwang e Yoo, 2015]	N/A (Pull)	Reordenamento	Heterogêneo
MPIS(2015) [Popat et al., 2015]	N/A (Híbrido)	Heterogeneidade	Homogêneo
OCPS(2015) [Ni et al., 2015]	N/A (Pull)	Reordenamento	Heterogêneo
ReMPTCP[Frömmgen et al., 2016]	N/A (Pull)	Heterogeneidade	Heterogêneo
S-EDPF[Garcia-Saavedra et al., 2015]	N/A (Pull)	Heterogeneidade	Heterogêneo
BLEST(2016) [Ferlin et al., 2016]	N/A (Pull)	Bloqueio HOL	Heterogêneo
CLS(2016) [Corbillon et al., 2016]	N/A (Push)	Bloqueio HOL	Heterogêneo
EDAM(2016) [Wu et al., 2016]	N/A (Push)	Consumo de energia	Heterogêneo
EEDLB(2016) [Morawski e Ignaciuk, 2017a]	N/A (Pull)	Energia	Heterogêneo
JAPS(2016) [Chan et al., 2016]	N/A (Pull)	Bloqueio HOL	Heterogêneo
MA ² (2016) [Ke et al., 2016]	N/A (Pull)	Reordenamento	Heterogêneo
MLCS(2016) [Hu et al., 2016]	N/A (Pull)	Reordenamento	Heterogêneo
MPTCP-L(2016) [Wang et al., 2016]	N/A (Pull)	N/A	Homogêneo
M-SEDPF[Partov e Leith, 2016]	N/A (Pull)	N/A	Heterogêneo
OJCS(2016) [Ou et al., 2016]	N/A (Push)	Reordenamento	Heterogêneo
rMTCP(2016) [Hunger e Klein, 2016]	N/A (Pull)	Heterogeneidade	Heterogêneo
RSS(2016) [Liu et al., 2016a]	N/A (Push)	N/A	Heterogêneo
SWTC(2016) [Tsai et al., 2016]	N/A (Push)	Reordenamento	Heterogêneo

A Tabela 3.1 apresenta a classificação das abordagens estudadas neste trabalho conforme as categorias: metodologia implementada; problema abordado; e o tipo de cenário relacionado. Em relação à classificação, algumas propostas não deixam claro os detalhes de implementação. Nestes casos, elas são classificadas como *Não Abordado* (N/A), e a estratégia utilizada é inferida

com base na descrição da proposta e colocada entre parênteses. Em relação ao problema, as abordagens são classificadas conforme o desafio principal que é abordado. Alguns desafios são consequências de outros, como do bloqueio HOL causado pelo problema do reordenamento. Deste modo fica implícita esta relação. Em relação ao cenário, o mesmo foi definido com base no tipo de rede descrito nos trabalhos. Por exemplo, se o trabalho não deixa claro o tipo de rede, mas lida com o problema de reordenamento devido à diversidade dos caminhos, estes são classificados como heterogêneos.

Além da classificação acima, as abordagens são apresentadas conforme a organização descrita na Figura 3.3, sendo: métodos de reconhecimento; métodos de seleção; políticas de alocação dos dados. Os métodos de reconhecimento e seleção possuem forte relação e muitas vezes são apresentados como parte de um único processo. Os métodos de reconhecimento têm como função identificar informações e características sobre o estado dos caminhos. Os métodos de seleção são responsáveis por classificar e selecionar os caminhos com base nas informações, que por vezes são obtidas através dos métodos de reconhecimento. As políticas de alocação consistem de regras para designação dos segmentos de dados através dos múltiplos caminhos. As próximas seções descrevem as abordagens estudadas conforme os métodos descritos acima.

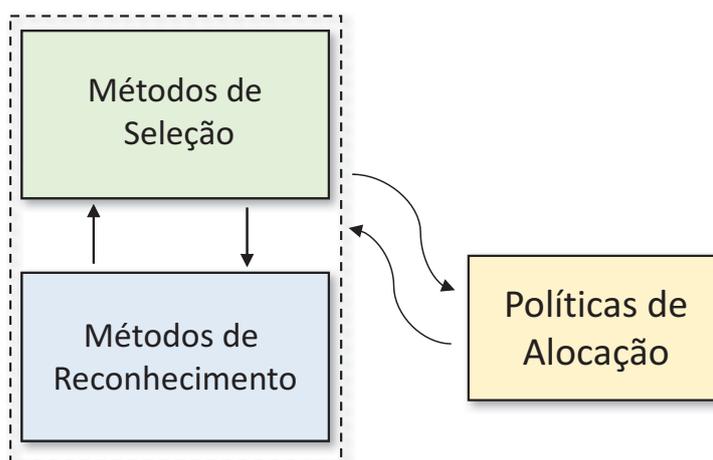


Figura 3.3: Organização das abordagens

3.3.1 Métodos de Reconhecimento

Os métodos de reconhecimento são responsáveis por identificar diferentes informações sobre os caminhos, como ilustrado na Figura 3.4. As informações mais comuns são o RTT e o tamanho da janela de congestionamento (CWND). Contudo, devido à heterogeneidade e à diversidade dos caminhos estas informações não têm sido suficientes ou obtidas com a precisão desejada. Neste sentido, novos métodos e métricas têm sido propostos para obter informações novas ou mais precisas, como métricas que estimam o atraso de uma via (ida/volta), a capacidade e qualidade dos caminhos. Além disso, alguns trabalhos propõem técnicas para estimar o consumo de energia do caminho, conforme a interface de rede sem fio utilizada, e também, identificar se os caminhos são compartilhados através de técnicas estatísticas (e.g., correlação). Por fim, são apresentados os métodos que propõem obter informações dos caminhos através de coleta de dados em outras camadas e métodos de sondagem que coletam informações dos caminhos.

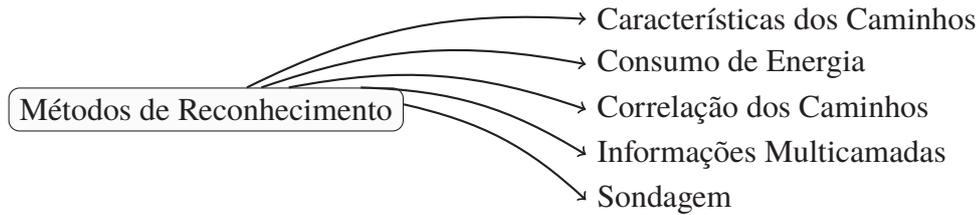


Figura 3.4: Organização dos Métodos de Reconhecimento

Características dos caminhos

Atraso: muitos trabalhos consideram a estimativa do atraso de diferentes formas. O caso mais trivial é considerar as medições do RTT, como feito pelo TCP. O TCP usa o RTT para calcular o RTO ou detectar um congestionamento, assumido ambos os atrasos de ida e volta como $RTT/2$. Neste caso o atraso medido se refere a um valor médio atribuído igualmente aos caminhos, assumindo que são simétricos. Na prática os caminhos tendem a ser assimétricos, com tempo de ida diferente do tempo de volta [Li et al., 2016]. Neste sentido muitas políticas têm se preocupado em determinar o atraso de uma via (OWD, do Inglês *One-Way Delay*) de ida e de volta de modo confiável. A estimativa precisa do OWD pode melhorar o desempenho da transmissão, uma vez que o escalonador pode optar sempre pelo caminho com menor atraso de ida. Teoricamente a sincronização de relógios entre origem-destino pode garantir a estimativa precisa do OWD. Porém isto não tem sido praticado devido à dificuldade de garantir a sincronização dos relógios em dispositivos conectados através da Internet. Um dos maiores objetivos de estimar o OWD se refere à possibilidade de reduzir o atraso de entrega dos dados. Em uma transmissão multicaminhos, além de conhecer o atraso de ida (OWD_f), conhecer o atraso de volta (OWD_b) pode permitir que os ACKs sejam retornados pelos caminhos com menor atraso e os pacotes sejam reconhecidos mais rapidamente.

Em [Coudron et al., 2015], os autores propõem um método de marcação diferencial para obter uma estimativa da diferença dos OWD s entre os diferentes caminhos. O método não requer a sincronização dos relógios e nem a cooperação da rede e faz a marcação do tempo exclusivamente no emissor. O objetivo é obter uma estimativa para o OWD_b mais precisa que o $RTT/2$. O RTT de um dado subfluxo S_i pode ser obtido com a Equação 3.1, sendo f_i e b_i o atraso de ida e de volta, respectivamente. A técnica utiliza dois caminhos, por definição, um caminho rápido (P_f), que possui o menor RTT, e um caminho lento (P_s), que possui um RTT maior. O atraso de ida e de volta de cada caminho é dado por $f_{P_f} = \min_i(f_i)$, $f_{P_s} = \max_i(f_i)$, $b_{P_f} = \min_i(b_i)$ e $b_{P_s} = \max_i(b_i)$. O OWD_f e OWD_b são obtidos por $\Delta^F OWD = (f_{P_s} - f_{P_f})$ e $\Delta^B OWD = (b_{P_s} - b_{P_f})$, respectivamente. O método proposto causa o recebimento fora de ordem de pacotes, contudo, tende a ocorrer menos se o método não fosse empregado.

$$RTT_i = f_i + \text{atraso de processamento} + b_i \quad (3.1)$$

O método proposto por Le e Bui (2015) [Le e Bui, 2015] estima apenas o atraso de ida (OWD_f) dos caminhos. O método consiste de enviar dois pacotes por dois caminhos diferentes (P_1 e P_2), marcando a diferença entre os tempos de envio na origem ($\Delta_o = t_{P_1}^o - t_{P_2}^o$) e a diferença entre os tempos de recebimento no destino ($\Delta_d = t_{P_1}^d - t_{P_2}^d$). Assim, o valor resultante entre $\Delta_d - \Delta_o$ é empregado para calcular o OWD_f dos caminhos P_1 e P_2 . Yang e Amer (2014) [Yang e Amer, 2014] propõem um método para estimar o OWD considerando também o tempo de retransmissão. Neste caso o atraso medido se refere ao tempo em que o PDU é enviado

pelo subfluxo pela primeira vez até o momento em que o PDU ou uma retransmissão do PDU chegue pela primeira vez em ordem na fila de recebimento do subfluxo.

Capacidade dos Caminhos: a largura de banda disponível (ABW, do Inglês *Available Bandwidth*) dos caminhos pode ser estimada a partir das mudanças nas amostras do RTT [Zhao et al., 2015]. Contudo, estas informações não podem prever a perda de pacote com confiabilidade, e conseqüentemente, não são precisas para estimar a capacidade dos caminhos [Liu et al., 2011]. Uma estimativa precisa previne ultrapassar a capacidade dos caminhos gerar congestionamentos [Yang et al., 2013b]. [Park et al., 2016] propõem um esquema para calcular a ABW dos caminhos modificando o algoritmo TFRC (do Inglês, *TCP-Friendly Rate Control*). O TRFC é um algoritmo adequado para transmissões de vídeo, porém, não se aplica em redes com perdas. Para resolver este problema, os autores modificam o TFRC para incluir o tempo de retransmissão dos pacotes no cálculo do RTT. Isto torna o algoritmo mais preciso para estimar a ABW. Em [Tsai et al., 2010], os autores propõem um esquema para estimar a ABW usando um método denominado PGM (do Inglês, *Packet Gap Mode*). O princípio do PGM é transmitir pacotes de sondagem em períodos de tempo contínuos (Δ_{in}), e conforme recebe confirmação, calcula a ABW. A estimativa precisa do PGM causa uma grande sobrecarga, devido ao grande número de pacotes de sondagem. Neste caso, o trabalho propõem ajustar o intervalo de sondagem, o PIT (do Inglês, *Probe Interval Time*), com base nos resultados estimados. Se o tempo estimado estiver próximo do anterior aumenta-se o PIT, do contrário, reduz-se.

Qualidade dos Caminhos: Em [Baidya e Prakash, 2014], Baydia e Prakash (2014) propõem uma técnica para classificar os caminhos em termos de qualidade. A qualidade do caminho é uma métrica dada pela razão entre o número de dados transmitidos entre duas perdas de pacote pelo RTT do caminho mais rápido. A métrica que define a qualidade é $\ell_P(t)/RTT(t)^2$, onde $\ell_P(t) = \max\{\ell_{1P}(t), \ell_{2P}(t)\}$, sendo $\ell_{1P}(t)$ o número de *bits* transmitidos com sucesso sobre o caminho P entre duas perdas, e ℓ_{2P} o número de *bits* transmitidos com sucesso pelo caminho P depois da última perda. Se a razão entre a qualidade do caminho P_i em relação ao caminho mais rápido, dado pela métrica R (Equação 3.2), for maior que um limiar pré-definido, o caminho então é marcado como ruim.

$$R = \frac{\ell_{P_{max}}/RTT_{min}^2}{\ell_{P_i}/RTT^2} \quad (3.2)$$

Em [Zhao et al., 2015], os autores propõem medir a qualidade do caminho com base no mecanismo de incremento aditivo e decremento multiplicativo (AIMD, do Inglês *Additive increase/multiplicative decrease*). O AIMD-PQ (do Inglês, *Path Quality Based Multipath AIMD*) utiliza o RTT de cada caminho para formular a estimativa de qualidade (Equação 3.3).

$$Q_p^{(i)} = 1/\log(1 + RTT_p^{(i)}) \quad (3.3)$$

O método proposto por [Liu et al., 2016b] estima a qualidade dos caminhos coletando estatisticamente informações do tamanho das filas de envio e recebimento. Para obter uma estimativa mais precisa das informações, os autores propõem que cada caminho tenha uma fila de recebimento individual, antes da fila única de recebimento. Com isto, a ABW de cada caminho pode ser estimada a partir da Eq. 3.4, onde $rcvsize_i$ e $sndsize_i$ se referem à quantidade de *bytes* que chega ao receptor e à quantidade de dados enviada através do caminho i no período atual, respectivamente. N_p indica o número de caminhos disponíveis, r_i a confiabilidade do caminho, T_{aver} a confiabilidade de toda a transmissão e ABW_i a largura de banda do caminho i . De acordo com a ABW_i , o método estima a largura de banda no próximo intervalo (bw'_i). Na primeira

medição $ABW'_i \leftarrow ABW_i$, subsequentemente $ABW'_i = (1 - \gamma) \times ABW'_i + \gamma \times ABW_i$, sendo γ um fator que indica o impacto da predição anterior na predição atual.

$$\begin{aligned}
 r_i &= \frac{rcvsize_i}{sndsize_i} \\
 T_{aver} &= \frac{\sum_{j=1}^{N_p} sndsize_j}{\sum_{j=1}^{N_p} rcvsize_j} \\
 ABW_i &= ABW_i \times \frac{r_i}{T_{aver}}
 \end{aligned} \tag{3.4}$$

Consumo de Energia

As interfaces de rádio diferem de forma significativa em termos de consumo de energia. Em [Pluntke et al., 2011], os autores propõem gerar externamente uma tabela com informações estáticas sobre o desempenho energético de cada interface e carregar no dispositivo para que seja utilizada nas decisões de escalonamento. [Morawski e Ignaciuk, 2017a] propõem um método que obtém informações a partir da observação dos parâmetros dinâmicos das interfaces de rede (vazão e consumo de energia). As informações são utilizadas para ajustar dinamicamente a capacidade de transmissão de cada caminho.

Correlação dos Caminhos

Muitas propostas assumem que os caminhos são disjuntos, mas isto é raramente válido em redes reais, como a Internet. Identificar os caminhos correlacionados (compartilhados) e o grau de correlação entre eles permite alcançar uma maior robustez na transmissão e também balancear o uso dos caminhos. Além disso, os caminhos correlacionados/compartilhados limitam o desempenho e reduzem a resiliência da transmissão, uma vez que podem compartilhar gargalos e falhas [Liao et al., 2011].

Em [Liu et al., 2011], os autores propõem um coeficiente de correlação para medir o quanto um caminho intersecta com outro. Assumindo valores entre $[0, 1]$, onde 1 indica que os caminhos são totalmente independentes e 0 indica que têm grande correlação entre si. Basicamente o coeficiente de correlação é dado por $(T_{P_i} / \sum T_{P_i})$, onde T_{P_i} se refere à vazão de um determinado caminho P_i . [Hu et al., 2016] também propõem um método para estimar a correlação (u) entre os caminhos assumindo valores entre $u = [0, 1]$. Se os caminhos são totalmente independentes, sendo $u = 1$, a transmissão multicaminhos pode alcançar uma vazão total próximo da soma da vazão dos caminhos. Do contrário, se o valor de u se aproxima de 0, os caminhos têm grande correlação e a vazão total fica limitada pelo gargalo. Diferente do trabalho anterior, o método acima calcula a correlação apenas para dois caminhos, sendo P_1 e P_2 . A correlação entre os caminhos pode ser obtida por $U_{P_1-P_2} = T_{total} / (T_{P_1} + T_{P_2})$, sendo T_{P_1} e T_{P_2} a vazão individual de cada caminho. Por fim, [Krupakaran et al., 2015] adaptaram o protocolo TraceFlow [Narasimhan et al., 2012] para rastrear o caminho de um subfluxo através da rede. Uma vez que o TraceFlow coleta informações sobre os enlaces de um determinado caminho na rede, isto é repassado ao protocolo de transporte multicaminhos para encaminhar os subfluxos por caminhos distintos.

[Kou et al., 2013] identificam a correlação dos caminhos utilizando a análise relacional cinza (IGA, do Inglês *Grey Relational Analysis*) para selecionar os caminhos para a transmissão multicaminhos. A análise relacional cinza trata-se de um método estatístico para estimar o

comportamento de um sistema incerto. A ideia básica foi estimar a amplitude do grau de proximidade entre os caminhos determinando o grau de similaridade entre uma sequência de dados de referência com outra sequência de dados de comparação. O método proposto compara dinamicamente a largura de banda e o RTT de forma quantitativa, com base no nível de similaridade e variabilidade para definir e analisar o grau de relacionamento entre eles.

Informações Multicamadas

A obtenção de informações de outras camadas visa melhorar as decisões de escalonamento referentes à seleção dos caminhos e alocação dos dados. A seleção de caminhos pode se beneficiar de informações das camadas mais baixas, como por exemplo, identificar previamente variação no estado do caminho, devido às oscilações do sinal de rádio (atenuação, sombreamento, interferências, etc.). Isto permite que o escalonador possa reagir rapidamente às mudanças. O método proposto por [Lim et al., 2014] destaca os benefícios em se obter informações de outras camadas. O método estima a qualidade dos caminhos associados com a interface WiFi através de informações obtidas da camada de enlace (MAC, do Inglês *Medium Access Control*). Os autores observam que quando ocorre uma desconexão com a rede WiFi, o número de retransmissões de quadros fica maior que o número de transmissões bem-sucedidas. Estas informações são coletadas e ficam disponíveis para o mecanismo de seleção de subfluxos do escalonador. As informações das camadas superiores também auxiliam as decisões de escalonamento. Em [Corbillon et al., 2016], os autores propõem um método que reúne informações da camada de transporte e aplicação para reordenar a transmissão dos dados e priorizar as partes mais significativas de uma transmissão de vídeo. Um dos objetivos do trabalho é possibilitar a interação entre a camada de transporte e aplicação, de modo que o escalonador esteja ciente sobre o conteúdo, os requisitos da aplicação e as condições da rede.

Métodos de Sondagem

Durante o tempo de vida de uma transmissão o desempenho dos caminhos pode variar em função das flutuações do estado da rede. Algumas estratégias sugerem remover os caminhos ruins a fim de reduzir o impacto do reordenamento [Oh e Lee, 2015]. Contudo, o estado do caminho ruim pode mudar, e ele novamente se tornar apto a fazer parte do conjunto de caminhos utilizados pela transmissão. Para identificar estas mudanças de estado dos caminhos que não estão sendo utilizados, são propostos alguns mecanismos de sondagem. Em [Chen et al., 2011], os autores propõem um mecanismo de sondagem ativa que encaminha pacotes de dados sobre os caminhos não utilizados. A decisão de realizar a sondagem em determinado caminho se baseia na qualidade do pior caminho a cada 4 RTTs. Para medir a diferença entre o pior caminho e os demais, se utiliza a métrica $Q = P_{pior}/\bar{P}$, sendo P_{pior} a vazão do pior caminho e \bar{P} a vazão média dos demais caminhos. Além da métrica Q , a sondagem de cada caminho ocorre com uma probabilidade $P = (1 - Q)(1 - p) + p$. Além de Q , a equação utiliza o fator p , que permite que mesmo se todos os caminhos tiverem uma vazão igualmente ruim, novos caminhos podem ser sondados. O método de sondagem proposto por [Oh e Lee, 2015] permite que pacotes de dados sejam encaminhados periodicamente sobre os caminhos que foram marcados como não utilizáveis. O método utiliza dois esquemas de sondagem, um estático e outro dinâmico. A sondagem estática estima o estado do caminho considerando um intervalo de tempo estático. Do contrário, o esquema dinâmico o intervalo de sondagem se ajusta conforme o número de interrupções do caminho.

3.3.2 Métodos de Seleção

A Figura 3.5 ilustra a organização dos tipos de informações ou requisitos utilizados para classificar e selecionar um ou mais caminhos.

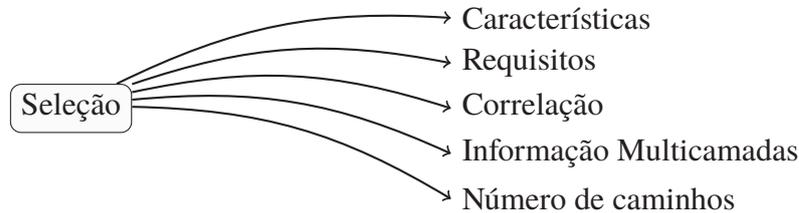


Figura 3.5: Organização dos Métodos de Seleção

Características

Os métodos de seleção baseados nas características dos caminhos, como atraso, largura de banda disponível e vazão, são os mais empregados pelas diferentes propostas de escalonadores de pacotes. O escalonador LRF (do Inglês, *Lowest RTT First*) emprega este tipo de método para classificar os caminhos em ordem crescente e selecionar sempre aquele com menor RTT. Assim como o LRF, as abordagens apresentadas em inúmeros trabalhos na literatura propõem métodos de seleção de caminhos com base no RTT [Liu et al., 2011], [Ni et al., 2014],[Chen et al., 2014],[Popat et al., 2015],[Ke et al., 2016], [Hunger e Klein, 2016],[Liu et al., 2016b]. As propostas que implementam o método *Pull* (Capítulo 2, Seção 2.4.1) consideram também o espaço disponível na CWND como requisito para selecionar ou não os caminhos. Como visto na seção anterior os caminhos podem ser assimétricos, e o RTT de ida ser muito diferente do RTT de volta. Neste sentido algumas propostas fazem a seleção utilizando informações assimétricas do atraso, calculando o *OWD* [Yang et al., 2014, Le e Bui, 2015]. Além de classificar os caminhos com base no RTT, [Hwang e Yoo, 2015] comparam os caminhos entre si e medem a diferença entre eles. Os caminhos são então selecionados apenas se a diferença entre o melhor e pior caminho for menor que um limiar pré-definido. Os trabalhos [Liu et al., 2014, Yang et al., 2013a, Tsai et al., 2010] selecionam os caminhos com base na ABW. Além das informações que são obtidas no emissor, existem métodos que selecionam os caminhos com base em informações do receptor. Em [Oh e Lee, 2015], os autores propõem um método que seleciona os caminhos com base no RTT e no tamanho residual da fila de recebimento.

O desafio enfrentado por estes métodos é garantir que as informações utilizadas para selecionar e classificar os caminhos não estejam desatualizadas. Em [Jin et al., 2015], os autores propõem um mecanismo de seleção com base em um modelo de predição da vazão e da largura de banda disponível. O mecanismo de seleção proposto é detalhado a seguir. Sendo $|P|$ o tamanho do conjunto de caminhos disponíveis, e cada caminho tendo os parâmetros (RTT, l, BW) , que denotam o RTT do caminho, a taxa de perda e a largura de banda disponível, respectivamente. A Equação 3.5 visa prever a vazão de cada caminho. O parâmetro b se refere ao número de pacotes reconhecidos em cada ACK (definido como $b = 1$). A partir desta equação, a vazão máxima pode ser definida como $T_{max} = \max_{i \in P} T_i$ se a largura de banda disponível dos melhores caminhos for $T_k = T_{max}$ e $T_k < BW_k$. Neste caso, o conjunto dado por $S = k | T_k = T_{max}, T_k < BW_k$ representa o melhor esquema de seleção de caminhos. Contudo, se existe uma limitação da largura de banda dos melhores caminhos, se assume $T_k = T_{max}$ e $T_k > BW_k$. Neste caso o conjunto $S_{init} = k | T_k = T_{max}, T_k > BW_k$ é utilizado como esquema de seleção inicial e os autores

empregam um método que mede a similaridade entre os caminhos para escolher o número de caminhos ideal para alcançar a largura de banda necessária. Este método de seleção de caminhos é estático. Se as características das redes mudam constantemente, ele precisa atualizar os parâmetros para executar novamente o método.

$$T = \frac{1}{RTT} * \sqrt{\frac{3}{2bp}} * packet_size \quad (3.5)$$

O método proposto por Badya e Prakash (2014) [Baidya e Prakash, 2014] classifica os caminhos a partir de uma métrica de qualidade. Conforme a métrica, um caminho (P_i) tem uma qualidade ruim se a razão entre ele e o caminho mais rápido (P_f) for maior que um limiar pré-definido. Em [Zhao et al., 2015], os autores classificam os caminhos com base na métrica AIMD-PQ, em que um RTT muito alto resulta numa classificação ruim do caminho.

Requisitos

Algumas abordagens utilizam os requisitos das aplicações (e.g., atraso, largura de banda, e outros) ou dos usuários (e.g., desempenho, custo, energia) como critério de seleção e classificação. Neste caso, os métodos de seleção podem melhorar o desempenho, selecionando os caminhos que atendem os requisitos de QoS das aplicações, e a tolerância a interrupções, otimizando o consumo de energia e prolongando a vida útil do dispositivo móvel, por exemplo.

Qualidade de Serviço: [Diop et al., 2011] propõem um método que seleciona o caminho a partir dos requisitos de QoS do fluxo de dados. Neste caso, os caminhos são selecionados conforme a largura de banda disponível que atenda a transmissão de um fluxo de vídeo, especificamente que suporte os requisitos mínimos para transmissão de quadros do tipo *I*. Os demais caminhos são selecionados para encaminhar os outros tipos de quadros (e.g., *P* e *B*). [Hu et al., 2016] propõem selecionar os caminhos que podem atender os requisitos de desempenho do fluxo de dados (Aplicação). Com base em informações do RTT, taxa de perda e correlação, os caminhos são classificados e, aqueles que podem atender a demanda dos serviços e melhorar a vazão, são selecionados como ativos. Os demais caminhos ficam como inativos.

Consumo de energia: além destes requisitos, os caminhos podem ser selecionados com base no custo (monetário) de transmissão e o consumo de energia de cada interface. [Pluntke et al., 2011] propõem um método para selecionar dinamicamente qual interface de rede será utilizada para transmissão com base em informações geradas previamente sobre a eficiência energética de cada interface. Os resultados apresentados por [Kaup et al., 2015] demonstram que utilizar apenas uma interface pode otimizar o consumo de energia desde que a seleção escolha sempre aquela mais eficiente. Contudo, se a agregação de múltiplas interfaces for inevitável, os autores sugerem priorizar a seleção da interface com menor RTT.

Custo: o IETF não cobre as questões práticas acerca do custo de transmissão de diferentes tecnologias de acesso [Ford et al., 2011]. Em muitos casos, dependendo dos requisitos das aplicações ou dos usuários, existe uma complexa relação entre custo e desempenho que tem grande impacto na seleção de caminhos. Em [Secci et al., 2014], os autores estudam a importância desta relação, entre custo e desempenho, utilizando teoria dos jogos não cooperativa. Embora não trate especificamente sobre a seleção do subfluxo, mas sim da distribuição, está implícito o quão importante deve ser o método empregado para selecionar os subfluxos conforme os requisitos de custo-desempenho.

Correlação

Após identificar a correlação entre os caminhos (Seção 3.3.1), o mecanismo de seleção tem como tarefa escolher aqueles que possuem a menor correlação entre si. Isto evita que determinados subfluxos compartilhem o mesmo gargalo no núcleo da rede e também permite determinar o número ideal de subfluxos que podem trazer benefícios para a transmissão. [Krupakaran et al., 2015] propõem um método que seleciona apenas os caminhos disjuntos entre origem-destino e remove os que são identificados como redundantes e com sobreposição. Além disso, o método proposto ajusta os parâmetros dos subfluxos para que eles utilizem sempre a mesma rota durante o tempo de vida da conexão. [Kou et al., 2013] propõem um método de seleção baseado no grau de similaridade dos caminhos, obtidos a partir do método estatístico IGA. O objetivo é selecionar os caminhos que têm pouca correlação para melhorar a vazão total.

Informação Multicamadas

A Seção 3.3.1 apresenta alguns métodos de identificação que obtêm informações de outras camadas para auxiliar nas decisões do escalonamento. Estas informações são consideradas pelos métodos de seleção para classificar e selecionar os caminhos. As informações podem ser locais, provenientes do próprio dispositivo, ou globais, coletadas por outros protocolos situados em diferentes pontos da rede. Em [Lim et al., 2014] e [Corbillon et al., 2016] os autores utilizam informações locais obtidas de camadas adjacentes, como dados da interface sem fio e dados da aplicação, respectivamente. [Liu et al., 2016a] obtêm informações globais a partir de uma abordagem centralizada. O método proposto centraliza as informações enviadas periodicamente pelos dispositivos móveis para que, a partir de informações globais do estado da rede, o MPTCP decida se abre ou fecha um determinado subfluxo. A seleção do segundo subfluxo (interface sem fio) se baseia na relação entre a melhor referência de qualidade de sinal recebido (RSRQ, do Inglês *Reference Signal Received Quality*) e o segundo melhor RSRQ coletados de modo centralizado e repassados ao escalonador. [Chen et al., 2011] propõem um mecanismo de seleção de subfluxos a partir de informações obtidas através da sondagem do estado dos caminhos. O mecanismo de sondagem coleta informações e encaminha através do protocolo XCP (do Inglês, *Explicit Congestion Control*) ao MPTCP. Com as informações, o mecanismo de seleção faz a troca do subfluxo que estão sobre os piores caminhos (menor capacidade) para os caminhos melhores, desde que a sondagem seja consistente em longa escala.

Número de caminhos

O número de caminhos entre origem-destino nem sempre corresponde ao número ativo de subfluxos. O MPTCP, por exemplo, possui o modo de operação (*Full-Mesh*), que cria um subfluxo para cada par de endereço disponível entre origem-destino [Paasch et al., 2012]. Deste modo, todos os caminhos disponíveis serão cobertos por um subfluxo. Contudo, em outros possíveis modos de operação, seja do próprio MPTCP (*backup*, *Single Path*) ou aqui não identificados, nem todos os caminhos existentes entre origem-destino são associados a um subfluxo ativo. As situações para isto são diversas. A Figura 3.6 ilustra três possíveis situações em relação a seleção de caminhos: (a) o número de subfluxos corresponde ao número de caminhos; (b) número de subfluxos é maior que o número de caminhos; (c) o número de subfluxos é menor que o número de caminhos.

O número de caminhos corresponde à combinação entre o número de endereços disponíveis entre origem-destino. Os caminhos identificados na figura com um número correspondem a uma 4-tupla (endereço de origem, porta de origem, endereço de destino, porta de destino).

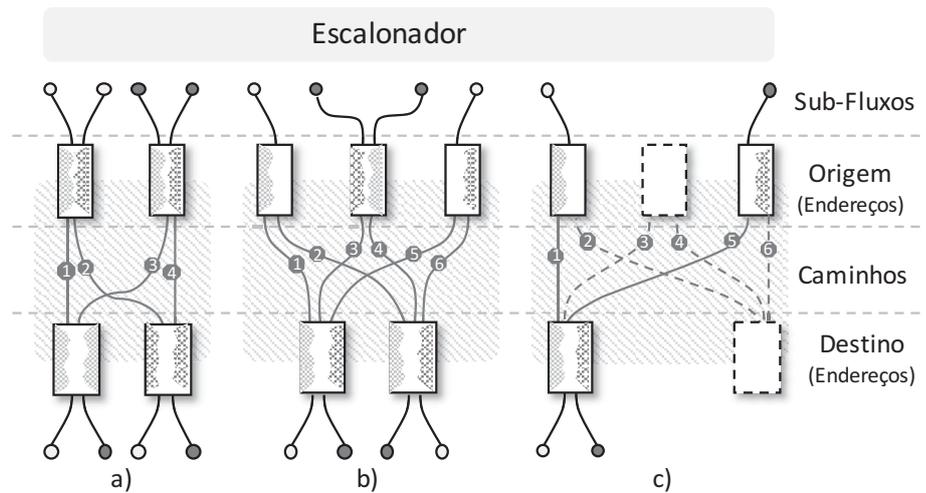


Figura 3.6: Número de caminhos vs. número de subfluxos

Nos casos mais triviais cada endereço corresponde a uma interface de rede. Na situação (a), o número de subfluxos pode ocasionar uma sobrecarga, sem trazer muitos benefícios, se o objetivo da transmissão é alcançar um maior vazão. Isto porque os subfluxos de uma mesma transmissão compartilham a mesma interface. Na situação (b), como o número de caminhos é maior, o mecanismo de seleção poderia optar em selecionar aqueles que são disjuntos ou que proveem maior tolerância a falhas. E por fim, na situação (c), o número de subfluxos não aproveita toda a capacidade disponível, uma vez que existem endereços e caminhos não utilizados. Neste caso, os subfluxos poderiam estar balanceados entre os caminhos, como por exemplo, o subfluxo que utiliza o caminho 5 poderia ser transferido para o caminho 6.

