



Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e Computação
DSIF - Departamento de Semicondutores,
Instrumentos e Fotônica

Dissertação de mestrado

**Arquitetura e Implementação de uma Rede em Anel de
Comutação Óptica de Pacotes**

Autor: Vinícius Garcia de Oliveira

Orientador: Prof. Dr. Peter Jürgen Tatsch

Co-Orientador: Dr. Marcos Rogério Salvador

Trabalho apresentado à Faculdade de Engenharia Elétrica e de Computação da UNICAMP como
parte dos requisitos exigidos para obtenção do título de Mestre em Engenharia Elétrica.

Comissão Examinadora:

Prof. Dr. Felipe Rudge Barbosa

Dr. Fabio Donati Simões

Este exemplar corresponde à redação final da Dissertação/Tese
defendida por VINÍCIUS GARCIA DE OLIVEIRA
e aprovada através da comissão julgada em 26 / 08 / 2009

Campinas

Peter Jürgen Tatsch
Orientador

Agosto 2009

UNIDADE BC
Nº CHAMADA
T/UNICAMP 019a
V
TOMBO BC/ 84609
PROC 16-134-10
C D X
PREÇO 11,00
DATA 14/01/10
CÓD TIT 471023

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

OL4a Oliveira, Garcia Vinícius
Arquitetura e implementação de uma rede em anel de comutação óptica de pacotes / Vinícius Garcia de Oliveira. --Campinas, SP: [s.n.], 2009.

Orientadores: Peter Jürgen Tatsch, Marcos Rogério Salvador.

Dissertação de Mestrado - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Telecomunicações. 2. OPS 5 (Linguagem de programação de computador). 3. Interconexão de redes (Telecomunicações). 4. Telecomunicações - Sistemas de comutação. 5. Comutação de pacotes (Transmissão de dados). I. Tatsch, Jürgen Peter. II. Salvador, Rogério Marcos. III. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Título em Inglês: Architecture and implementation of an optical packet switched ring network

Palavras-chave em Inglês: Telecommunication, OPS 5 (Computer systems), Internetworking (Telecommunication), Switching systems telecommunication, Switching, packet data transmission

Área de concentração: Telecomunicações

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Felipe Rudge Barbosa, Fabio Donati Simões

Data da defesa: 26/08/2009

Programa de Pós Graduação: Engenharia Elétrica

C₁
R-25

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Vinícius Garcia de Oliveira

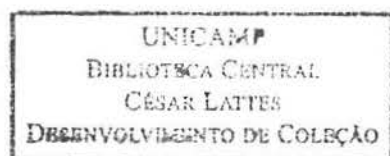
Data da Defesa: 26 de agosto de 2009

Título da Tese: "Arquitetura e Implementação de uma Rede em Anel de Comutação Óptica de Pacotes"

Prof. Dr. Peter Jürgen Tatsch (Presidente): Peter Jürgen Tatsch
Dr. Fabio Donati Simões: Fabio Donati Simões
Dr. Felipe Rudge Barbosa: Felipe Rudge Barbosa

2010-1008

Aos meus pais, por tudo.



Resumo

A convergência das aplicações, dos serviços e das redes de telecomunicação está mudando o mundo e o jeito como as pessoas agem e interagem. Embora com limitações, computação distribuída, teleconferência, tele-educação, radio, televisão e jogos já são possíveis na Internet. É razoável propor que em um futuro próximo estas limitações tendam a desaparecer e será possível suportar aplicações bem mais sofisticadas e complexas, tais como realidade virtual e tele-presença.

Para que este futuro se torne realidade é necessário que as arquiteturas das redes de telecomunicação evoluam para um novo paradigma de rede que abranja as várias camadas de transporte, de serviço e de gerência. Portanto a abordagem padrão de aumentar a capacidade de transmissão, que até então foi predominante, não será suficiente para atender as demandas acima mencionadas. É preciso muito mais.

Esta evolução deve acontecer dentro dos próximos dez anos e, essencialmente, começando pelas redes metropolitanas e de acesso dada a sua proximidade física e lógica com os usuários finais. Nestas condições as necessidades e a dinâmica das aplicações se fazem mais presentes do que nas redes interestaduais, nacionais e internacionais, onde a agregação de tráfego ocorre e suaviza os impactos das aplicações, sendo hoje a rede óptica a melhor candidata para realizar esta função dado o fato que esta possui a maior banda passante e a maior taxa de transmissão.

Assim, este trabalho propõe uma nova arquitetura de rede em anel de comutação óptica de pacotes. Com foco bem definido nas camadas de transporte, o seu objetivo é conceber uma tecnologia inovadora, mesclando a tecnologia de comutação de pacotes com as características favoráveis das redes de comutação de comprimento de onda e de rajadas. Além disso, a rede é proposta de forma a ser construída com os componentes já disponíveis nos dias de hoje tendo assim potencial para assumir posição de destaque em relação a outras propostas de redes similares porém que apresentam dificuldades na implementação ou são ainda inviáveis.

Abstract

The convergence of the services, applications and telecommunication networks is changing the way people do their things and interact to each other. With limitations, distributed computing, teleconference, tele-education, radio, television and games are already possible through the Internet. In a close future these limitations probably will disappear and it will be possible to provide new and more sophisticated applications as virtual reality and telepresence.

In order to make it possible in the near future, it is necessary to evolve the current network architecture to a new paradigm that covers several network layers including transport, services and management. The current approach of increasing the transmission bandwidth to provide new services is no longer enough, a lot more is needed.

This evolution should happen in the next ten years beginning from the metropolitan and access networks given their physical and logical closeness to the end users. In these networks the dynamic needs of the applications result in a bursty traffic that is extremely hard to deal with, differently from the long distance networks where the traffic aggregation occurs and minimizes the impact of the applications.

Having this in mind this work proposes a new optical packet switched ring network architecture. Aiming the transport layers as it focus the idea of this network is to combine the Optical Packet Switching technique (OPS) with the Optical Circuit Switching (OCS) and Optical Bust Switching (OBS) in order to create an innovative architecture that provides means to face the new challenges above described.

It is also objective of this work to conceive a feasible network architecture, possible to be implemented with the already available devices and optical components.

Agradecimentos

- A Deus por tudo o que tenho e sou.
- À minha família que me proporcionou todas as oportunidades para chegar até aqui.
- À minha esposa Simone pelo companheirismo e compreensão.
- Ao Prof. Peter pela imensurável paciência e apoio ao longo de todo este trabalho.
- Ao Marcos Salvador por todo o trabalho de revisão e valiosas considerações.
- Ao CPqD e ao governo brasileiro pelo suporte financeiro.
- Aos colegas do CPqD sem os quais este trabalho seria impossível de ser realizado, em especial os amigos Luis Renato Monte, Rodrigo Bernardo e Eduardo Mobilon.
- Aos membros da banca pelas inúmeras correções e excelentes sugestões.

Índice

Resumo	VII
Abstract.....	IX
Agradecimentos	XI
Índice	XIII
Índice de Figuras	XVII
Índice de Tabelas.....	XIX
Lista de Abreviaturas	XXI
1. Introdução.....	23
1.1. Redes ópticas	24
1.1.1. WDM	24
1.1.2. Componentes ópticos	25
1.1.3. Comutação eletrônica versus comutação óptica	25
1.1.4. Redes ópticas de comutação de comprimento de onda.....	27
1.1.5. Redes ópticas de comutação de rajada	28
1.1.6. Redes ópticas de comutação de pacote	29
1.2. Motivação e Objetivo do Trabalho.....	30
1.3. Organização deste trabalho.....	32
2. Arquiteturas de redes em anel de comutação óptica de pacotes.....	33
2.1. Anel MOPS – Multiple-wavelength Optical Packet-Switched	34
2.1.1. Divisão temporal no acesso ao anel	35
2.1.2. Configuração de transmissão e recepção	36
2.2. Arquitetura PIPELINE	37
2.3. Arquitetura BORN	39
2.4. Arquitetura MAWSON	40
2.5. Arquitetura HORNET	41
2.6. Arquitetura RINGO	43
2.7. Arquitetura FLAMINGO	44
2.8. Discussão	45

3. Arquitetura da RCPO	47
3.1. Diretrizes	48
3.2. Especificação geral da arquitetura.....	49
3.3. Serviço de transporte	53
3.4. Controle de acesso ao meio	59
3.4.1. Formato do cabeçalho	63
3.4.2. Inserção e sincronização dos cabeçalhos	64
3.4.3. Formato do pacote óptico.....	65
3.5. Tolerância a falhas.....	65
3.5.1. Serviço de proteção de caminho	68
3.6. Resultados simulados de desempenho	71
3.7. Interconexão	76
3.8. Discussão.....	77
4. Arquitetura do Nó.....	81
4.1. Unidade de controle.....	83
4.2. Unidade de transporte de dados.....	84
4.2.1. Bloco eletrônico	84
4.2.2. Bloco de comutação óptica	87
4.3. Estados dos anéis no nó.....	89
4.4. Discussão.....	90
5. Projeto de implementação	93
5.1. Análise das tecnologias candidatas	94
5.1.1. Componentes eletrônicos	94
5.1.2. Componentes ópticos	98
5.1.3. Tecnologias adotadas	99
5.2. Projeto de implementação	101
5.2.1. Placa base e placas filhas	101
5.2.2. Blocos digitais.....	105
5.3. Atual estado do nó prototipado	107
5.3.1. Diagrama em blocos do nó prototipado	109
5.4. Cenários de teste e resultados.....	110

5.4.1. Transmissão entre dois nós ligados ponto a ponto.....	111
5.4.2. Transmissão entre três nós ligados em anel	113
5.5. Discussão.....	114
6. Conclusões e Trabalhos Futuros.....	117
Referências Bibliográficas	121
Glossário de Componentes Ópticos	125