Em [Li et al., 2015], os autores propõem um método para ajustar dinamicamente o número de subfluxos de acordo com a demanda da aplicação. O objetivo é minimizar o número de subfluxos e a sobrecarga de escalonamento sem sacrificar a vazão. O método utiliza um limite para a vazão (T) e cada subfluxo possui um contador para registrar o número de *bytes* transmitidos em um intervalo de tempo. O método zera o contador no início de cada intervalo. A vazão de um determinado subfluxo é dada por $r_{ij} = b_{ij}/\Delta$, sendo Δ o tamanho do intervalo, b_{ij} o número de bytes transferidos durante o j^{esimo} intervalo para o i^{esimo} subfluxo. Se todos os subfluxos alcançam o limite definido para a vazão $r_{ij} > T$ o mecanismo aumenta o número de subfluxos. Do contrário, se a vazão de um subfluxo for menor que $T/2$ por um longo período de tempo ($10 \times \Delta$), o método reduz o número de subfluxos. Além do número de subfluxo, a escolha de qual será utilizado inicialmente para abrir a conexão tem consequências no desempenho da transmissão multicaminhos. A partir de análises numéricas e um modelo dirigido, [Arzani et al., 2014] demonstram que a seleção do subfluxo inicial pode reduzir o problema do reordenamento. Os resultados do trabalho indicam que iniciar com o subfluxo que possui um RTT mais alto pode melhorar o desempenho da transmissão.

3.3.3 Políticas de Alocação

As políticas de alocação estabelecem as regras para designar os pacotes através dos múltiplos caminhos. Estas regras variam conforme os objetivos das políticas, que por sua vez, podem estar ligados a diferentes requisitos, como por exemplo, reduzir os efeitos da heterogeneidade em uma transmissão multicaminhos. Neste sentido, esta seção organiza os principais trabalhos com base no objetivo da política de alocação, ilustrados pela Figura 3.7.

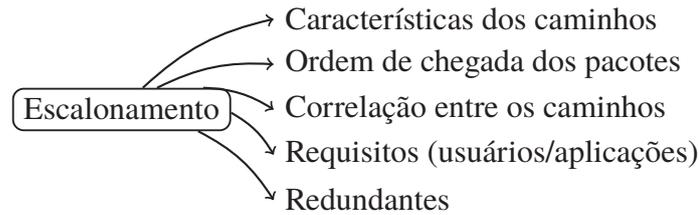


Figura 3.7: Organização das Políticas de Alocação

Características dos caminhos

As abordagens TSC [Park et al., 2014], MA^2 [Ke et al., 2016], AMPTCP [Li et al., 2015], SWTC [Tsai et al., 2016] e APS [Chan et al., 2016] designam os pacotes para os subfluxos com menor atraso ou menor tempo de entrega. Estes métodos são voltados para um cenário com caminhos heterogêneos. O MPIS [Popat et al., 2015] também designa os pacotes para os caminhos com menor RTT, contudo tem como objetivo otimizar o desempenho de aplicações em uma rede de centro de dados (cenário homogêneo). A política empregada pelo EAPC [Yang et al., 2013a] favorece os caminhos com maior largura de banda disponível (ABW).

O MA^2 inicialmente monitora informações dos caminhos, como RTT e CWND. Em seguida ordena de forma ascendente pelo RTT. Se existem caminhos com mesmo RTT utiliza-se também a CWND. De forma adaptativa transmite os pacotes conforme ordem dos caminhos. Os dados são escalonados para o primeiro caminho da lista ordenada ou para o próximo, caso não tenha espaço na CWND. O AMPTCP também possui uma abordagem adaptativa. Neste caso, o escalonador adapta o número de subfluxos conforme a ocupação dos mesmos. Se a ocupação for grande (excede um limiar) o escalonador abre um novo subfluxo. Se a ocupação for pequena, ele remove um subfluxo.

O SWTC considera um cenário colaborativo de compartilhamento de banda. Esta proposta se baseia no estado do caminho entre o nó adjacente, que compartilha a banda, e o servidor proxy, utilizado para coordenar as conexões MPTCP. O método proposto pelo JAPS [Chan et al., 2016] tem como objetivo reduzir as variações de atraso devido às flutuações do estado dos caminhos. Quando rajadas de atrasos são detectadas pelo emissor, o JAPS retransmite imediatamente os pacotes atrasados através do caminho mais rápido. Para minimizar o impacto das retransmissões, apenas os pacotes que têm possibilidade de chegar antes no receptor são retransmitidos. Embora voltado para o cenário mais homogêneo, o MPIS escolhe o melhor subfluxo com base na disponibilidade, situação de congestionamento e RTT.

O método proposto pelo CPS [Oh e Lee, 2015] realiza o escalonamento de pacotes conforme o tamanho residual da fila de recebimento e o atraso de cada caminho (RTT_i). O $RWND_r$ é um valor preditivo sobre as condições dos caminhos e pode estar superestimado ou subestimado. Por isso opera em modo radical ou conservativo. O modo radical inicia designando pacotes para o subfluxo com menor RTT, assim como o LRF, padrão do MPTCP. Os pacotes são designados para o próximo caminho apenas se $RWND_r$ for maior que o requerido. No modo conservativo, se o caminho mais rápido está na fase de partida lenta ou tem $2 \times PCT_p > RWND_r$, onde PCT_p se refere ao número de pacotes pendentes de reconhecimento, se utiliza apenas o caminho mais rápido para transmissão. Se $2 \times PCT_p \leq RWND_r$, os pacotes são encaminhados pelo próximo caminho apenas quando $RWND_r > RWND_r + 2 \times PCT_p$.

O esquema proposto em M-SEDPF [Partov e Leith, 2016] se baseia no EDPF [Garcia-Saavedra et al., 2015] e DPF [Chebrolu e Rao, 2006]. Ele tem como objetivo estimar os valores do atraso dos caminhos com base nos valores prévios observados. Os dados

estimados são inseridos em uma matriz e utilizados pelo escalonador EDPF. O EDPF encaminha os dados pelo caminho com menor atraso de entrega estimado. O SEDPF é uma variante estocástica do EDPF que tenta minimizar o atraso de reordenamento escalonando os pacotes pelos caminhos de forma apropriada.

Por outro lado, algumas abordagens buscam reduzir as diferenças entre os atrasos dos caminhos a fim de evitar a degradação de desempenho causada pela diversidade dos caminhos. Neste sentido, o RTT-CTS [Chen et al., 2014] tem uma proposta adaptativa. Se os RTT dos caminhos são iguais, usa-se um simples *Round Robin*, do contrário, aloca os dados para o caminho com menor RTT. Ao alocar um segmento para o caminho, o escalonador incrementa o tempo estimado de transmissão. Isto é feito até o atraso dos caminhos ficarem compatíveis. Assim, os pacotes podem ser encaminhados usando o *Round Robin*. O FDPS [Le e Bui, 2015] estima o atraso de encaminhamento de cada caminho, calculando a diferença entre eles. Para escalonar os pacotes, o FDPS seleciona os caminhos com menor diferença de atraso de encaminhamento. O FrPS [Hwang e Yoo, 2015] escalona os pacotes com base na diferença entre o atraso dos caminhos. Contudo, se a diferenças entre os caminhos são significativas, ou autores entendem que não precisam utilizar o caminho mais lento.

Ordem de chegada dos pacotes

A diversidade de caminhos notavelmente incorre no problema do reordenamento quando os pacotes são escalonados de forma indiscriminada, ou seja, sem considerar o número de sequência dos pacotes (ordem). Na tentativa de minimizar este problema, algumas propostas tentam estimar a chegada dos pacotes no receptor e alocar os dados de forma discriminada a fim de que cheguem em ordem. Estas propostas implementam dois modos diferentes de envio, em ordem e fora de ordem. O envio em ordem, como ilustrado pela Figura 3.8, retira os pacotes da fila de envio e os encaminha na mesma ordem que foram recebidos da aplicação. O envio fora de ordem, como ilustrado pela Figura 3.9, retira os dados da fila de envio não respeitando esta ordem. Estes dois modos precisam resolver o seguinte problema: quais e quantos pacotes serão escalonados e em qual subfluxo. Os pacotes precisam ser escalonados pelo subfluxo que fará a entrega em ordem no menor tempo possível.

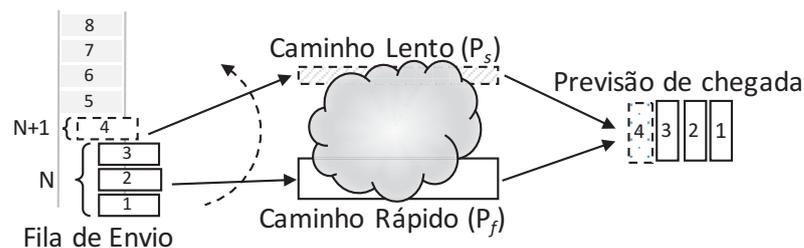


Figura 3.8: Processo de envio em ordem

O método FPS [Mirani et al., 2010] foi um dos precursores em estimar o tempo de chegada dos pacotes no receptor e transmiti-los de modo discriminado para que sejam recebidos em ordem. A dificuldade para este tipo de método se trata da estimativa do tempo de chegada, que depende das características individuais de cada caminho. Se as características dos caminhos, como largura de banda, atraso e taxa de perda sofrem poucas variações, os métodos conseguem alcançar bons resultados, do contrário, podem até agravar o problema do reordenamento [Ni et al., 2015].

As diferentes estratégias de implementação (*Pull*, *Push* ou Híbrida) também tem impacto na estimativa do tempo de chegada dos pacotes. Abordagens que implementam uma estratégia do tipo *Push*, como o DAPS [Kuhn et al., 2014], CWA-MPTCP [Zhou et al., 2013] ou Híbrida como

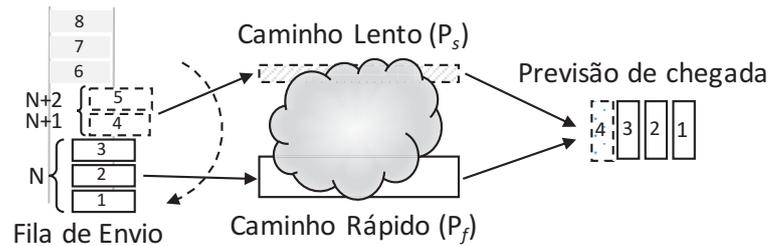


Figura 3.9: Processo de envio fora de ordem

o HDDS [Bhat e Talmale, 2014], precisam estimar o tempo de entrega considerando também o atraso na fila dos subfluxos. Cabe ressaltar que neste caso o escalonador não controla quando os pacotes serão enviados e assim não pode alterar as decisões de escalonamento. As abordagens baseadas na estratégia *Pull*, como $F^2P - DPS$ [Ni et al., 2014], não tem esta dificuldade, o escalonador designa o pacote para o subfluxo no momento em que ele pode ser encaminhado pela rede. Contudo, ele precisa se esforçar para estimar o tempo de entrega somente a partir do recebimento de um ACK. Uma abordagem Híbrida, como HDDS, tem as mesmas dificuldades da abordagem *Push*, contudo, como os pacotes permanecem na fila de envio compartilhada até serem reconhecidos, o escalonador pode alterar as decisões de escalonamento de forma dinâmica em função das variações do estado dos caminhos.

Os métodos de escalonamento apresentados em CWA-PTCP, HDDS, DAPS e $F^2P - DPS$ enviam os pacotes respeitando a ordem que são recebidos da aplicação. O funcionamento básico destes métodos pode ser descrito pelo seguinte exemplo. Considere 2 subfluxos com similar capacidade (largura de banda e perdas), porém, um subfluxo é mais lento (P_s) e possui um atraso 3 vezes maior que o subfluxo mais rápido (P_f). Considerando que os dados são escalonados em ordem, os pacotes de 1 a 3 são designados para o P_f , e o 4º pacote para P_s . Com isto, se espera que o 4º pacote chegue após o 3º. O RCPS [Kim et al., 2012] possui uma tabela para cada caminho com o tempo de entrega estimado de cada pacote. O escalonador então aloca o pacote para o caminho com menor tempo de entrega.

Os métodos empregados pelo OTIAS [Yang et al., 2014] e OCPS [Ni et al., 2015] transmitem os pacotes de dados fora de ordem, ou seja, sem preservar a ordem de sequenciamento dos dados recebidos da aplicação. Isto porque os algoritmos tentam prever a ordem de recebimento conforme os atrasos dos caminhos. Sendo assim, aloca-se os pacotes de modo que, enquanto os caminhos com menor atraso transmitem os pacotes com números de sequência baixos, os caminhos com maior atraso, transmitem os pacotes com números de sequência mais altos. Com isso, espera-se que os pacotes sejam recebidos em ordem, minimizando o problema do reordenamento. O método utilizado pelo OTIAS decide qual subfluxo vai utilizar baseado em cada pacote. Ele leva em conta o RTT e o tamanho da fila dos subfluxos em um dado momento, o que está mais próximo do escalonador padrão do MPTCP, embora considere mais informações.

Os métodos propostos pelo NTS [Liu et al., 2014] e BLEST [Ferlin et al., 2016] empregam a noção de escalonamento por bloco de pacotes. Um bloco se refere a um determinado número de pacotes que será transmitido em cada caminho. Neste caso, o escalonador determina previamente o tamanho dos blocos com base nas informações dos caminhos. As decisões de escalonamento são tomadas por bloco e não por pacote. Enquanto um subfluxo envia um bloco, outro bloco pode ser encaminhado em um subfluxo mais rápido, não aguardando um novo ciclo.

Correlação entre os caminhos

A política empregada pelo DRA [Liu et al., 2011] e MLCS [Hu et al., 2016] escalona os pacotes para os caminhos a partir do grau de correlação entre eles, calculado a partir de diferentes métodos estatísticos. O grau de correlação visa indicar se os caminhos são compartilhados, ou seja, possuem pontos (enlaces) em comum entre origem-destino. O OJCS [Ou et al., 2016] propõem um método adaptativo, que decide inicialmente qual escalonador deve ser utilizado se os caminhos compartilham um gargalo. O método propõem utilizar uma versão do OTIAS, caso os caminhos compartilhem um gargalo, ou um novo método proposto, caso contrário. O novo método calcula para cada caminho a ocupação da fila e quantos RTTs devem ser aguardados até que um determinado pacote seja enviado. O OCPS obtém informações dos caminhos através do reconhecimento seletivo (SACK do Inglês *Selective Acknowledgments*) [Mathis et al., 1996] para decidir se escalona muitos ou poucos pacotes no próximo ciclo. O mecanismo de escalonamento básico empregado pode ser FPS, $F^2P - DPS$ ou qualquer outro. Ao utilizar o FPS ou $F^2P - DPS$, não se despreza o caminho com maior atraso (P_s). Ao invés disto, o escalonador prediz o tempo de chegada (t') do pacote pelo caminho P_s no próximo ciclo e calcula a quantidade de dados (N) que pode ser entregue pelos demais caminhos antes de (t'). O escalonador então pula os N primeiros pacotes na fila de envio e escolhe o $(N + 1)$ -ésimo pacote para enviar pelo caminho P_s .

Requisitos

Alguns métodos de escalonamento propostos têm considerado o uso de informações sobre as aplicações para designar os dados sobre os caminhos disponíveis, com objetivo de atender os requisitos de QoS. O MPTCP-LB [Dugue et al., 2012] propõem direcionamentos para construção de um método baseado no paradigma de Computação Autônoma, a fim de escalonar os pacotes para o caminho que melhor se encaixa com as prioridades do fluxo de dados. O esquema proposto em MTRFC [Park et al., 2016] lida com o problema de escalonar pacotes de vídeo sobre os caminhos com perdas, comuns no ambiente de redes sem fio. Os autores consideram o atraso ocasionado pelas retransmissões no cálculo da largura de banda disponível. O escalonador calcula os tempos de recebimento dos pacotes nos diferentes caminhos e utiliza esta informação para alocar os pacotes. O método proposto pelo CMT-SR [Liu et al., 2016b] propõem otimizar as retransmissões para o fluxo de dados de vídeo implementando um mecanismo de retransmissão seletiva. O mecanismo proposto retransmite os dados de vídeo mais importante sobre um caminho alternativo considerando que o mesmo será entregue antes do tempo limite para decodificação no receptor.

Para atender os requisitos de QoS, algumas propostas de escalonadores têm implementado a transmissão parcialmente confiável, que se refere à transmissão de dados multimídia com tolerância a perdas [Diop et al., 2011]. Os escalonadores que implementam este conceito tem a possibilidade de não recuperar perdas aceitáveis a fim de melhorar outros parâmetros de QoS, como atraso e taxa de transmissão. O método proposto pelo escalonador EDAM [Wu et al., 2016], similar também ao apresentado em [Dugue et al., 2012], descarta os quadros considerados de baixa prioridade para alcançar uma taxa de transferência que satisfaça um limite, ou seja, que mantenha a qualidade da transmissão com um limite aceitável de distorções. O problema destas abordagens é que não descrevem como o escalonador vai identificar os parâmetros de QoS no modelo de camada de transporte atualmente empregado.

No trabalho realizado por [Corbillon et al., 2016], é introduzido o conceito de escalonamento multicamadas, o qual utiliza informações das camadas de aplicação e transporte para ordenar a transmissão dos dados e priorizar as partes mais significativas em uma transmissão de vídeo, maximizando a quantidade de dados de vídeo recebidos pelo cliente. O trabalho

considera a implementação escalonador na camada de aplicação. O esquema proposto em EEDLB [Morawski e Ignaciuk, 2017a] direciona ao problema da distribuição de carga ótima entre as interfaces disponíveis em um dispositivo de rede, como um telefone móvel, com objetivo de minimizar o consumo de energia total. O objetivo é projetar um algoritmo de balanceamento de carga para o escalonador de tráfego. O escalonador adquire as informações necessárias a partir das interfaces (NICs) e através da observação das atividades da camada de enlace.

Consumo de Energia: o consumo de energia tem sido um fator muito importante para a transmissão multicaminhos no cenário de redes móveis. O uso de múltiplas interfaces (e.g., WiFi, LTE) naturalmente aumenta o consumo de energia e reduz a vida útil da bateria. Neste sentido, os métodos de escalonamento têm como objetivo reduzir o consumo de energia direcionando o tráfego para os caminhos mais eficientes. O desafio destes escalonadores está justamente em identificar os caminhos mais eficientes, uma vez que existem muitas variáveis que determinam a eficiência do consumo, como o tipo do tráfego, taxa de transferência, tecnologia, atraso, mobilidade, etc.

Em [Pluntke et al., 2011], os autores sugerem um modelo baseado em cadeias de Markov (MDP) para geração de diferentes escalonadores baseados nas variáveis que influenciam o consumo de energia. O método proposto pelo eMTCP [Chen et al., 2013] move o tráfego para a interface que tem menor consumo de energia. O EDAM [Wu et al., 2016] propõem um esquema que apenas retransmite o pacote que pelo caminho com menor atraso e menor consumo de energia. O esquema proposto em EEDLB [Morawski e Ignaciuk, 2017a] tem como objetivo reduzir o consumo de energia. Para isto, o esquema proposto ajusta de forma dinâmica o uso dos caminhos com base em características da interface, como distância do ponto de acesso e o modo de operação (802.11a/b/g/n/ac).

Redundantes

Enquanto o escalonamento dos dados sobre os caminhos de modo concorrente visa melhorar o desempenho da transmissão, outras propostas de escalonadores visam aumentar a resiliência da transmissão. Os escalonadores redundantes tentam compensar as flutuações no estado dos caminhos que causam problemas na transmissão de dados. O objetivo é prover tolerância a falhas e buscar obter uma menor variação de latência. Em [Lopez et al., 2015] e [Lopez et al., 2016], os autores propuseram uma estratégia de redundância para o MPTCP em que os pacotes de dados são replicados em múltiplos subfluxos. A estratégia focou em aumentar a confiabilidade na transmissão de dados em sistemas SCADA (*Supervisory Control and Data Acquisition*). Este escalonador atualmente é distribuído juntamente com a implementação do MPTCP [Paasch et al., 2013]. O rMPTCP [Hunger e Klein, 2016] e MPTCP-L [Wang et al., 2016] replica todos os pacotes recebidos da aplicação e distribui sobre os diferentes subfluxos. Fazendo isto, os dados são enviados em todos os caminhos de forma concorrente. Em uma versão mais recente do rMPTCP [A. Hunger and P. A. Klein and M. H. Verbunt, 2016], os autores apresentam diferentes níveis de redundância. Neste caso, o objetivo é controlar o número de réplicas que serão criadas pelo escalonador. O ReMPTCP [Frömmgen et al., 2016] usa um marcador para diferenciar retransmissões de pacotes redundantes. Possui uma abordagem adaptativa com os modos de operação conservativo e agressivo. O primeiro apenas retransmite o pacote perdido. O segundo usa o mesmo número de sequência do subfluxo (SSN) para transmitir um diferente número de sequência da conexão (DSN).

3.4 Resumo

Este capítulo apresentou uma revisão dos desafios enfrentados pelas transmissões multicaminhos e as principais abordagens empregadas. É notável que o número de pesquisas e contribuições têm crescido nos últimos anos, em especial, aquelas que lidam com os desafios do ambiente de redes sem fio heterogêneas. As soluções existentes, especialmente se vistas de forma isoladas, não são definitivas e não resolvem todos os problemas. Os problemas são muitos e necessitam de uma estratégia capaz de coordenar e gerenciar a aplicação eficiente das soluções. Neste sentido, esta tese provê uma estratégia de resiliência a fim de identificar os desafios e prover a solução adequada. Esta proposta é apresentada no próximo capítulo.

4 Uma Estratégia de Resiliência para a Transmissão Multicaminhos

Este capítulo apresenta uma estratégia de resiliência para as transmissões de dados multicaminhos derivada dos princípios de resiliência para redes de comunicação. O objetivo da estratégia é aumentar o desempenho e a resiliência na transmissão de dados multicaminhos. A fim de reforçar a validade da estratégia proposta são apresentados diferentes estudos de casos, descritos a partir do Capítulo 6. O Capítulo atual está organizado da seguinte forma. A Seção 4.1 descreve as motivações e os trabalhos relacionados. A Seção 4.2 reúne os principais requisitos para uma transmissão multicaminhos resiliente. A Seção 4.3 descreve a estratégia de resiliência proposta. Por fim, a Seção 4.4 apresenta um breve resumo do capítulo.

4.1 Contextualização e Trabalhos Relacionados

A estratégia de resiliência proposta neste capítulo segue os princípios do comportamento auto-organizado e autônomo [Computing et al., 2006a], de modo que os componentes tenham um ciclo evolutivo [Smith et al., 2011] e operem de forma adaptativa, com a mínima ou nenhuma intervenção humana. A definição da estratégia tem as seguintes motivações: (i) os desafios que impactam na resiliência da transmissão multicaminhos, descritos nos Capítulos 2 e 3; (ii) os princípios e requisitos para alcançar uma maior resiliência, abordados de forma geral no Capítulo 2; (iii) e a revisão da literatura, apresentada no Capítulo 3.

A estratégia proposta também tem como objetivo à empregabilidade dos princípios de resiliência e que não são completamente atendidos pelos protocolos de transmissão de dados atualmente [Exposito, 2013]. Por exemplo, o MPTCP não considera diretamente os requisitos das aplicações e serviços para o transporte de dados fim-a-fim, aplicando basicamente uma política de melhor esforço para estes casos. Neste caso, a estratégia proposta visa conhecer estes requisitos a fim de manter a operação normal ou o estado esperado da conexão diante dos principais desafios que afetam a transmissão dos dados. Além disso, cabe ressaltar que uma única abordagem não é capaz de atender aos diferentes desafios enfrentados pela transmissão, citados no Capítulo 3. Por isso, a estratégia proposta emprega uma abordagem adaptativa, capaz de reconhecer os requisitos, identificar os desafios e prover soluções adequadas.

A estratégia de resiliência proposta tem como base a arquitetura de resiliência designada para a Internet, proposta por Sterbenz *et al.* [Sterbenz et al., 2010] e Smith *et al.* [Smith et al., 2011]. Esta arquitetura foi descrita com detalhes no Capítulo 2. Além deste trabalho, Lima *et al.* [Lima et al., 2008a] propõe uma arquitetura de sobrevivência para redes ad hoc e Mesh denominada SAMNAR (do Inglês, *Survivable Ad hoc and Mesh Network Architecture*). O objetivo da arquitetura SAMNAR é oferecer um nível mínimo de serviço ou de operação da rede em face de ataques ativos ou passivos. A arquitetura SAMNAR tem como base a cooperação entre três linhas de defesa, a preventiva, a reativa e a tolerante. Além disso, as linhas

de defesa possuem a capacidade adaptativa em função do tipo de ameaça. Essas características, bem como a construção modular da arquitetura, que compreende os módulos de comunicação, coleta de dados e sobrevivência, norteiam a construção da estratégia descrita nesta tese.

4.2 Requisitos

Esta seção descreve os requisitos de resiliência considerados na construção da estratégia de resiliência para a transmissão multicaminhos. Estes requisitos são definidos a partir da estratégia de resiliência para redes de comunicação [Schaeffer-Filho et al., 2012, Sterbenz et al., 2010, Smith et al., 2011] e para redes em malha [Lima et al., 2008a].

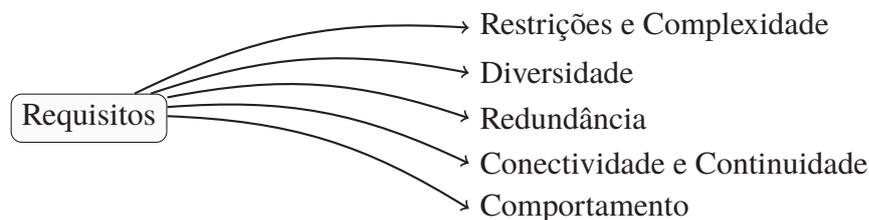


Figura 4.1: Requisitos de resiliência

4.2.1 Restrições e Complexidade

O consumo de energia [Pluntke et al., 2011] e custo [Secci et al., 2014] impõem restrições para a resiliência da transmissão multicaminhos, especialmente no cenário de redes móveis. Estes requisitos são importantes porque a conectividade do dispositivo é dependente da quantidade de banda larga contratada e da carga de bateria do dispositivo. A complexidade também é um fator importante, uma vez que lidar com desafios, como a diversidade e a heterogeneidade dos caminhos, exige técnicas otimizadas para cada problema. Como por exemplo, determinar o número ótimo de caminhos ou alocar os pacotes de dados de forma a reduzir o problema do reordenamento no cenário das HetNets.

4.2.2 Diversidade

A diversidade de caminhos é essencial para projetar uma rede resiliente. Na transmissão multicaminhos a diversidade é alcançada utilizando diferentes caminhos estabelecidos sobre as interfaces de rede. Alcançar a diversidade não é trivial, pois depende de técnicas capazes de identificar os caminhos compartilhados ou protocolos capazes de prover rotas distintas para os caminhos de uma mesma conexão. Em uma transmissão multicaminhos, é desejável caminhos disjuntos a fim de alcançar um maior desempenho com a agregação de banda e uma maior resiliência não compartilhando pontos comuns. Os caminhos P_1 e P_2 , ilustrados na Figura 4.2, percorrem caminhos disjuntos entre os nós de origem e destino. Embora a situação dos caminhos P_1 e P_2 seja desejável, em nível de camada de transporte não é possível prover esta diversidade. No entanto, através das informações coletadas dos subfluxos é possível calcular estatisticamente a relação ou a correlação entre eles, como apresentado no Capítulo 3. Nestes casos, o desafio é prover uma abordagem que calcule com precisão a correlação entre os caminhos.

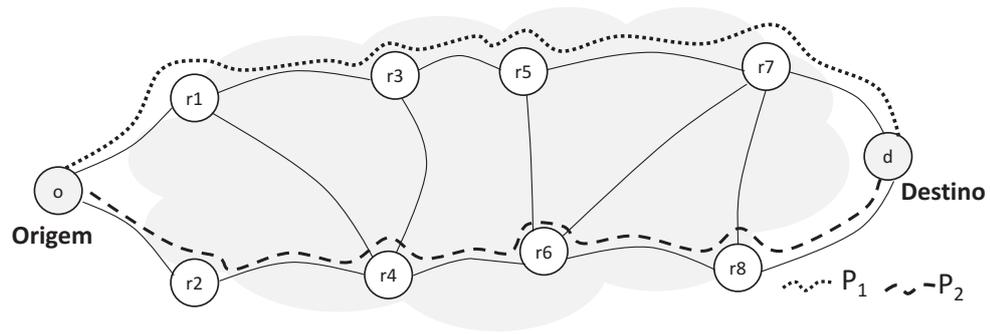


Figura 4.2: Diversidade de Caminhos

4.2.3 Redundância

O uso de múltiplos caminhos provê diferentes tipos de redundância que aumentam a resiliência contra os desafios que afetam a transmissão. Estes diferentes tipos podem ser categorizados em:

- *Redundância temporal* - se refere à ação de entregar a informação mais que uma vez. Neste caso, podem ser consideradas abordagens que otimizem a retransmissão de pacotes [Yang et al., 2014, Chen et al., 2014].
- *Redundância física* - a transmissão multicaminhos tem como objetivo prover este tipo de redundância utilizando diferentes interfaces de rede. Contudo, múltiplos endereços podem ser definidos sobre a mesma interface de rede. Neste caso, medidas preventivas podem reforçar o uso de caminhos sobre diferentes interfaces (conectadas a diferentes redes de acesso).
- *Redundância lógica* - pode ser vista como uma generalização da redundância física, na qual múltiplos caminhos são estabelecidos através de uma única interface. Neste caso, pode ser reforçado o uso de caminhos com um menor grau de correlação.
- *Redundância de informação* - definido como o uso redundante de dados a fim de permitir a reconstrução pelo receptor em caso de perdas. Neste caso podem ser empregadas técnicas de codificação [Garcia-Saavedra et al., 2015] e replicação de pacotes [Hunger e Klein, 2016].

A estratégia de resiliência pode considerar os diferentes tipos de redundância, fazendo uso da diversidade para aumentar a robustez da transmissão.

4.2.4 Conectividade e Continuidade

A conectividade e a continuidade da transmissão devem ser mantidas quando possível, mesmo se não existir um caminho estável que atenda às medidas de desempenho. A conectividade deve suportar múltiplas interfaces para conexão em mais de uma rede de acesso deve ser mantida em caso de perda de conectividade. O MPTCP emprega um mecanismo para assegurar a continuidade em caso de perda de conectividade, descrito no Capítulo 2.

4.2.5 Comportamento

O comportamento auto-organizado e autônomo, a adaptabilidade e a evolutividade fazem parte de uma estratégia de resiliência, como descrito na Subseção 2.2.4. No contexto da transmissão multicaminhos, uma estratégia resiliente deve se preocupar em prover estes diferentes comportamentos, como adaptar-se conforme as características conhecidas e desconhecidas, estáticas e dinâmicas dos caminhos e operar com mínima ou nenhuma intervenção humana [Computing et al., 2006a, Sterbenz et al., 2014]. Neste caso, os componentes, métricas e modelos devem ser ajustados conforme as condições e ao contexto da transmissão e também passar por um processo de refinamento, ou seja, precisam evoluir para que se adequem às novas condições [Sterbenz et al., 2010].

4.3 Estratégia de resiliência

A transmissão multicaminhos em si não compreende todos os pré-requisitos necessários para um sistema de comunicação resiliente. Os requisitos ausentes incluem, por exemplo, identificar o tipo de fluxo e os seus requisitos de QoS e as restrições de consumo de energia [Pluntke et al., 2011] e de custo (monetário) [Secci et al., 2014]. Estes requisitos são importantes para determinar o comportamento normal ou o comportamento esperado da transmissão. Conhecendo o comportamento normal é possível empregar métricas e modelos que detectem, por exemplo, desvios de comportamento ou ameaças [Sterbenz et al., 2014] e determinar o desempenho da transmissão. Os protocolos de transmissão atuais, como o MPTCP, dificulta obter os requisitos das aplicações devido ao uso da API padrão do TCP [Scharf e Ford, 2013], que não provê acesso a essas informações. No entanto, a eficiência de uma estratégia de transmissão resiliente depende do conhecimento dos requisitos aplicações [Nikraves et al., 2016].

O reflexo deste desconhecimento dos requisitos das aplicações se traduz na diminuição da robustez de toda a transmissão. A transmissão multicaminhos vislumbra uma complexa relação entre cenários, desafios, requisitos e objetivos. Isto demonstra, por exemplo, que uma única abordagem não é suficiente para atender a toda esta diversidade. Alguns trabalhos propõem abordagens adaptativas, a fim de alterar um determinado componente do protocolo de transmissão, como o controle de congestionamento e o escalonador, conforme as condições dos caminhos e os requisitos das aplicações [Hu et al., 2016, Ni et al., 2014, Chen et al., 2014, Li et al., 2015, Ni et al., 2015]. Isto vai de encontro com os trabalhos baseados em Computação Autônoma [Diop et al., 2011, Dugue et al., 2012]. As observações apresentadas nesta seção servem de base para definição da estratégia, que conforme ilustra a Figura 4.3, incorpora um arcabouço de resiliência, uma base de informação e um módulo gerenciador de contexto. Estes componentes são descritos com detalhes a seguir.

4.3.1 Arcabouço de Resiliência

O arcabouço de resiliência visa garantir a resiliência da transmissão multicaminhos. Ele possui cinco módulos: (1) módulo de coleta, responsável por adquirir todas as informações utilizadas pelos demais módulos; (2) módulo de reconhecimento, responsável pela análise dos desafios e avaliação da resiliência; (3) módulo de resistência, responsável pelas medidas defensivas; (4) o módulo de recuperação, que aplica medidas reativas para recuperar a transmissão caso seja identificado que ela desviou de seu comportamento esperado; (5) e o módulo de controle, que é responsável pelo gerenciamento do arcabouço.



Figura 4.3: Componentes da estratégia de resiliência

A organização e o relacionamento entre os módulos do arcabouço de resiliência tem como base as estratégias de resiliência apresentadas em [Smith et al., 2011, Sterbenz et al., 2010, Lima et al., 2008b]. Conforme ilustra a Figura 4.4, os módulos estão organizados para seguir um ciclo adaptativo, similar ao empregado pela estratégia D^2R^2+DR [Sterbenz et al., 2010]. O relacionamento entre os módulos ocorre da seguinte forma: A partir da coleta dos dados, o arcabouço inicialmente busca identificar os requisitos das aplicações, os desafios, e o estado normal da transmissão. A fim de aumentar a robustez da transmissão multicaminhos, são empregadas medidas de resistência. Contudo, estes mecanismos podem não ser suficientes, devido à ocorrência de uma situação adversa ou desconhecida. Neste sentido, as informações reunidas pelo módulo de coleta e repassadas ao módulo de reconhecimento, são utilizadas para determinar as medidas de recuperação. Estas medidas têm como objetivo remediar ou minimizar os efeitos dos danos causados pela situação adversa e tentar manter a transmissão com uma capacidade aceitável, em face dos requisitos das aplicações. Todas estas ações são avaliadas, utilizando parâmetros definidos pelas políticas armazenadas na base de informação, e caso necessário, são adaptadas para incorporar lições aprendidas sobre uma determinada situação adversa, por exemplo. As próximas seções descrevem com mais detalhes esses componentes.

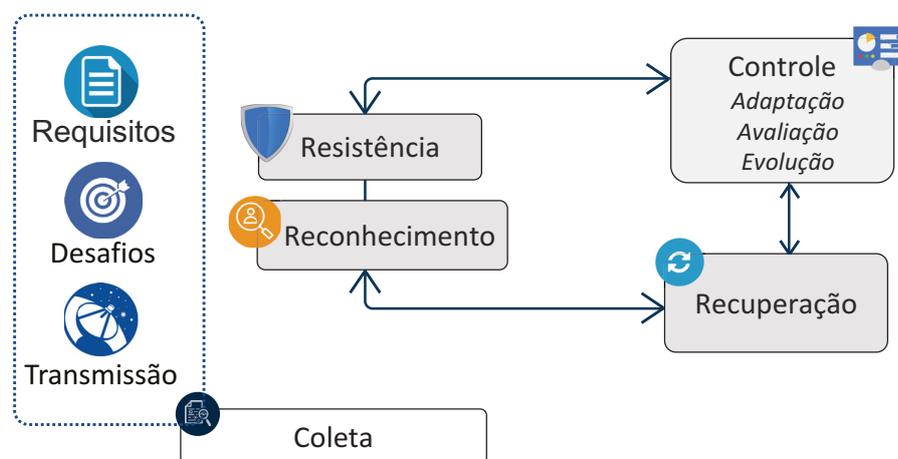


Figura 4.4: Arcabouço de Resiliência

Módulo de Coleta

O módulo de coleta tem como função reunir as informações que serão utilizadas pelo arcabouço de resiliência. Estas informações podem ser coletadas em diferentes camadas. Na camada de transporte, é possível obter informações sobre o estado de cada transmissão multicaminhos ativa, ou seja, de cada instância que esteja sendo executada em um determinado instante. Uma única transmissão possui informações em nível de conexão e em nível de subfluxo, sendo possível monitorar dados das variáveis utilizadas pelo controle de congestionamento e pelo escalonador, como o atraso (RTT), largura de banda disponível, retransmissões e fila de envio (e.g., CWND), informações das métricas, como correlação, taxa de perda, atraso de uma via (OWD), número de pacotes fora de ordem (RFO), entre outras. As informações obtidas de outras camadas têm como objetivo aumentar a capacidade de atuação do módulo de reconhecimento. Por exemplo, através de informações obtidas sobre a aplicação pode-se determinar os requisitos (QoS) ou o tipo (curta ou longa duração) do fluxo de dados e as informações obtidas da camada de enlace (taxa de retransmissões), podem indicar uma possível perda de conectividade ainda não percebida pelo protocolo de transporte. A relação das variáveis e métricas monitoradas podem ser determinadas conforme requerido pelo módulo de reconhecimento.

Módulo de Reconhecimento

O módulo de reconhecimento inclui as funcionalidades para analisar e caracterizar um desafio, usando uma variedade de fontes de informação, e também, avaliar as medidas de resiliência empregadas. O módulo de reconhecimento usa informações da base de informação, onde estão armazenados os modelos que descrevem os sintomas de um desafio, no processo de análise. O processo de análise emprega um modelo incremental, assim como proposto em [Smith et al., 2011], onde o entendimento sobre a natureza de um desafio evolui à medida que mais informações se tornam disponíveis. Uma das vantagens aparentes do modelo incremental é o uso de mecanismos leves de detecção, baseados apenas em determinadas informações que estejam sendo monitoradas. Por exemplo, considere um dispositivo móvel conectado em duas redes de acesso, em que uma das redes sofre um aumento repentino no volume tráfego. O simples monitoramento do RTT dos caminhos, de forma correlacionada, pode indicar que o problema é comum a uma determinada rede de acesso. Isto pode ser utilizado para desencadear o direcionamento parcial ou total do tráfego das conexões ativas sobre os caminhos da interface de rede menos congestionada. Os modelos mais complexos podem ser utilizados se existir a necessidade de identificar as causas do problema, como identificar previamente um *bufferbloat*.

Com base na saída do componente de análise de desafio e no estado da rede, o componente de avaliação determina se a resiliência requerida está sendo alcançada. Esta medida tem como base as métricas de resiliência e é influenciada pela efetividade dos mecanismos de defesa e remediação para responder aos desafios. Alguns desafios possuem sintomas semelhantes, como o problema do reordenamento e os congestionamentos, em que ambos são percebidos pelo aumento do atraso [Xu et al., 2016, Alheid et al., 2016]. Existem muitas formas de detectar e diferenciar estes sintomas. Contudo, deve-se conhecer os requisitos mínimos e os objetivos da transmissão. Com isto, se pode determinar o comportamento normal ou esperado da transmissão, ou seja, o comportamento que deve ser mantido. A partir de métricas existentes (atraso, taxa de perda, largura de banda, e outras) pode-se coletar informações sobre os caminhos e o fluxo de dados a fim de detectar um possível desvio do comportamento normal ou esperado da transmissão. E através de um modelo pré-definido, pode-se reconhecer a origem do problema, por exemplo, se o aumento do atraso foi causado por um congestionamento, pelo problema do reordenamento ou

até devido a mobilidade do dispositivo. Este reconhecimento e caracterização das causas de um desvio (e.g., reordenamento) são importantes para definição da estratégia de remediação.

Módulo de Resistência

O módulo de resistência tem como função prover um conjunto de medidas defensivas que diminuam a probabilidade de que uma ameaça afete a operação normal da transmissão multicaminhos. As medidas defensivas são proativas, ou seja, tem como objetivo prevenir, por exemplo, a degradação da transmissão na ocorrência de uma ameaça. As medidas de resistência podem ser determinadas com base nos requisitos requeridos e nos objetivos de uma determinada transmissão. Neste sentido, diferentes medidas preventivas podem ser empregadas ou alteradas de forma adaptativa para cada instancia de uma transmissão multicaminhos em um mesmo dispositivo. As medidas podem ser empregadas para prover diferentes formas de tolerância a desafios, tais como: a **sobrevivência** da transmissão, utilizando um número de caminhos ideal e/ou com baixa correlação (diversidade de caminhos); **tolerância a falhas** correlacionadas, empregando diferentes níveis de redundância com apoio da diversidade; **tolerância a interrupções**, empregando técnicas para manter a conectividade e continuidade da conexão; **tolerância a tráfego** inesperado, empregando técnicas de balanceamento, redundância e diversidade, entre outras.

Módulo de Recuperação

O módulo de recuperação tem como funcionalidade prover medidas para recuperar e remediar uma transmissão multicaminhos durante e após a ocorrência de eventos ou condições adversas. O controle de congestionamento é um exemplo de mecanismo que, na ocorrência de uma perda, faz uma degradação na taxa de envio a fim de minimizar o impacto de um possível congestionamento [Xu et al., 2016]. Contudo, assim como existem diferentes estratégias de controle de congestionamento, como algoritmos baseados em perdas e algoritmos baseados no atraso, existem também inúmeras estratégias que podem ser empregadas para aliviar e recuperar o impacto na transmissão. A identificação precisa (i.e., correta) da ameaça pelo módulo de reconhecimento é determinante para o sucesso de uma estratégia de recuperação. Por exemplo, não adianta determinar a suspensão apenas de um caminho se os demais compartilham um gargalo (congestionamento) entre a origem e destino. A análise da correlação entre os caminhos pode determinar, quando possível, a abertura de novos caminhos com uma maior diversidade. Além disso, uma análise da extensão do dano ajuda a determinar também a extensão ou duração da medida de recuperação. Isto serve para determinar o término da medida de recuperação e retomar a operação normal, o que em muitos casos pode ser necessário para que se justifique o uso da transmissão multicaminhos [Sterbenz et al., 2010].

A recuperação precisa ocorrer de forma automática, ou seja, a aplicação das medidas deve ocorrer automaticamente a partir da detecção do desafio. Por exemplo, quando um caminho falhar, o fluxo de dados deste caminho deve ser redirecionado para outro caminho com capacidade suficiente para receber o tráfego de modo a não causar grande impacto nos requisitos mínimos de serviço requeridos. Enquanto isto, o mecanismo de recuperação deve aplicar medidas para restabelecer o funcionamento normal da transmissão. Após a falha ser reparada, a transmissão deve redirecionar os dados para o caminho original (supondo ser o melhor caminho disponível). O módulo de recuperação pode empregar diversas medidas, como abrir novos caminhos [Li et al., 2015], penalizar fluxos concorrentes [Paasch et al., 2014] ou empregar técnicas de transmissão parcialmente confiável [Diop et al., 2011], em fluxos que suportam, com objetivo de recuperar a transmissão e alcançar o estado de operação normal da transmissão.

As operações da etapa de recuperação requerem a caracterização das operações normais e as propriedades (requisitos) dos serviços, assim como a quantificação da degradação dos serviços no que se refere à degradação das operações. Essa caracterização remete à definição de modelos e métricas. As métricas ajudam a determinar, por exemplo, a duração do desafio e quando a transmissão foi restaurada para seu estado original [Jabbar, 2010].

Módulo de Controle

O módulo de controle é responsável pela gerência do arcabouço de resiliência. Ele tem como função aplicar e avaliar as políticas de resiliência, controlar a adaptação e evolução das medidas, modelos e métricas empregados. O módulo de controle visa preservar o estado normal ou esperado para a transmissão, levando em conta os requisitos e objetivos requeridos. Uma política se refere ao conjunto de medidas, métodos e métricas empregados de forma correlacionada entre os módulos a fim de alcançar um determinado objetivo ou requisito. As políticas são dirigidas usando o conhecimento de resiliência contido na base de informação.

O processo de adaptação se refere ao diagnóstico e refinamento das medidas empregadas pelos módulos de reconhecimento, resistência e recuperação. O diagnóstico visa aprender com os resultados obtidos através da avaliação da resiliência, realizada pelo módulo de reconhecimento, e refletir isto nos demais módulos, modo que possa ser continuamente melhorado. O módulo de controle pode, por exemplo, substituir uma medida preventiva do módulo de resistência caso a avaliação da resiliência demonstre que não tem sido efetiva [Lima et al., 2008b]. Além disso pode também trabalhar em conjunto com o gerenciador de contexto para ajustar e refinar parâmetros específicos em um determinado escalonador ou algoritmos de controle de congestionamento. Em síntese, todos os módulos podem ser reconfigurados, se necessário. Os ajustes realizados devem ser registrados na base de informação, a fim de criar um histórico das alterações e para comparar a efetividade daquelas que foram aplicadas. É desejável que as medidas de diagnóstico e refinamento, as quais compreendem um laço contínuo [Sterbenz et al., 2010], sejam realizados sem a intervenção humana. Através de informações coletadas das camadas inferiores (enlace e física), o módulo de controle pode inferir que, a partir de um determinado limiar de taxa de perda e força de sinal, os caminhos sobre uma determinada interface de rede afetam o desempenho da transmissão [Partov e Leith, 2016]. Neste caso, o módulo pode ajustar os parâmetros do escalonador para considerar estes limiares no processo de seleção dos caminhos.

4.3.2 Base de Informações

A base de informações, conforme ilustra a Figura 4.5, tem como função guardar os modelos de desafios, medidas de resistência e recuperação, métricas, políticas, avaliações e históricos das medidas empregadas, entre outros. Os modelos de desafios, por exemplo, são armazenados para identificar os desafios conhecidos. Além disso pode conter inúmeras medidas de resistência e recuperação para os diferentes desafios, requisitos e objetivos. Ela pode expressar também um grafo de dependência entre as diferentes medidas, uma vez que a aplicação de uma política pode envolver o relacionamento entre diferentes módulos. Por exemplo, uma determinada medida de resistência pode depender de informações específicas do método de coleta e medidas específicas do módulo de reconhecimento. Além disso, podem estar associados métricas e modelos para avaliar a efetividade da política empregada.

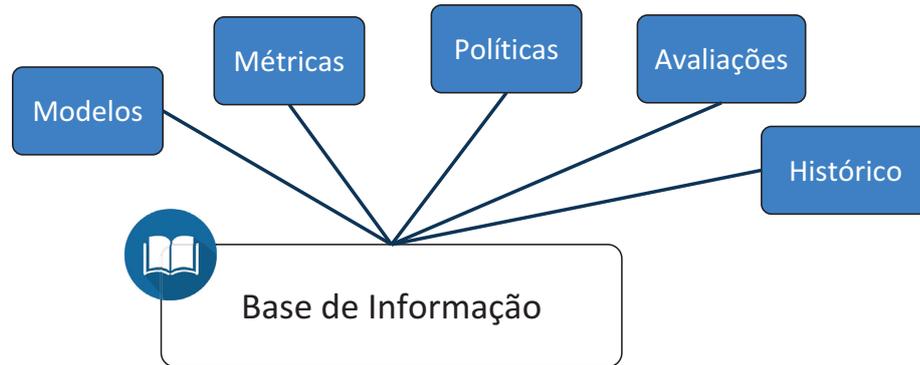


Figura 4.5: Funções da base de informações

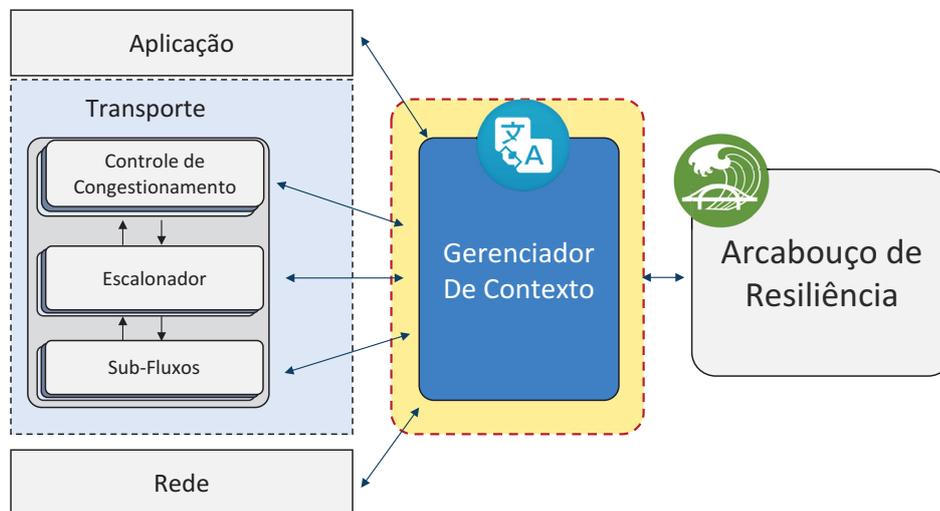


Figura 4.6: Posicionamento do gerenciador de contexto

4.3.3 Gerenciador de Contexto

O gerenciador de contexto (GC), conforme ilustra a Figura 4.6, tem como função refletir as medidas empregadas pelo arcabouço de resiliência relacionadas aos protocolos e serviços oferecidos pelas diferentes camadas. O GC pode ser visto como uma interface de comunicação entre o arcabouço e as outras camadas. De um modo prático, conforme a política empregada pelo módulo de controle, o GC pode alocar um determinado componente para um ou mais protocolos, como por exemplo, um algoritmo de escalonamento ou de controle de congestionamento. Ele ainda pode ser mais específico, podendo refletir mudanças internas em um determinado componente do protocolo, alterando por exemplo, uma métrica ou método específico do algoritmo de escalonamento.

4.4 Resumo

Este capítulo apresentou uma estratégia de resiliência para a transmissão multicaminhos resiliente. A proposta considerou os princípios de resiliência, apresentados na Seção 2.2, e a revisão do estado da arte, apresentada no Capítulo 3. A estratégia tem como objetivo prover uma abordagem adequada aos requisitos e/ou desafios que afetam a transmissão multicaminhos. Para

validar a estratégia proposta, são realizados diferentes estudos de casos relacionados aos módulos de reconhecimento e de resistência. Os estudos de caso são apresentados nos próximos capítulos e abordam os principais desafios que afetam as transmissões multicaminhos, a heterogeneidade e a falta de diversidade de caminhos.

5 Análise da Diversidade em Transmissões Multicaminhos

Este capítulo apresenta um estudo de caso relacionado ao reconhecimento da diversidade de caminhos no contexto da transmissões multicaminhos no cenário de redes heterogêneas. Esse estudo de caso está diretamente ligado às atividades do módulo de reconhecimento da estratégia de resiliência proposta no Capítulo 4. O capítulo está organizado da seguinte forma. A Seção 5.1 descreve o problema da diversidade para as transmissões multicaminhos. A Seção 5.2 apresenta uma taxonomia das abordagens estudadas. A Seção 5.3 descreve a metodologia de avaliação e os resultados alcançados. Por fim, um resumo do capítulo é apresentado na Seção 5.5.

5.1 Contextualização

A diversidade de caminhos é um termo comumente empregado no contexto de camada de redes, especificamente no roteamento de fluxos de dados (e.g. fluxos TCP). Neste contexto, a diversidade de caminhos se refere a uma variedade (multiplicidade) de caminhos, ou seja, de rotas diferentes em que os fluxos de dados podem ser encaminhados pela rede [Rohrer et al., 2014]. Os protocolos de roteamento têm informações sobre os enlaces e sobre os nós que compõe um caminho, e com isso podem criar caminhos distintos. No entanto, estas informações não estão disponíveis aos protocolos de camada de transporte. Mesmo assim, os protocolos multicaminhos, como o MPTCP, tentam prover a diversidade de caminhos. Para este caso, o protocolo associa uma conexão a múltiplos subfluxos, geralmente associados a diferentes redes de acesso, na expectativa que os protocolos de rede se encarreguem de prover diversidade. Porém, como os protocolos de transporte não têm controle sobre a diversidade dos caminhos, ou seja, a capacidade de tomar decisões sobre suas rotas, os subfluxos de uma mesma conexão podem ser encaminhados por caminhos (rotas) que compartilham um ou mais enlaces (ou nós) em sua trajetória. Em um contexto de redes móveis, os subfluxos de uma conexão multicaminhos enviados por caminhos compartilhados estão sujeitos a diferentes tipos de ameaças: falhas ligadas ao ambiente, como perda de conectividade ou interferências no sinal sem fio; desastres naturais, como terremotos, furacões, rompimento de fibra; tráfego não usual, com congestionamentos repentinos comuns em grandes eventos ou ataques maliciosos e outros. Os caminhos compartilhados limitam o desempenho e a resiliência das transmissões multicaminhos. Neste sentido, este capítulo apresenta uma taxonomia sobre as abordagens empregadas para prover diversidade para as transmissões multicaminhos. O objetivo desta taxonomia é agrupar as abordagens em categorias conforme suas características comuns, a fim de identificar os pontos positivos e negativos que podem contribuir ou não com os objetivos da tese.

5.2 Taxonomia

Esta seção apresenta uma taxonomia das abordagens as quais visam identificar ou prover diversidade para o transporte de dados multicaminhos. A classificação sugere duas classes principais: *intracamada* e *intercamada*. A classe intracamada compreende abordagens que empregam visam identificar a diversidade dos caminhos através de informações disponíveis apenas na camada de transporte; a classe intercamada inclui abordagens baseadas na interação entre múltiplas camadas, que trabalham de forma cooperativa e não cooperativa para prover a diversidade de caminhos. A Figura 5.1 ilustra a organização geral da taxonomia proposta. A classe intracamada compreende as categorias de *correlação ativa* e *correlação passiva*. A categoria de correlação corresponde às abordagens que trocam informações adicionais entre os nós para obter os dados necessários e calcular a correlação dos caminhos. A correlação passiva inclui abordagens que empregam métodos *estatísticos* e de *proporcionalidade*. A classe intercamada compreende as abordagens das categorias *cooperativa* e *não-cooperativa*. As próximas seções detalham esta taxonomia e apresenta as principais abordagens da literatura.

Além disso, a classe intercamada inclui abordagens que provêm diversidade de caminhos aos protocolos de transporte, como MPTCP, de forma cooperativa e não-cooperativa através de abordagens como SDN [Nakasan et al., 2017, Raza et al., 2017] e protocolos como o MPLS [Lopez et al., 2016]. Por outro lado, a classe intracamada inclui abordagens que utilizam apenas dados disponíveis na camada de transporte para determinar a diversidade entre os caminhos. Ambas as classes têm vantagens e desvantagens, as quais são discutidas neste capítulo. As abordagens da classe intracamada são diretamente afetadas pela dinâmica das redes HetNet, pois as características heterogêneas e dinâmicas dos caminhos interferem na acurácia dos métodos empregados. Neste sentido, apresenta-se uma avaliação dessas abordagens a fim de averiguar a acurácia no cenário HetNet.

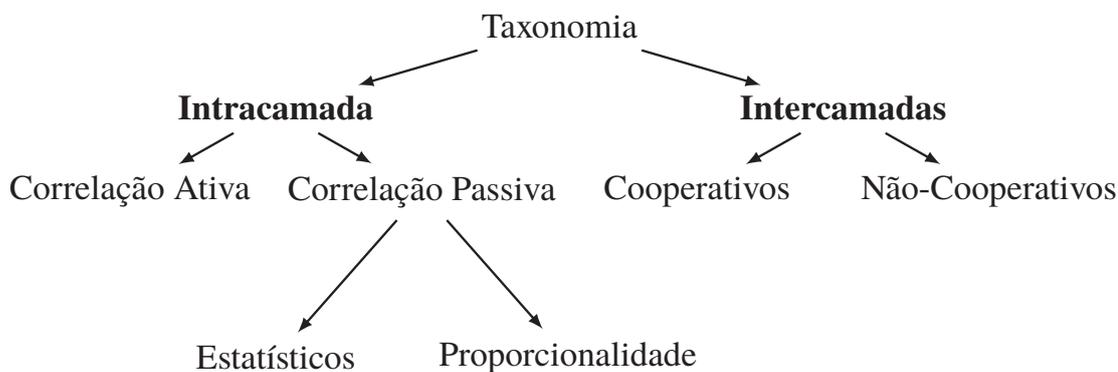


Figura 5.1: Taxonomia das Abordagens

5.2.1 Intracamada

Este grupo se subdivide nas categorias Correlação Ativa e Correlação Passiva. As abordagens da categoria **correlação ativa** tentam estimar a correlação dos caminhos proativamente, exigindo menos tempo e sem depender de informações prévias (dados históricos) dos caminhos. Em [Liao et al., 2011], os autores propõem um método de correlação ativa para estimar o atraso de uma via e determinar se os caminhos compartilham um gargalo. O método aqui denominado de MSCM (do Inglês, *Multipath Selection for Concurrent Multipath Transfer*), seleciona o conjunto de caminhos que possui a menor correlação entre si. O método sonda

a variação de atraso de todos os caminhos e agrupa os compartilhados e não-compartilhados. Além disso, para medir a correlação entre dois caminhos, o método emprega o coeficiente de correlação Pearson.

Na **Categoria de correlação passiva**, as abordagens estimam a correlação de caminhos através de dados coletados dos subfluxos, como RTT e largura de banda. As abordagens são classificadas como passivas porque não geram sobrecarga de tráfego através das medições. A categoria se subdivide em:

- *Métodos estatísticos* - baseiam-se na ideia de que os dados coletados dos subfluxos são atributos estatísticos e podem indicar, quando existe, a correlação entre eles. Em [Singh e Reddy, 2013], os autores monitoram o tamanho das filas com base nos valores do RTT e acumulam dados a fim de indicar os subfluxos que compartilham um gargalo. O método é denominado MPERT (do Inglês, *Multipath Probabilistic Early response TCP*). O MPERT assume que a probabilidade de diferentes roteadores apresentarem a mesma dinâmica de fila em um determinado instante de tempo é altamente improvável. Assim, o MPERT deduz que a correlação entre os tamanhos das filas para subfluxos compartilhados deve ser maior do que os subfluxos não compartilhados. Assim, o método seleciona apenas os subfluxos que não compartilhados porque oferecem mais vantagens em termos de largura de banda agregada. Para calcular a correlação entre os caminhos, o MPERT emprega o coeficiente de correlação de Pearson. Pelo melhor de nosso conhecimento um dos trabalhos percursoros em empregar o coeficiente de Pearson para detectar caminhos compartilhados foi realizado por [Rubenstein et al., 2002]. No entanto, Hayes et al. [Hayes et al., 2014] sugeriu que o atraso (RTT) em situações de congestionamento não segue uma distribuição normal e as variáveis dos caminhos podem não ter uma relação linear entre si. Isso tem um grande impacto nos métodos baseados no coeficiente de correlação de Pearson. Desta forma, Hayes et al. propuseram três medidas estatísticas a fim de superar as limitações de Pearson, sendo a disparidade, variabilidade e frequência. A primeira conta o número de medições do OWD abaixo ou acima da média. A segunda quantifica o desvio padrão absoluto da média. E a terceira conta e normaliza o número de vezes que uma média de curto prazo intercepta uma média de longo prazo. Os resultados das medidas estatísticas são usados para agrupar subfluxos semelhantes.
- *Métodos de Proporcionalidade* - calculam a relação de proporcionalidade entre os caminhos a partir de dados como a taxa de transferência, RTT e taxa de perda. As abordagens estudadas que empregam este método comumente usam uma regra simples de proporção. Dado os valores médios A e B referentes a um par de caminhos, calcula-se a relação de proporcionalidade dividindo o maior valor pela soma de ambos ($\max(A, B)/(A + B)$). Um resultado próximo de 0,5 indica que os caminhos têm características semelhantes, ou seja, possuem relação (correlação). Do contrário, um resultado próximo a 1 sugere que os caminhos são disjuntos e não estão relacionados. Em [Liu et al., 2011] e [Hu et al., 2016], propuseram métodos baseados em proporcionalidade para medir a relação (correlação) entre dois caminhos, os quais são denominados PACC (do Inglês, *Path Associativity Congestion Control*) e MLCS (do Inglês, *Multi-level Correlation Scheduling*), respectivamente. Ambos os métodos usam a vazão dos caminhos como parâmetro para calcular a relação de proporcionalidade.

5.2.2 Intercamadas

Esta classe inclui abordagens baseadas na interação entre protocolos de diferentes camadas, especificamente a camada de rede e de transporte, a fim de prover a diversidade de caminhos. Esta interação pode ocorrer de duas formas: cooperativa e não-cooperativa. A subcategoria **cooperativa** engloba as abordagens em que existe cooperação entre os protocolos das camadas adjacentes para determinar o conjunto de caminhos disjuntos que são mapeados para os subfluxos. O trabalho proposto por [Krupakaran et al., 2015] propõem utilizar o protocolo TraceFlow para mapear os caminhos entre os nós de origem e destino, repassando essas informações ao MPTCP. Neste caso, o TraceFlow e MPTCP trabalham de forma cooperativa para identificar os caminhos compartilhados.

Na subcategoria **não-cooperativa** o protocolo da camada de rede provê a diversidade de caminhos sem o conhecimento ou cooperação com o protocolo de transporte. Em [Lopez et al., 2016], os autores propõem um método não cooperativo que explora a integração entre os protocolos MPTCP e MPLS a fim de aumentar a resiliência das transmissões multicaminhos. O MPLS separa os subfluxos marcando individualmente cada pacote que entra na rede. Os pacotes com o mesmo rótulo são encaminhados pelo mesmo caminho (rota) na rede. O MPTCP não tem conhecimento do que ocorre na camada inferior (rede). Em [Sandri et al., 2015] os autores empregam um método similar. No entanto os subfluxos de uma mesma conexão MPTCP são mapeados e encaminhados por rotas distintas utilizando as Redes Definidas por Software (SDN, do Inglês *Software Defined Network*).

5.2.3 Vantagens e Desvantagens

Esta seção apresenta as vantagens e as desvantagens para cada classe/categoria apresentada anteriormente. A Tabela 5.1 resume esta análise conforme cinco métricas empíricas descritas a seguir:

- Escalabilidade - indica o grau de escalabilidade (baixo, médio, alto) em termos de número de caminhos ou alcance da rede;
- Dependência entre camadas - classifica as abordagens que são dependentes (múltiplas camadas) ou não-dependentes (única camada) de protocolos de outras camadas;
- Sobrecarga de tráfego - indica se as abordagens necessariamente provocam sobrecarga ou não do tráfego de dados nas transmissões;
- Custo computacional - indica se as abordagens têm um custo computacional alto ou baixo em termos de ciclos de CPU ou uso de memória;
- Aplicável em HetNet - classifica as abordagens aplicáveis ou não para os cenários de redes com caminhos heterogêneos, tais como as HetNets.

Um dos principais problemas das abordagens da categoria cooperativa se refere à interação entre os protocolos de diferentes camadas. Isto retrocede para o antigo problema de dependência entre os protocolos TCP/IP, que entre outros problemas, dificulta a mobilidade dos nós que possuem uma conexão aberta e retém a evolução da Internet [Computing et al., 2006b]. Além disso, a dependência dificulta as alterações ou atualizações de ambas os protocolos, dificultando que este tipo de abordagem seja empregada na comunicação fim-a-fim através da Internet. Embora não crie o problema da dependência, os métodos cooperativos também possuem limitações. Nesse caso, o protocolo de transporte permanece inconsciente ao roteamento

Tabela 5.1: Vantagens e Desvantagens

Classe	Categoria	Método	Prós	Contras
Intercamadas	Cooperativa	TraceFlow	Aplicável em HetNet; custo computacional baixo; sem sobrecarga de tráfego	baixa escalabilidade; múltiplas camadas
	Não-cooperativa	SDN, MPLS	Aplicável em HetNet; camada única; custo computacional baixo; sem sobrecarga de tráfego	baixa escalabilidade
Intracamada	Ativa	MSCM	Aplicável em HetNet; camada única	sobrecarga de tráfego; baixa escalabilidade; custo computacional alto
	Estatísticos	MPERT, MSCM	sem sobrecarga de tráfego; camada única	Não Aplicável em HetNet; baixa escalabilidade; custo computacional alto
	Proporcionalidade	PACC, MLCS	custo computacional baixo; sem sobrecarga de tráfego; alta escalabilidade; camada única	Não Aplicável em HetNet

dos caminhos e, portanto, não tem garantias que os dados são alocados para caminhos não compartilhados. Além disso, as abordagens não são escaláveis, uma vez que atualizar toda a infraestrutura da Internet não é uma tarefa simples e rápida. Do mais, abordagens como MPLS [Lopez et al., 2016], em que os subfluxos são marcados tendo como base nas cinco tuplas que identificam uma conexão (IP de origem, porta de origem, IP de destino, porta de destino, protocolo), pode ser mal interpretada por dispositivos de rede (*middleboxes*), como *proxys*, *IDS* e *firewalls*. Estes dispositivos podem reescrever estas tuplas e apagar o identificador adicionado pelas abordagens.

As abordagens da classe intracamada também possuem algumas desvantagens. Os métodos que utilizam sondagem ativa através do envio de informações nos cabeçalhos dos pacotes aumentam a sobrecarga de dados de controle nas transmissões. No entanto, eles podem reduzir esta sobrecarga aplicando um controle mais explícito sobre as técnicas de amostragem, como por exemplo, otimizando o intervalo (tempo) e a frequência em que são realizadas as sondagens. Por outro lado, os métodos passivos estão sujeitos à dinâmica dos caminhos heterogêneos, que introduzem ruídos nos dados e causam falsos positivos. A correlação ativa pode ser mais adequada para HetNets, uma vez que o envio de diferentes sequências de sondagem mitigam as perdas e as variações de atraso que afetam a precisão das abordagens passivas. Contudo, os métodos passivos mais suscetíveis a ruídos podem alcançar uma maior taxa de verdadeiros positivos e evitar a sobrecarga da transmissão. Nesta tese, defendemos que os métodos intracamadas são mais adequados à transmissão de dados multicaminhos no ambiente de HetNets, uma vez que o dispositivo móvel tem acesso a diferentes redes, que podem abranger diferentes domínios dispersos geograficamente, o que dificulta a implantação das abordagens intercamadas.

5.3 Avaliação

Esta seção apresenta uma análise sobre os métodos da classe intracamadas. O MSCM representa a categoria ativa, enquanto os demais a categoria passiva. O método PERT corresponde aos métodos estatísticos e o MCLS e o PACC aos métodos baseados em proporcionalidade. O

objetivo principal desta análise foi verificar a precisão dos métodos em um cenário HetNet. Para isso, os dados foram coletados de uma transmissão multicaminhos em um cenário controlado, emulado através do Mininet [Antonenko e Smelyanskiy, 2013]. Conforme ilustra a Figura 5.2, o dispositivo móvel ou nó de origem (N_s), tem acesso a duas diferentes redes, WiFi (r_1) e Celular (r_2), e se conecta com o nó de destino (N_d). Ambos os nós, de origem e destino, possuem um par de endereços IP ($\{1, 2\}$), o que lhes permite criar uma malha de 2×2 subfluxos entre si. No entanto, para facilitar a categorização e simplificar a discussão acerca dos caminhos compartilhados e não compartilhados, consideram-se apenas os caminhos A , B e C . Os caminhos B e C compartilham o gargalo entre os roteadores r_3 e r_4 . O gargalo tem uma capacidade de transmissão de 10% da capacidade dos demais enlaces da rede. Além disso, as características dos caminhos, em termos de largura de banda, atraso e taxa de perda, são configuradas para três diferentes cenários: (i) *homogêneo*, onde os caminhos têm a mesma largura de banda (10Mb/s), atraso (5ms) e taxa de perda (0%); (ii) *baixa heterogeneidade*, onde varia-se apenas a taxa de perda (de 0 a $0,01\%$) do caminho A ; (iii) *alta heterogeneidade*, em que o caminho A possui diferentes valores de largura de banda (100Mb/s), atraso (50ms) e taxa de perda ($0,01\%$), enquanto os caminhos B e C mantêm os valores do cenário homogêneo. Cabe ressaltar que embora os caminhos sejam definidos com as mesmas características no cenário homogêneo, o gargalo ocasiona perdas nos caminhos B e C . Não se faz o controle explícito das perdas no gargalo, uma vez que o ambiente é emulado com nós reais. Sendo assim, o cenário com baixa heterogeneidade pode apresentar uma maior similaridade entre os caminhos do que o homogêneo.