Índice de Figuras

Figura 1-1 Sistema WDM	25
Figura 1-2 Comutação eletrônica versus óptica	26
Figura 1-3 Comutação por comprimento de onda.....	28
Figura 1-4 Comutador de uma rede de comutação de rajadas	29
Figura 1-5 Comutador de uma rede de comutação de pacotes ópticos	30
Figura 2-1 Anel MOPS	34
Figura 2-2 Divisão por slots em um canal óptico	36
Figura 2-3 Acesso aos slots na arquitetura Pipeline.....	37
Figura 2-4 Diagrama de blocos do nó da rede Pipeline	38
Figura 2-5 Estrutura dos <i>buffers</i> e <i>transceivers</i> na rede Born.....	39
Figura 2-6 Nó da rede Born	40
Figura 2-7 Arquitetura MAWSON	41
Figura 2-8 Nó da rede Hornet	42
Figura 2-9 Nó da rede Ringo.....	44
Figura 2-10 Nó da rede Flamingo	45
Figura 3-1 Arquitetura macro da RCPO	49
Figura 3-2 Inter-relacionamento entre os planos funcionais.....	52
Figura 3-3 Solicitação do cliente e configuração da rede	52
Figura 3-4 Classe de serviço de melhor esforço	54
Figura 3-5 Classe de serviço reservada.....	55
Figura 3-6 Utilização da tabela de alocação de <i>slots</i> para a distribuição do tráfego	56
Figura 3-8 Fluxograma do protocolo SAT+	61
Figura 3-9 Cabeçalhos e <i>slots</i>	62
Figura 3-10 Fluxograma do processamento do cabeçalho e a transmissão dos pacotes ópticos para a classe de serviço de melhor esforço.....	62
Figura 3-12 Formato do pacote óptico	65
Figura 3-13 Mecanismo de reinserção contínua	66
Figura 3-14 Tolerância a múltiplas falhas de fibra	67
Figura 3-15 Tolerância a falha de nó	68

Figura 3-16 Distribuição de tráfego a fim de prover proteção de caminho	69
Figura 3-17 Serviço de proteção de caminho.....	70
Figura 3-18 Vazão total da rede versus tráfego no cenário simétrico.....	73
Figura 3-19 Grau de justiça no acesso versus carga no cenário simétrico.....	74
Figura 3-20 Vazão total da rede versus tráfego no cenário assimétrico	75
Figura 3-21 Grau de justiça no acesso versus carga no cenário assimétrico	76
Figura 3-22 Formato dos quadros Ethernet IEEE 802.3 e IEEE 802.1q/p	77
Figura 4-1 Visão arquitetônica do nó.....	82
Figura 4-2 Unidade de controle	83
Figura 4-3 Bloco eletrônico de classificação de quadros e montagem de rajadas	85
Figura 4-4 Bloco de comutação óptica	88
Figura 4-5 Comutador óptico.....	89
Figura 4-6 Estados do anel no nó da rede	90
Figura 5-1 IXF 1104 e interconexões	100
Figura 5-2 Encapsulamento do SOA	101
Figura 5-3 Placa do fabricante PLD.....	102
figura 5-4 Placa MAC	102
Figura 5-5 Placa Serdes.....	103
Figura 5-6 Placa SOA	104
Figura 5-7 Ligação entre a placa base e as placas filhas.....	104
Figura 5-8 Blocos digitais	105
Figura 5-9 Nó prototipado.....	108
Figura 5-10 Blocos digitais implementados.....	109
Figura 5-11 Diagrama em blocos do nó prototipado	110
Figura 5-12 Cenário de teste com dois nós ligados ponto a ponto	111
Figura 5-13 Cenário de teste com três nós ligados em anel.....	113
Figura A-1 Acoplador óptico	125
Figura A-2 Circulador óptico.....	126
Figura A-4 Filtro de Fabry-Perot	127
Figura A-5 Roteador de comprimentos de onda	128
Figura A-6 Add and Drop	128

Índice de Tabelas

Tabela 3-1 Requisitos das redes de nova geração	48
Tabela 3-2 Classes de serviço de transporte da rede	59
Tabela 3-3 Parâmetros de entrada	73
Tabela 3-4 Atraso médio dos pacotes	74
Tabela A-1 Tempo de atuação das chaves em função da aplicação	130
Tabela A-2 Características das chaves ópticas	130

Lista de Abreviaturas

ASE	Amplified Spontaneous Emission
ASIC	Application Specific Integrated Circuit
AWG	Arrayed Waveguide Grating
BP	Bytes of Padding
CERN	Organização Europeia para Estudos Nucleares
CoS	Class of Service
CPLD	Complex Programmable Logic Device
CRC	Cyclic Redundancy Check
CVPD	Chemical Vapour Phase Deposition
CW	Carrier Wave
DA	Destination Address
DCN	Data Communication Network
DDR-2	Double Data Rate 2
DWDM	Dense Wavelength Division Multiplexing
EDFA	Erbion-Doped Fiber Amplifier
FDL	Fiber Delay Line
FPGA	Field Programmable Gate Array
FR	Faraday Rotator
GMPLS	Generalized Multi-Protocol Label Switching
HWP	Half-Wave Plate
IP	Internet Protocol
LASER	Light Amplification by Stimulated Emission of Radiation
LED	Light Emitting Diode
MAC	Medium Access Control
MIPS	Millions Instructions Per Second
MOPS	Multiple-wavelength Optical Packet-Switched
MSS	Multiple Subcarrier Signalling
NIU	Network Interface Unit

NP	Network Processor
OADM	Optical Add and Drop Multiplexers
OBS	Optical Burst Switching
OCS	Optical Circuit Switching
OPS	Optical Packet Switchin
OSA	Optical Spectrum Analyzer
OXC	Optical cross-Connects
PAT	Packet Aggregate Transport
PCI	Peripheral Component Interconnect
PDFA	Praseodymium-Doped Amplifier
PLL	Phase Locked Loop
PS	Padding Size
QoS	Quality of Service
RTOS	Real Time Operating System
SA	Source Address
SDH	Synchronous Digital Hierarchy
SERDES	SERialazer/DESerializer
SFP	Small Form Plugable
SOA	Semiconductur Optical Amplifier
SOC	Switch On a Chip
SOP	State Of Polarization
SPI-3	System Packet Interface Level 3
SWP	Spatial Walk-off Polarazier
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TFF	Thin-Film Fileter
TTL	Time-To-Live
UDP	User Datagram Protocol
VHDL	VHSIC Hardware Description Language
WDM	Wavelength Division Multiplexing

Capítulo 1

Introdução

O reconhecimento das limitações das arquiteturas das redes atuais tem levado muitas universidades e centros de pesquisa de excelência a investirem na pesquisa de novas arquiteturas de rede. Dentre as arquiteturas em estudo destacam-se as que maximizam o uso das tecnologias ópticas para reduzir o chamado gargalo eletrônico. Estas arquiteturas representam uma mudança no paradigma reinante, centrado em soluções que otimizam os aspectos de transmissão, o grande gargalo do passado, para um novo paradigma centrado em soluções que reduzem o processamento eletrônico, o grande gargalo atual.

São várias as arquiteturas inovadoras de redes ópticas em estudo, porém, de uma forma geral, elas podem ser classificadas de acordo com as técnicas de comutação óptica que utilizam, sendo elas: a comutação de comprimento de onda (OCS - *Optical Circuit Switching*), a comutação de rajada (OBS - *Optical Burst Switching*) e a comutação de pacote (OPS - *Optical Packet Switching*). Todas elas apresentam vantagens e desvantagens, porém as arquiteturas de rede mais inovadoras têm se concentrado nas duas últimas em vista da maior flexibilidade que elas oferecem.

Este capítulo apresenta uma introdução sobre as redes ópticas a fim de contextualizar o leitor sobre o cenário atual e seus desafios. Ao final são apresentados a motivação e o objetivo desse trabalho em conjunto com a organização dos capítulos que seguem.

1.1. Redes ópticas

O desafio das redes de telecomunicações estava na transmissão de longa distância e no transporte dos dados no núcleo da rede, onde ocorre a agregação do tráfego.

Com o amadurecimento da optoeletrônica, dos amplificadores ópticos e da própria fibra óptica esta questão deixou de se tornar um desafio, sendo que conexões com capacidade de dezenas de gigabits por segundo passaram a ser comuns, atendendo a qualquer demanda presente.

Entretanto uma vez sanada a dificuldade da transmissão de longa distância outro desafio surgiu – como realizar a comutação desse tráfego colossal no domínio elétrico. Essa questão se agrava ainda mais na região da rede próxima ao usuário final, onde a característica do tráfego é de rajadas, ou seja, onde existe grande demanda de banda por curtos períodos de tempo.

Os próximos itens descrevem em maiores detalhes as tecnologias e propostas de redes ópticas que visam solucionar este desafio.

1.1.1. WDM

As redes ópticas se tornaram altamente atrativa através utilização da técnica da divisão por comprimento de onda (WDM - *Wavelength Division Multiplexing*). Através dela é possível realizar a transmissão de diversos comprimentos de ondas em uma mesma fibra, sem que haja interação entre as transmissões, similar a divisão por frequência nas transmissões por radiodifusão.

No WDM cada transmissão, além de co-existir com as demais dentro de uma mesma fibra, pode atingir uma capacidade de transmissão tão alta quanto uma transmissão simples, ou seja, sem a utilização desta técnica.

Atualmente os sistemas comerciais são capazes de atingir banda de transmissão de até 40 Gbps, desta forma, utilizando-se os sistemas WDM de 32 a 64 canais, ou ainda os sistemas DWDM (*Dense Wavelength Division Multiplexing*) com 160 canais, pode-se atingir uma capacidade de transmissão na ordem de terabits por segundo com uma única fibra. A figura 1-1 ilustra o conceito do WDM. Em [1] é apresentada a maior rede WDM experimental do Brasil.

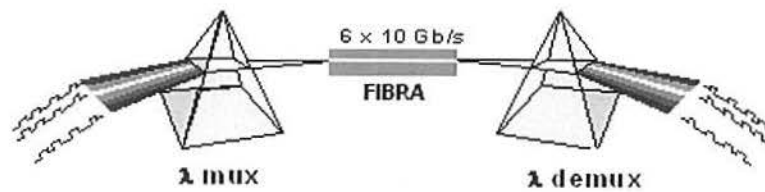


Figura 1-1 Sistema WDM

1.1.2. Componentes ópticos

O surgimento das redes ópticas não seria possível se os componentes ópticos não sofressem significativa evolução ao longo dos últimos anos. Componentes como filtros, circuladores, acopladores, chaves entre muitos outros possibilitaram, por exemplo, a realização da comutação de dados no domínio óptico.

É certo que estes componentes se tornarão ainda mais sofisticados no futuro mas muito já é possível com as tecnologias disponíveis. Para uma descrição mais detalhada dos principais componentes ópticos e seu princípio de funcionamento consulte o apêndice A.

1.1.3. Comutação eletrônica versus comutação óptica

Como visto os sistemas de transmissão ponto a ponto atingem hoje capacidades de dezenas de gigabits por segundo em uma única transmissão podendo atingir até a taxa de terabit através da técnica WDM. Desta forma o gargalo das redes que até então estava na transmissão passa para a comutação nos nós, que deve lidar com o encaminhamento do tráfego que chega por inúmeras conexões de alta capacidade.

A figura 1-2(A) mostra o nó típico das redes atuais de comutação de pacotes. Nela o tráfego chega ao nó por suas diversas interfaces ópticas e, após a conversão para o domínio elétrico, será comutado para as interfaces de destino, sendo novamente convertido para o domínio óptico na transmissão.

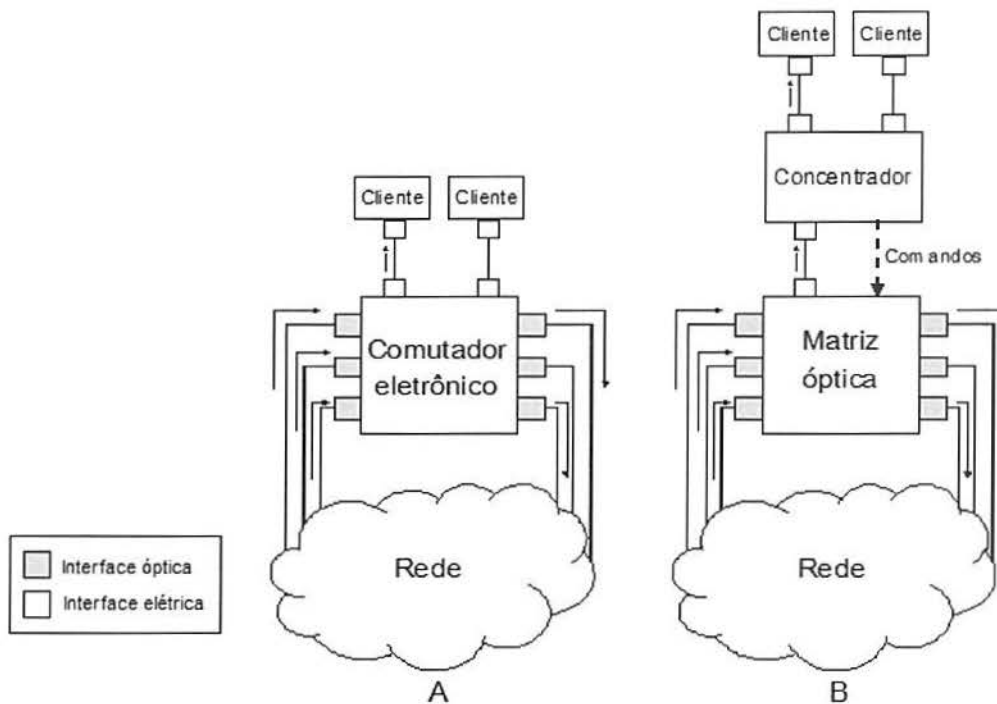


Figura 1-2 Comutação eletrônica versus óptica

Notam-se nesse processo diversos problemas:

- **Custo:** o equipamento que realiza a conversão do domínio óptico para o elétrico e vice-versa possui valor elevado especialmente nas taxas de gigabit.
- **Ineficiência:** apenas uma fração do tráfego é destinada aos clientes deste nó, o restante apenas o utiliza como salto para o seu destino final.
- **Complexidade:** o nó deve ser multiprotocolos, sendo apto a entender todos os protocolos de transporte utilizados no tráfego que passa por ele.
- **Capacidade:** as matrizes de comutação eletrônica tendem a atingir o seu limite físico dentro dos próximos anos. É importante notar que cada gigabit aumentado nas interfaces reflete-se em N gigabits no aumento de capacidade da matriz eletrônica, onde N é o número de interfaces no nó.
- **Consumo:** além das inúmeras conversões eletro-ópticas a alta complexidade e capacidade das matrizes eletrônicas resultam em um consumo que pode chegar a milhares de watts em apenas um nó da rede.

Assim o novo modelo exibido na figura 1-3(B) é proposto e apresenta inúmeras vantagens. Neste é utilizada uma matriz óptica para a comutação, o que evita que o tráfego que não se

destina aos clientes do nó seja convertido para o domínio elétrico e novamente para o óptico. Desta maneira o sinal que chega é comutado de forma totalmente óptica. Isso também isola o nó dos diversos protocolos de transporte trafegados na rede, necessitando implementar apenas os protocolos utilizados pelos clientes diretamente conectados a ele.

Entretanto frente a tantas vantagens uma questão se mostra desafiadora – como executar os comandos na matriz óptica de forma a atender todos os tipos de demanda tráfego. Nesta questão está o foco da pesquisa em redes ópticas.

Ao longo dos últimos anos diversos modelos foram propostos, porém dentre todos estes três se destacam: a comutação por comprimento de onda, a comutação de rajadas e a comutação de pacotes ópticos. Os próximos itens abordam cada um destes três modelos.

1.1.4. Redes ópticas de comutação de comprimento de onda

Nas redes OCS [2] a transmissão dos dados ocorre somente após o estabelecimento de um caminho óptico. O estabelecimento deste ocorre mediante a configuração das matrizes de comutação ópticas nos nós onde passará o tráfego através do envio de uma mensagem de solicitação pelo transmissor em direção ao receptor e da confirmação do receptor para o transmissor.

As mensagens de sinalização podem trafegar através de fibra óptica ou cabo metálico e são processadas eletronicamente em cada nó ao longo do caminho. Uma vez estabelecida a conexão, o tráfego segue o caminho óptico sem qualquer atraso (fora o próprio atraso da luz na fibra), processamento, ou conversão entre os domínios elétrico e óptico nos nós de transporte da rede. Ao término das transmissões o caminho óptico deve ser desfeito explicitamente tal como no estabelecimento. A criação e remoção dos caminhos ocorrem na ordem de milissegundos.

A figura 1-3 ilustra uma rede OCS. Os principais elementos desta rede são os OADMs (*Optical Add and Drop Multiplexers*) que realizam a adição e remoção dos canais em cada um dos nós, os OXCs (*Optical Cross-connects*) que realizam a comutação dos canais ao longo da rede, os *transponders*, que transmitem e recebem o sinal em um comprimento de onda específico e os amplificadores ópticos, que permitem a amplificação do sinal sem a necessidade de qualquer conversão eletro-óptica.

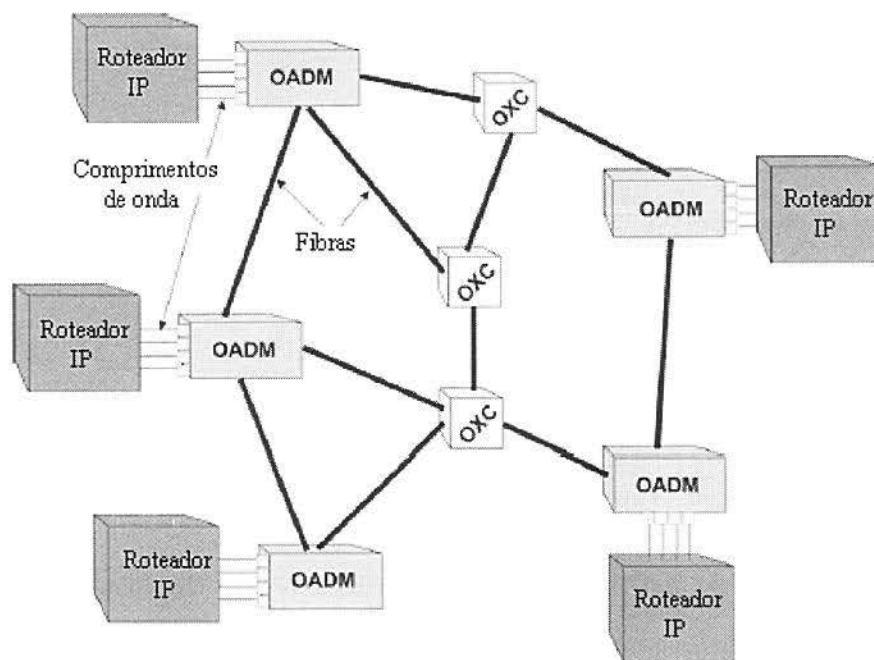


Figura 1-3 Comutação por comprimento de onda

1.1.5. Redes ópticas de comutação de rajada

As redes OBS [3] são similares às redes de comutação de comprimento de onda, entretanto esta nova técnica se diferencia da primeira por não precisar de mensagem de confirmação de estabelecimento de caminho óptico para início de transmissão e nem de mensagem de destruição de caminho óptico após o final da transmissão. Assim o transmissor periodicamente agrega um número de pacotes cujo tamanho total seja inferior ao tamanho da rajada permitida na rede e envia uma mensagem de solicitação de caminho óptico em direção ao receptor. Ciente do tempo de processamento da mensagem em cada nó e do tempo de propagação no meio o transmissor envia a mensagem de reserva de caminho e inicia a transmissão assim que vencer esse tempo de guarda.

Após o termino do tempo da rajada o caminho óptico é automaticamente destruído. Neste cenário uma falha na reserva do caminho leva ao descarte das informações contidas na rajada. O tempo de existência dos caminhos criados nas redes OBS é da ordem de centenas de microssegundos.

A figura 1-4 ilustra um comutador de uma rede OBS, neste um processador recebe as mensagens de solicitação de caminho e atua na matriz de comutação óptica para atender esta demanda.

A principal diferença entre as redes OCS e OBS está na eficiência da utilização do meio. Na primeira o caminho é criado porém não é certo se o tráfego gerado será constante por todo o tempo que o caminho existir. Na segunda o caminho é criado para dar vazão a um tráfego que já está aguardando pela transmissão. Em contrapartida a sua eficiência na utilização do meio de transmissão as redes OBS introduzem atrasos por causa do armazenamento tráfego.