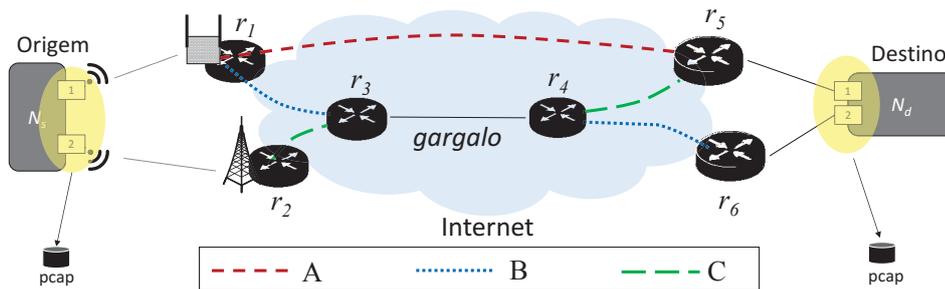


Figura 5.2: Cenário de coleta dos dados

O MSCM implementa um método de sondagem ativa para extrair o atraso de uma via (OWD, do Inglês *One Way Delay*). Uma vez que são utilizados traços da simulação, o *OWD* foi calculado extraindo informações em ambos os nós de origem e destino. Com isso, foi possível fornecer ao MSCM o *OWD* empregado no cálculo da correlação. Para o MPERT, se utilizou o valor do RTT suavizado (SRTT), calculado conforme [Paxson et al., 2011]. Ambos MCLS e PACC empregam a taxa de transferência (vazão) para calcular a relação entre os caminhos. Para o cálculo da vazão, utilizam-se informações sobre o tamanho da janela de congestionamento, SRTT e tamanho dos pacotes [Constantine et al., 2011]. Os métodos MSCM e MPERT calculam a correlação entre os pares de caminhos através do coeficiente de correlação de Pearson. E os métodos MLCS e PACC empregam uma regra proporcional simples ($\max(x, y)/(x + y)$). Para todos os métodos, os dados coletados dos caminhos, como OWD, RTT e vazão, são armazenados em uma janela deslizante, parametrizadas com um tamanho fixo (padrão de 50 valores). O funcionamento da janela é bem simples: quando a janela enche, conforme um valor é inserido, se remove o mais antigo. Adicionalmente, todos os métodos calculam a relação/correlação entre todos os pares de caminhos conforme um intervalo de tempo parametrizado (padrão de 100ms). Essas definições ajudam a criar um cenário comum e controlado para comparar a acurácia dos diferentes métodos. Por fim, a avaliação dos resultados foi agrupada conforme a subcategoria dos métodos, ou seja, estatísticos ou de proporcionalidade. Emprega-se c como variável referente

aos valores obtidos pelos métodos estatísticos, através do coeficiente de correlação de Pearson, e pelos métodos de proporcionalidade, através da sua referida equação.

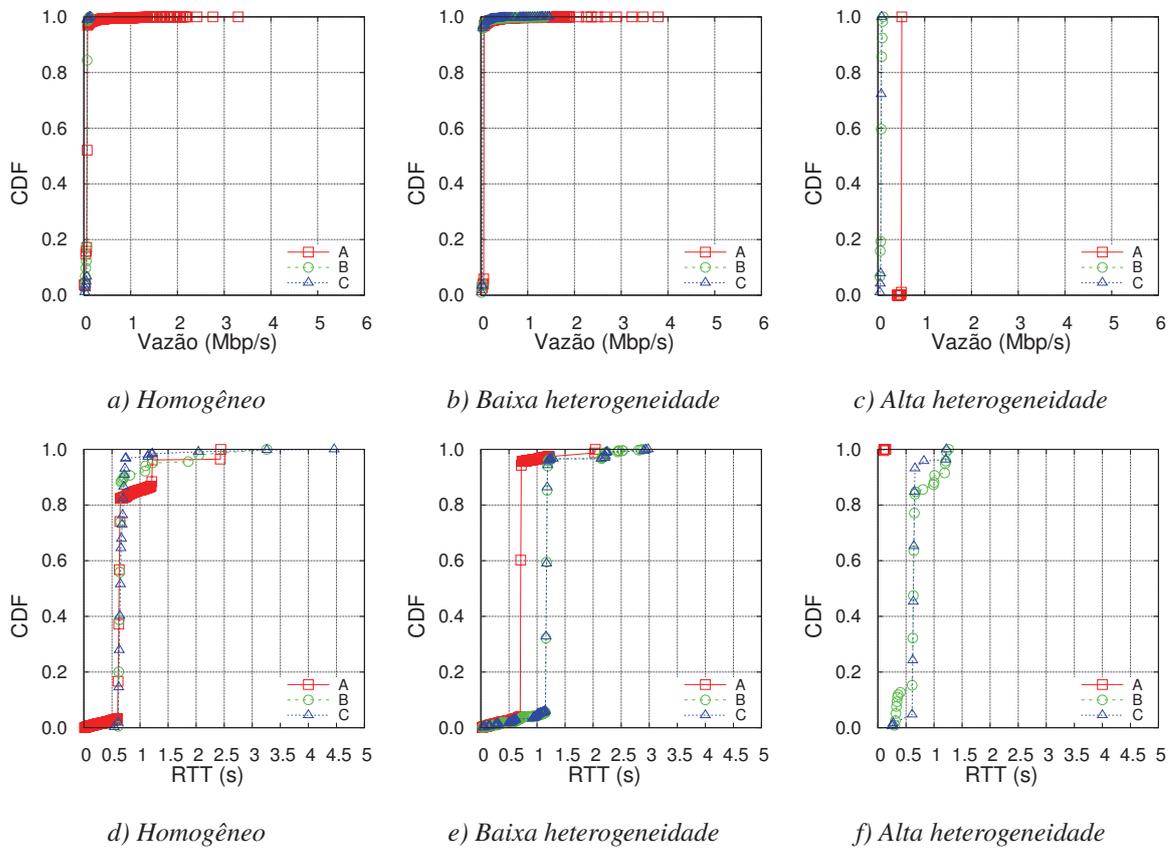


Figura 5.3: Distribuição cumulativa do atraso e da vazão dos caminhos

Antes de discutir os resultados da acurácia de cada método, é importante observar o atraso e a vazão dos caminhos nos três diferentes cenários. Os gráficos das Figura 5.3a, 5.3b e 5.3c representam a distribuição acumulada da vazão (CDF¹, do Inglês *Cumulative Distribution Function*) para os caminhos A, B e C. Embora o caminho A alcance uma taxa de transferência (vazão) superior aos demais caminhos, a maior parte dos dados são entregues abaixo dos 0,2Mbps/s em ambos os cenários homogêneos e de baixa heterogeneidade. Observa-se uma maior diferença no cenário com alta heterogeneidade, em que o caminho A alcança uma taxa de transferência muito distinta dos caminhos B e C. Em relação ao atraso, se observa um comportamento similar ao da vazão descrito anteriormente. Nos cenários homogêneos e de baixa heterogeneidade a distribuição os três caminhos possuem curvas de distribuição similares. No cenário com alta heterogeneidade os caminhos apresentam uma maior diferença no atraso.

A Figura 5.4 apresenta a distribuição dos valores de c (na figura descrito como r), coletados durante o experimento, para os métodos MCLS e PACC. Como descrito previamente, um valor de c próximo de 0,5 indica que os caminhos têm forte correlação. Por outro lado, um valor de c mais próximo de 1,0 indica que os caminhos têm baixa correlação. Os resultados mostram que o MCLS e PACC têm dificuldade para identificar os caminhos não compartilhados nos cenários homogêneo e de baixa heterogeneidade, onde as características dos caminhos, neste caso a vazão, são similares. Mesmo assim, a distribuição dos valores de c dos pares A – B e

¹O gráfico CDF tem como objetivo demonstrar a distribuição acumulada dos valores do eixo y em função do eixo x em um único experimento.

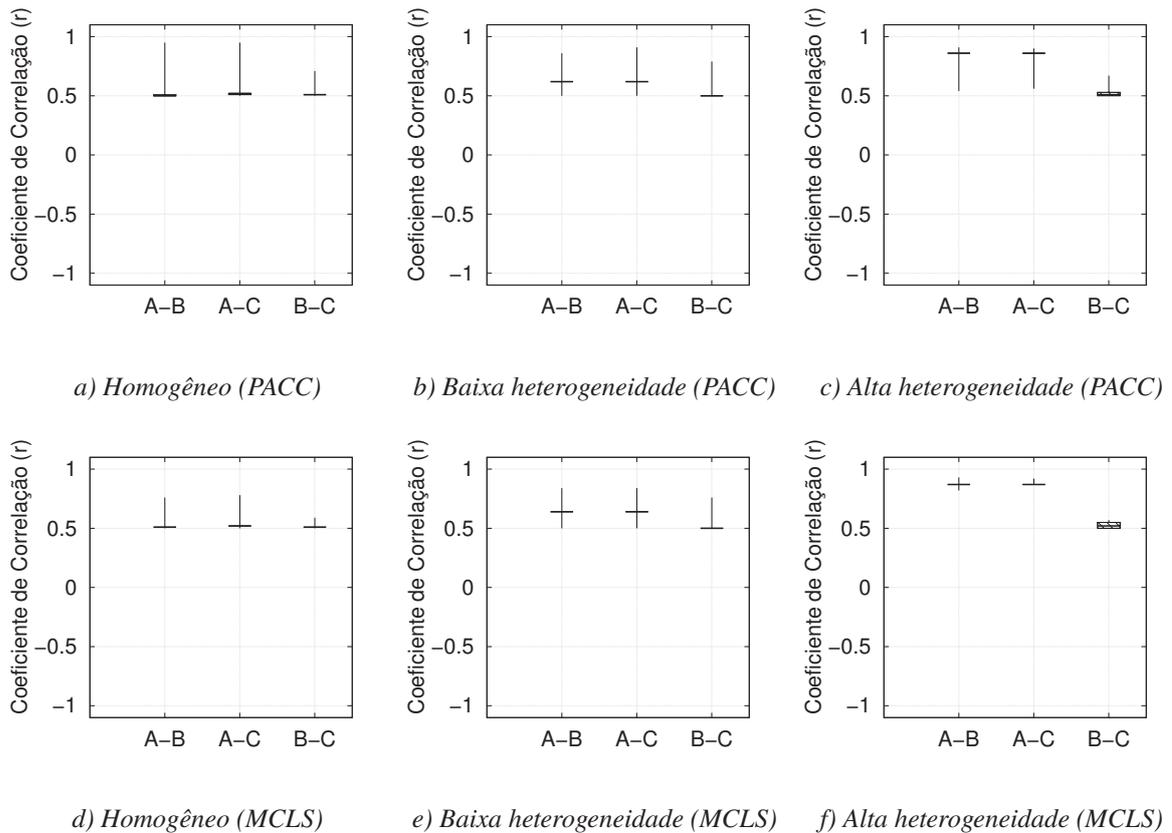


Figura 5.4: Distribuição dos valores do coeficiente c

$A - C$ apresentam uma maior dispersão em direção a 1 do que $B - C$, como ilustrado na transição de Figuras 5.4a e 5.4d para 5.4b e 5.4e. Conforme aumenta a heterogeneidade, especialmente a diferença entre os caminhos compartilhados e não compartilhados, a mediana dos pares $A - B$ e $A - C$ se aproximam de 1. Isto pode ser observado na transição entre os cenários de baixa heterogeneidade para o de alta heterogeneidade. Seguindo esta tendência, a distribuição dos valores de c referentes aos caminhos $B - C$ ficam mais precisas, tendo uma maior aglomeração em torno da mediana. Esse comportamento, salvo poucas diferenças, é comum para ambos os métodos MCLS e PACC.

A Figura 5.5 apresenta os valores de c para os métodos estatísticos MSCM e MPERT. Neste caso, 0 indica baixa correlação, -1 indica uma correlação negativa e 1 uma correlação positiva. Considera-se que os valores abaixo de 0 também indicam baixa correlação, uma vez que os valores negativos representam uma relação inversamente proporcional entre as variáveis. Para o cenário homogêneo, a mediana dos pares de caminhos não compartilhados $A - B$ e $A - C$ fica próxima de 0. Por outro lado, a mediana do par $B - C$ se distancia de 0, ficando próxima de 0,5.

A heterogeneidade dos caminhos causa uma grande variabilidade na distribuição dos dados, como pode ser observado para os cenário de baixa e alta heterogeneidade (Figura 5.5). No entanto, o cenário com alta heterogeneidade apresenta uma melhor distribuição dos dados, ficando a mediana dos pares $A - B$ e $A - C$ próxima de 0 e do par $B - C$ próxima de 1, para os dois métodos. O que não ocorreu para o cenário de baixa heterogeneidade. Neste caso, a mediana para todos os pares de caminhos ficou próxima de 0. A Figura 5.6 apresenta os valores de c durante o experimento ilustrado. Tendo como exemplo o método MPERT, se observa uma grande variação nos valores de c para todos os pares de caminho no cenário de baixa heterogeneidade. Isto não ocorre nos demais cenários, o que propiciou um melhor desempenho dos métodos

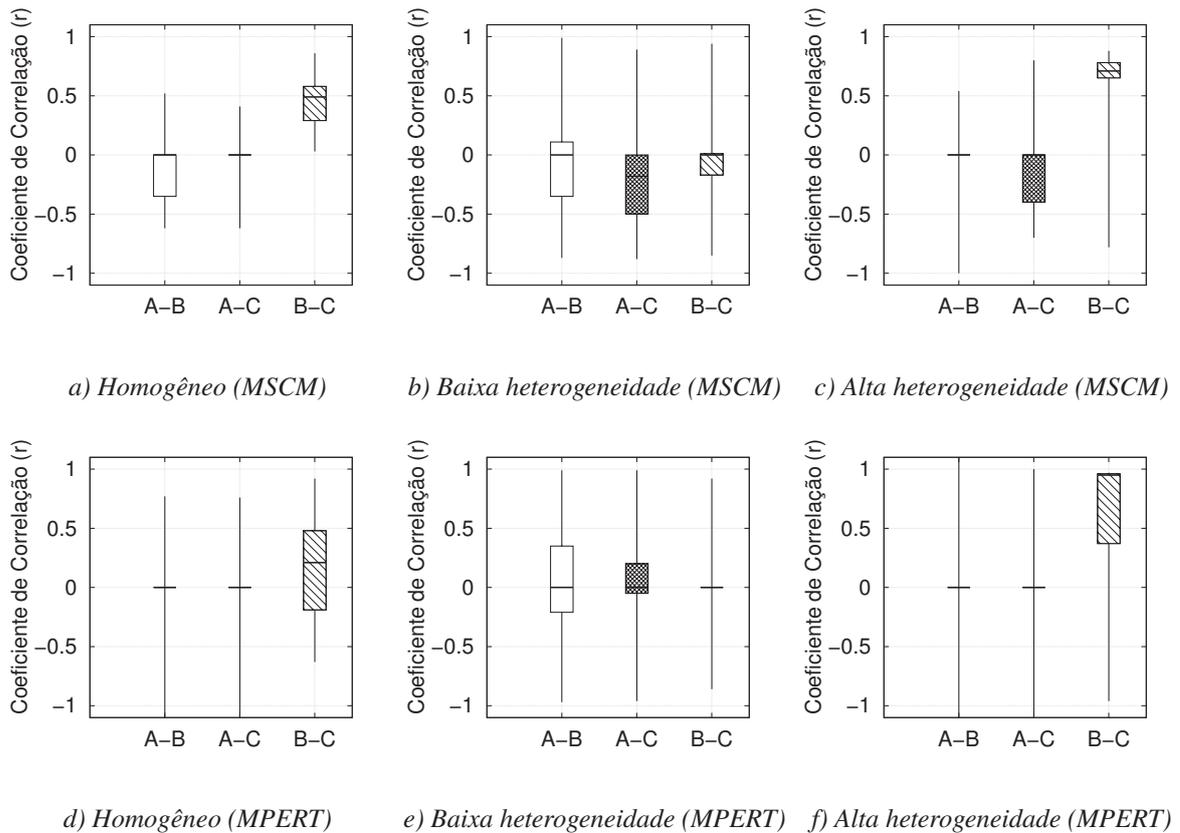


Figura 5.5: Distribuição dos valores do coeficiente c

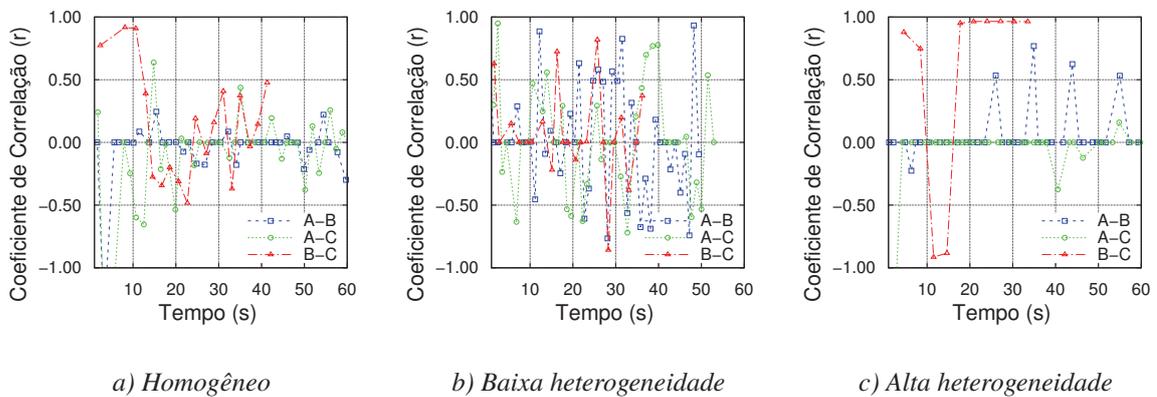


Figura 5.6: Valores de r durante o experimento para o método MPERT

observado na Figura 5.5. Embora os métodos baseados em estatísticas indiquem com mais precisão a correlação entre os caminhos $B - C$, as variações dos valores de c provocam muitos falsos positivos.

As Figuras 5.7a, 5.7b e 5.7c mostram a taxa de falsos positivos (FPR, do Inglês *False Positive Rate*) para os métodos MCLS e PACC. As Figuras 5.7d, 5.7e e 5.7f mostram a FPR para os métodos MSCM e MPERT. A FPR se refere ao percentual de falsos positivos obtidos conforme o limiar φ definido, que varia de 0 a 0,5 para os métodos de proporcionalidade e de 0 a 1 para os métodos estatísticos. Um falso positivo ocorre quando um par de caminhos não-correlacionada é classificado como correlacionado. Neste caso, consideram-se apenas os pares de caminhos não

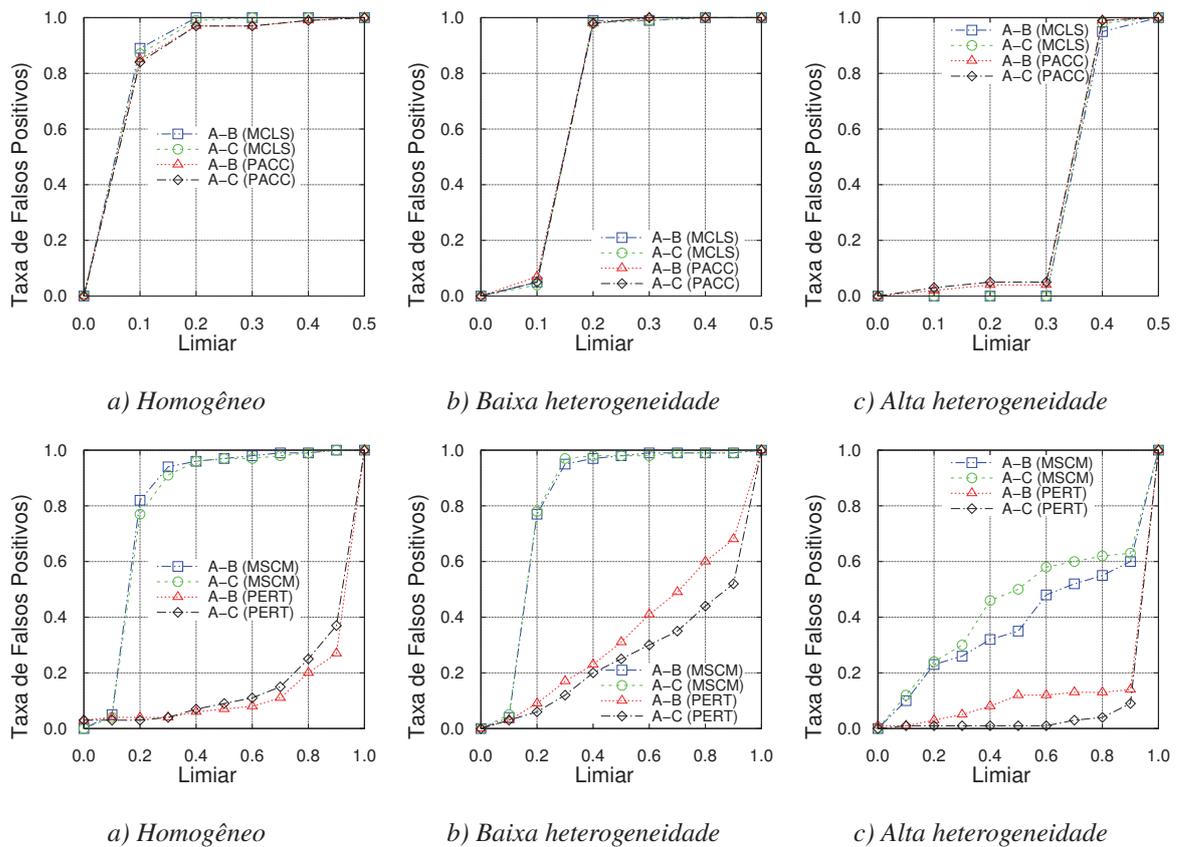


Figura 5.7: Acurácia dos métodos

compartilhados $A - B$ e $A - C$. Os métodos MCLS e PACC apresentam valores muito próximos em todos os cenários e ambos apresentam uma redução na FPR conforme o cenário se torna mais heterogêneo. A FPR do MCLS e do PACC ficaram abaixo de 0,1 para um limiar de 0,3 no cenário de alta heterogeneidade. Por outro lado, o MSCM e MPERT apresentam diferentes valores para a FPR, sendo que o segundo tem valores bem inferiores ao primeiro em cenários mais homogêneos. Esta diferença reduz no cenário com alta heterogeneidade, onde o MPERT alcança uma FPR abaixo de 0,2 até um limiar de 0,9.

Os resultados acima demonstram que ambos os métodos (proporcionalidade e estatísticos) conseguem indicar os caminhos que compartilham o gargalo quando as características entre os caminhos compartilhados e não-compartilhados são diferentes. No entanto, em um contexto geral, o MPERT teve uma menor taxa de falsos positivos. O cálculo da vazão beneficia os métodos de proporcionalidade, uma vez que os valores médios acabam funcionando como um filtro, que suaviza as flutuações. Esperava-se um melhor desempenho do método MSCM, uma vez que utiliza o atraso de uma via. No entanto, os dados utilizados não passaram por nenhum filtro, como feito para obtenção do SRTT, empregado pelo MPERT. Um ajuste mais fino nos parâmetros e também o emprego de um filtro na aquisição de dados pode reduzir os ruídos e melhorar a acurácia dos métodos. Por fim, não foi considerado o custo computacional de cada método. Em cenários com uma taxa de transmissão muito alta isto pode ter grande influência no intervalo definido para o cálculo da correlação. Uma opção seria deixar este intervalo dinâmico, como por exemplo, definindo um fator de ajuste baseado no tamanho da janela de congestionamento.

5.4 Importância dos resultados para a tese

Os resultados da análise dos métodos estatísticos e de proporcionalidade demonstram que ambas as soluções são cabíveis em determinar a diversidade de caminhos no cenário de alta heterogeneidade. A diferença entre os métodos está na complexidade do cálculo. Os métodos proporcionais são mais simples e com, como analisado empiricamente, requerem menos recursos de computação e armazenamento, sendo mais aplicáveis em dispositivos com recursos restritos. Por outro lado, os métodos estatísticos oferecem uma maior acurácia, sendo mais indicado para cenários com características dinâmicas, comum em redes móveis. O controle de conflitos como este (requisitos dos usuários e dispositivos vs. soluções) é uma das funções da estratégia de resiliência. Isto reitera a tese de que a diversidade de soluções é necessária para manter a resiliência diante os diferentes desafios e requisitos da transmissão.

5.5 Resumo

O capítulo apresentou um estudo de caso sobre as abordagens que provêm diversidade de caminhos para a transmissão de dados multicaminhos. Identificar ou prover a diversidade de caminhos é essencial para garantir que a transmissão possa usufruir dos benefícios da agregação de banda dos caminhos e também possa ser mais resiliente em caso de falhas em um dos caminhos. A taxonomia apresentada classificou as abordagens em duas categorias principais: intracamada e intercamadas. As abordagens intracamadas utilizam apenas informações dos caminhos para identificar a diversidade entre os caminhos. No segundo caso, as abordagens intercamadas se baseiam no relacionamento cooperativo ou não-cooperativo entre o protocolo de transporte multicaminhos e protocolos das camadas inferiores a fim de prover a diversidade dos caminhos. Esta última categoria pode criar uma dependência indesejada entre as camadas, especialmente quando existe um relacionamento cooperativo entre as abordagens. Por outro lado, as técnicas intracamadas são mais fáceis de implementar porque requerem mudanças apenas no protocolo de transporte. Mesmo sendo capazes de apenas identificar a diversidade, essas técnicas são úteis para melhorar o processo de alocação dos dados sobre os múltiplos caminhos. Neste sentido, apresentou-se uma análise das abordagens da classe intracamada. Os resultados mostram que os métodos analisados são capazes de determinar a diversidade entre os caminhos, especialmente quando existe uma diferença explícita entre as características dos caminhos. Os resultados colhidos neste capítulo são considerados para elaboração do estudo de caso apresentado no próximo capítulo.

6 Escalonador Redundante com Seleção de Caminhos Diversos

O Protocolo de Controle de Transmissão multicaminhos (MPTCP) emprega um escalonador redundante para replicar os pacotes de dados em múltiplos caminhos. Entretanto, a existência de enlaces compartilhados entre os caminhos fim-a-fim reduzem os benefícios do escalonador. Desta forma, este capítulo apresenta o RED (*REdundant Diversity scheduler*), um escalonador que prioriza a replicação dos pacotes por caminhos diversos (não-correlacionados). Os resultados deste estudo foram publicados em [Felix et al., 2018a, Felix et al., 2018c, Felix e Nogueira, 2018]. O capítulo está organizado da seguinte forma. A Seção 6.1 descreve o problema específico. A Seção 6.2 apresenta os trabalhos relacionados. A Seção 6.3 detalha a proposta do escalonador RED. A Seção 6.4 detalha a metodologia de avaliação empregada. A Seção 6.5 são apresentados os resultados. Por fim, a Seção 6.7 apresenta um resumo do capítulo.

6.1 Contextualização

A redundância é uma alternativa que tem sido explorada na literatura com objetivo de diminuir o impacto negativo causado pela heterogeneidade [Wang et al., 2016, Frömmgen et al., 2016, Hunger e Klein, 2016, Nguyen et al., 2017]. Em geral, o uso redundante dos caminhos depende do escalonador replicar os pacotes de dados nos caminhos disponíveis. O escalonador desconhece a diversidade dos caminhos e segue a premissa de que são disjuntos. Contudo isto nem sempre é verdade em redes reais, especialmente sobre a Internet [Rohrer et al., 2014]. Além disso, o MPTCP utiliza subfluxos TCP, o que força os protocolos de roteamento a tratá-los como fluxos TCP distintos e não restringir que sejam roteados por caminhos compartilhados. Os caminhos compartilhados limitam os benefícios da redundância e o desempenho da transmissão, uma vez que ambos estão sujeitos a falhas correlacionadas e gargalos em comum, respectivamente. Os métodos da classe intercadas, abordados no Capítulo 5, fornecem uma maior diversidade de caminhos para as transmissões multicaminhos [Lopez et al., 2016]. No entanto, como abordado previamente, isto cria um engessamento indesejável entre protocolos de camadas adjacentes.

Os métodos estatísticos, pertencentes à classe intracamadas buscam determinar a diversidade dos caminhos comparando estatisticamente suas características [Rubenstein et al., 2002]. No entanto, esses métodos sofrem com as variações dos caminhos no cenário das HetNets, que introduzem ruídos nos dados coletados, provocando muitos falsos positivos [Liao et al., 2011]. A fim de reduzir estes ruídos, os métodos da categoria ativa medem o atraso de uma via (OWD). Contudo, as técnicas empregadas para calcular o OWD adicionam informações aos cabeçalhos dos pacotes, aumentando a sobrecarga da rede com dados de controle. Isto não é desejável pois o MPTCP também adiciona informações ao cabeçalho dos pacotes. Com o escalonador redundante, as réplicas multiplicam esta sobrecarga. As técnicas da categoria passiva, embora

sejam afetadas pela heterogeneidade dos caminhos, evitam esta sobrecarga. No entanto, Hayes *et. al* [Hayes et al., 2014] sugere que as amostras de atraso observadas nos caminhos quando indicam um congestionamento podem apresentar um comportamento não monotônico, ou seja, as variáveis dos caminhos, como o atraso (RTT), tendem a mudar em conjunto, mas não necessariamente a uma taxa constante. Isto reduz a eficácia das técnicas passivas baseadas no coeficiente de correlação de Pearson, uma vez que este avalia uma relação linear entre duas variáveis contínuas.

Com objetivo de sobrepor estas limitações, este capítulo apresenta um novo escalonador redundante para MPTCP. O escalonador RED emprega uma técnica passiva para identificar os subfluxos não compartilhados. Para reduzir os efeitos da heterogeneidade ele utiliza o coeficiente de correlação de Spearman, menos suscetível às variações que ocorrem entre as amostras dos dados dos caminhos. Os resultados mostram que o RED potencializa os ganhos obtidos com o escalonador redundante usado pelo MPTCP ao replicar os pacotes prioritariamente pelos caminhos não correlacionados. Além disso, reduz o atraso e aumenta a vazão da transmissão multicaminhos. Por fim, o RED provê maior resiliência e melhor desempenho para as transmissões multicaminhos, o que tende a beneficiar as aplicações sensíveis ao atraso.

6.2 Trabalhos Relacionados

Na literatura existem diferentes escalonadores de pacotes para o MPTCP que empregam uma estratégia redundante, ou seja, replicam os pacotes de dados sobre os múltiplos caminhos. Em [Lopez et al., 2015], os autores propuseram um escalonador redundante a fim de aumentar a confiabilidade das transmissões de dados em sistemas SCADA (*Supervisory Control and Data Acquisition*). Outros exemplos, como rMPTCP [Hunger e Klein, 2016], ReMPTCP [Frömmgen et al., 2016] e MPTCP-L [Wang et al., 2016], também aplicam a técnica de replicar os pacotes de dados através dos múltiplos caminhos com objetivo de equalizar as variações de atraso. No entanto, isto ainda causa o impacto do problema do reordenamento devido à heterogeneidade dos caminhos. Particularmente, o ReMPTCP [A. Hunger and P. A. Klein and M. H. Verbunt, 2016] tenta reduzir este impacto usando as retransmissões, que são engatilhadas pelo problema do reordenamento, para transmitir um novo segmento. Com isso, o escalonador tenta evitar que um dado já replicado em outro caminho seja novamente enviado. O escalonador redundante pode causar uma grande sobrecarga de dados na rede conforme o número de caminhos utilizados. Em alguns casos, o número de réplicas pode ser otimizado para reduzir esta sobrecarga. Neste sentido, a versão mais recente do rMPTCP [A. Hunger and P. A. Klein and M. H. Verbunt, 2016] utiliza um parâmetro para controlar o número de réplicas e assim reduzir a sobrecarga.

De um modo geral, os caminhos compartilhados reduzem os benefícios alcançados com os escalonadores redundantes [Rohrer et al., 2014]. Por isso, [Lopez et al., 2016] propuseram tratar a diversidade de caminhos integrando os protocolos MPTCP e MPLS (*Multiprotocol Label Switching*), a fim de melhorar o escalonador redundante anteriormente apresentado em [Lopez et al., 2015]. Para alcançar a diversidade de caminhos, o MPLS se encarrega de identificar os subfluxos do MPTCP e encaminhá-los através de caminhos disjuntos. No entanto, o MPTCP não tem garantias de que os caminhos são realmente disjuntos. Neste caso, a abordagem é dependente da infraestrutura de rede, especificamente de alterações no protocolo MPLS. Esta dependência dificulta a escalabilidade da abordagem na Internet, uma vez que atualizar toda a infraestrutura de rede é uma tarefa custosa.

Outros trabalhos têm proposto abordagens para identificar os caminhos compartilhados confiando apenas em informação da camada de transporte [Liu et al., 2011, Singh et al., 2013,

Liao et al., 2011, Hayes et al., 2014, Hu et al., 2016]. Mesmo aplicando diferentes técnicas, a maioria das abordagens tem como base o trabalho de [Rubenstein et al., 2002], o qual emprega o coeficiente de correlação de Pearson para calcular a autocorrelação e a correlação cruzada entre dois caminhos. [Liao et al., 2011] aplicaram o mesmo coeficiente, mas sugeriram um esquema mais leve para obter o atraso de uma via (OWD). [Singh et al., 2013] propuseram um novo protocolo de transporte multicaminhos baseado em [Rubenstein et al., 2002]. O protocolo mostrou um melhor desempenho que o MPTCP. No entanto, [Hayes et al., 2014] defendem que as variações do RTT em cenários congestionados não seguem uma distribuição normal e que os dados coletados não possuem uma relação linear, como normalmente se espera.