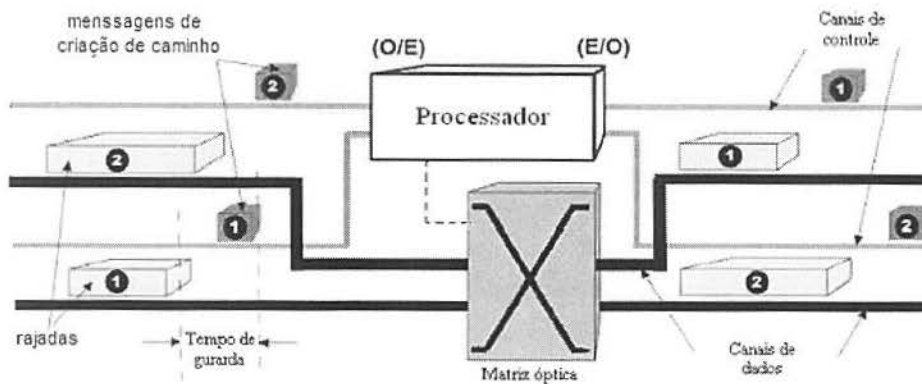


Figura 1-4 Comutador de uma rede de comutação de rajadas

1.1.6. Redes ópticas de comutação de pacote

Na rede OPS [4] a sinalização de reserva de caminho trafega em conjunto com os dados. Assim os dados e a sinalização constituirão o pacote óptico, sendo os dados usualmente chamados de carga útil do pacote e a sinalização de cabeçalho do pacote.

Ao chegar no nó o cabeçalho é removido e convertido para o domínio elétrico sendo analisado por um processador que, baseado nas informações do cabeçalho, atuará na matriz de comutação óptica de forma a encaminhar os dados corretamente. Na saída do nó um novo cabeçalho contendo a informação necessária para a comutação no próximo nó será adicionado à frente dos dados.

Durante o tempo de processamento do cabeçalho os dados circulam por um trecho de fibra contido em um carretel comumente chamado de atrasador FDL (*Fiber delay Line*). O nó

possui esse tempo de atraso para realizar todas as operações necessárias. A figura 1-5 ilustra um exemplo de comutador de rede OPS.

A rede OPS difere das redes OBS e OCS principalmente por nesta não existir a alocação de um caminho fim a fim. Em cada nó os dados serão comutados individualmente até atingir o seu destino final.

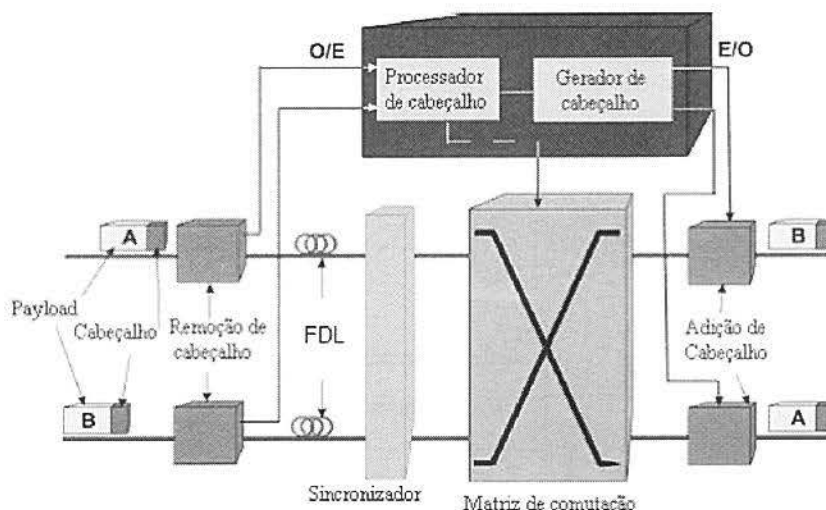


Figura 1-5 Comutador de uma rede de comutação de pacotes ópticos

1.2. Motivação e Objetivo do Trabalho

Entendemos que o próximo passo das redes de telecomunicação está na migração da comutação eletrônica para a comutação óptica por diversos fatores como capacidade, custo, simplicidade e eficiência. Entretanto não é claro qual será o modelo que predominará nesta nova fase, existindo hoje três promissores candidatos: OCS, OBS e OPS. Todos estes modelos apresentam vantagens e desvantagens.

O modelo de comutação por comprimento de onda é o mais simples, porém é orientado a conexões. Esta característica é interessante em redes de núcleo onde o tráfego requer alta capacidade de transmissão se mantendo constante. Entretanto em redes de acesso ou metropolitanas as características de tráfego são bem diferentes, tendendo ao comportamento de rajadas. Esse comportamento está se tornando cada vez mais comum com a popularização de serviços como voz e vídeo sobre IP (*internet protocol*). Assim os modelos de comutação óptica por rajada e pacote visam atender esta demanda.

Entretanto são grandes os desafios para a construção destas redes. Por exemplo, o controle de acesso ao meio é o principal responsável por determinar a eficiência no transporte de dados. Permitir o acesso dos nós aos recursos da rede evitando colisões e ao mesmo tempo garantindo alta vazão e justiça é fundamental para se obter uma arquitetura de rede apta a atender as novas demandas de tráfego, porém não é algo fácil de alcançar. Além disto, definir o mecanismo de controle de acesso ao meio com suporte a transmissão de um para muitos na camada óptica, definido como *multicast* óptico, não é apenas uma característica desejável, mas cada vez mais necessária para as redes de acesso e metropolitana, sendo mais um desafio considerável.

Outro aspecto importante que apresenta alto desafio nas redes de comutação de pacote óptico é o tempo para a decisão na comutação do pacote. Em geral este modelo de rede necessita de tomada decisões de comutação na ordem de poucas dezenas de micro segundos, o que aumenta a complexidade de implementação. Outras características críticas destas redes são a recuperação de relógio na recepção, devida a não existência de conexão fixa na camada de enlace, e a sincronização entre nós.

Além destas dificuldades, questões como engenharia de tráfego, diferenciação de serviços, tolerância a falhas, proteção, escalabilidade e segurança, que são necessárias em toda arquitetura de rede de nova geração, também devem ser previstos para as redes de comutação óptica [5].

Desta forma, este trabalho tem o objetivo de especificar e implementar uma rede em anel de chaveamento de pacotes ópticos denominada de RCPO (Rede de Chaveamento de Pacotes Ópticos). A RCPO busca resolver os problemas acima mencionados não apenas através da comutação óptica de pacotes, mas também utilizando os conceitos das redes OCS e OBS, protocolos de agregação de tráfego, além de mecanismos de controle de acesso e utilização de recursos, criando assim uma arquitetura de rede completa e inovadora.

Este trabalho foi financiado com recursos do Fundo para o Desenvolvimento Tecnológico das Telecomunicações (FUNTTEL) e executado no Centro de Pesquisa e Desenvolvimento (CPqD) em parceria com as universidades Unicamp e PUC-Rio.

1.3. Organização deste trabalho

Este trabalho é organizado em seis capítulos. O primeiro capítulo apresenta uma introdução sobre as redes ópticas e os seus atuais desafios. Também é colocada neste a motivação, o objetivo e organização deste trabalho. O capítulo 2 expõe algumas arquiteturas de redes de comutação óptica de pacotes em anel similares a RCPO e descritas na literatura ou prototipadas em laboratório. O capítulo 3 apresenta a arquitetura da RCPO com as suas diretrizes gerais, serviço de transporte, tolerância a falhas, proteção, controle de acesso ao meio, interconexão e resultados de simulação. O capítulo 4 mostra a arquitetura do nó da rede com os seus planos funcionais. O capítulo 5 descreve o projeto de implementação analisando as tecnologias dos componentes possíveis de serem utilizados no projeto, a rede prototipada e os testes executados. O capítulo 6 encerra o trabalho com as conclusões e os trabalhos futuros.

Capítulo 2

Arquiteturas de redes em anel de comutação óptica de pacotes

Este capítulo tem o objetivo de descrever o modelo de comutação óptica de pacotes através dos exemplos das principais arquiteturas propostas na literatura ou prototipadas em laboratório. Ao final é realizada uma discussão sobre estas arquiteturas comparando-se os seus pontos positivos e negativos.

2.1. Anel MOPS – Multiple-wavelength Optical Packet-Switched

O anel MOPS (*Multiple-wavelength Optical Packet-Switched*) é utilizado na maioria das arquiteturas de redes de comutação óptica de pacotes e visa, como o nome indica, a transmissão de pacotes ópticos entre nós ligados em uma topologia em anel através de diversos comprimentos de onda.

Um anel MOPS pode ser composto por diversos nós concatenados por diversas fibras ópticas em anel. Cada anel de fibra possui um número de canais, sendo um canal um dado comprimento de onda especificado pelo padrão WDM. Os mesmos canais utilizados em uma fibra podem ser reutilizados nas demais. Para realizar o acesso aos canais os nós são construídos com OADMs.

A figura 2-1 ilustra um anel MOPS composto de quatro nós e quatro canais circulando através de uma única fibra.

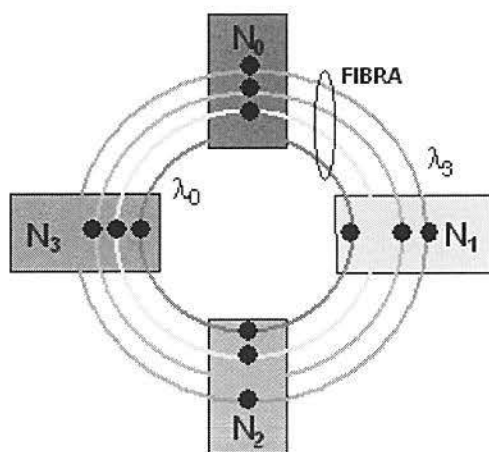


Figura 2-1 Anel MOPS

Neste exemplo cada nó recebe um número N_n que o identifica dentro do anel. No caso o nó N_0 utiliza o canal λ_0 , N_1 o canal λ_1 e assim por diante. A cada nó é associado o canal que possui o mesmo número identificador. Assim é possível a um nó apenas receber dados pelo canal de mesmo número, porém a transmissão pode ocorrer através de qualquer um dos três canais restantes de acordo com o nó que se deseja alcançar.

Porém, esta associação de comprimentos de ondas aos nós não necessariamente se dá desta forma. Outros esquemas podem ser assumidos de acordo com as características da rede que se pretende obter como desempenho, robustez e simplicidade. Todos os exemplos de redes descritos neste capítulo utilizam o conceito do anel MOPS.

2.1.1. Divisão temporal no acesso ao anel

Em uma rede óptica de comutação de pacotes um canal pode ser acessado por diversos nós ao longo da rede. Desta forma ocorre, inevitavelmente, concorrência pelo acesso ao canal. Assim dois ou mais transmissores podem tentar utilizar o canal compartilhado em um dado instante causando uma colisão. As colisões impendem que a informação transmitida seja recuperada no receptor.

São propostos diversos protocolos para que o acesso ao canal compartilhado seja controlado afim de não apenas evitar colisões, mas garantir que o número de acessos e o tempo de utilização do canal pelos nós sejam semelhantes, impedindo que um único nó monopolize a utilização do canal.

Um dos modelos propostos mais promissores é a divisão do canal em intervalos de tempo definidos como *slots* temporais ou simplesmente *slots*. Ele utiliza o mesmo conceito do TDM (*Time Division Multiplexing*), onde um nó recebe a permissão para utilizar o recurso por um período de tempo limitado. Entretanto a divisão por *slots* possui característica mais dinâmica que a divisão temporal tradicional. Neste conceito um nó não recebe um tempo fixo e periódico no qual pode utilizar o canal, ao contrário, o acesso é realizado sobre a demanda da transmissão. Desta forma o nó irá utilizar um *slot* apenas existindo a necessidade de transmitir dados.

Em contrapartida, mesmo garantindo um excelente desempenho na transmissão de dados a utilização de *slots*, em especial no caso de *slots* com comprimentos variáveis, traz uma grande complexidade na implementação dos transmissores e receptores dos nós, principalmente no que se diz respeito à sincronização entre *slots* e recuperação de relógio quando a informação é novamente convertida para o domínio elétrico. A figura 2-2 ilustra o conceito de divisão por *slots* em uma rede óptica.

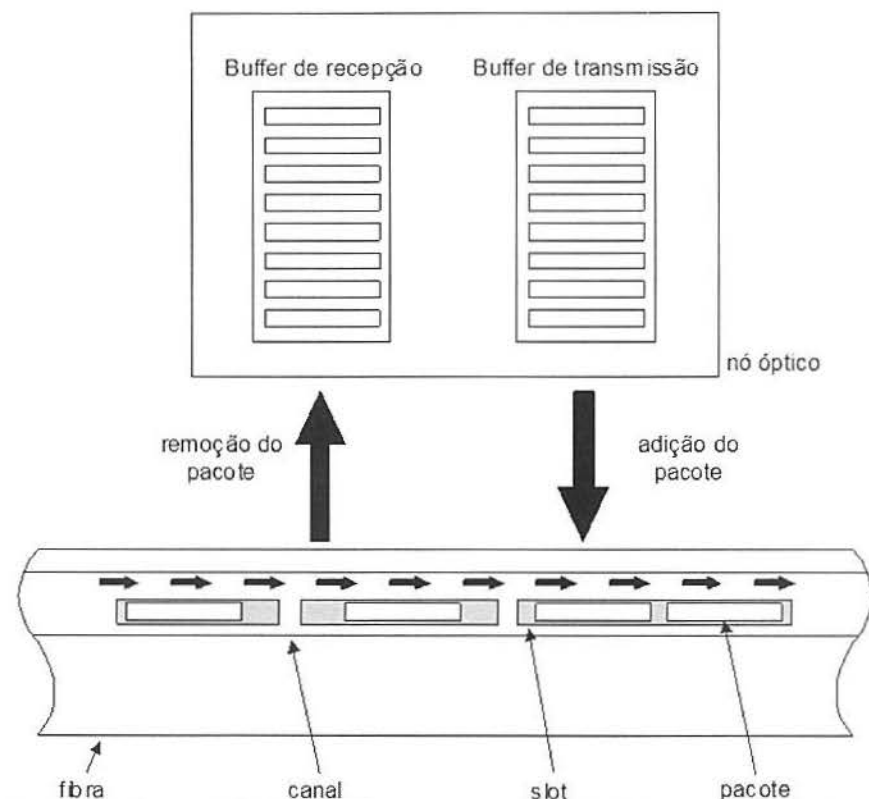


Figura 2-2 Divisão por slots em um canal óptico

Neste exemplo um nó da rede pode realizar três operações sobre o *slot*: a adição de dados, definida como *add*, a remoção, definida como *drop*, ou simplesmente a não alteração do estado corrente do *slot*, definido como *bypass*.

2.1.2. Configuração de transmissão e recepção

Para a construção do nó é possível utilizar transmissores com laser sintonizável ou fixo, assim como receptores com capacidade de recepção em apenas um comprimento de onda ou seletivo dentro de uma faixa espectral.

Para descrever a configuração é utilizada a notação $FTx^i TTx^j - FRx^m TRx^n$ [6]. Nesta notação i representa o número de transmissores fixos (*fixed Tx*), j o número de transmissores sintonizáveis (*tunable Tx*), m o número de receptores fixos (*fixed Rx*), e n o número de receptores sintonizáveis (*tunable Rx*).

As arquiteturas descritas a seguir seguem esta notação para a descrição de suas configurações.

2.2. Arquitetura PIPELINE

Uma das primeiras arquiteturas a utilizar o anel MOPS com acesso temporal por *slots* é a PIPELINE [7]. Ela utiliza uma única fibra com C números de canais em uma rede com N números de nós, onde é definido que N é igual a C.

Os *slots* de tempo de cada canal são sincronizados paralelamente entre si. O cabeçalho do *slot* traz consigo a informação sobre sua disponibilidade, ou seja, se está livre para ser utilizado ou não. Nesta arquitetura o cabeçalho é um sinal senoidal transmitido através da técnica MSS (*multiple subcarrier signalling*)[8] sendo o sinal modulado sobre o próprio comprimento de onda utilizado para transmitir os dados dos usuários da rede. A frequência do sinal modulado é definida como sub-portadora.

A detecção da sub-portadora de um *slot* implica na condição de indisponibilidade deste para a transmissão, ou seja, o *slot* já está sendo utilizado para o transporte da carga útil de um usuário ligado a outro nó. Além da informação sobre a disponibilidade a sub-portadora também define o nó destino da carga útil contida no *slot*. Para tal denomina-se n_d como o nó de destino, e sc_d como a única sub-portadora associada à n_d . Para se transmitir dados para n_d é modulada a sub-portadora sc_d no *slot* utilizado.

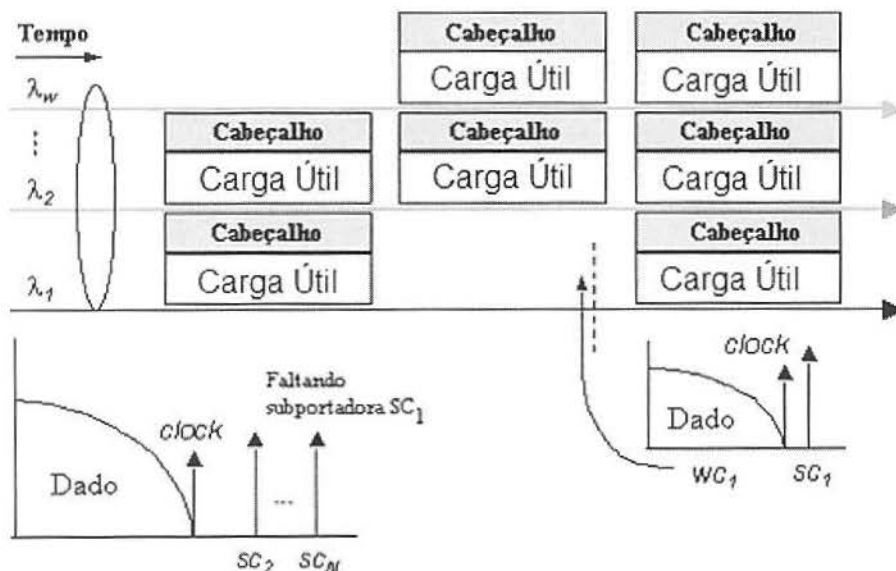


Figura 2-3 Acesso aos slots na arquitetura Pipeline

A figura 2-3 ilustra como se dá o acesso aos *slots*. O nó detecta se existe alguma sub-portadora no *slot* e caso haja verifica se ela é associada a ele. Nesta condição o nó realiza uma operação de *drop*. Caso a portadora seja qualquer outra o nó permite que o *slot* siga de forma inalterada para o próximo nó (*bypass*). Caso não detecte nenhuma sub-portadora o nó assume que o *slot* está livre e pode ser utilizado para a transmissão de seus dados. Assim o nó realiza uma operação de *add* inserindo os dados em conjunto com a sub-portadora correspondente ao nó de destino.

Em conjunto com os dados e o cabeçalho também é modulado um sinal com frequência $1/T$, onde T é o período de duração de um bit transmitido. Este sinal, definido como relógio, é transmitido para que o receptor obtenha a taxa de transmissão dos bits da carga útil e possa assim recebê-los corretamente. A figura 2-4 apresenta o diagrama de blocos do nó da rede PIPELINE.

Para que seja possível a detecção do cabeçalho antes da chegada dos dados é utilizado um *splitter* posto antes de um atrasador FDL. O nó tem este tempo de atraso para realizar as decisões necessárias.

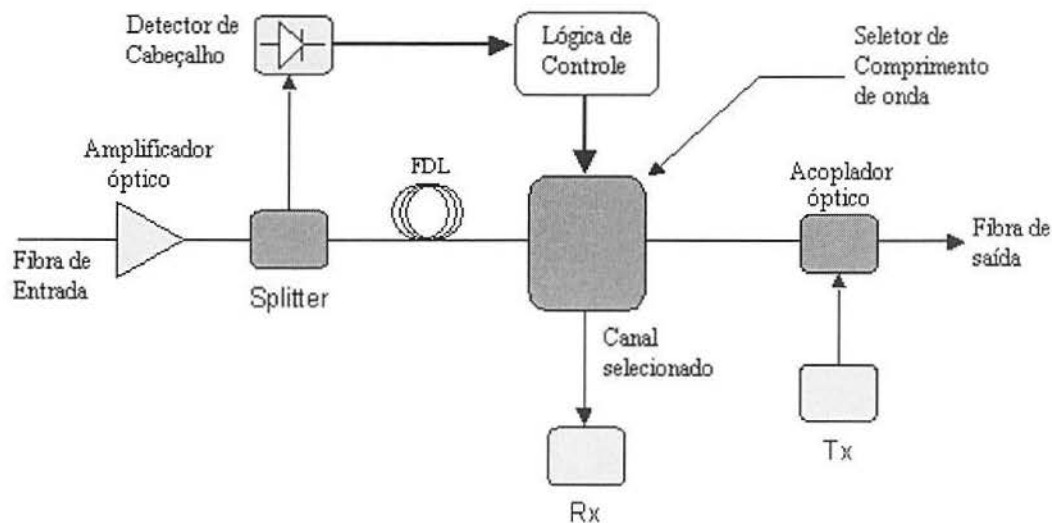


Figura 2-4 Diagrama de blocos do nó da rede Pipeline

2.3. Arquitetura BORN

A arquitetura BORN (*Broadband optical ring network*) [9] segue a configuração FTx^c-FRx^c. Os nós se conectam à rede através de interfaces chamadas de NIUs (*Network Interface Units*). Uma NIU consiste de C componentes OPS e C transmissores que compartilham os mesmos armazenadores temporários de transmissão e recepção, denominados, respectivamente, de *buffer* de transmissão e *buffer* de recepção. Também existem C canais na rede. A figura 2-5 mostra a estrutura dos *buffers* e transmissores, e a figura 2-6 mostra a organização da NIU.