Embora o OWD reduza os ruídos nas medições do RTT, as técnicas aplicadas para medir o OWD adicionam um sobrecarga de dados de controle às transmissões, como por exemplo, quando adicionam-se informações, como um *timestamp*, no cabeçalho dos pacotes [Liao et al., 2011, Hayes et al., 2014]. Particularmente, os escalonadores redundantes geram uma sobrecarga de dados de controle na rede com a replicação de pacotes. Portanto, medir OWD tem o custo de aumentar ainda mais esta sobrecarga na rede. Neste sentido, o escalonador proposto neste capítulo emprega uma técnica de correlação que não utiliza o atraso de uma via e também não depende da relação linear entre as amostras do RTT dos caminhos.

6.3 Escalonador RED

Os escalonadores redundantes fornecem uma comunicação altamente confiável, alocando os segmentos de forma redundante. A expectativa é que, no cenário das HetNet, os escalonadores redundantes equalizem as diferenças entre os caminhos. Embora o escalonador redundante do MPTCP tenha este objetivo, ele desconsidera que os caminhos podem ser compartilhados e estes benefícios podem ser anulados. O RED visa sobrepor as limitações do escalonador redundante do MPTCP, tolerando falhas e provendo a replicação de pacotes por caminhos com uma maior diversidade. A Figura 6.1 apresenta uma visão geral do RED e sua relação com o MPTCP. Ele situa-se no nó de origem N_s , especificamente como um componente do MPTCP, na camada de transporte. Os pacotes provenientes da camada de aplicação (em preto) são sequenciados e ordenados na fila de envio. Quando o RED possui dois ou mais caminhos disponíveis, ele aloca os pacotes e suas réplicas (pacotes vermelhos). O nó de destino N_d recebe os pacotes, descarta os duplicados, reordena e entrega os mesmos à camada de aplicação.

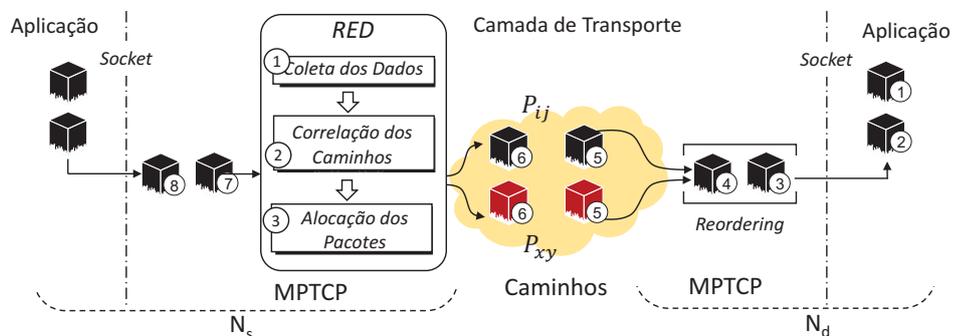


Figura 6.1: Escalonador RED com MPTCP

De uma forma mais detalhada, a Figura 6.2 ilustra as fases do RED: (i) *Coleta de dados*, quando o RED coleta e armazena os valores do atraso dos caminhos (RTT); (ii) *Correlação dos caminhos*, quando o RED calcula a correlação entre os pares de caminhos disponíveis; e (iii)

Alocação dos pacotes, quando o RED emprega os valores do coeficiente de correlação para alocar os pacotes e suas réplicas. A complexidade do tempo para (i) é $O(1)$, uma vez que ele apenas captura o SRTT (O RTT suavizado) de cada caminho; por (ii) é $O(n \log n)$, uma vez que usa um algoritmo *heapsort* para ordenar os valores do SRTT e calcular o coeficiente de correlação de Spearman; e por (iii) é $O(1)$, uma vez que apenas aloca os pacotes para os caminhos previamente selecionados. Todas as fases são detalhadas a seguir.

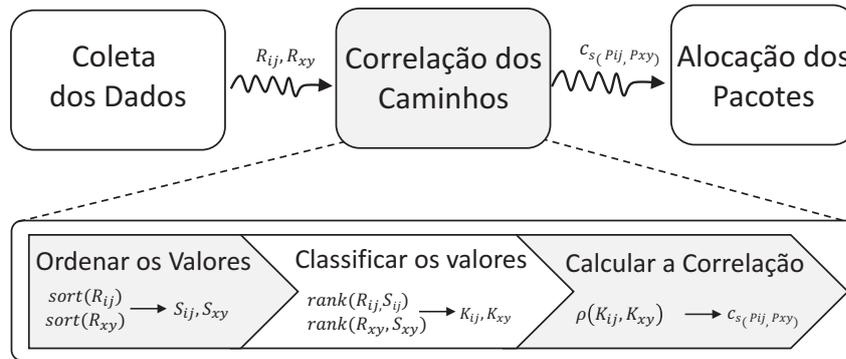


Figura 6.2: Fases do escalonador RED

6.3.1 Coleta dos Dados

A primeira fase corresponde à captura dos dados de cada caminho de forma passiva. Particularmente, se refere à coleta dos valores do SRTT medidas pelo MPTCP [Paxson et al., 2011]. O RTT suavizado (SRTT) [Paxson et al., 2011], computado internamente pelo MPTCP, tem menos ruído do que o RTT bruto. O RED obtém os SRTTs da seguinte maneira: dado um par de caminhos $\pi = \{P_{i,j}, P_{x,y}\}$, ele cria uma fila, que segue um esquema de janela deslizante, para cada caminho a fim de armazenar os valores do SRTT. Neste caso, as filas R_{ij} e R_{xy} correspondem às janelas deslizantes dos caminhos $P_{i,j}$ e $P_{x,y}$, respectivamente. Conforme ilustrado na Figura 6.3, ambas as janelas R_{ij} e R_{xy} têm tamanho N e são atualizadas, respectivamente, com os valores capturados do SRTT (r), onde $r_{i,j} \in R_{ij}$ e $r_{x,y} \in R_{xy}$, medidos a cada instante t . As janelas são atualizadas de modo independente e, portanto, as amostras não estão sincronizadas por tempo.

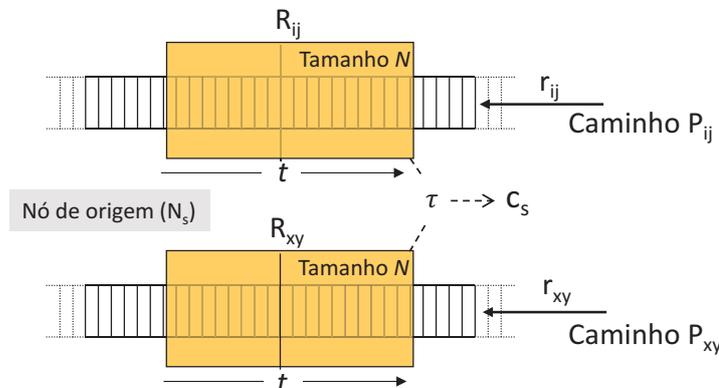


Figura 6.3: Esquema de janela deslizante

No entanto, o RED não depende que as amostras dos SRTTs dos caminhos estejam sincronizadas, uma vez que parametriza um intervalo de tempo τ para calcular a correlação. Isto

porque as diferenças de tempo, existentes entre os valores coletados, têm um menor impacto no método de correlação de Spearman.

6.3.2 Determinando a Correlação dos Caminhos

A segunda fase corresponde ao cálculo e atualização do grau de correlação c_s para cada par de caminho $\{P_{i,j}, P_{x,y}\}$. O RED emprega um método estatístico para calcular o grau de correlação. Para cada par de caminhos, tendo como entrada os dados da Fase de Coleta, calcula o coeficiente de correlação de Spearman (c_s). Assim como os métodos baseados no coeficiente de correlação de Pearson, o coeficiente c_s indica o grau de correlação entre um par de caminhos, sendo $c_s(P_{ij}, P_{xy})$. O grau de correlação c_s pode assumir qualquer valor entre -1 e 1 . Na Figura 6.4, c_s indica uma correlação positiva, quando o coeficiente tende para 1 ; nulo (quando não há correlação) quando o coeficiente tende a 0 ; e negativo, quando há uma relação inversamente proporcional entre a variável e o coeficiente tende a -1 .

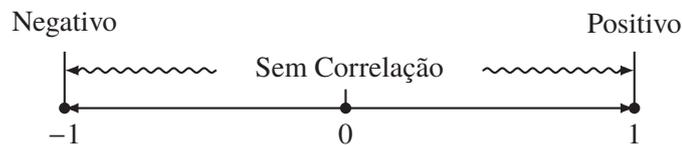


Figura 6.4: Coeficiente de Correlação (c_s)

O coeficiente de correlação de Spearman avalia a relação monotônica (linear ou não linear) entre duas variáveis contínuas ou ordinais. Em uma relação monotônica, as variáveis tendem a mudar proporcionalmente, mas não necessariamente a uma taxa constante. Levando isso em consideração, tendo em conta as variações frequentes causadas por eventos como a mobilidade dos nós no cenário das HetNets, este trabalho pretende demonstrar que o coeficiente de correlação de Spearman tende a proporcionar uma maior precisão na correlação entre caminhos do que outros coeficientes.

O RED executa as seguintes etapas para calcular o coeficiente de correlação de Spearman (Figura 6.2): (1) ordenam os valores em ambas as janelas deslizantes; (2) classificam (*ranking*) os valores ordenados, mantendo suas posições; e (3) calcula o coeficiente de correlação usando apenas valores classificados. Na Etapa 1, os valores R_{ij} e R_{xy} são colocados em ordem decrescente, sendo S_{ij} e S_{xy} o conjunto de saída dos valores ordenados, respectivamente. Na Etapa 2, os valores ordenados de cada conjunto S_{ij} e S_{xy} são classificados seguindo uma ordem decrescente. Como saída, tem-se $k_{ij} \in K_{ij}$ e $k_{xy} \in K_{xy}$, que são os índices referentes à classificação. Os índices em K_{ij} e K_{xy} são colocados na mesma posição (ou seja, na mesma ordem) que seus respectivos valores em R_{ij} e R_{xy} . A classificação ocorre da seguinte forma. O maior valor recebe o índice 1, o segundo maior o 2 e assim por diante. Isso é feito individualmente em cada conjunto S_{ij} e S_{xy} . Em seguida, os índices (k_{ij} e k_{xy}) são colocados na ordem original dos conjuntos R_{ij} e R_{xy} . Este passo tem como saída os conjuntos K_{ij} e K_{xy} . Uma vez que o SRTT pode se repetir, os empates na classificação são inevitáveis. Quando isso acontece, a sequência de valores empatados recebe um valor médio. Por exemplo, suponha que os sete primeiros valores ordenados sejam iguais. Como eles compartilham os índices 1, 2, 3, 4, 5, 6 e 7, recebem como classificação o índice médio desta sequência, que neste caso seria $28 \div 7 = 4$. Assim, o próximo índice a ser aplicado é 8, uma vez que as posições de 1 a 7 já estão em uso. Para a Etapa 3, calcula-se o coeficiente de correlação c_s usando os índices em K_{ij} e K_{xy} (Equação 6.1), onde ρ

se refere ao usual coeficiente de Pearson (ρ), sendo $cov(K_{ij}, K_{xy})$ a covariância (Equação 6.2) e $\sigma_{K_{ij}}, \sigma_{K_{xy}}$ (Equação 6.3) o desvio padrão.

$$c_s = \rho(K_{ij}, K_{xy}) = \frac{cov(K_{ij}, K_{xy})}{\sigma_{K_{ij}} \sigma_{K_{xy}}} \quad (6.1)$$

$$cov = \frac{1}{n} \left[\sum_{i=1}^n k_{ij} k_{xy} - \frac{1}{n} \left(\sum_{i=1}^n k_{ij} \right) \left(\sum_{i=1}^n k_{xy} \right) \right] \quad (6.2)$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (k_{ij} - \bar{k}_{ij})^2} \quad (6.3)$$

6.3.3 Alocação dos Pacotes

O RED aloca os segmentos e suas réplicas para os caminhos com baixa correlação. Neste caso, ele aloca os segmentos e suas réplicas para os caminhos que tenham um coeficiente c_s mais próximo de 0. Além disso, usa um limiar (φ) para controlar o número de falsos positivos. Estes passos são detalhados no algoritmo de alocação (Algoritmo 1) do RED. O algoritmo inicia a partir do conjunto de caminhos disponíveis (Φ). Em seguida, define um conjunto vazio C para separar os caminhos selecionados. A seguir, o RED adiciona a C o par de caminhos com o menor grau de correlação, sendo $\pi(P_{ij}, P_{xy}) = \min(c_s(P_{i,j}, P_{x,y})) \& c_s < \varphi \mid \forall P \in \Phi$. Para exemplificar, suponha que estejam disponíveis em um determinado instante os caminhos 1, 2, 3 e 4. O par de caminhos {1, 3}, hipoteticamente, tem o menor valor de correlação entre os demais. Portanto, no início, C recebe {1, 3} enquanto Φ mantém os caminhos 2 e 4. O próximo passo, verifica quais dos caminhos em Φ tem menor correlação com 1 e 3 e que esteja abaixo de φ . Considerando que o caminho 2 tem a menor correlação entre ambos, ele é adicionado a C . No próximo ciclo, são avaliados os demais caminhos em Φ . Neste caso, o caminho 4 possui baixa correlação com o 2, mas não com 1 e 3. Então, o caminho 4 não é utilizado. Por fim, o conjunto C contém os caminhos {1, 2, 3}, os quais possuem entre si o menor índice de correlação.

Algoritmo 1 Algoritmo de Alocação do RED

Input: Conjunto de caminhos Φ

- 1: $C \leftarrow \emptyset$
 - 2: $C \leftarrow \pi((P_{i,j}, P_{x,y}))$ with $\min(c_s(P_{i,j}, P_{x,y}))$ and $c_s > \varphi$
 - 3: **if** $\min(c_s(P_{i,j}, P_{x,y})), \forall P_{ij} \in P$ and $\forall P_{xy} \in C$ **then**
 - 4: $C \leftarrow P_{ij}$
 - 5: **end if**
-

6.4 Avaliação

Esta seção descreve a metodologia de avaliação, incluindo o cenário, parâmetros e limiares, e detalhes de implementação no kernel do Linux. O RED foi implementado como um módulo do MPTCP [Paasch et al., 2013], versão 0.92, Kernel Linux v.4.4.95 + (Debian 8). Esta implementação foi empregada para realizar os experimentos em uma topologia de rede emulada com o Mininet [Antonenko e Smelyanskiy, 2013]. A topologia de rede, conforme ilustra a Figura 6.5, representa um cenário HetNet em que um nó móvel N_s , com acesso a múltiplas redes sem fio, troca dados com um nó de destino N_d . O nó N_s estabelece uma conexão MPTCP com N_d através da rede de acesso WiFi (r_1) e posteriormente adiciona um segundo

subfluxo através da rede celular (r_2). Como ambos os nós N_s e N_d possuem múltiplas interfaces, um terceiro subfluxo é estabelecido combinando os endereços das interfaces $N_s(1)$ e $N_d(2)$. Poderia existir um quarto subfluxo entre $N_s(2)$ e $N_d(1)$, no entanto, não existe rota que ligue ambos. Com objetivo de simplificar, os subfluxos são referidos como caminhos $P_{1,1}$ (A), $P_{1,2}$ (B) e $P_{2,1}$ (C), sendo que B e C compartilham um gargalo entre roteadores r_3 e r_4 (linha vermelha).

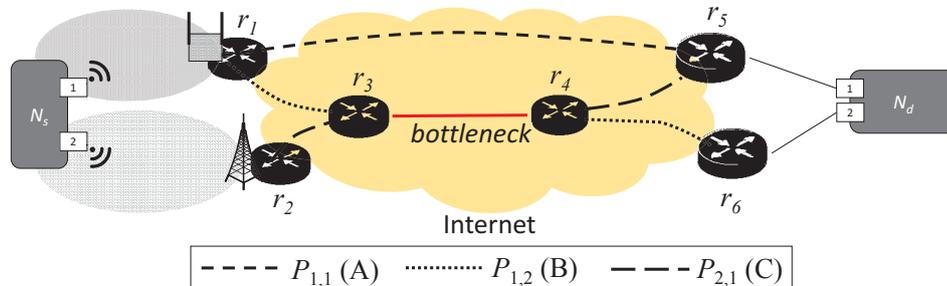


Figura 6.5: Cenário de Avaliação

Os principais parâmetros definidos para os experimentos são: porcentagem de perda do enlace (PP_e), largura de banda do enlace (LB_e) e atraso do enlace (A_e) (Tabela 6.1). Os valores para os parâmetros têm como objetivo representar um cenário controlado de HetNet, como em [Ferlin et al., 2014]. Além desses parâmetros, os experimentos são realizados com o algoritmo de controle de congestionamento (CC) desacoplado (TCP Cubic) [Xu et al., 2016], a fim de evitar o acoplamento dos caminhos, realizado pelo CC tradicional do MPTCP; utiliza-se o gerenciador de caminhos *full-mesh* [Paasch et al., 2012], o qual cria uma malha de subfluxos combinando os endereços IP acessíveis entre N_s e N_d . Para comparar o desempenho do método de correlação do RED, utiliza-se como base o método MSCM [Liao et al., 2011], que emprega o coeficiente de correlação de Pearson. O MSCM é um dos métodos estatísticos proeminentes baseado no trabalho precursor de [Rubenstein et al., 2002]. Ele usa o OWD para minimizar os ruídos e calcular a correlação entre os caminhos. Ambos são comparados usando os mesmos parâmetros, descritos posteriormente.

Tabela 6.1: Parâmetros de simulação

Parâmetros	$r_1 \iff r_5$	$r_2 \iff r_6$	$r_3 \iff r_4$	outros
PP_e (% pacotes)	0,001	0,0001	0,001	0,00
LB_e (Mbit/s)	150	100	20	100
A_e (ms)	5	15	40	1

A implementação padrão do MPTCP no Kernel do Linux possui uma estrutura modular, que inclui diferentes escalonadores e algoritmos de controle de congestionamento. Deste modo, o RED foi adicionado como um módulo do MPTCP, sendo opcional sua seleção. Foram mantidas algumas práticas de implementação as quais são empregadas no desenvolvimento do Kernel Linux. Uma das principais foi evitar o uso de cálculos de ponto flutuante, uma vez que existem arquiteturas que não suportam esse tipo de operação. Isto aumenta a compatibilidade do RED. No mais, foi realizado um ajuste na replicação de pacotes. De uma forma geral, todos os pacotes são replicados quando existem dois ou mais caminhos disponíveis e que atendem os requisitos de baixa correlação (Algoritmo 1). No entanto, a replicação é restrita somente aos pacotes na fila de envio que ainda não foram escalonados.

6.5 Resultados

Esta seção compreende duas partes: (i) a primeira apresenta os resultados sobre a acurácia do coeficiente de Spearman em analisar a correlação entre os caminhos (Fase 2 do RED); (ii) a segunda parte apresenta os resultados de desempenho do escalonador RED (todas as fases) e o compara com o escalonador redundante do MPTCP. Ao apresentar os resultados do método de correlação emprega-se o termo RED-CM (do Inglês, *RED Correlation Method*), uma vez que se avalia apenas a Fase 2 do RED. Na seção posterior, usa-se normalmente o termo RED, que engloba todas as fases, e o termo MPTCP-Redundante, para se referir ao escalonador redundante do MPTCP.

Desempenho do método de correlação

Inicialmente, são coletados os valores do RTT e do OWD para cada um dos caminhos A , B e C a partir dos traços de simulação (pcaps) do cenário da Figura 6.5. As simulações empregam os mesmos parâmetros descritos na Seção 6.4, juntamente com o escalonador redundante do MPTCP. Estes valores do RTT e OWD são usados como entrada para comparar o desempenho dos métodos MSCM e RED-CM, respectivamente. A fim de calcular o OWD, são obtidos os tempos de recebimento dos pacotes de dados em N_d , e o tempo de recebimento (RTT) dos pacotes de reconhecimento (ACK) em N_s . Para o cálculo da correlação emprega-se uma janela deslizante de tamanho $N = 50$ e um intervalo de tempo de $t = 0,01ms$.

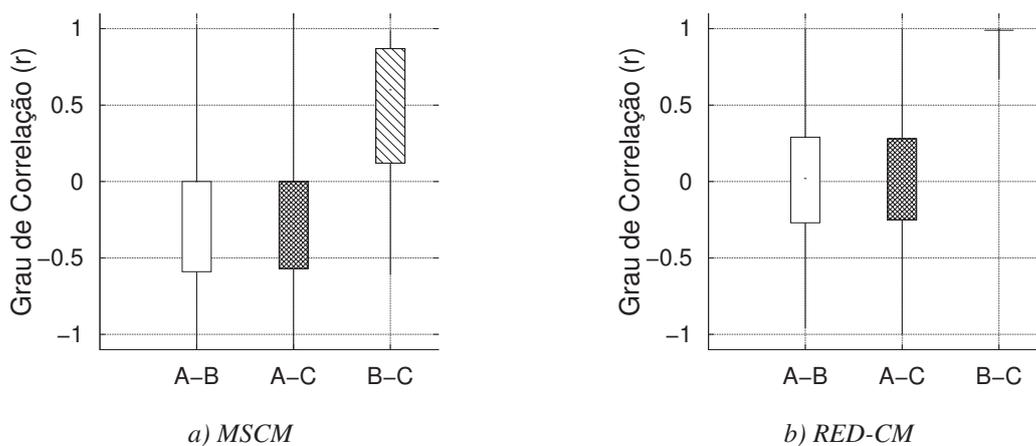


Figura 6.6: Valores de c_s durante um experimento

Conforme ilustra a Figura 6.6, a qual apresenta a distribuição dos valores do coeficiente de correlação c_s calculado entre os pares de caminhos $\{A, B\}$, $\{A, C\}$ e $\{B, C\}$, durante o tempo de experimento $T_e = 60s$. A distribuição de c_s para o MSCM indica fortemente que os pares $\{A, B\}$ e $\{A, C\}$ possuem uma correlação nula ou negativa (inversamente proporcional), ou seja, não são correlacionados. Para o par de caminhos $\{B, C\}$, embora tenha uma grande dispersão entre 0, 2 e 0, 8, indica corretamente a correlação positiva entre os caminhos. O RED-CM tem dificuldade em indicar a correlação dos pares $\{A, B\}$ e $\{A, C\}$, ficando a mediana da distribuição dos dados sobre 0. No entanto, tem uma maior eficácia em indicar a correlação entre os pares $\{B, C\}$. Neste caso, os valores se concentram perto de 1, o que indica que existe uma forte correlação.

O RED-CM mostra-se mais eficiente que o MSCM para detectar os caminhos correlacionados. Conforme ilustra a Figura 6.7.a, que apresenta a taxa de falsos positivos (FPR) para os caminhos não compartilhados, conforme um limiar (φ) variando entre 0 e 1. Para medir os falsos

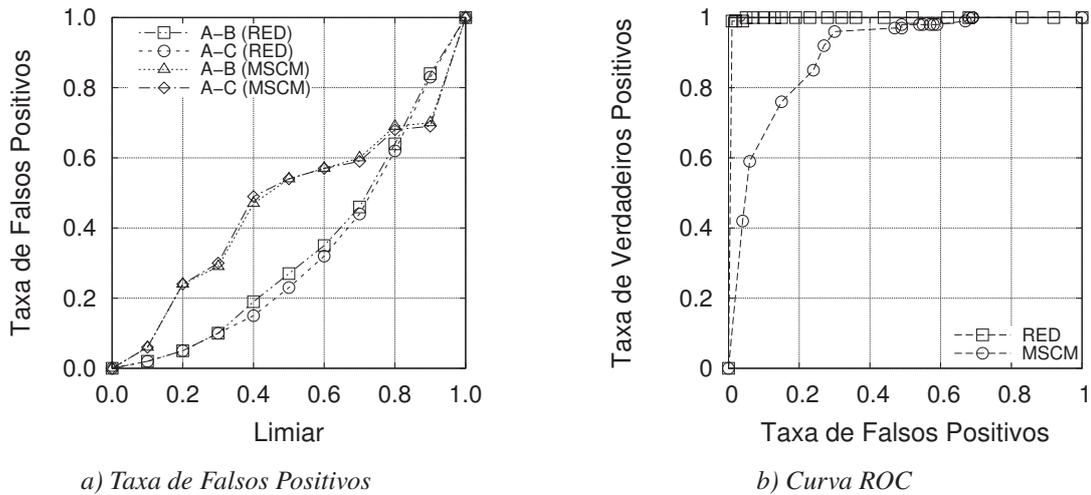


Figura 6.7: Acurácia do RED-CM vs. MSCM

positivos, foi considerado apenas o 0 como ponto central e que indica correlação nula. Neste sentido, para cada valor de φ , o FPR indica a taxa de caminhos não correlacionados indicados como correlacionados. O RED-CM apresenta um FPR menor que o MSCM para $\varphi \leq 0,8$. A curva ROC, ilustrada pela Figura 6.7.b, demonstra as variações de sensibilidade (positivos verdadeiros) no eixo y e 1-especificidade (falsos positivos) no eixo x, para diferentes valores de φ . A curva do MSCM indica uma maior variação na taxa de falsos positivos em relação à taxa de verdadeiros positivos (TPR). Nesse caso, a partir de 0,8 TPR, o MSCM aumenta o FPR para 0,2, enquanto o RED-CM atinge um TPR maior que 0,9 e um FPR inferior a 0,1.

Os resultados apresentados na Figura 6.7 também mostram que um limiar dentro do intervalo $0,8 \geq \varphi \geq 0,9$ pode ser usado para indicar com grande precisão se os caminhos estão correlacionados ou não. A partir desses resultados, o limiar de $\varphi = 0,8$ foi empregado como parâmetro na implementação do RED em Kernel Linux.

Desempenho do Escalonador

A Figura 6.8 apresenta a distribuição acumulada (CDF) do atraso (RTT) e da taxa de transferência (vazão) para os caminhos A, B e C. Pode-se dizer que o caminho A não tem gargalos e, portanto, não possui variação expressiva do atraso nos experimentos. Neste sentido, a comparação do atraso é mais expressiva para os caminhos compartilhados B e C. Conforme ilustra a Figura 6.8.(b), correspondente ao RED, o mesmo apresenta uma redução significativa no atraso, comparado aos resultados do MPTCP-Redundante, na Figura 6.8.(a). Isso acontece porque o RED prioriza o par de caminhos com baixa correlação. O RED também reduz a vazão dos caminhos B e C, movendo o tráfego para o caminho A, que tem valores acumulados a partir de 1,3, maior que o MPTCP-Redundante, que inicia em 0,8.

Além do atraso e da taxa de transferência, outras duas métricas são definidas a fim de comparar o RED e MPTCP-Redundante, como: (i) o **número cumulativo de ACKs duplicados (CNDA)**; e o **número acumulado de retransmissões (CNR)**. O CNDA conta o número de ACKs duplicados (Dup-ACK) e servem para mostrar o recebimento acumulado de segmentos recebidos fora de ordem em N_d durante o experimento. Por outro lado, o CNR mostra o número acumulado de retransmissões que ocorrem devido à perda de pacotes ou recebimento de Dup-ACKs durante o experimento. Como ilustrado na Figura 6.9.(b), o RED apresenta diferentes valores de CNDA entre os caminhos B e C. Enquanto o CNDA do MPTCP-Redundante, ilustrado pela Figura 6.9.(a), mostra que estes caminhos apresentam valores acumulados bem próximos.

Isto acontece também com o CNR, ilustrado pelas Figuras 6.9.(c) e 6.9.(d). Isso ocorre porque o RED replica menos pacotes pelos caminhos com menor grau de correlação, otimizando o uso dos caminhos *A* e *B*, o que reduz tanto o CNDA e o CNR.

Definiu-se também uma métrica para avaliar a sobrecarga causada pelo número de réplicas criadas pelo RED e MPTCP-Redundante, denominada **número médio de réplicas (MNR)**. O MNR calcula o número médio de segmentos encaminhados (replicados) através de mais de um caminho. Os valores do MNR são apresentados com um intervalo de confiança de 95% para 35 amostras, ou seja, repetições do experimento. A Figura 6.10 mostra que o RED reduz o MNR total em comparação ao MPTCP. No entanto, o MNR do caminho *C* é muito menor. Isso novamente ocorre devido ao RED priorizar os caminhos com baixa correlação. Nesse caso, o caminho *A*, que não está correlacionado com os demais, acaba sendo o mais priorizado. Como o RED escolhe o segundo caminho, no caso o *B*, com base no menor grau de correlação com o caminho *A*, este recebe uma quantidade similar de réplicas. Sendo assim, o caminho *C* apenas recebe um número menor de réplicas.

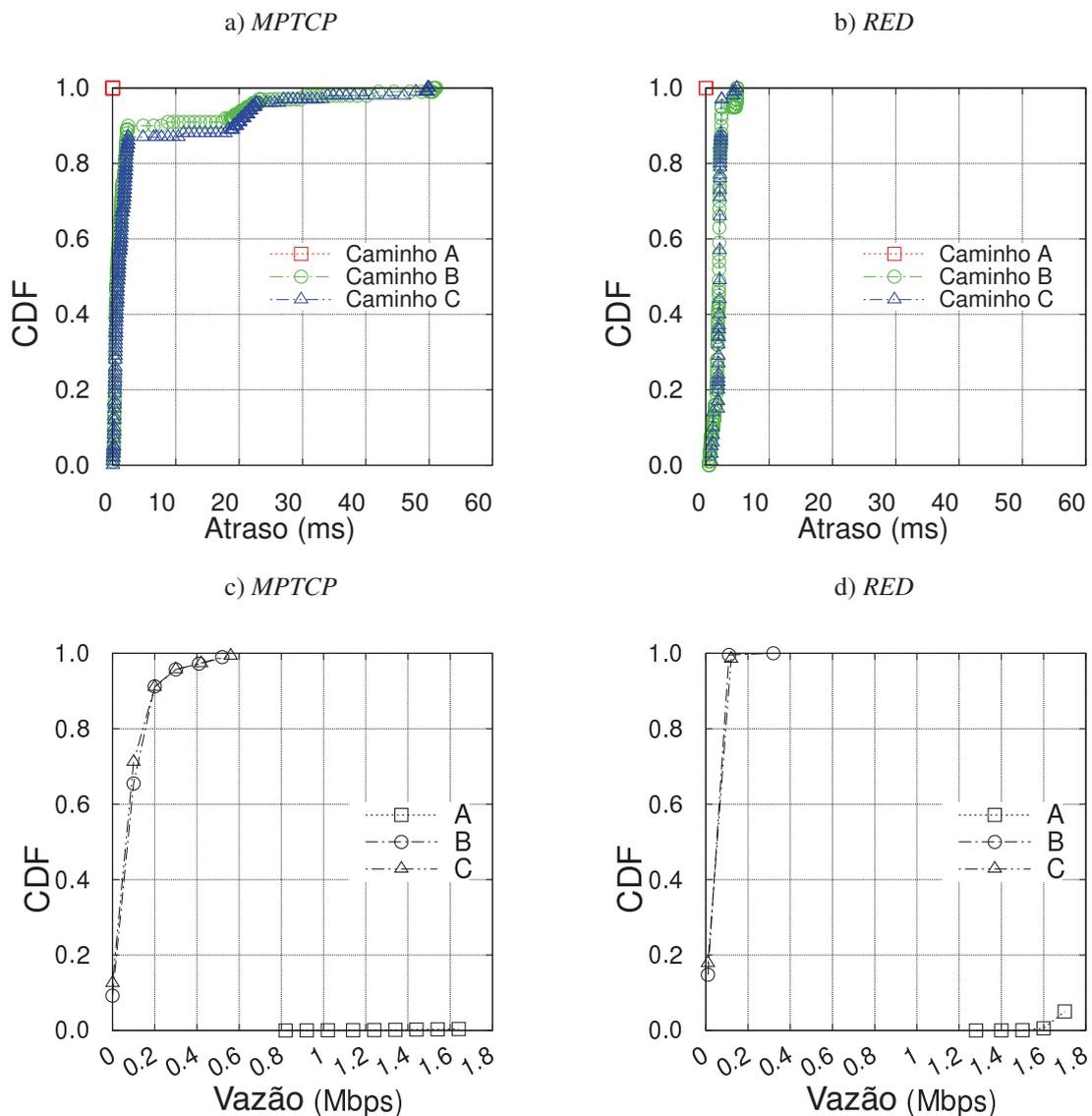


Figura 6.8: CDF do Atraso e Vazão (*Throughput*)

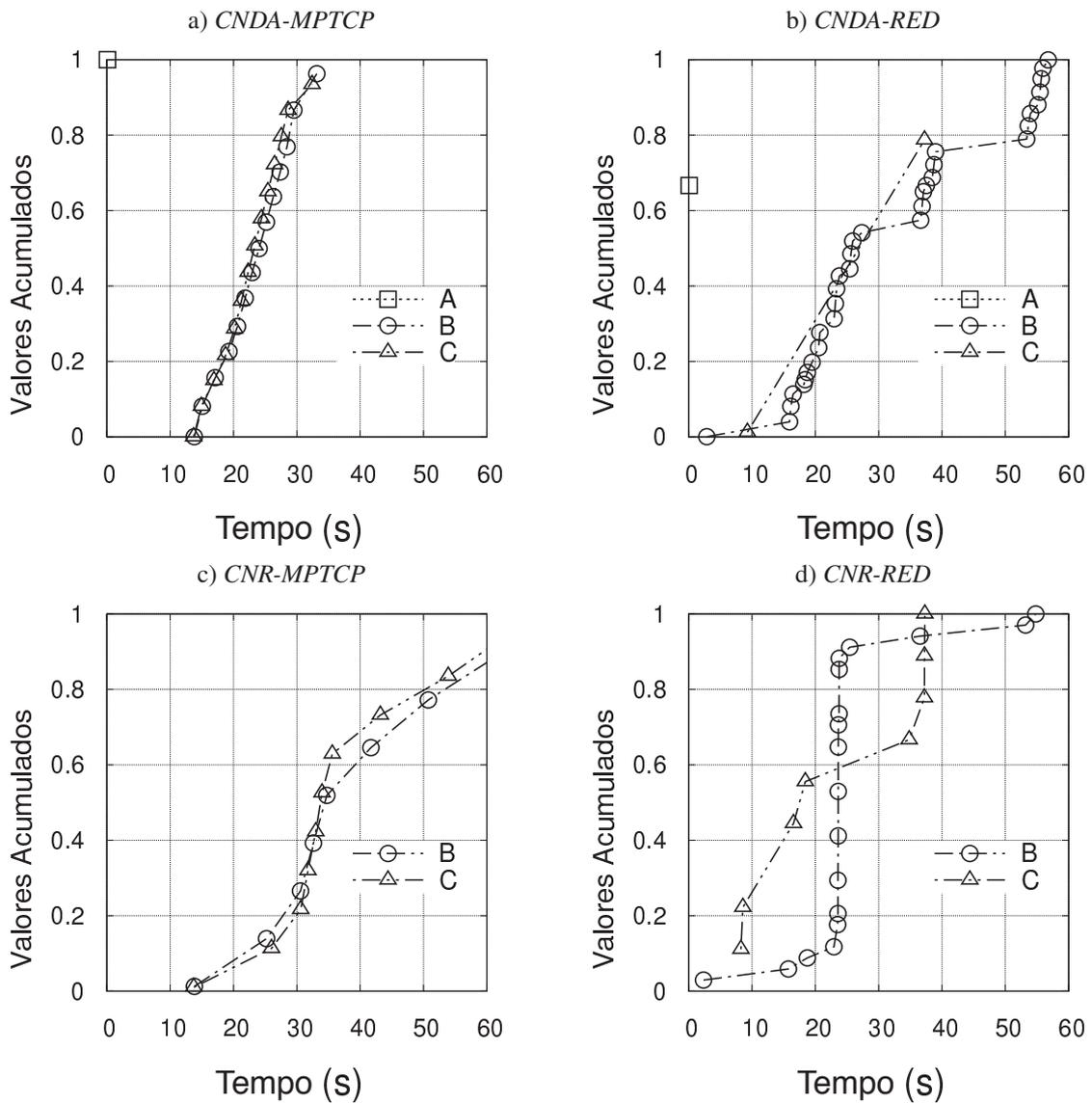


Figura 6.9: CNDA vs. CNR

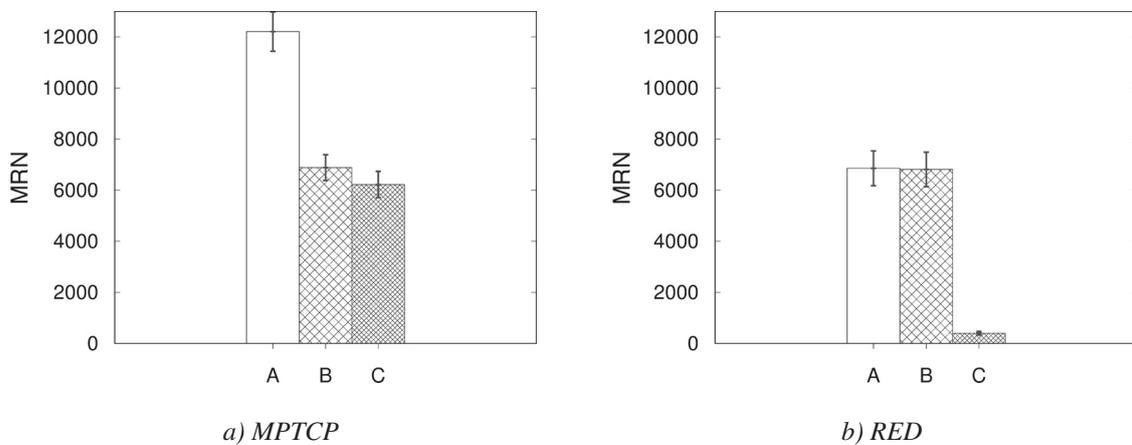


Figura 6.10: MPTCP Original Scheduler vs. RED: Overhead

6.6 Importância dos resultados para a tese

Ao observar os princípios para construção de redes resilientes, a redundância deve estar associada à diversidade a fim de prover uma maior resiliência [Sterbenz et al., 2010]. Na transmissão multicaminhos, a redundância e a diversidade estão associadas ao uso dos múltiplos caminhos. Neste caso, a falta de diversidade, que se refere aos caminhos compartilhados, reduz os benefícios da redundância, limitando o desempenho e afetando a resiliência da transmissão.