A NIU também inclui um demultiplexador e um multiplexador WDM para que cada componente OPS opere dados em um canal específico. Um acoplador remove uma pequena fração do sinal para assim obter o cabeçalho do *slot*. O cabeçalho consiste do endereçamento de destino e o estado do *slot*. Antes do processador de endereçamento receber o cabeçalho o sinal sofre uma amplificação e durante o tempo de processamento do endereço o *slot* é atrasado com um atrasador FDL, para assim compensar o tempo gasto no processamento do cabeçalho. A figura 2-6 ilustra a composição do nó da rede BORN.

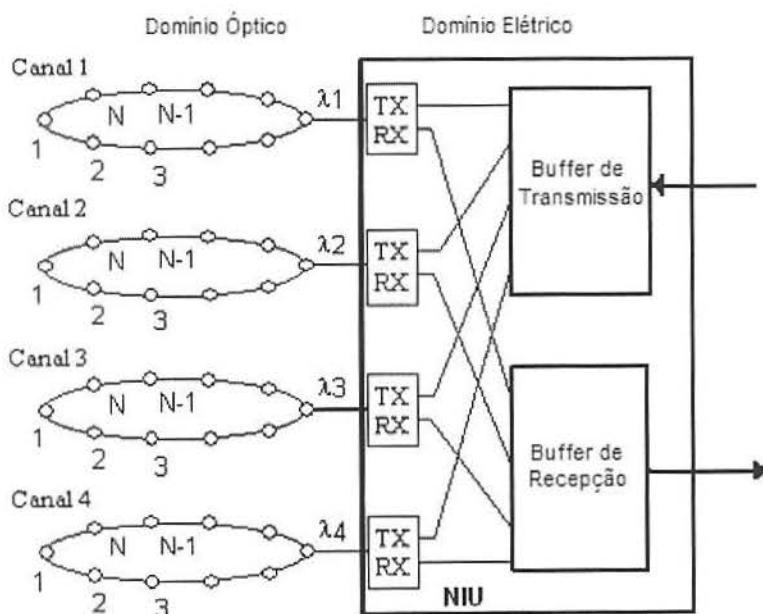


Figura 2-5 Estrutura dos *buffers* e *transceivers* na rede Born

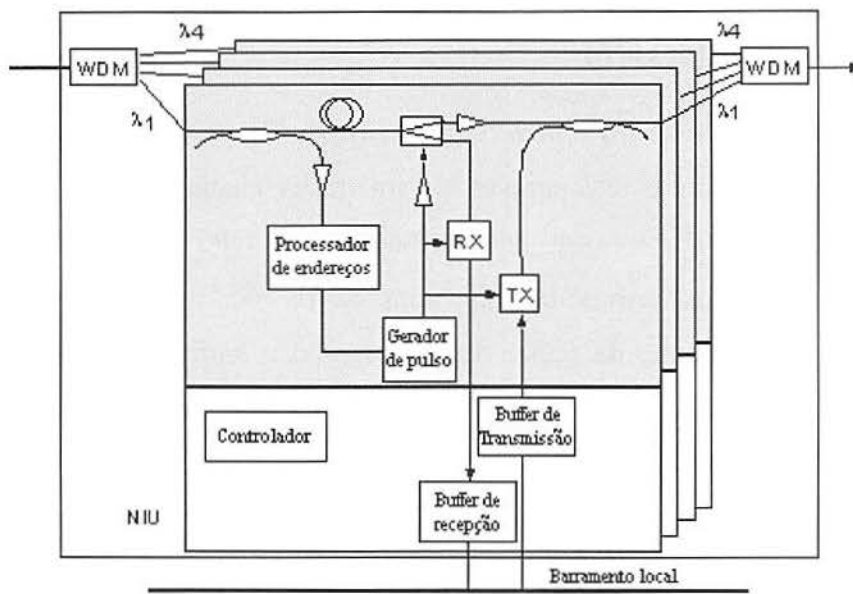


Figura 2-6 Nó da rede Born

2.4. Arquitetura MAWSON

A arquitetura MAWSON (*Metropolitan area wavelength switched optical network*) [10] tem por objetivo ser simples e de baixo custo, utilizando apenas componentes disponíveis comercialmente.

A MAWSON, como mostrado na figura 2-7, consiste de uma única fibra óptica, N nós OADMs e C canais, onde N é igual C . A configuração da rede é FTx^C-FRx^1 .

Cada nó é equipado com um conjunto de lasers que não só são capazes de chavear rapidamente entre os canais, mas também transmitir a mesma informação em diversos canais simultaneamente, possibilitando que um nó envie o mesmo dado para mais de um nó receptor. Essa operação é definida como *multicast*.

O nó é associado a um único canal para a recepção, e o endereçamento ao nó destino se dá exatamente pela escolha do canal associado a este. Desta maneira para transmitir para o nó de destino d , o transmissor deve selecionar o canal que este nó recebe, também chamado aqui de d .

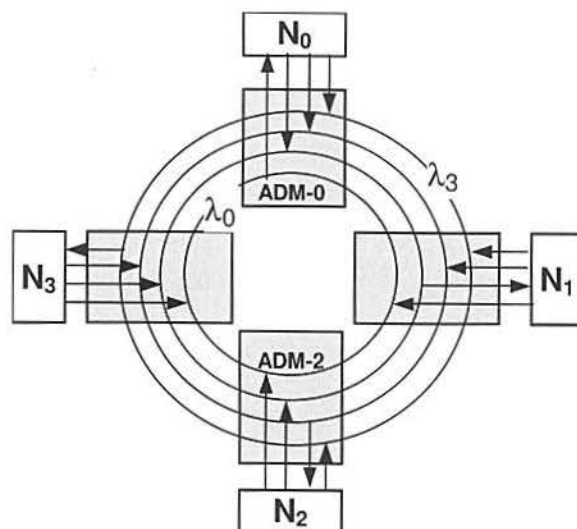


Figura 2-7 Arquitetura MAWSON

A estrutura do OADM segue o exemplo mostrado no item A.4, onde é utilizado um conjunto composto de um circulador óptico e uma grade de *Bragg* para a realização da operação de *drop*, e um acoplador para a operação de *add*.

Os nós desta rede não são capazes de obter informação sobre o meio de transmissão para realizar a decisão da transmissão. Desta forma a rede MAWSON conta com protocolos de acesso ao meio que não dependem desta verificação.

2.5. Arquitetura HORNET

A arquitetura HORNET (*Hybrid opto-electronic ring network*) [11] visa a composição de uma anel de até 100 nós de acesso e tamanho máximo de 100 quilômetros. Na especificação corrente a taxa de transmissão de cada canal é de 2,5 Gbps.

A configuração de transmissão e recepção é TTx^1-FRx^1 , e existem dois modos de operação: único salto e múltiplos saltos. Se o número N de nós é menor que o número C de canais, então dois ou mais nós podem compartilhar dinamicamente um canal de recepção. Nestas condições a rede trabalha no modo múltiplos saltos.

Em qualquer caso, único salto ou múltiplos saltos, é utilizada a sinalização de sub-portadora para levar o endereço do nó destino. Cada canal é associado a uma única frequência de sub-portadora, em um modelo parecido com a rede PIPELINE, descrita na seção 2.2.

Um nó consiste de três blocos funcionais: gerenciador de *slot*, remoção inteligente (*smart drop*) e adição inteligente (*smart add*). A figura 2-8 mostra a estrutura do nó.

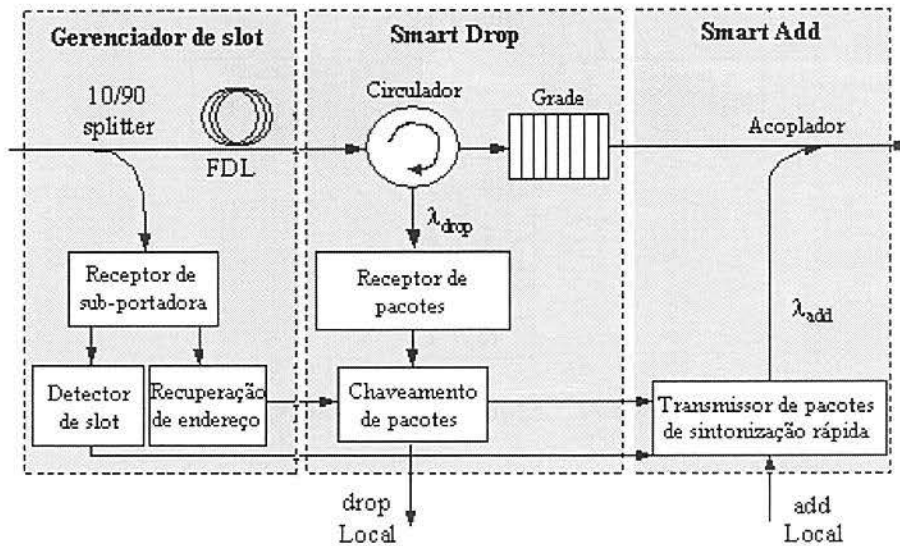


Figura 2-8 Nó da rede Hornet

O gerenciador de *slot* executa duas funções paralelas:

- Informar o *smart drop* sobre a existência de um pacote que deve ser removido pelo nó. Neste caso existem duas possibilidades: o pacote se destinar para o próprio nó ou o pacote ter como destino outro nó, usando o próprio nó para o encaminhamento (múltiplo salto).
- Informar o *smart add* da existência de canais ociosos. Neste caso o módulo detector de *slot* informa o transmissor sintonizável no *smart add* sobre qual canal está livre e o tamanho do *slot*, para assim fazer o melhor uso do recurso disponível.