Os resultados obtidos neste capítulo mostram que prover redundância com diversidade em transmissões multicaminhos melhora o desempenho e a resiliência da transmissão, uma vez que a solução proposta prioriza os caminhos disjuntos. Fazendo uma relação com a estratégia de resiliência proposta, identificar a diversidade de caminhos é uma das funções do módulo de reconhecimento, assim como prover o uso redundante dos caminhos com diversidade é uma das medidas a serem empregadas pelo módulo de resistência. Isto reforça a ideia do relacionamento entre estes dois módulos com o módulo de controle, responsável por aplicar as medidas. Neste caso, o módulo de controle determina a política a ser empregada com base nos requisitos da transmissão, repassando ao módulo de reconhecimento e resistência as informações e os métodos que seriam empregados para identificar e prevenir o uso de caminhos compartilhados. O escalonador RED e seu método para identificar a correlação entre os caminhos se encaixam como medidas dos módulos de reconhecimento e resistência.

6.7 Resumo

Este capítulo apresentou o RED (*REdundant Diversity Scheduling*), um novo escalonador redundante para o MPTCP. O escalonador RED emprega uma técnica para determinar a diversidade dos caminhos e alocar as réplicas dos pacotes apenas para os caminhos que tem o menor grau de correlação, ou seja, a menor chance de serem compartilhados. A técnica empregada está alinhada com os objetivos do módulo de reconhecimento da estratégia de resiliência proposta. Outras técnicas podem fazer parte a fim de atender as diferentes políticas de resiliência adotadas. Em relação ao RED, os resultados mostraram uma baixa taxa de falsos positivos na identificação dos caminhos não correlacionados, quando comparado com o método MSCM, um dos métodos representativos da literatura. Além disso, o RED possui uma maior taxa de verdadeiros positivos. Os resultados dos experimentos utilizando a implementação do RED em Kernel Linux confirmam a eficiência do RED. Além disso, a abordagem de replicação de pacotes do RED reduz o uso dos caminhos compartilhados e melhora o uso dos demais caminhos. Por fim, o escalonador RED reduz o atraso e aumenta a vazão mínima dos caminhos não compartilhados.

7 Resistindo ao *Bufferbloat* em Caminhos Compartilhados

Este capítulo apresenta um estudo sobre o impacto causado pelo fenômeno *bufferbloat* na transmissão multicaminhos em um enlace compartilhado pelos subfluxos. Este fenômeno, evidenciado por [Nichols e Jacobson, 2012], ocorre devido à existência de grandes filas que absorvem enorme quantidade de tráfego em um canal congestionado, acarretando alta latência com baixa perda de pacotes. Neste capítulo, os resultados obtidos evidenciam um *trade-off* entre a vazão e o atraso em função da disciplina de controle de fila utilizada na rede. Para reduzir este *trade-off*, foi proposta uma estratégia para mitigar o problema em nível de rede. Os resultados deste estudo foram publicados em [Felix et al., 2016, Felix et al., 2017]. O capítulo está organizado da seguinte forma. A Seção 7.1 descreve o contexto, o problema abordado e os trabalhos relacionados. A Seção 7.2 detalha o cenário e metodologia de avaliação. A Seção 7.3 apresenta uma discussão do impacto do fenômeno do *bufferbloat* nas transmissões multicaminhos. A Seção 7.4 descreve a solução proposta para mitigar o problema. Os resultados alcançados são descritos na Seção 7.5. Por fim, a Seção 7.7 faz um resumo do capítulo.

7.1 Contextualização Trabalhos Relacionados

O MPTCP tem experienciado uma redução no desempenho em cenários heterogêneos [Ferlin et al., 2014, Khalili et al., 2012, Arzani et al., 2014], com aumento do atraso e redução da vazão. Isto se deve à diferença entre as características dos caminhos, que ocasionam o problema do reordenamento e degrada seriamente o desempenho. A degradação aumenta quando os caminhos sofrem grandes variações do atraso causadas pelo fenômeno *bufferbloat* [Ferlin-Oliveira et al., 2014]. Uma vez que o controle de congestionamento do MPTCP utiliza a perda de pacotes como sinal para identificar o congestionamento, o *bufferbloat* não percebido acaba interferindo no desempenho dos demais caminhos.

Estudos recentes acerca do *bufferbloat* atestam a existência do problema em redes sem fio [Alfredsson et al., 2013]. As abordagens empregadas para mitigar o problema envolvem algoritmos de controle de congestionamento (CC) ou disciplinas de controle ativo de filas (AQM, do Inglês *Active Queue Management*). O MPTCP utiliza um algoritmo de CC baseado em perdas, o que leva a uma detecção tardia do fenômeno. As disciplinas de fila dependem dos administradores da rede para que sejam implementadas nos pontos de encaminhamento. No melhor de nosso conhecimento, a disciplina AQM *Controlled Delay* (CoDel) [K. Nichols, 2016] e suas variações, como *Flow Queue Controlled Delay* (FQ-CoDel) [Hoeiland-Joergensen et al., 2015], possuem bons resultados frente ao fenômeno com redução expressiva do tempo de permanência na fila.

A maioria dos roteadores atuais empregam um mecanismo FIFO (*First In, First Out*) para encaminhamento dos pacotes, e por simplicidade, descartam aqueles que chegam quando a fila está cheia (*DropTail*). A disciplina *DropTail* frente às grandes filas presentes nos roteadores e a

um congestionamento persistente tem acarretado grandes variações no atraso. Sobre a perspectiva deste problema, a disciplina FIFO é ineficiente. O problema das grandes filas (*bufferbloat*) resulta no aumento excessivo da latência e degradação da vazão que acabam afetando a disponibilidade da rede. A disciplina CoDel visa mitigar o problema controlando o tempo de permanência do pacote na fila, e consequentemente limitando o tamanho da fila. A disciplina CoDel guarda o instante enq_i que cada pacote p_i é colocado na fila. Ao retirar o pacote da fila no tempo deq_i , ele calcula o tempo de permanência do pacote, conforme Equação 7.1).

$$\delta_i = deq_i - enq_i \quad (7.1)$$

Este intervalo é utilizado pelo mecanismo empregado pelo CoDel para alternar entre dois estados: descarte e não-descarte. A disciplina CoDel inicia no estado de não-descarte e pode mudar seu estado conforme valor do intervalo (μ). O valor inicial de μ é λ , porém o valor é ajustado conforme o número de descartes consecutivos (n_{drop}). Se entre o tempo $t_1 = t$ e $t_2 = t + \mu$ um pacote p_i é desenfileirado tal que, $deq_i \in [t_1, t_2]$ e $\delta_i > \tau$, a CoDel entra em estado de descarte e os próximos pacotes desenfileirados são descartados enquanto $\delta_i > \tau$, conforme ilustra o diagrama de fluxo da Figura 7.1. A variável que controla do número de descartes consecutivos é incrementada em $n_{drop+} = 1$ e o valor de μ é atualizado para $\mu\sqrt{n_{drop}}$. Caso $\delta_i < \tau$, a disciplina CoDel permanece ou entra no estado de não-descarte e a variável n_{drop} é reinicializada com 1 e μ para λ [Kulatunga et al., 2015].

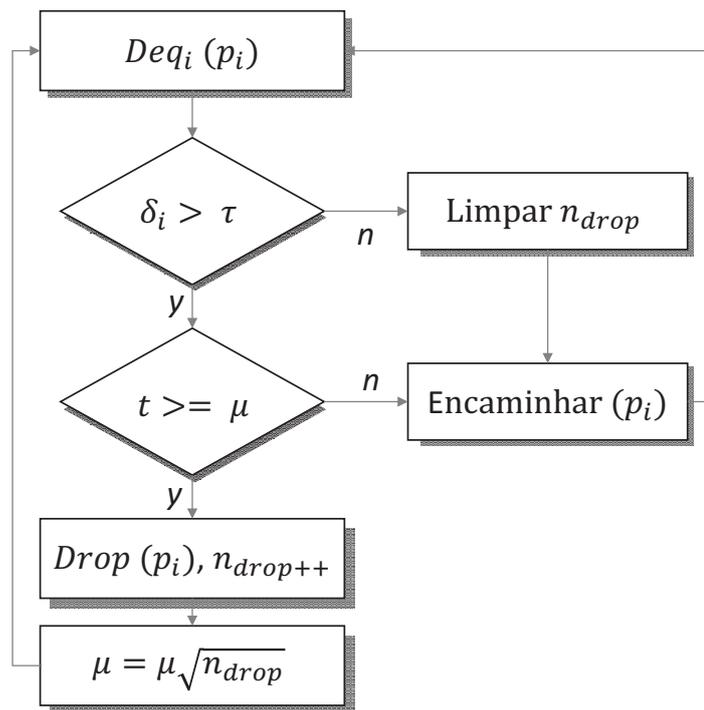


Figura 7.1: Diagrama de Fluxo do CoDel

A variação FQ-CoDel emprega a mesma abordagem do CoDel para controlar o atraso do pacote na fila. A diferença é que ele separa os fluxos em n filas para aplicar o mecanismo do CoDel de forma independente. Para tornar a utilização justa, os pacotes são escalonados com um algoritmo *Deficit Round Robin* (DRR) [Hoeiland-Joergensen et al., 2015]. O algoritmo DRR configura uma quantidade de Q_i (*quantum*) bytes para ser retirado da fila n_i a cada ciclo. Caso a fila n_i não tenha Q_i bytes, este déficit é acumulado para o próximo ciclo. O CoDel e FQ-CoDel sugerem valores $\tau = 5ms$ e $\lambda = 100ms$ como padrão para maioria das situações [Ali et al., 2015].

Os autores em [Chen e Towsley, 2014, Ferlin-Oliveira et al., 2014] investigaram o impacto do fenômeno do *bufferbloat* no desempenho do MPTCP. Eles mostraram que, no cenário de redes celulares, o fenômeno é responsável por uma grande variação e um grande aumento do RTT, prejudicando as transmissões com o MPTCP. [Schwarzkopf et al., 2016] avaliaram três diferentes disciplinas de fila, o CoDel, PIE e CoDel-ACT, sendo o último uma variante do CoDel que propõe uma alteração no tempo de descarte. Os autores mostraram que o CoDel pode levar a uma subutilização do enlace devido ao mecanismo de controle do tempo de permanência dos pacotes na fila, enquanto as disciplinas PIE e CoDel-ACT provêm uma melhor utilização do enlace. [Kulatunga et al., 2015] avaliaram o CoDel e o FQ-CoDel sobre redes com capacidade limitada e que apresentam grandes atrasos. A contribuição do trabalho foi de otimizar os parâmetros do CoDel a fim de reduzir a latência para este cenário específico. Esses trabalhos avaliaram o *bufferbloat* em HetNets com o MPTCP ou o desempenho do CoDel frente a outros AQMs. No entanto, nenhum deles propôs reduzir o impacto causado pelo descarte de pacotes nos algoritmos de controle de congestionamento baseado em perdas em transmissões onde os caminhos atravessam um gargalo compartilhado.

7.2 Cenário de Avaliação

A fim de compreender o impacto do fenômeno *bufferbloat* na transmissão multicaminhos foram realizadas simulações utilizando o NS-3 [Kheirkhah, 2015], comparando os algoritmos de CC e a disciplina de fila CoDel. Este estudo está em conformidade com as recomendações da RFC 7567 [Baker e Fairhurst, 2015], que sugere uma maior investigação da interação entre as disciplinas de fila e os novos protocolos de transporte. O cenário de simulação combina as disciplina de fila, DropTail e CoDel com os algoritmos de controle de congestionamento (LIA, RTT-Compensator ou Uncoupled) disponível na implementação do MPTCP [Kheirkhah, 2015]. Conforme ilustra a Figura 7.2, o cenário representava um ambiente comum de transferência multicaminhos para um dispositivo móvel (e.g., smartphone), empregado em trabalhos recentes [Ferlin et al., 2014, Ferlin-Oliveira et al., 2014, Chen e Towsley, 2014]. Neste cenário, um dispositivo com múltiplas interfaces (UE, do Inglês *User Equipment*) se conecta a duas redes de acesso heterogêneas e transmite os dados para o receptor. O roteador funciona como um *gateway* de borda e provê acesso entre as redes LTE (enlace A), WiFi (enlace B), nó UDP (enlace D) com o receptor (enlace C). A UE utiliza ambos os caminhos de forma concorrente.

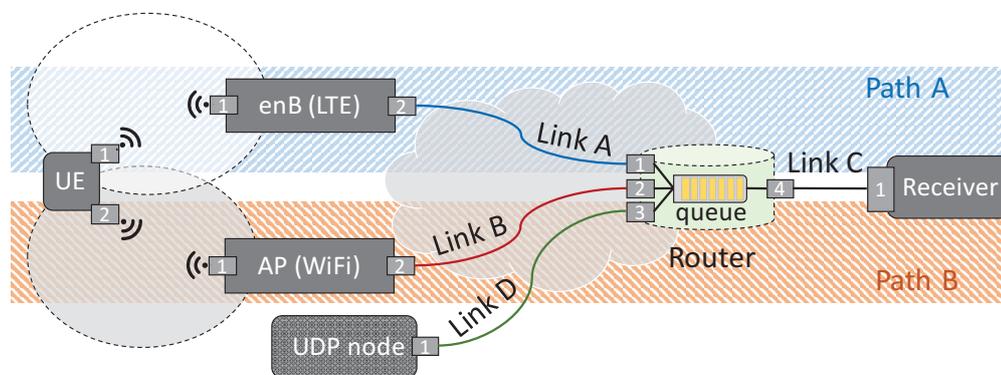


Figura 7.2: Topologia do cenário de simulação

Os pacotes trafegando pelo caminho A e B compartilham a mesma fila do roteador e o gargalo, ou seja, o enlace C entre o roteador e o receptor. O nó UDP gera uma carga de trabalho

(pacotes UDP) e envia para o receptor com uma taxa constante de 0,25 Mbps. Seu único objetivo é gerar tráfego para a fila. Em todas as simulações, a UE estabelece a conexão com receptor pela interface LTE. Posteriormente, a UE abre um segundo caminho através da interface WiFi. Após o estabelecimento do segundo subfluxo, os dados são transmitidos pelos dois caminhos.

Tabela 7.1: Parâmetros de Simulação

Parâmetros	Valores
Taxa de Dados	1.0
Atraso - Caminho A	1, 10, 100, 300 ms
Tamanho da fila	100 pacotes
Tamanho dos pacotes	1458 bytes
CWND, RCWD	64 Kb
Controle de Congestionamento	LIA, RTT_Compensator, Uncoupled

As simulações são realizadas com uma carga de trabalho fixa de 4 Mbytes para cada execução. A UE transfere essa quantidade de dados para o receptor simultaneamente pelos caminhos *A* e *B*. Inicialmente, os links *A*, *B* e *D* têm 1 Gbps de largura de banda e 1 ms de atraso. Estes valores são mantidos fixos para manter um cenário controlador de simulação. No entanto, em ambientes reais, especialmente em redes sem fio, os atrasos mudam constantemente. Esta variação é simulada variando o atraso do enlace *A*, entre 1,10,100 e 300 ms. A diferença de atraso força o escalonador (i.e., LRF) a dar preferência para o caminho mais rápido.

As métricas a seguir foram utilizadas para avaliar o desempenho dos algoritmos de controle de congestionamento e das disciplinas de fila: (i) *Goodput* médio (Equação 7.2); (ii) RTT médio por caminho (Equação 7.3); (iii) número médio de pacotes perdidos (Equação 7.4); (iv) comprimento médio da fila e tempo médio de permanência (Equação 7.5). Nas equações, *n* refere-se ao número de repetições (isto é, mudanças de semente) e *t* ao tempo de simulação necessário para transferir a carga de trabalho em cada execução. O tempo de permanência e o comprimento da fila são calculados apenas para disciplinas baseadas no CoDel.

$$\overline{Gdp} = \frac{1}{n} \sum_{i=1}^n Gdp_i \quad (7.2)$$

$$Gdp_i = \frac{\text{bytes_received} \times 8}{\text{TimeLastPkt} - \text{TimeFirstPkt}}$$

$$\overline{RttPath_k} = \frac{1}{n} \sum_{i=1}^n RttRep_i \quad (7.3)$$

$$RttRep_i = \frac{1}{t} \sum_{j=1}^t Rtt_j$$

$$\overline{DropsQueue} = \frac{1}{n} \sum_{i=1}^n DropRep_i \quad (7.4)$$

$$DropRep_i = \sum Drop$$

$$\begin{aligned}\overline{Mean} &= \frac{1}{n} \sum_{i=1}^n Rep_i \\ Rep_i &= \frac{1}{t} \sum_{j=1}^t V_j\end{aligned}\tag{7.5}$$

7.3 Impacto do *bufferbloat* na Transmissão multicaminhos

Os resultados apresentados abaixo referem-se à comparação entre os algoritmos LIA, *RTT_Compensator* e *Uncoupled* com as disciplinas CoDel e DropTail. Esta última foi empregada como base para comparações. A Figura 7.3 apresenta os resultados do tamanho médio da janela de congestionamento (CWND) e o valor médio do RTT para o CoDel e DropTail, provenientes de simulações realizadas com o algoritmo LIA. Em todos os gráficos, o eixo x representa o atraso atribuído ao caminho A . Com ambos os caminhos com atraso de $1ms$, o RTT médio observado, tendo o CoDel como disciplina de fila do roteador, foi de $85ms$ para o caminho A (LTE) e $67ms$ para o B (WiFi). Os resultados indicam que a variação do atraso do caminho A influencia o RTT do caminho B . Com atraso de $100ms$, o RTT médio do caminho A foi de $276ms$ e o caminho B variou entre $69ms$ a $71ms$. Tendo como disciplina de fila o DropTail, o congestionamento tem um impacto considerável no RTT médio, atingindo $691ms$ para o caminho A e $478ms$ para o B com atraso de $100ms$. Estes valores são quase 9 vezes maiores que os obtidos pelo CoDel. Os gráficos com os resultados para *RTT_Compensator* e *Uncoupled* não são apresentados, uma vez que tiveram uma pequena variação em relação ao LIA de aproximadamente 5% a 8%, respectivamente.

Ainda na Figura 7.3, pode-se observar que o comportamento da janela de congestionamento (CWND) reflete o funcionamento do algoritmo de CC do MPTCP. Conforme o atraso aumenta no caminho A , a sua CWND é reduzida a fim de priorizar a CWND do caminho B , que possui um menor RTT. Contudo, este balanceamento do algoritmo não ocorre de forma adequada quando o CoDel é empregado, o que pode ser observado no gráfico quando o caminho A tem atraso de $10ms$. Neste caso, ambas as janelas tiveram um redução de tamanho e os resultados apresentados na Figura 7.4.(a), referentes ao descarte médio na fila, ajudam a compreender o porquê. O CoDel (CoDel-FIFO) descarta um número muito maior de pacotes que o DropTail, impedindo a CWND de crescer. Os descartes têm grande impacto no desempenho da transmissão multicaminhos, como pode ser observado no *goodput* apresentado pela Figura 7.4.(b). O DropTail possibilita que toda a capacidade do canal disponível para as transmissões com MPTCP seja alcançada (i.e., 75%), enquanto o CoDel um pouco mais de 50%.

A variação do *goodput* do CoDel está diretamente relacionada ao número de descartes de pacotes na fila. Conforme o número de pacotes descartados aumenta, diminui o *goodput* da transmissão. Contudo, o RTT médio é expressivamente menor que o observado com o DropTail. Como o DropTail não descarta pacotes enquanto a fila é suficiente para absorver o tráfego, este atraso acarreta um RTT médio muito alto, no entanto, como a fila está sempre cheia, a utilização do canal é mais eficiente. Com o CoDel, a fila não atinge 20% de sua capacidade. Isto pode ser observado nos resultados apresentados na Figura 7.5.(a), referentes ao tamanho médio da fila alcançado com os diferentes algoritmos de CC. O algoritmo *Uncoupled* possui resultados mais divergentes com o algoritmo padrão (LIA) porque não acopla os subfluxos. Deste modo o atraso no caminho A interfere menos no caminho B e o mesmo consegue encaminhar mais pacotes. Mas isto também interfere no tempo de permanência do pacote na fila. Conforme Figura 7.5.(b), este tempo é maior que os demais na maioria dos casos observados.

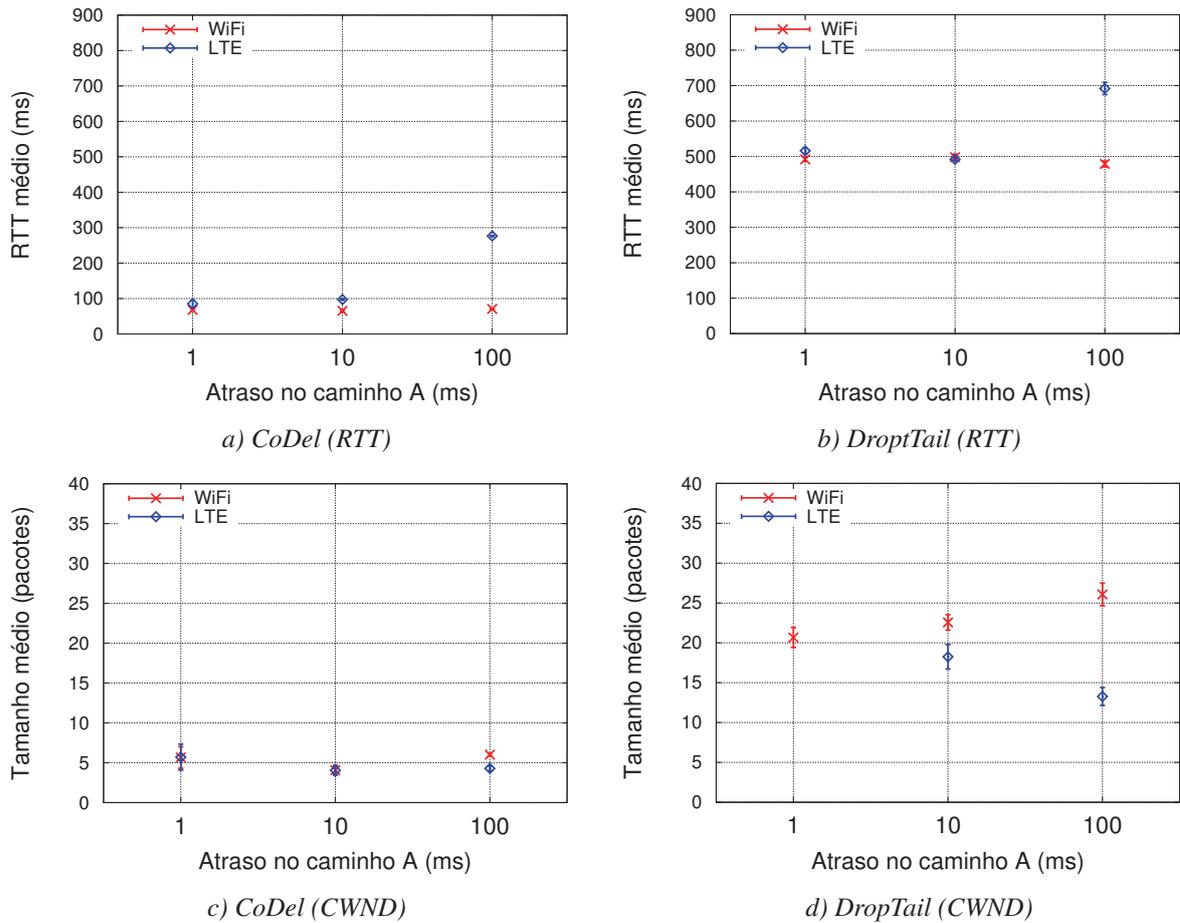


Figura 7.3: Desempenho da transmissão

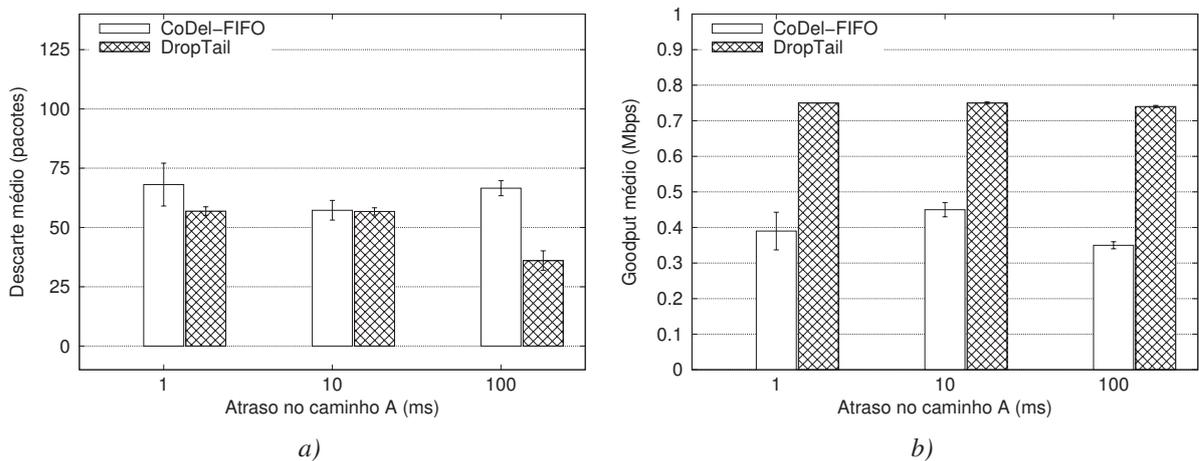


Figura 7.4: Descartes vs. Goodput

O CoDel reduz drasticamente o RTT dos caminhos em relação à disciplina padrão DropTail, contudo, penaliza o *goodput* da transmissão devido ao elevado número de descartes na fila. Este *trade-off* entre o *goodput* e a latência é percebido desde a proposta inicial da disciplina FIFO-CoDel em [Nichols e Jacobson, 2012], porém com o MPTCP, se demonstra mais expressivo. Com intuito de reduzir o *trade-off*, este trabalho projetou e avaliou uma nova

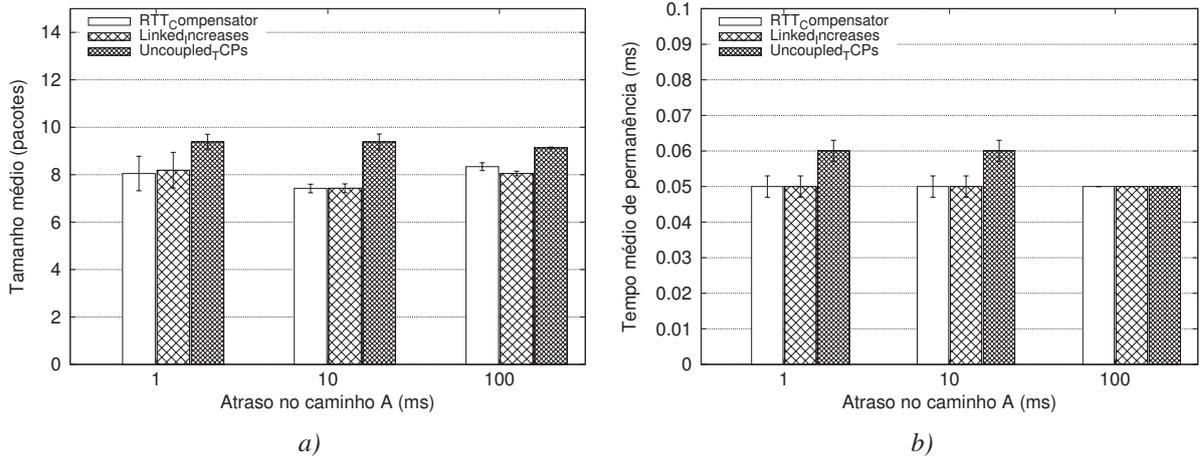


Figura 7.5: Estado da fila com CoDel vs. MPTCP CC

disciplina de fila que controla os descartes na fila para reduzir o impacto das perdas de pacotes nas transmissões multicaminhos.

7.4 Mitigando o problema em nível de rede

Esta seção apresenta o LIFO-CoDel [Felix et al., 2016], uma nova proposta para reduzir o impacto causado pelo mecanismo de controle de fila no CC do MPTCP. A estratégia empregada pela proposta altera o escalonamento dos pacotes FIFO (*First In, First Out*), utilizada originalmente pelo FIFO-CoDel, para LIFO (*Last In, First Out*) além dos demais ajustes detalhados a seguir. Esta proposta foi inspirada em [Maurer, 2015], o qual utiliza uma estratégia denominada LIFO adaptativo com CoDel para controlar a sobrecarga de requisições nos servidores do Facebook. Neste trabalho, a proposta LIFO-CoDel é projetada para um contexto original de redes heterogêneas sem fio, as quais precisam de soluções específicas capazes de tratar suas características peculiares em conjunto com os requisitos de suas aplicações.

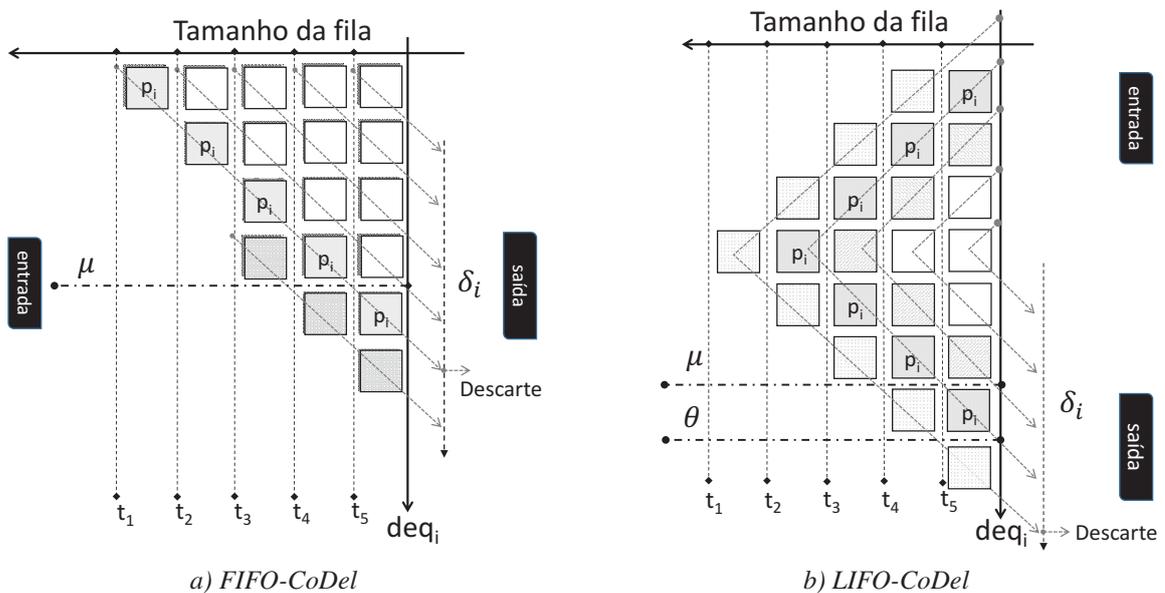


Figura 7.6: Atraso na fila vs. pilha

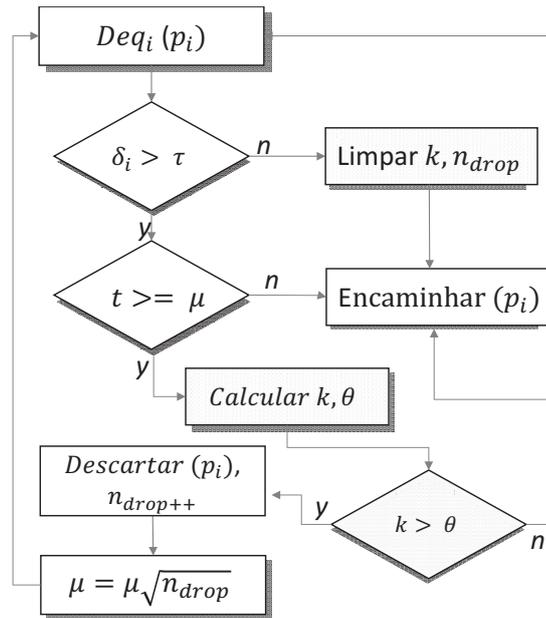


Figura 7.7: Diagrama de Fluxo do LIFO-CoDel

A fim de mostrar a diferença entre FIFO e LIFO, a Figura 7.6 ilustra em um plano cartesiano de tamanho da fila vs. tempo, a entrada, permanência e saída de pacotes das estruturas fila e pilha. Na disciplina FIFO-CoDel, conforme a fila cresce, o atraso do pacote aumenta. O parâmetro τ controla o atraso do pacote e o μ o tamanho da fila. Quando o pacote p_i é desenfileirado, ele será descartado se δ_i for maior que τ e o instante $t = deq_i$ for maior que μ . Este último se ajusta dinamicamente conforme o número de descartes para controlar o tamanho da fila. Isto permite que pacotes com um tempo de permanência δ_i maiores que τ sejam encaminhados até que se entre novamente em estado de descarte. O LIFO-CoDel propõe controlar este estado de descarte com duas principais alterações: alterar a fila para pilha, a fim de priorizar os pacotes com menor tempo de permanência; e um novo parâmetro para controlar os descartes na fila.

A alteração da estrutura fila para pilha permite priorizar a retirada dos pacotes com menor tempo de permanência, ou seja, os pacotes mais recentes. Considerando que o mecanismo esteja em estado de descarte, na Figura 7.6.(b) o pacote p_i seria descartado porque seu tempo de permanência é maior que τ e $deq_i > \mu$. Contudo, LIFO-CoDel utiliza o parâmetro θ para determinar se o pacote será descartado. Conforme ilustra o diagrama de fluxo da Figura 7.7, o parâmetro θ representa a razão entre o tempo de permanência máximo δ_{max} (Equação 7.6) e o tempo de permanência médio $\bar{\delta}$ (Equação 7.7) para o n -ésimo pacote retirado da pilha, dado pela Equação 7.8. Se $\bar{\delta} = 0$, então se assume $\theta = 0$. Juntamente com o θ , é calculado Γ_t , que se refere a diferença entre o tempo de permanência atual e o anterior (Equação 7.9). O valor de Γ_t é utilizado para incrementar ou reiniciar o valor da variável k (Equação 7.10). Quando $\Gamma_t > 0$, assumindo a estrutura pilha, refere-se a um pacote já na pilha antes da última retirada, porém pode indicar também um atraso maior do sistema em relação à última retirada. No entanto, o valor de k e θ é reinicializado quando $\Gamma_t < 0$, ou seja, quando o pacote atual possuir um tempo de permanência menor que o anterior. Por fim, o pacote p_i somente é descartado se $k > \theta$.

$$\delta_{max} = \max(\delta_1, \dots, \delta_n) \quad (7.6)$$

$$\bar{\delta} = \frac{1}{n} \sum_{i=1}^n \delta_i \quad (7.7)$$

$$\theta = \frac{\delta_{max}}{\bar{\delta}}, \bar{\delta} \neq 0 \quad (7.8)$$

$$\Gamma_t = \delta_i - \delta_{i-1} \quad (7.9)$$

$$k = \begin{cases} k + 1, & \text{if } \Gamma_t > 0 \\ 0, & \text{if } \Gamma_t \leq 0 \end{cases} \quad (7.10)$$

7.5 Resultados

Nesta seção apresentamos os resultados da avaliação da LIFO-CoDel em comparação com os resultados do CoDel e FQ-CoDel. O cenário, as métricas e os parâmetros utilizados nas simulações são os mesmos apresentados no início do capítulo. Os resultados mostram que o LIFO-CoDel causa um impacto menor no CC do MPTCP, e com isto, melhora o desempenho da transmissão multicaminhos. Um dos motivos desta melhoria é a redução no número de descartes na fila, ilustrado pela Figura 7.8.(a). Esta redução contribui com o aumento do *goodput*, conforme Figura 7.8.(b). O LIFO-CoDel tem um *goodput* cerca de 40% maior que FIFO-CoDel, para atraso de 1ms no caminho A. O aumento do atraso ocasiona um descarte maior de pacotes, e isto reduz o *goodput* do LIFO-CoDel. Contudo, com atraso de 100ms os resultados são discretamente melhores que o FIFO-CoDel.

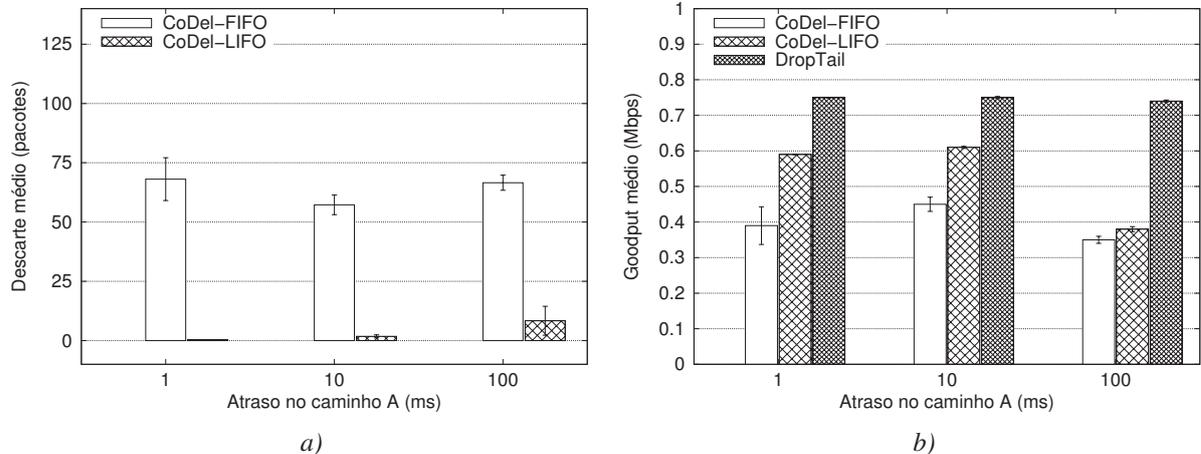


Figura 7.8: Descartes vs. Goodput

A Figura 7.9 apresenta os resultados para a variação do tamanho da CWND e RTT do LIFO-CoDel. Assim como CoDel, a janela de congestionamento se manteve em torno de 5 pacotes. Porém, com atraso de 100ms. O algoritmo CC não priorizou o caminho com menor RTT, como observado nos resultados diante do CoDel e DropTail. As Figuras 7.10.(a) e 7.10.(b) apresentam respectivamente o tamanho médio e o tempo médio de permanência na fila sob as disciplinas FIFO-CoDel e LIFO-CoDel. O LIFO-CoDel reduz o tamanho da fila e consequentemente o tempo de permanência. Estes resultados se referem às simulações com o algoritmo LIA. Com o algoritmo *RT_Compensator* os resultados variam de 2% a 5%. Com o algoritmo *Uncoupled*, a variação é mais expressiva para o LIFO-CoDel, com valores até 50% acima dos apresentados com LIA.

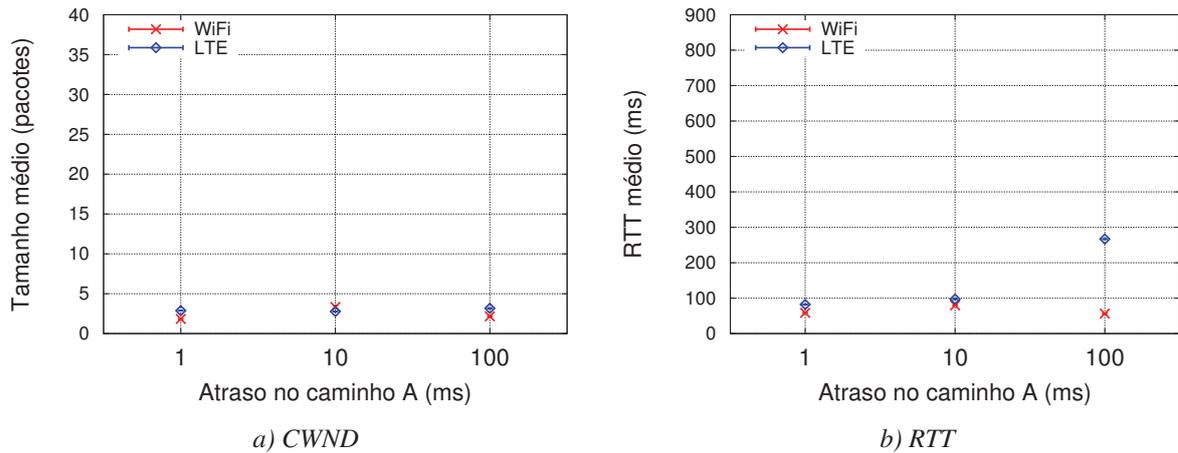


Figura 7.9: LIFO-CoDEL com LIA

A fim de compreender o impacto do FIFO e LIFO em cada caminho, realizamos simulações com o FQ-CoDel, que separa os subfluxos em filas independentes. Os resultados obtidos para ambas as disciplinas tiveram uma variação muito discreta, para o *goodput* ficou entre 2% a 3%. Estes resultados indicam que a separação em filas distintas não muda o fato que ambos os subfluxos compartilham o mesmo gargalo. Os resultados comparando as disciplinas DropTail, CoDel e FQ-CoDel em relação aos algoritmos *Uncoupled*, *LIA* e *RTT_Compensator* demonstram que existe um *trade-off* entre RTT e vazão. Diante de filas com grande capacidade, o DropTail não descarta pacotes e a fila cresce acarretando um alto valor de RTT (fenômeno *bufferbloat*). Porém a transmissão alcança a utilização máxima do canal. O CoDel resolve o problema dos atrasos, porém, os descartes impactam no CC do MPTCP e isto reduz a utilização do canal de transmissão. No cenário de redes heterogêneas sem fio, este aspecto tem impacto ainda maior. Ao aumentar o atraso em um dos caminhos se observou uma redução na vazão e um aumento no RTT de ambos os caminhos.

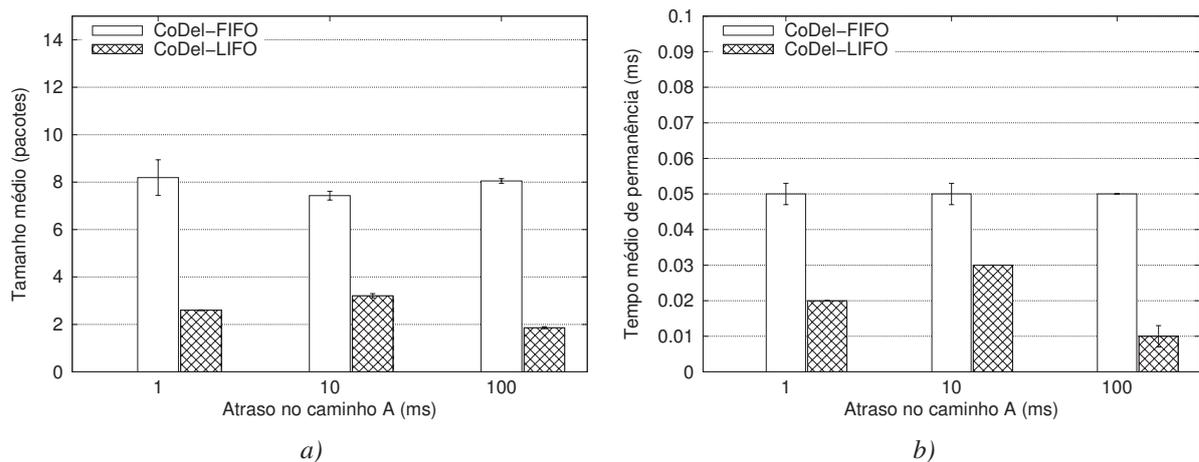


Figura 7.10: Estado da fila

7.6 Importância dos resultados para a tese

O uso de múltiplos caminhos é uma das principais características que tornam os protocolos multicaminhos mais resilientes que os protocolos tradicionais, com o TCP e o UDP. Isto porque uma falha pode ser facilmente superada movendo o tráfego para um outro caminho. No entanto, os protocolos que empregam o controle de congestionamento baseado em perdas identificam tardiamente o congestionamento causado pelo fenômeno do *bufferbloat*, devido as perdas só ocorrem quando a fila do gargalo enche. Os resultados apresentados demonstram que isto tem grande impacto no desempenho e na resiliência da transmissão quando os caminhos compartilham um gargalo que sofre com o *bufferbloat*. Neste caso, o reconhecimento dos caminhos compartilhados pode reduzir este impacto. Por outro lado, em um cenário em que o gargalo é inevitável, como um proxy MPTCP¹, outras medidas são necessárias para garantir a resiliência da transmissão. Uma opção é prover medidas de resistência junto ao protocolo de transporte, como empregar um controle de congestionamento baseado em atraso, como o wVegas [Cao et al., 2012]. No entanto, este capítulo propôs uma solução em nível de rede a fim de minimizar o problema. Isto estende as possibilidades de aplicação da estratégia de resiliência além dos protocolos de transporte multicaminhos. Ela pode, por exemplo, ser aplicada em um proxy de camada de rede, similar ao proposto pela Tessares². Neste caso, como medida de resistência, a estratégia de resiliência pode definir a disciplina de fila que causa um menor impacto sobre as transmissões multicaminhos em caso de congestionamento.

7.7 Resumo

O capítulo apresentou um estudo de caso sobre o impacto do fenômeno *bufferbloat* nas transmissões multicaminhos. Este estudo de caso é coerente com as medidas a serem empregadas pelo módulo de resistência do arcabouço proposto pela estratégia de resiliência. Os resultados da avaliação mostraram que em um cenário heterogêneo onde os caminhos compartilham um congestionamento, o desempenho do MPTCP pode ser limitado não apenas pelo impacto do congestionamento, mas também pela disciplina de fila. Para este último caso, foi proposta a disciplina LIFO-CoDel a fim de controlar os descartes na fila e minimizar o impacto no controle de congestionamento do MPTCP. Os resultados obtidos demonstram que isto minimiza o *trade-off* entre *goodput* e latência causado pelas disciplinas DropTail e CoDel.

¹Operadoras como a Tessares, oferecem um proxy MPTCP para interligar os usuários domésticos utilizando duas diferentes redes de acesso, como uma Linha Digital de Assinante (DSL, do Inglês *Digital Subscriber Line*) e LTE.

²Informações disponíveis em <https://www.tessares.net>.

8 Conclusões

Este capítulo apresenta um síntese das conclusões desta tese, as questões em aberto e as direções futuras. O objetivo é reforçar as contribuições apresentadas, relacionando-as com a tese, apresentar as questões em aberto e os motivos pelos quais não foram abordadas pela tese e por fim, destacar as publicações, os trabalhos que estão sendo desenvolvidos e as direções futuras. O Capítulo está organizado da seguinte forma. A Seção 8.1 apresenta as contribuições e resultados alcançados. A Seção 8.2 descreve as questões em aberto, os trabalhos oriundos da tese e os trabalhos futuros. Por fim, a Seção 8.3 detalha as publicações relacionadas direta e indiretamente com a tese.

8.1 Problemas tratados e resultados alcançados

As transmissões de dados multicaminhos no cenário das HetNets estão sujeitas a inúmeros desafios. Dentre estes desafios, a heterogeneidade e a falta de diversidade entre os caminhos tem grande impacto no desempenho e na resiliência das transmissões multicaminhos. A heterogeneidade causa o problema do reordenamento, que reduz a vazão e aumenta o atraso das transmissões, afetando principalmente aplicações de tempo real. Além disso, a falta de diversidade limita os benefícios alcançados com a agregação de banda dos caminhos e reduz a resiliência das transmissões, uma vez que os caminhos podem compartilhar gargalos e falhas. A revisão bibliográfica apresentada no Capítulo 3 destacou os principais desafios para as transmissões multicaminhos e as abordagens empregadas na literatura. As abordagens foram apresentadas conforme sua contribuição com métodos de reconhecimento, métodos de seleção e políticas de alocação. O objetivo desta organização foi evidenciar que são etapas complementares, porém, não dependentes. Neste caso, um ou mais métodos de reconhecimento podem ser utilizados por diferentes métodos de seleção ou políticas de alocação em uma transmissão multicaminhos. A revisão destacou que uma única solução não é capaz de atender a todos os desafios e requisitos para garantir o desempenho e a resiliência das transmissões multicaminhos. Isto porque os problemas são diversos, as soluções são específicas e muitas vezes conflitantes, como por exemplo: *priorizar os requisitos do fluxo de dados, agregando o tráfego sobre múltiplas interfaces sem fio, ou priorizar o consumo de energia, limitando o uso concorrente das interfaces de rede.*

A partir da revisão bibliográfica, definiu-se a tese de que *a diversidade, em termos de soluções, é capaz de prover um melhor desempenho e uma maior resiliência para as transmissão de dados multicaminhos.* A fim de comprovar a tese, foi proposto no Capítulo 4 uma estratégia de resiliência com objetivo prover uma diversidade de soluções para os diferentes problemas enfrentados pela transmissões multicaminhos. O arcabouço de resiliência é um dos principais componentes da estratégia e possui diferentes módulos, que de modo geral, coletam os dados da transmissão, como requisitos e estado dos caminhos, a fim de prover diferentes medidas de resistência e recuperação para garantir o desempenho e a resiliência das transmissões.

Esta trabalho apresentou diferentes estudos de caso a fim de validar a estratégia proposta e reiterar a tese de que a diversidade é capaz de prover um melhor desempenho e uma maior resiliência. Os estudos de caso foram apresentados nos Capítulos 5, 6 e 7. O estudo de caso desenvolvido no Capítulo 5, sobre a diversidade de caminhos, demonstrou que diferentes métodos podem ser aplicados para um determinado problema, no entanto, podem gerar conflitos com os requisitos dos usuários ou dispositivos. Isto contribui para reforçar a tese de que a diversidade de soluções, provida pela arquitetura de gerenciamento de resiliência, é necessária para manter o desempenho e a resiliência das transmissões. O estudo de caso desenvolvido no Capítulo 6 estendeu a análise do estudo de caso sobre a diversidade, provendo redundância com diversidade de caminhos. Os resultados alcançados demonstraram que prover diversidade para a transmissão de dados multicaminhos resulta em um melhor desempenho e uma maior resiliência. Este estudo de caso demonstrou como a interação entre os módulos de reconhecimento e resistência, mediados pelo módulo de controle, poderiam ser aplicados, reforçando a tese defendida. Por fim, o estudo de caso desenvolvido no Capítulo 7 apresentou uma solução em nível de rede, estendendo as possibilidades de aplicação da estratégia para além da camada de transporte (dispositivo móvel). Um exemplo citado é a aplicação em um Proxy MPTCP, que provê uma conexão multicaminhos entre o dispositivo móvel e a rede da operadora. Para este exemplo, a diversidade de soluções providas pela estratégia se estende para a camada de rede.

Além desses estudos de caso, foram realizados dois trabalhos colaborativos associados ao módulo de resistência e que trataram questões de segurança contra dois tipos de ataque, o ataque *Low-rate DoS* e o ataque do homem do meio (do Inglês, *Man-in-the-Middle*). Estes estudos não estão descritos nesta tese, no entanto, estão disponíveis para consulta em [Felix et al., 2018b] e [Nguyen et al., 2017], respectivamente.

Os estudos de caso buscaram demonstrar que a tese definida é válida. Embora a diversidade de soluções foi tratada de forma isolada nos estudos de caso, a proposta é que seja aplicada de forma adaptativa. Para isto, as soluções propostas nos estudos de caso foram associados com medidas empregadas pelos módulos do arcabouço de resiliência. Assim, como resultado final, as medidas aplicadas melhoraram o desempenho e a resiliência das transmissões multicaminhos, validando a tese definida.

8.2 Questões em Aberto e Direções Futuras

Neste trabalho não realizou-se a implementação de nenhum dos componentes da estratégia de resiliência, como o arcabouço de resiliência, o gerenciador de contexto e a base de informações. Além da falta dessa implementação, foram identificadas outras questões em aberto que direcionam para trabalhos futuros. Essas questões estão divididas em: (i) estudos de caso; (ii) realização da arquitetura; (iii) e protocolos e arquiteturas emergentes. Em relação ao primeiro, a realização de novos estudos de caso envolvendo os desafios que ainda não foram cobertos, como de restrição de energia e custo, segurança e equidade, e os módulos da arquitetura, como o de reconhecimento, resistência, recuperação, e outros. O estudo de caso realizado no Capítulo 6 não avaliou a complexidade e o custo computacional do método empregado para calcular a correlação e selecionar os caminhos disjuntos. No cenário de redes móveis em que o número de caminhos está associado ao número de interfaces de rede sem fio isto pode não ser um problema. No entanto, isto requer uma avaliação mais precisa, especialmente para determinar o tempo de resposta e o custo de processamento. No mais, novas métricas, métodos de reconhecimento, medidas de resistência e recuperação podem ser propostos e comparados com os métodos existentes.

As categorias apresentadas no Capítulo 5, sendo a intracamada e intercamadas, indicam diferentes direções de pesquisa. A principal desvantagem das abordagens da classe inter-camada

se refere à dependência entre camadas, a baixa escalabilidade e a possível interferência dos *middle-boxes*, como firewalls, IDS, etc. Estas abordagens possuem muitos pontos a se explorar, especialmente para fornecer diversidade de caminhos em redes privadas, inter-domínios e em cenários restritos, como a Internet das Coisas (IoT, do Inglês *Internet of Things*). Por outro lado, as abordagens da classe intracamada são mais compatíveis com o atual modelo de camadas. Neste caso, a concepção de um método ativo, que tenha baixa sobrecarga, ou um método passivo que apresente uma boa acurácia, é um campo ainda aberto para contribuições. Por fim, para ambos os casos, é importante considerar as dinâmicas do cenário HetNet, como as transições comuns em um cenário de redes densas e que podem causar desconexões e flutuações no desempenho e resiliência das transmissões. Sobre a IoT, dada a restrição de recursos de alguns dispositivos, uma medida de resistência que faz uso da diversidade de caminhos podem garantir uma maior privacidade dos dados sem a necessidade de consumir recursos dos dispositivos, como as técnicas de criptografia.

A implementação da estratégia de resiliência é um dos próximos passos a serem realizados. Par isto, estão sendo consideradas questões práticas como, definição do funcionamento dos módulos e da troca de mensagens entre eles, e de uma linguagem de interação com os protocolos de transporte, aplicações e outros. A partir destas definições, são requeridas as definições formais de políticas de resiliência que empregam diferentes técnicas, como por exemplo, cognitivas [Almasi, 2013] e bio-inspiradas [Gong et al., 2017]. Essas técnicas envolvem desde a aplicação das medidas de resistência e recuperação, até as definições acerca do ciclo evolutivo e adaptativo da estratégia.

Neste trabalho foi considerado apenas o protocolo MPTCP para realização dos estudos de caso. Contudo, novos protocolos de transmissão multicaminhos tem emergido, como o QUIC Multicaminhos (MPQUIC) [De Coninck e Bonaventure, 2017]. O QUIC foi proposto pelo Google como um protocolo de aplicação e que tem sido largamente empregados pelo navegador Chrome, um dos mais populares atualmente¹. O MPQUIC adiciona a capacidade de transmissão multicaminhos ao protocolo QUIC. O fato de ser um protocolo de camada de aplicação pode facilitar a implementação e avaliação da estratégia de resiliência.

8.3 Publicações

Esta seção apresenta uma lista das publicações realizadas durante o desenvolvimento da tese e que estão diretamente e indiretamente relacionadas com o tema estudado. Além disso, são apresentadas as publicações em revisão, que foram submetidas e estão aguardando um retorno, e as publicações que estão sendo preparadas para submissão. Por fim, destaca-se os prêmios recebidos durante o curso de doutoramento.

Diretamente Relacionadas

1. FELIX, B.; SANTOS, A.; NOGUEIRA, M. **Reduzindo os Efeitos do Bufferbloat sobre multicaminhos em Redes Sem Fio Heterogêneas**. SBRC, Salvador, 2016.
2. FELIX, B.; SANTOS, A.; KANTARCI, B.; NOGUEIRA, M. **CD-ASM: A New Queuing Paradigm to Overcome Bufferbloat Effects in HetNets**. PIMRC, Montreal, 2017.

¹Segundo o tráfego contabilizado pela W3C, o Chrome corresponde a mais de 58% do tráfego de todos os sites que usam o contador W3Counter. <https://www.w3counter.com/globalstats.php>

3. NGUYEN, H. D. D.; PHUNG, C. D.; SECCI, S.; FELIX, B. and NOGUEIRA, M. **Can MPTCP secure Internet communications from man-in-the-middle attacks?**. CNSM, Tokyo, 2017.
4. FELIX, B.; STEUCK, I.; SANTOS, A.; NOGUEIRA, M. **Escalonador de Pacotes para Transmissão por multicaminhos Não-Correlacionados em Redes Heterogêneas sem Fio**. SBRC, Campos do Jordão, 2018.
5. FELIX, B.; STEUCK, I.; SANTOS, A.; NOGUEIRA, M. **Redundant Packet Scheduling by Uncorrelated Paths in Heterogeneous Wireless Networks**. ISCC, Natal, 2018.
6. POSSATI, D.; FELIX, B.; SANTOS, A.; NOGUEIRA, M. **Mitigação de Ataque Low-Rate DoS no Protocolo MPTCP**. WGRS, Campos do Jordão, 2018.
7. FELIX, B.; NOGUEIRA, M. **Packet Replication on Disjoint MPTCP Paths**. IEEE ICC/N2Women and WICE Workshop, Kansas City, 2018.

Indiretamente Relacionadas

1. FELIX, B.; SANTOS, A.; NOGUEIRA, M. **Redes Sem fio Definidas por Software sobre Mini-PC: Uma Avaliação de Desempenho**. WGRS, 2015.
2. Wangham, M. S.; Nogueira, M.; Fernandes, C. P.; Paviani, O.; Felix, B. **Segurança em Redes Veiculares: Inovações e Direções Futuras**. SBSEG, 2014.

Em revisão

1. POSSATI, D.; FELIX, B.; SANTOS, A.; NOGUEIRA, M. **Exploiting and Mitigating Low-Rate DoS Attacks on MPTCP**. GLOBECOM, 2018.
2. NGUYEN, H. D. D.; PHUNG, C. D.; SECCI, S.; FELIX, B. and NOGUEIRA, M. **MPTCP Robustness Against Large-Scale Man-in-the-Middle Attacks**. Computer Networks, 2018.
3. FELIX, B. and NOGUEIRA, M. **The Power of Diversity for Resilience on IoT End-to-End Communication in Smart Cities**. IEEE Network Magazine, 2018.

Em preparação

1. Periódico. **Multipath Transport Resiliency in HetNet**. IEEE Transaction on Network and Service Management.

Prêmios Recebidos

1. Prêmio de melhor artigo na conferência SBRC, 2016.
2. Mérito UFPR em função do Prêmio de melhor artigo no SBRC2016, Universidade Federal do Paraná (UFPR), 2016.
3. Certificado de Apreciação - pela contribuição ao grupo de pesquisa CCSC, Centro de Ciência de Segurança Computacional (CCSC) - UFPR, 2017.

REFERÊNCIAS

- [app, 2018] (2018). Improving Network Reliability Using Multipath TCP. <https://goo.gl/vu7EEv>. Acessado: 2018-10-04.
- [mpt, 2018] (2018). Multipath TCP in the Linux Kernel. <https://www.multipath-tcp.org>. Acessado: 2018-13-04.
- [A. Hunger and P. A. Klein and M. H. Verbunt, 2016] A. Hunger and P. A. Klein and M. H. Verbunt (2016). Evaluation of the Redundancy-Bandwidth Trade-Off and Jitter Compensation in rMPTCP. Em *International Conference on New Technologies, Mobility and Security (NTMS)*, páginas 1–5. IEEE.
- [Afzal e Lindskog, 2015] Afzal, Z. e Lindskog, S. (2015). *Multipath TCP IDS Evasion and Mitigation*, páginas 265–282. Springer International Publishing, Cham.
- [Alfredsson et al., 2013] Alfredsson, S., Del Giudice, G., Garcia, J., Brunstrom, A., De Cicco, L. e Mascolo, S. (2013). Impact of TCP Congestion Control on Bufferbloat in Cellular Networks. Em *World of Wireless Mobile and Multimedia Networks WoWMoM*, páginas 1–7.
- [Alheid et al., 2016] Alheid, A., Doufexi, A. e Kaleshi, D. (2016). A Study on MPTCP for Tolerating Packet Reordering and Path Heterogeneity in Wireless Networks. Em *Wireless Days (WD)*, páginas 1–7. IEEE.
- [Ali et al., 2015] Ali, H., Khan, S. e Quaid, M. (2015). Comparative Analysis of Controlled Delay (CoDel) with Deficit Round Robin (DRR) to Overcome Bufferbloat Problem in Wired Network. *International Journal of Current Engineering and Technology*, 5(5):3378–3386.
- [Almasi, 2013] Almasi, B. (2013). Multipath communication—A new basis for the Future Internet Cognitive Infocommunication. Em *Cognitive Infocommunications (CogInfoCom)*, páginas 201–204. IEEE.
- [Antonenko e Smelyanskiy, 2013] Antonenko, V. e Smelyanskiy, R. (2013). Global network modelling based on mininet approach. Em *ACM SIGCOMM*, páginas 145–146.
- [Arzani et al., 2014] Arzani, B., Gurney, A., Cheng, S., Guerin, R. e Loo, B. T. (2014). Deconstructing MPTCP Performance. Em *International Conference on Network Protocols (ICNP)*, páginas 269–274. IEEE.
- [Aydin et al., 2012] Aydin, I., Iyengar, J., Conrad, P., Shen, C.-C. e Amer, P. (2012). Evaluating TCP-friendliness in light of Concurrent Multipath Transfer. *Computer Networks*, 56(7):1876–1892.
- [Bagnulo, 2011] Bagnulo, M. (2011). Threat Analysis for TCP Extensions for Multipath Operation with Multiple Addresses. RFC 6181, IETF.
- [Bagnulo et al., 2015] Bagnulo, M., Paasch, C., Gont, F., Bonaventure, O. e Raiciu, C. (2015). Analysis of Residual Threats and Possible Fixes for Multipath TCP (MPTCP). RFC 7430, IETF.