Neste meio de tempo o atrasador FDL situado no gerenciador de *slot* atrasa o sinal para que as operações do *smart drop* e do *smart add* sejam propriamente preparadas.

Assim como na MAWSON é utilizado um conjunto composto de um circulador e uma grade de Bragg para realização de uma operação de *drop*, e um acoplador para a operação de *add*. Quando um pacote é recebido ele pode ser encaminhado para o próprio nó, ou ser enviado para o transmissor, que escolherá outro canal disponível para realizar a

retransmissão, caracterizando assim a funcionalidade de múltiplo salto anteriormente descrita.

2.6. Arquitetura RINGO

A arquitetura RINGO (*ring optical network*) [12] surgiu como uma alternativa em relação aos anéis de comutação temporal como o SDH (*Synchronous Digital Hierarchy*) [13], desta forma, esta é uma rede WAN (*Wide Area Network*), ou seja, de abrangência ampla em geral interligando cidades e países.

A arquitetura consiste de um anel de fibra unidirecional com N nós e C canais, onde N é igual a C . Cada canal é dividido em *slots* de tempo fixo, sendo os *slots* de todos os canais sincronizados paralelamente entre si.

A RINGO segue a configuração FTx^C-FRx^1 . Para a transmissão cada nó é equipado com um conjunto de lasers, para a recepção é utilizado um único receptor fixo que receberá os dados no canal destinado a recepção do nó.

Um nó pode transmitir em qualquer canal com exceção do seu próprio canal de recepção. Para transmitir para um nó de destino d , o nó de origem s seleciona o laser do canal, também chamado de d . A figura 2-9 ilustra o nó da rede.

O sinal entrante passa por uma amplificação de ganho fixo, realizado por um amplificador óptico (ver seção A.5), para assim compensar todas as perdas intrínsecas dos componentes passivos utilizados para a construção do nó. Um demultiplexador separa os canais. Como um dos canais é sempre destinado a um nó, este é sempre removido em seu nó associado sem a necessidade de analisar se existe ou não informação.

Parte da potência dos demais canais é removida e analisada por um comparador de nível. Caso seja detectado que um canal está livre, o controlador do nó será notificado de sua disponibilidade para a transmissão. Todos os canais são agrupados novamente com a utilização de um multiplexador. À frente, após um atrasador FDL, ocorre a transmissão no canal disponível.

A utilização da série de lasers em conjunto com um multiplexador permite que um modulador externo seja compartilhado entre todos os canais. Entretanto, não é possível realizar transmissões simultâneas de diferentes informações com esta arquitetura.

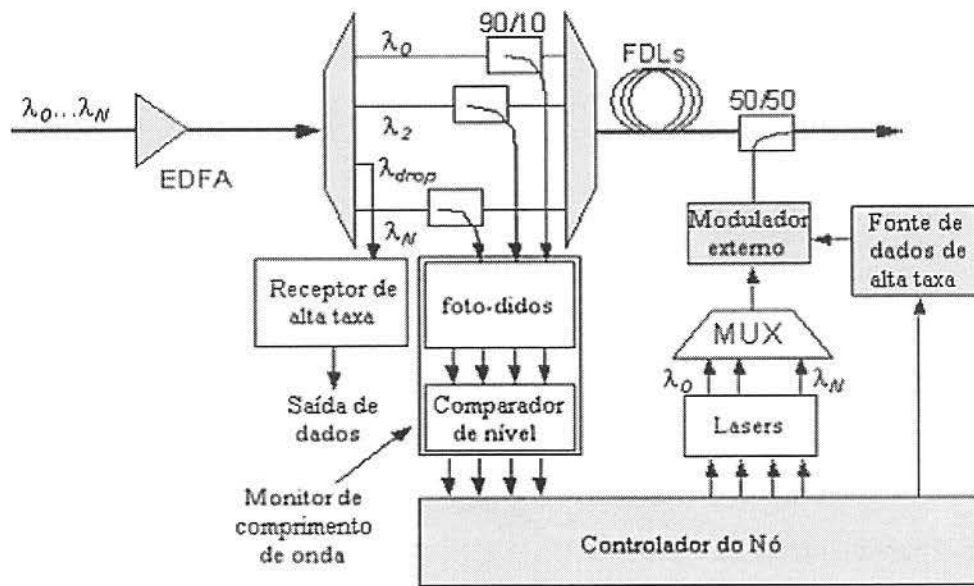


Figura 2-9 Nó da rede Ringo

2.7. Arquitetura FLAMINGO

A arquitetura FLAMINGO [14] tem a configuração $FTxC-FRxC$, onde C é o número de canais de dados. Porém os nós não necessariamente precisam ter a capacidade de transmitir em todos os canais. De fato a rede FLAMINGO suporta a configuração mínima FTx^1-FRxC , onde um nó possui apenas um canal para a transmissão, podendo ser este canal compartilhado entre outros nós.

Os canais são divididos em *slots* de tamanho fixo, sendo os *slots* de todos os canais sincronizados de forma a chegarem juntos em um nó. Um dos canais é utilizado exclusivamente para a transmissão da informação de controle. O *slot* do canal de controle, também conhecido como *slot* de controle, envia um pacote à frente dos seus pacotes correspondentes. Um *slot* de controle carrega a informação sobre o acesso de todos os *slots* de dados.

A rede consiste de N nós OADMs. Se os nós seguem a configuração FTx^1-FRxC , então a unidade transmissora é associada a um canal de transmissão. Se N for menor ou igual a C , cada nó é associado a um canal que será de uso exclusivo deste. Se N for maior que C ocorrerá o compartilhamento de um canal entre dois ou mais nós. A figura 2-10 mostra a arquitetura do nó da rede na configuração completa $FTxC-FRxC$.

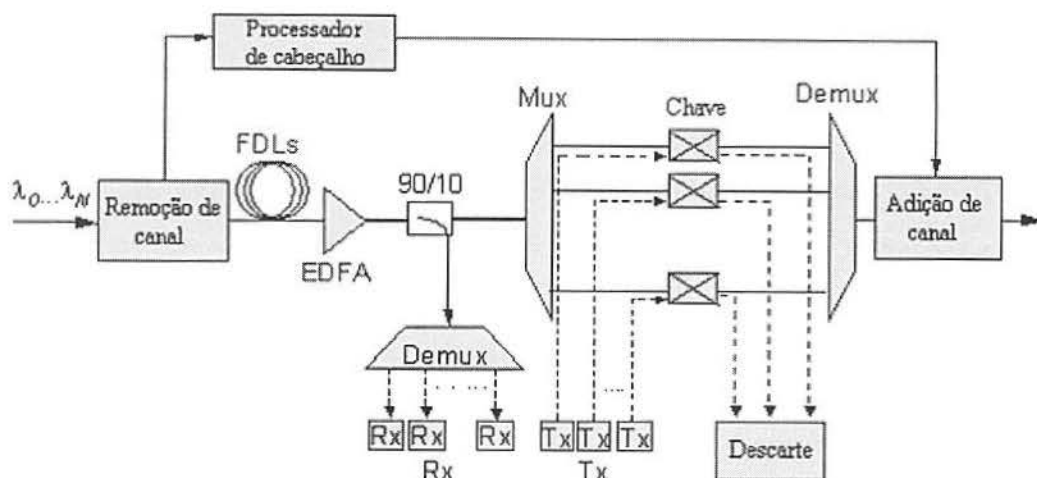


Figura 2-10 Nó da rede Flamingo

Cada nó também é equipado com uma unidade de transmissão e uma unidade de recepção dedicado ao canal de controle. O comprimento de onda do canal de controle é o mesmo para todos os nós da rede. Na entrada do nó o canal de controle é separado dos demais, convertido para o domínio elétrico e enviado para a unidade de processamento de cabeçalho.

Um divisor de sinal é utilizado para captar 10% da potência do sinal de todos os canais e destiná-los para a recepção. Isso acontece independentemente dos dados se destinarem ao nó ou não. Isso facilita a operação do processador de cabeçalho pois, se a informação tem por destino o nó, ela já estará disponível, caso contrário, ela será automaticamente descartada.

Para a inserção são utilizadas chaves ópticas. Desta forma, ao inserir novos dados no *slot* o nó automaticamente retira do anel a antiga informação.

Como a arquitetura do nó permite que a informação do *slot* seja lida sem remoção é possível que a mesma informação seja distribuída para vários receptores.

2.8. Discussão

As arquiteturas de redes ópticas em anel descritas neste capítulo seguem o mesmo conceito: transportar cabeçalhos e carga útil de forma independente a fim de separá-los posteriormente. O cabeçalho é convertido para o domínio elétrico para a realização do processamento de endereço enquanto a carga útil é atrasada em um FDL. Baseado no

processamento do cabeçalho o nó configura o seu OADM para receber o pacote ou encaminhá-lo para o próximo nó.

As diferenças entre as diversas arquiteturas concentram-se na implementação do mecanismo de sinalização, no compartilhamento de canais e nas diversas tecnologias de componentes ópticos utilizadas. Tais diferenças influenciam não apenas na funcionalidade mas também na desempenho, escalabilidade, complexidade e custo da rede.

Tomando por exemplo MAWSON e RINGO. Nestas redes um nó pode remover apenas um canal que é fixamente associado a este. Isto favorece a simplicidade e o custo pois facilita as operações de *drop* nos nós e o endereçamento na rede, porém prejudica o desempenho, por exemplo, no caso de tráfego *multicast* onde deve ocorrer a replicação dos dados em diversos canais, o que gera um grande desperdício de banda.

O mesmo problema ocorre na arquitetura BORN. Apesar do nó desta rede ter a capacidade de transmitir em todos os canais disponíveis não é possível realizar a operação de remoção e encaminhamento de um pacote simultaneamente. Novamente um tráfego com múltiplos destinatários deve ser convertido em vários tráfegos, um por destinatário.

Por outro lado a rede FLAMINGO possibilita operações de múltiplos destinatários, mas requer um multiplexador óptico adicional além de uma chave óptica para cada canal, o que encarece a solução.

Outro exemplo de rede de mais alta complexidade é a HORNET, que possibilita operações de múltiplos saltos, onde um pacote pode ser transmitido por diversos canais até a chegada no nó destino. Esta funcionalidade diminui a probabilidade de contenção na transmissão, porém dificulta a implementação do plano de controle.