- [Baidya e Prakash, 2014] Baidya, S. H. e Prakash, R. (2014). Improving the Performance of Multipath TCP over Heterogeneous Paths Using Slow Path Adaptation. Em *International Conference on Communications (ICC)*, páginas 3222–3227. IEEE.
- [Baker e Fairhurst, 2015] Baker, F. e Fairhurst, G. (2015). IETF Recommendations Regarding Active Queue Management. RFC 7567, IETF.
- [Barré et al., 2011a] Barré, S., Bonaventure, O., Raiciu, C. e Handley, M. (2011a). Experimenting with Multipath TCP. *ACM SIGCOMM*, 41(4):443–444.
- [Barré et al., 2011b] Barré, S., Paasch, C., Bonaventure, O. et al. (2011b). MultiPath TCP-Guidelines for implementers. Relatório técnico, IETF.
- [Becke, 2014] Becke, M. (2014). *Revisiting the IETF Multipath Extensions on Transport Layer*. Tese de doutorado, University of Duisburg-Essen, Duisburg - Germany.
- [Becke et al., 2013] Becke, M., Adhari, H., Rathgeb, E. P., Fa, F., Yang, X. e Zhou, X. (2013). Comparison of Multipath TCP and CMT-SCTP based on Intercontinental Measurements. Em *Global Communications Conference (GLOBECOM)*, páginas 1360–1366. IEEE.
- [Becke et al., 2012] Becke, M., Dreibholz, T., Adhari, H. e Rathgeb, E. P. (2012). On the Fairness of Transport Protocols in a Multi-path Environment. Em *International Conference on Communications (ICC)*, páginas 2666–2672. IEEE.
- [Bhat e Talmale, 2014] Bhat, P. A. e Talmale, G. (2014). MPTCP Combining Congestion Window Adaptation and Packet Scheduling for Multi-homed Device. Em *International Conference for Convergence of Technology (I2CT)*, páginas 1–6. IEEE.
- [Bonaventure et al., 2015] Bonaventure, O., De Coninck, Q., Baerts, M., Duchene, F. e Hesmans, B. (2015). Improving Multipath TCP Backup Subflows. Relatório técnico, IETF.
- [Bonaventure et al., 2012] Bonaventure, O., Handley, M., Raiciu, C. et al. (2012). An Overview of Multipath TCP. *login: The Usenix Magazine*, 37(5):17–23.
- [Brualdi Timmins, 1996] Brualdi Timmins, A. C. (1996). Multiple intelligences: Gardner’s theory. *Practical Assessment, Research & Evaluation*, 5(10):3.
- [Cao et al., 2012] Cao, Y., Xu, M. e Fu, X. (2012). Delay-based congestion control for multipath TCP. Em *Network Protocols (ICNP)*, páginas 1–10. IEEE.
- [Cetinkaya e Sterbenz, 2013] Cetinkaya, E. K. e Sterbenz, J. P. (2013). A Taxonomy of Network Challenges. Em *Design of Reliable Communication Networks (DRCN)*, páginas 322–330. IEEE.
- [Chan et al., 2016] Chan, M.-C., Tseng, C.-C. e Yen, L.-H. (2016). Jitter-Aware Packet Scheduler for Concurrent Multipath Transmission in Heterogeneous Wireless Networks. Em *Wireless Communications and Networking Conference (WCNC)*, páginas 1–7. IEEE.
- [Chebrolu e Rao, 2006] Chebrolu, K. e Rao, R. R. (2006). Bandwidth aggregation for real-time applications in heterogeneous wireless networks. *IEEE Transactions on Mobile Computing*, 5(4):388–403.

- [Chen et al., 2014] Chen, R., Wang, W.-N., Zhu, J.-L. e Wang, B. (2014). A Round-Trip-Time Based Concurrent Transmission Scheduling for MPTCP. Em *Wireless Communications and Signal Processing (WCSP)*, páginas 1–5. IEEE.
- [Chen et al., 2013] Chen, S., Yuan, Z. e Muntean, G. M. (2013). An Energy-Aware Multipath-TCP-Based Content Delivery Scheme in Heterogeneous Wireless Networks. Em *Wireless Communications and Networking Conference (WCNC)*, páginas 1291–1296.
- [Chen et al., 2011] Chen, Y., Wu, X. e Yang, X. (2011). MAPS: Adaptive Path Selection for Multipath Transport Protocols in the Internet. Acessado: 2016-11-07.
- [Chen e Towsley, 2014] Chen, Y.-C. e Towsley, D. (2014). On Bufferbloat and Delay Analysis of Multipath TCP in Wireless Networks. Em *IFIP Networking Conference and Workshops*, páginas 1–9.
- [Cisco, 2017a] Cisco (2017a). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021. *Online*.
- [Cisco, 2017b] Cisco (2017b). The Zettabyte Era: Trends and Analysis. *Online*.
- [Comer, 2008] Comer, D. E. (2008). *Computer Networks and Internets*. Prentice Hall Press, Upper Saddle River, NJ, USA, 5th edition.
- [Computing et al., 2006a] Computing, A. et al. (2006a). An Architectural Blueprint for Autonomic Computing. *IBM White Paper*, 31.
- [Computing et al., 2006b] Computing, A. et al. (2006b). An Architectural Blueprint for Autonomic Computing. *IBM White Paper*, 31.
- [Constantine et al., 2011] Constantine, B., Forget, G., Geib, R. e Schrage, R. (2011). Framework for TCP Throughput Testing. RFC 6349, IETF.
- [Corbillon et al., 2016] Corbillon, X., Aparicio-Pardo, R., Kuhn, N., Texier, G. e Simon, G. (2016). Cross-layer Scheduler for Video Streaming over MPTCP. Em *International Conference on Multimedia Systems*, página 7. ACM.
- [Coudron et al., 2015] Coudron, M., Secci, S. e Pujolle, G. (2015). Differentiated Pacing on Multiple Paths to Improve One-way Delay Estimations. Em *Integrated Network Management (IM)*, páginas 672–678. IEEE.
- [De Coninck e Bonaventure, 2017] De Coninck, Q. e Bonaventure, O. (2017). Multipath QUIC: Design and Evaluation. Em *International Conference on Emerging Networking EXperiments and Technologies (CoNEXT)*, páginas 160–166. ACM.
- [Deng et al., 2014] Deng, S., Netravali, R., Sivaraman, A. e Balakrishnan, H. (2014). WiFi, LTE, or Both?: Measuring Multi-homed Wireless Internet Performance. Em *Internet Measurement Conference (IMC)*, páginas 181–194. ACM.
- [Diop et al., 2011] Diop, C., Dugué, G., Chassot, C. e Exposito, E. (2011). Qos-Aware Multipath-TCP Extensions For Mobile And Multimedia Applications. Em *International Conference on Advances in Mobile Computing and Multimedia*, páginas 139–146. ACM.

- [Diop et al., 2012] Diop, C., Gómez-Montalvo, J., Dugué, G., Chassot, C. e Exposito, E. (2012). Towards a Semantic and MPTCP-based Autonomic Transport Protocol for Mobile and Multimedia Applications. Em *Multimedia Computing and Systems (ICMCS)*, páginas 496–501. IEEE.
- [Domżał et al., 2015] Domżał, J., Duliński, Z., Kantor, M., Rząsa, J., Stankiewicz, R., Wajda, K. e Wójcik, R. (2015). A Survey on Methods to Provide Multipath Transmission in Wired Packet Networks. *Computer Networks*, 77:18–41.
- [Dugue et al., 2012] Dugue, G., Diop, C., Chassot, C. e Exposito, E. (2012). Towards Autonomic Multipath Transport for Infotainment-Like Systems. Em *International Conference on Communications (ICC)*, páginas 6453–6457. IEEE.
- [Exposito, 2013] Exposito, E. (2013). Transport Protocols State of the Art. *Advanced Transport Protocols*, páginas 7–24.
- [Felix e Nogueira, 2018] Felix, B. e Nogueira, M. (2018). Packet Replication on Disjoint MPTCP Paths. Em *IEEE ICC/N2Women and WICE: Professional Development Workshop*.
- [Felix et al., 2017] Felix, B., Santos, A., Kantarci, B. e Nogueira, M. (2017). CD-ASM: A new queuing paradigm to overcome bufferbloat effects in HetNets. Em *Personal, Indoor, and Mobile Radio Communications (PIMRC)*, páginas 1–6.
- [Felix et al., 2016] Felix, B., Santos, A. e Nogueira, M. (2016). Reduzindo os Efeitos do Bufferbloat sobre Multi-Caminhos em Redes Sem Fio Heterogêneas. Em *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, páginas 718–731. SBC.
- [Felix et al., 2018a] Felix, B., Steuck, I., Santos, A. e Nogueira, M. (2018a). Escalonador de Pacotes para Transmissão por Multi-caminhos Não-Correlacionados em Redes Heterogêneas sem Fio. Em *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, páginas 718–731. SBC.
- [Felix et al., 2018b] Felix, B., Steuck, I., Santos, A. e Nogueira, M. (2018b). Mitigação de Ataque Low-Rate DoS no Protocolo MPTCP. Em *Workshop de Gerência e Operação de Redes e Serviços (WGRS)*. SBC.
- [Felix et al., 2018c] Felix, B., Steuck, I., Secci, S., Santos, A. e Nogueira, M. (2018c). Redundant Packet Scheduling by Uncorrelated Paths in Heterogeneous Wireless Networks. Em *IEEE Symposium on Computers and Communications (ISCC)*.
- [Ferlin et al., 2016] Ferlin, S., Alay, O., Mehani, O. e Boreli, R. (2016). BLEST: Blocking Estimation-Based MPTCP Scheduler for Heterogeneous Networks. Em *IFIP Networking Conference and Workshops*, páginas 431–439. IEEE.
- [Ferlin et al., 2014] Ferlin, S., Dreiholz, T. e Alay, O. (2014). Multi-path Transport Over Heterogeneous Wireless Networks: Does It Really Pay Off? Em *Global Communications Conference (GLOBECOM)*, páginas 4807–4813. IEE.
- [Ferlin-Oliveira et al., 2014] Ferlin-Oliveira, S., Dreiholz, T. e Alay, O. (2014). Tackling the Challenge of Bufferbloat in Multi-Path Transport over Heterogeneous Wireless Networks. Em *International Symposium of Quality of Service (IWQoS)*, páginas 123–128. IEEE.

- [Ford et al., 2011] Ford, A., Raiciu, C., Handley, M., Barre, S. e Iyengar, J. (2011). Architectural Guidelines for Multipath TCP Development. RFC 6182, IETF.
- [Ford et al., 2013] Ford, A., Raiciu, C., Handley, M. e Bonaventure, O. (2013). TCP Extensions for Multipath Operation with Multiple Addresses. RFC 6824, IETF.
- [Frömmgen et al., 2016] Frömmgen, A., Erbschäuber, T., Zimmermann, T., Wehle, K. e Buchmann, A. (2016). ReMP TCP: Low latency multipath TCP. Em *ICC*.
- [Fry et al., 2010] Fry, M., Fischer, M., Karaliopoulos, M., Smith, P. e Hutchison, D. (2010). Challenge Identification for Network Resilience. Em *Next Generation Internet (NGI)*, páginas 1–8. IEEE.
- [Garcia-Saavedra et al., 2015] Garcia-Saavedra, A., Karzand, M. e Leith, D. J. (2015). Low Delay Random Linear Coding and Scheduling Over Multiple Interfaces. *arXiv*.
- [Ge et al., 2016] Ge, X., Tu, S., Mao, G., Wang, C. X. e Han, T. (2016). 5G Ultra-Dense Cellular Networks. *IEEE Wireless Communications*, 23(1):72–79.
- [Gong et al., 2017] Gong, W., Yang, X., Zhang, M. e Long, K. (2017). An adaptive traffic distribution scheme for CMT based on Lotka-Volterra model in multihomed networks. *China Communications*, 14(2):70–89.
- [Hayes et al., 2014] Hayes, D. A., Ferlin, S. e Welzl, M. (2014). Practical Passive Shared Bottleneck Detection using Shape Summary Statistics. Em *IEEE LCN*, páginas 150–158. IEEE.
- [He e Rexford, 2008] He, J. e Rexford, J. (2008). Toward Internet-Wide Multipath Routing. *IEEE Network*, 22(2).
- [Hoeiland-Joergensen et al., 2015] Hoeiland-Joergensen, T., McKenney, P., Taht, D., Gettys, J. e Dumazet, E. (2015). FlowQueue-CoDel. Internet-draft, IETF.
- [Hu et al., 2016] Hu, B., Xing, L., Wang, Z. e Liu, N. (2016). MLCS: A Multi-level Correlation Scheduling Algorithm for Multipath Transport. Em *International Conference on Information Networking (ICOIN)*, páginas 166–171. IEEE.
- [Hunger e Klein, 2016] Hunger, A. e Klein, P. (2016). Equalizing Latency Peaks Using a Redundant Multipath-TCP scheme. Em *International Conference on Information Networking (ICOIN)*, páginas 184–189. IEEE.
- [Hwang e Yoo, 2015] Hwang, J. e Yoo, J. (2015). Packet Scheduling for Multipath TCP. Em *International Conference on Ubiquitous and Future Networks*, páginas 177–179. IEEE.
- [Iyengar et al., 2006] Iyengar, J. R., Amer, P. D. e Stewart, R. (2006). Concurrent Multipath Transfer using SCTP Multihoming over Independent End-to-End Paths. *IEEE/ACM Transactions on Networking*, 14(5):951–964.
- [Jabbar, 2010] Jabbar, A. (2010). *A Framework to Quantify Network Resilience and Survivability*. Tese de doutorado, University of Kansas, Lawrence, KS, USA. AAI3417968.

- [Jin et al., 2015] Jin, X., Liu, Y., Fan, W., Wu, F. e Zhang, H. (2015). A Throughput Improved Path Selection Method Based on Throughput Prediction Model and Available Bandwidth for MPTCP. *International Journal of Future Generation Communication and Networking*, 8(2):105–114.
- [K. Nichols, 2016] K. Nichols, V. J. (2016). Controlled Delay Active Queue management. Rfc draft, IETF.
- [Kaup et al., 2015] Kaup, F., Wichtlhuber, M., Rado, S. e Hausheer, D. (2015). Can Multipath TCP save energy? A Measuring and Modeling study of MPTCP Energy Consumption. Em *Local Computer Networks (LCN)*, páginas 442–445.
- [Ke et al., 2016] Ke, F., Huang, M., Liu, Z., Liu, Q. e Cao, Y. (2016). Multi-Attribute Aware Multipath Data Scheduling Strategy for Efficient MPTCP-based Data Delivery. Em *Asia-Pacific Conference on Communications (APCC)*, páginas 248–253. IEEE.
- [Khalili et al., 2012] Khalili, R., Gast, N., Popovic, M., Upadhyay, U. e Le Boudec, J.-Y. (2012). MPTCP is Not Pareto-Optimal: Performance Issues and a Possible Solution. Em *International Conference on Emerging Networking Experiments and Technologies*, páginas 1–12. ACM.
- [Kheirkhah, 2015] Kheirkhah, M. (2015). MultiPath TCP in NS-3. Código fonte do projeto disponível em <http://dx.doi.org/10.5281/zenodo.32691>.
- [Kim et al., 2012] Kim, H. A., Oh, B.-h. e Lee, J. (2012). Improvement of MPTCP Performance in Heterogeneous Network Using Packet Scheduling Mechanism. Em *Asia-Pacific Conference on Communications (APCC)*, páginas 842–847. IEEE.
- [Kou et al., 2013] Kou, L., Chen, S. R. e Wang, R. (2013). A MPTCP Path Selection Strategy Based on Improved Grey Relational Analysis. Em *Applied Mechanics and Materials*, volume 401, páginas 1766–1771. Trans Tech Publ.
- [Krupakaran et al., 2015] Krupakaran, K., Sridharan, A. P. e Venkatesan, S. M. (2015). Optimized Multipath TCP subflows using Traceflow. Relatório técnico, IETF.
- [Kuhn et al., 2014] Kuhn, N., Lochin, E., Mifdaoui, A., Sarwar, G., Mehani, O. e Boreli, R. (2014). DAPS: Intelligent Delay-Aware Packet Scheduling for Multipath Transport. Em *International Conference on Communications (ICC)*, páginas 1222–1227. IEEE.
- [Kulatunga et al., 2015] Kulatunga, C., Kuhn, N., Fairhurst, G. e Ros, D. (2015). Tackling Bufferbloat in Capacity-Limited Networks. Em *European Conference on Networks and Communications (EuCNC)*, páginas 381–385.
- [Kurose e Ross, 2013] Kurose, J. F. e Ross, K. W. (2013). *Redes de Computadores e a Internet: Uma abordagem top-down*. Pearson, São Paulo, trad. 6 ed. edition.
- [Le e Bui, 2015] Le, T.-A. e Bui, L. X. (2015). Forward Delay-based Packet Scheduling Algorithm for Multipath TCP. *arXiv*.
- [Li et al., 2015] Li, L., Hu, N., Liu, K., Fu, B., Chen, M. e Zhang, L. (2015). AMPTCP: An Adaptive Multi-Path Transmission Control Protocol. Em *International Conference on Computing Frontiers*, página 29. ACM.

- [Li et al., 2016] Li, M., Lukyanenko, A., Ou, Z., Yla-Jaaski, A., Tarkoma, S., Coudron, M. e Secci, S. (2016). Multipath Transmission for the Internet: A Survey. *IEEE Communications Surveys & Tutorials*.
- [Liao et al., 2011] Liao, J., Wang, J., Li, T. e Zhu, X. (2011). Introducing Multipath Selection for Concurrent Multipath Transfer in the Future Internet. *Computer Networks*, 55(4):1024–1035.
- [Lim et al., 2014] Lim, Y.-s., Chen, Y.-C., Nahum, E. M., Towsley, D. e Lee, K.-W. (2014). Cross-layer Path Management In Multi-path Transport Protocol For Mobile Devices. Em *IEEE Conference on Computer Communications (INFOCOM)*, páginas 1815–1823. IEEE.
- [Lima et al., 2008a] Lima, M. N., da Silva, H. W., dos Santos, A. L. e Pujolle, G. (2008a). An Architecture for Survivable Mesh Networking. Em *Global Telecommunications Conference (GLOBECOM)*, páginas 1–5. IEEE.
- [Lima et al., 2008b] Lima, M. N., da Silva, H. W., dos Santos, A. L. e Pujolle, G. (2008b). An Architecture for Survivable Mesh Networking. Em *Global Telecommunications Conference (GLOBECOM)*, páginas 1–5. IEEE.
- [Liu et al., 2016a] Liu, J., Rayamajhi, A. e Martin, J. (2016a). Using MPTCP Subflow Association Control for Heterogeneous Wireless Network Optimization. Em *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, páginas 1–8. IEEE.
- [Liu et al., 2014] Liu, K., Huang, Q., Liu, Y., Fan, Y. e Xie, G. (2014). A Novel Trunk Scheduling for Concurrent Multipath Transfer in Heterogeneous Wireless Networks. Em *Wireless Communications, Networking and Mobile Computing (WiCOM)*, páginas 398–403. IET.
- [Liu et al., 2016b] Liu, S., Lei, W., Zhang, W. e Guan, Y. (2016b). CMT-SR: A Selective Retransmission Based Concurrent Multipath Transmission Mechanism for Conversational Video. *Computer Networks*.
- [Liu et al., 2011] Liu, Y., Wang, B., Xu, K. e Ma, Z. (2011). PACC: A Path Associativity Congestion Control and Throughput Model For Multi-path TCP. *Procedia Computer Science*, 4:1278 – 1287.
- [Lopez et al., 2015] Lopez, I., Aguado, M., Pinedo, C. e Jacob, E. (2015). Scada systems in the railway domain: enhancing reliability through redundant multipathtcp. Em *IEEE ITSC*, páginas 2305–2310.
- [Lopez et al., 2016] Lopez, I., Aguado, M., Ugarte, D., Mendiola, A. e Higuero, M. (2016). Exploiting Redundancy and Path Diversity for Railway Signalling Resiliency. Em *International Conference on Intelligent Rail Transportation (ICIRT)*, páginas 432–439. IEEE.
- [Mathis et al., 1996] Mathis, M., Mahdavi, J., Floyd, S. e Romanow, A. (1996). TCP Selective Acknowledgment Options. RFC 2018, IETF.
- [Maurer, 2015] Maurer, B. (2015). Fail at Scale. *ACM Queue*, 13(8):30:30–30:46.
- [Mauthe et al., 2016] Mauthe, A., Hutchison, D., Cetinkaya, E. K., Ganchev, I., Rak, J., Sterbenz, J. P., Gunkelk, M., Smith, P. e Gomes, T. (2016). Disaster-Resilient Communication Networks: Principles and Best practices. Em *Resilient Networks Design and Modeling (RNDM)*, páginas 1–10. IEEE.

- [Mirani et al., 2010] Mirani, F. H., Boukhatem, N. e Tran, M. A. (2010). A Data-Scheduling Mechanism for Multi-homed Mobile Terminals With Disparate Link Latencies. Em *Vehicular Technology Conference (VTC)*, páginas 1–5. IEEE.
- [Morawski e Ignaciuk, 2017a] Morawski, M. e Ignaciuk, P. (2017a). Energy Efficient Dynamic Load Balancing in Multipath TCP for Mobile Devices. Em *International Conference on Information Systems Architecture and Technology (ISAT)*, páginas 187–197. Springer.
- [Morawski e Ignaciuk, 2017b] Morawski, M. e Ignaciuk, P. (2017b). Energy efficient MPTCP transmission: Scheduler implementation and evaluation. Em *International Conference on System Theory, Control and Computing (ICSTCC)*, páginas 654–659.
- [Nakasan et al., 2017] Nakasan, C., Ichikawa, K., Iida, H. e Uthayopas, P. (2017). A simple multipath OpenFlow controller using topology-based algorithm for multipath TCP. *Concurrency and Computation: Practice and Experience*, 29(13).
- [Narasimhan et al., 2012] Narasimhan, J., Venkataswami, B. V., Groves, R. e Hoose, P. (2012). Traceflow. Relatório técnico, IETF.
- [Nguyen et al., 2017] Nguyen, H. D. D., Phung, C. D., Secci, S., Felix, B. e Nogueira, M. (2017). Can MPTCP secure Internet communications from man-in-the-middle attacks? Em *IEEE CNSM*.
- [Ni et al., 2014] Ni, D., Xue, K., Hong, P. e Shen, S. (2014). Fine-Grained Forward Prediction Based Dynamic Packet Scheduling Mechanism for Multipath TCP in Lossy Networks. Em *International Conference on Computer Communication and Networks (ICCCN)*, páginas 1–7. IEEE.
- [Ni et al., 2015] Ni, D., Xue, K., Hong, P., Zhang, H. e Lu, H. (2015). OCPS: Offset Compensation Based Packet Scheduling Mechanism for Multipath TCP. Em *International Conference on Communications (ICC)*, páginas 6187–6192. IEEE.
- [Nichols e Jacobson, 2012] Nichols, K. e Jacobson, V. (2012). Controlling Queue Delay. *ACM Queue*, 10(5):20:20–20:34.
- [Nikraves et al., 2016] Nikraves, A., Guo, Y., Qian, F., Mao, Z. M. e Sen, S. (2016). An In-Depth Understanding of Multipath TCP on Mobile Devices: Measurement and System Design. Em *International Conference on Mobile Computing and Networking*, páginas 189–201. ACM.
- [Oh e Lee, 2015] Oh, B.-H. e Lee, J. (2015). Constraint-based Proactive Scheduling for MPTCP in Wireless Networks. *Computer Networks*, 91:548–563.
- [Ou et al., 2016] Ou, S.-H., Huang, C.-W., Lee, T.-K. e Huang, C.-Y. (2016). Out-of-Order Transmission Enabled Congestion and Scheduling Control for Multipath TCP. Em *International Wireless Communications and Mobile Computing Conference (IWCMC)*, páginas 1069–1073. IEEE.
- [Paasch et al., 2013] Paasch, C., Barré, S. et al. (2013). Multipath TCP in the Linux kernel. Available from {www.multipath-tcp.org}.
- [Paasch e Bonaventure, 2014] Paasch, C. e Bonaventure, O. (2014). Multipath TCP. *Queue*, 12(2):40–51.

- [Paasch et al., 2012] Paasch, C., Detal, G., Duchene, F., Raiciu, C. e Bonaventure, O. (2012). Exploring Mobile/wifi Handover With Multipath TCP. Em *Workshop on Cellular Networks: Operations, Challenges, and Future Design*, páginas 31–36. ACM.
- [Paasch et al., 2014] Paasch, C., Ferlin, S., Alay, O. e Bonaventure, O. (2014). Experimental Evaluation Of Multipath TCP Schedulers. Em *SIGCOMM Workshop on Capacity Sharing Workshop*, páginas 27–32. ACM.
- [Park et al., 2016] Park, S., Kim, D., Yang, C. M. e Song, C.-J. (2016). A Packet Scheduling Scheme for Seamless Transmission of Life Media Contents. Em *International Conference on Ubiquitous and Future Networks (ICUFN)*, páginas 478–480. IEEE.
- [Park et al., 2014] Park, S.-Y., Joo, C., Park, Y. e Bank, S. (2014). Impact of Traffic Splitting on the Delay Performance of MPTCP. Em *International Conference on Communications (ICC)*, páginas 1204–1209. IEEE.
- [Partov e Leith, 2016] Partov, B. e Leith, D. J. (2016). Experimental Evaluation of Multi-path Schedulers for LTE/WiFi devices. Em *International Workshop on Wireless Network Testbeds, Experimental Evaluation, and Characterization (WiNTECH)*, páginas 41–48. ACM.
- [Paxson et al., 2011] Paxson, V., Allman, M., Chu, J. e Sargent, M. (2011). Computing TCP’s Retransmission Timer. RFC 6298, IETF.
- [Pearce e Thomas, 2014] Pearce, C. e Thomas, P. (2014). Multipath TCP—Breaking Today’s Networks with Tomorrow’s Protocol. <http://www.economist.com/node/2724397>.
- [Peng et al., 2014] Peng, Q., Chen, M., Walid, A. e Low, S. (2014). Energy Efficient Multipath TCP for Mobile Devices. Em *International Symposium on Mobile Ad Hoc Networking and Computing*, páginas 257–266. ACM.
- [Pluntke et al., 2011] Pluntke, C., Eggert, L. e Kiukkonen, N. (2011). Saving Mobile Device Energy With Multipath TCP. Em *International Workshop on MobiArch*, páginas 1–6. ACM.
- [Popat et al., 2015] Popat, K. J., Raval, J., Johnson, S. e Patel, B. (2015). An Efficient Scheduling Scheme of Multipath TCP for MPI. Em *International Symposium on Women in Computing and Informatics*, páginas 533–538. ACM.
- [Raiciu et al., 2011a] Raiciu, C., Barre, S., Pluntke, C., Greenhalgh, A., Wischik, D. e Handley, M. (2011a). Improving Datacenter Performance and Robustness With Multipath TCP. Em *SIGCOM Computer Communication Review*, páginas 266–277. ACM.
- [Raiciu et al., 2011b] Raiciu, C., Handly, M. e Wischik, D. (2011b). Coupled Congestion Control for Multipath Transport Protocols. RFC 6356, IETF.
- [Raiciu et al., 2010] Raiciu, C., Pluntke, C., Barre, S., Greenhalgh, A., Wischik, D. e Handley, M. (2010). Data Center Networking With Multipath TCP. Em *Workshop on Hot Topics in Networks*, página 10. ACM.
- [Rak, 2015] Rak, J. (2015). *Principles of Communication Networks Resilience*, páginas 11–43. Springer International Publishing, Cham.
- [Ramaboli et al., 2012] Ramaboli, A. L., Falowo, O. E. e Chan, A. H. (2012). Bandwidth aggregation in heterogeneous wireless networks A survey of current approaches and issues. *Journal of Network and Computer Applications*.

- [Raza et al., 2017] Raza, A., Gohar, A. e Lee, S. (2017). MPTCP based in-band controlling for the software defined networks. Em *International Conference on Information and Communication Technology Convergence (ICTC)*, páginas 163–167.
- [Rohrer et al., 2014] Rohrer, J. P., Jabbar, A. e Sterbenz, J. P. (2014). Path Diversification for Future Internet End-to-End Resilience and Survivability. *Telecommunication Systems*, 56(1):49–67.
- [Rubenstein et al., 2002] Rubenstein, D., Kurose, J. e Towsley, D. (2002). Detecting shared congestion of flows via end-to-end measurement. *IEEE/ACM Trans. on Networking (TON)*, 10(3):381–395.
- [Sandri et al., 2015] Sandri, M., Silva, A., Rocha, L. A. e Verdi, F. L. (2015). On the benefits of using multipath TCP and Openflow in Shared Bottlenecks. *Advanced Information Networking and Applications (AINA)*, páginas 9–16.
- [Sarwar et al., 2013] Sarwar, G., Boreli, R., Lochin, E., Mifdaoui, A. e Smith, G. (2013). Mitigating Receiver’s Buffer Blocking by Delay Aware Packet Scheduling in Multipath Data Transfer. Em *Workshop on Advanced Information Networking and Applications (WAINA)*, páginas 1119–1124. IEEE.
- [Schaeffer-Filho et al., 2012] Schaeffer-Filho, A., Smith, P., Mauthe, A., Hutchison, D., Yu, Y. e Fry, M. (2012). A Framework for the Design and Evaluation of Network Resilience Management. Em *Network Operations and Management Symposium*, páginas 401–408. IEEE.
- [Scharf e Ford, 2013] Scharf, M. e Ford, A. (2013). Multipath TCP (MPTCP) Application Interface Considerations. RFC 6897, IETF.
- [Schwarzkopf et al., 2016] Schwarzkopf, F., Veith, S. e Menth, M. (2016). Performance analysis of CoDel and PIE for saturated TCP sources. Em *ITC*, volume 1, páginas 175–183. IEEE.
- [Secci et al., 2014] Secci, S., Pujolle, G., Nguyen, T. M. T. e Nguyen, S. C. (2014). Performance-Cost Trade-Off Strategic Evaluation of Multipath TCP Communications. *IEEE Transactions on Network and Service Management*, 11(2):250–263.
- [Singh et al., 2012] Singh, A., Goerg, C., Timm-Giel, A., Scharf, M. e Banniza, T.-R. (2012). Performance Comparison of Scheduling Algorithms for Multipath Transfer. Em *Global Communications Conference (GLOBECOM)*, páginas 2653–2658. IEEE.
- [Singh e Reddy, 2013] Singh, A. e Reddy, A. L. N. (2013). Multi Path PERT. Em *International Conference on Computer Communication and Networks (ICCCN)*.
- [Singh et al., 2013] Singh, A., Xiang, M., Konsgen, A., Goerg, C. e Zaki, Y. (2013). Enhancing Fairness And Congestion Control In Multipath TCP. Em *Wireless and Mobile Networking Conference (WMNC)*, páginas 1–8. IEEE.
- [Smith et al., 2011] Smith, P., Hutchison, D., Sterbenz, J. P., Schöller, M., Fessi, A., Karaliopoulos, M., Lac, C. e Plattner, B. (2011). Network Resilience: A Systematic Approach. *Communications Magazine*, 49(7):88–97.
- [Song et al., 2017] Song, J., Dong, P., Zhou, H., Zheng, T., Du, X. e Guizani, M. (2017). A Performance Analysis Model for TCP over Multiple Heterogeneous Paths in 5G Networks. Em *IEEE Global Communications Conference (GLOBECOM)*, páginas 1–6.

- [Sterbenz et al., 2014] Sterbenz, J. P., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M. e Smith, P. (2014). Redundancy, Diversity, and Connectivity to Achieve Multilevel Network Resilience, Survivability, and Disruption Tolerance. *Telecommunication Systems*, 56(1):17–31.
- [Sterbenz et al., 2010] Sterbenz, J. P. G., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M. e Smith, P. (2010). Resilience and Survivability in Communication Networks: Strategies, Principles, and Survey of Disciplines. *Computer Networks*, 54(8):1245–1265.
- [Stewart, 2007] Stewart, R. (2007). Stream Control Transmission Protocol. RFC 4960, IETF.
- [Tsai et al., 2010] Tsai, M.-F., Chilamkurti, N., Park, J. H. e Shieh, C.-K. (2010). Multi-path Transmission Control Scheme Combining Bandwidth Aggregation and Packet Scheduling for Real-Time Streaming in Multi-Path Environment. *IET communications*, 4(8):937–945.
- [Tsai et al., 2016] Tsai, M.-H., Chou, C.-M. e Lan, K.-c. (2016). Avoiding Biased-Feeding in the Scheduling of Collaborative Multipath TCP. *PloS one*, 11(8):e0161213.
- [Wang et al., 2016] Wang, W., Zhou, L. e Sun, Y. (2016). Improving Multipath TCP for Latency Sensitive Flows in the Cloud. Em *International Conference on Cloud Networking (Cloudnet)*, páginas 45–50. IEEE.
- [Wischik et al., 2008] Wischik, D., Handley, M. e Braun, M. B. (2008). The Resource Pooling Principle. *SIGCOMM Computer Communication Review*, 38(5):47–52.
- [Wu et al., 2016] Wu, J., Cheng, B. e Wang, M. (2016). Energy Minimization for Quality-Constrained Video with Multipath TCP over Heterogeneous Wireless Networks. Em *International Conference on Distributed Computing Systems (ICDCS)*, páginas 487–496. IEEE.
- [Xu et al., 2016] Xu, C., Zhao, J. e Muntean, G.-M. (2016). Congestion Control Design for Multipath Transport Protocols: A Survey. *IEEE Communications Surveys & Tutorials*.
- [Yang e Amer, 2014] Yang, F. e Amer, P. (2014). Work in Progress: Using One-Way Communication Delay for In-Order Arrival MPTCP Scheduling. Em *Communications and Networking in China (CHINACOM)*, páginas 122–125. IEEE.
- [Yang et al., 2013a] Yang, F., Amer, P. D. e Ekiz, N. (2013a). A Scheduler for Multipath TCP. Em *International Conference on Computer Communications and Networks (ICCCN)*, páginas 1–7.
- [Yang et al., 2013b] Yang, F., Amer, P. D. e Ekiz, N. (2013b). A Scheduler for Multipath TCP. Em *International Conference Computer Communications and Networks (ICCCN)*, páginas 1–7.
- [Yang et al., 2014] Yang, F., Wang, Q. e Amer, P. D. (2014). Out-of-order Transmission for In-order Arrival Scheduling for Multipath TCP. Em *Workshop on Advanced Information Networking and Applications (WAINA)*, páginas 749–752. IEEE.
- [Yedugundla et al., 2016] Yedugundla, K., Ferlin, S., Dreibholz, T., Alay, Ö., Kuhn, N., Hurtig, P. e Brunstrom, A. (2016). Is Multi-path Transport Suitable for Latency Sensitive Traffic? *Computer Networks*, 105:1–21.
- [Zakon, 1997] Zakon, R. (1997). Hobbes’ Internet Timeline. RFC 2235, IETF.

- [Zhao et al., 2015] Zhao, J., Xu, C., Guan, J. e Zhang, H. (2015). AIMD-PQ: A Path Quality Based TCP-Friendly AIMD Algorithm for Multipath Congestion Control in Heterogeneous Wireless Networks. Em *Wireless Communications and Networking Conference (WCNC)*, páginas 1678–1683. IEEE.
- [Zhou et al., 2013] Zhou, D., Song, W. e Shi, M. (2013). Goodput Improvement for Multipath TCP by Congestion Window Adaptation in Multi-Radio Devices. Em *Consumer Communications and Networking Conference (CCNC)*, páginas 508–514. IEEE.
- [Zhou et al., 2015] Zhou, D., Song, W., Wang, P. e Zhuang, W. (2015). Multipath TCP for user cooperation in LTE networks. *IEEE Network*, 29(1):18–24.