Assim o grande desafio das redes de comutação óptica está em balancear os fatores complexidade, desempenho e custo, a fim de se obter uma rede que apresente vantagens significativas em relação às redes atuais, porém que seja passível de implementação e com custo baixo o suficiente para justificar a sua implantação. Como será visto no próximo capítulo a RCPO é um passo nesta direção.

Capítulo 3

Arquitetura da RCPO

Este capítulo descreve as diretrizes, os serviços e a especificação arquitetônica da RCPO – Rede de Chaveamento de Pacotes Ópticos. Esta é uma arquitetura de rede óptica inovadora pois combina as técnicas de comutação de rajadas e de comutação de comprimentos de onda, apesar de ter o seu foco no transporte de pacotes ópticos.

Além disto, o desenvolvimento da RCPO buscou uma solução viável tecnicamente, de baixa complexidade de operação e passível de ser implementada mesmo com as tecnologias atuais, como mostrado no capítulo 5.

Inicialmente apresentam-se as diretrizes que nortearam a concepção da rede. Em seguida é descrita a arquitetura da rede em conjunto com as suas funcionalidades, o serviço de transporte, o serviço de proteção e o controle de acesso ao meio. Também é mostrada como ocorre a interconexão entre os nós da RCPO e como os clientes são ligados à rede.

No final é realizada uma discussão que apresenta uma visão crítica da arquitetura da RCPO, apontando as suas principais funcionalidades em função da complexidade e do desempenho.

3.1. Diretrizes

Como qualquer arquitetura de rede, a arquitetura da RCPO deve seguir as diretrizes genéricas definidas na Tabela 3-1.

Requisito	Objetivos
Desempenho	Atender à demanda de tráfego atual com baixa latência, independentemente da característica e padrão de distribuição do tráfego
Multiserviço	Suportar múltiplos serviços para satisfazer os requisitos das várias aplicações
Flexibilidade	Suportar novos serviços sem mudanças na arquitetura
Expansividade	Crescer em cobertura, em número de nós e em taxa de transmissão sem mudanças na arquitetura
Simplicidade	Facilitar e baratear o projeto, a manutenção, a operação e a evolução da rede
Robustez	Recuperar-se rapidamente de falhas variadas
Eficiência	Usar os recursos da rede de forma eficiente, independentemente da característica e padrão de distribuição do tráfego
Segurança	Dificultar o acesso às informações em trânsito e facilitar a inclusão de mecanismos de criptografia
Custo	Ter baixo custo para facilitar aceitação

Tabela 3-1 Requisitos das redes de nova geração

Embora de grande importância em qualquer arquitetura, o baixo custo e a segurança não estão entre as prioridades mais altas deste trabalho. Com relação ao custo a RCPO, sendo uma rede óptica de comutação de pacotes ópticos, utiliza componentes que hoje apresentam um alto valor. Espera-se que ao longo dos próximos anos o custo de muitos dos componentes ópticos utilizados caia significativamente.

Por se tratar de um protótipo de pacotes ópticos a segurança da informação não é foco do trabalho, entretanto nada impede que novas versões desta arquitetura incluam métodos de proteção como criptografia e autenticação.

Para satisfazer aos demais requisitos da tabela 3-1 a arquitetura deve:

- ser aberta e modular visando flexibilidade, expansão e evolução;
- se basear na tecnologia WDM visando capacidade, expansão gradativa e suporte a multiserviços;

- usar comunicação do tipo *single-hop* (sem qualquer conversão entre os domínios óptico e elétrico durante o transporte), visando desempenho, segurança, transparências de protocolo de transmissão, de taxa de transmissão e de formato de sinal.
- suportar comunicação do tipo um-para-muitos (*multicast*) e um para todos (*broadcast*) no domínio óptico, visando eficiência e desempenho tanto quantitativo quanto qualitativo;
- ter controle distribuído visando robustez.

A seguir são descritas as especificações arquitetônicas da RCPO que buscam atender as diretrizes aqui colocadas.

3.2. Especificação geral da arquitetura

A RCPO é uma arquitetura de rede em anel que combina técnicas de montagem de rajadas com comutação óptica de pacotes e comutação de comprimento de ondas. Esta arquitetura é auto denominada como *BlasteRing* (*Burst and Lambda Services for a Transparent and Ethernet services Ring*). Ela apóia-se em tecnologias ópticas de rede disponíveis comercialmente. A figura 3-1 ilustra a visão arquitetônica da rede.

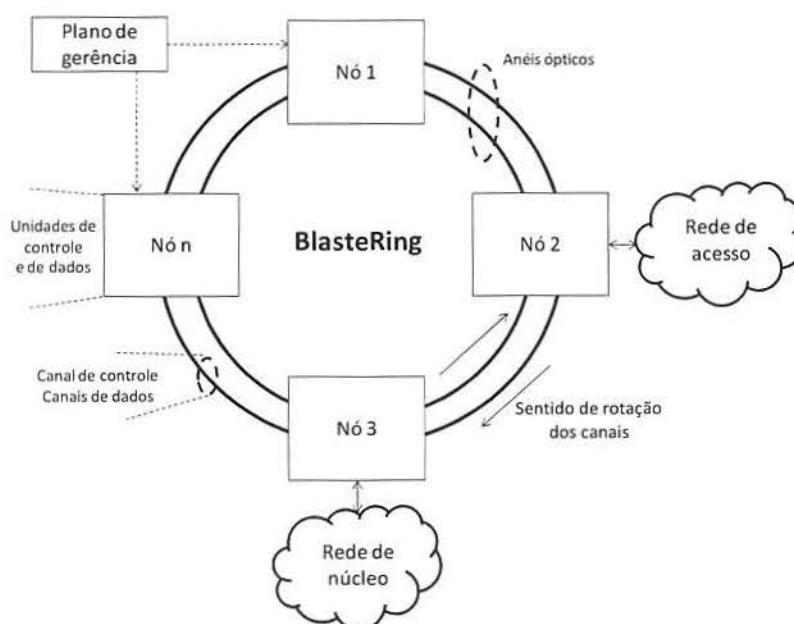


Figura 3-1 Arquitetura macro da RCPO

A rede consiste de um anel MOPS com acesso por *slots* de tamanho fixo, onde um *slot* pode levar apenas um pacote óptico por vez. O pacote óptico possui tamanho variável, porém é necessário respeitar o tamanho máximo dado pelo tamanho do *slot*. Uma vez transmitido o pacote óptico é convertido novamente para o domínio elétrico no seu nó de destino, caracterizando a RCPO como uma rede óptica *single-hop*. A rede também suporta transmissão *multicast* e *broadcast* realizados no domínio óptico. O modelo de transmissão e recepção é FTx^C - FRx^C , onde C é o número de canais de dados. São utilizados $F \geq 1$ fibra(s) óptica(s) de tamanho não superior a 100km dada a sua aplicação de rede metropolitana. Cada fibra óptica contém $L \geq 2$ comprimentos de onda unidirecionais, sendo $L-1$ comprimento(s) de onda (na faixa de 1550 nanômetros) para o transporte de dados e 1 comprimento de onda (na faixa de 1310 nanômetros ou também na faixa de 1550 nanômetros) para o transporte das informações do plano de controle, este comprimento de onda transporta os cabeçalhos que controlam o acesso aos demais canais na fibra através da divisão temporal por *slots*.

Para propiciar uma melhor distribuição do tráfego na rede as fibras ópticas de número ímpar devem propagar seus canais em um sentido de transmissão e as fibras de número par devem propagar seus canais no sentido contrário.

Um nó de rede não necessariamente faz acesso a todos os comprimentos de onda de dados de um anel. Por motivos de simplicidade na operação e custo no desenvolvimento são utilizadas apenas unidades de transmissão e recepção fixas.

Um dos nós é eleito como responsável pela inserção dos cabeçalhos na fase de inicialização da rede. Como dito os cabeçalhos contêm as informações de disponibilidade dos *slots* e são projetados de forma a buscar a justiça no acesso dos recursos da rede entre todos os nós. Como em um *slot* podem existir diversos canais de dados o nó da rede pode realizar a transmissão de mais de um pacote óptico mediante o recebimento de um cabeçalho caso exista mais de um canal de dados livres neste *slot* de tempo. O item 3.5 descreve com mais detalhes a questão do uso dos cabeçalhos e o controle de acesso ao meio.

A rede fornece aos seus clientes três tipos de classes de serviço para o transporte de dados: de melhor esforço, reservada e transparente. As duas primeiras classes implementam diferenciação de qualidade de serviço além de possuírem serviço de proteção ativa contra

falhas de múltiplas conexões e múltiplo nós. Os serviços de transporte e proteção são detalhados nas seções 3.3 e 3.4 respectivamente.

A RCPO opera com três planos funcionais bem definidos: planos de gerência, de controle e de transporte de dados. O plano de gerência é centralizado e responsável pela coordenação e proteção da rede como um todo; ele atua diretamente no plano de controle podendo utilizar o plano de transporte de dados para isto ou ainda uma rede de dados secundária dedicada a este fim. Este tipo de abordagem é comum em muitas operadoras de telecomunicações que nomeiam esta rede de apoio como DCN (Data Communication Network).

O plano de controle é distribuído e atua diretamente nos elementos do plano de transporte de dados a fim de realizar a sua função principal: o controle de acesso ao meio. Além desta função o plano de controle, em conjunto com o plano de gerência, também é responsável pelo serviço de proteção da rede.

O plano de transporte de dados é distribuído, sendo responsável pelo serviço de transporte tanto no domínio elétrico quanto óptico. Através dele os clientes enviam os seus dados, que serão classificados na entrada do nó e enviados para filas de transmissão de acordo com os seguintes critérios: o nó que se pretende alcançar, a classe de serviço e a qualidade de serviço do tráfego. Após a classificação os dados são agregados em rajadas que serão traduzidas em pacotes ópticos durante a transmissão. O objetivo da agregação é otimizar a utilização do *slot* através de um pacote óptico que se aproxime o máximo do tamanho do *slot*. Em contrapartida a agregação aumenta o tempo médio de espera para transmissão dos pacotes, o que não é suportado por muitos tipos de demanda de tráfego. Assim, o tempo de montagem de uma rajada deve ser parametrizado de acordo com as necessidades dos clientes, como será visto na próxima seção.

Cada nó da rede é constituído de uma unidade do plano transporte de dados e uma unidade do plano de controle. A figura 3-2 ilustra a inter-relação entre os planos funcionais.

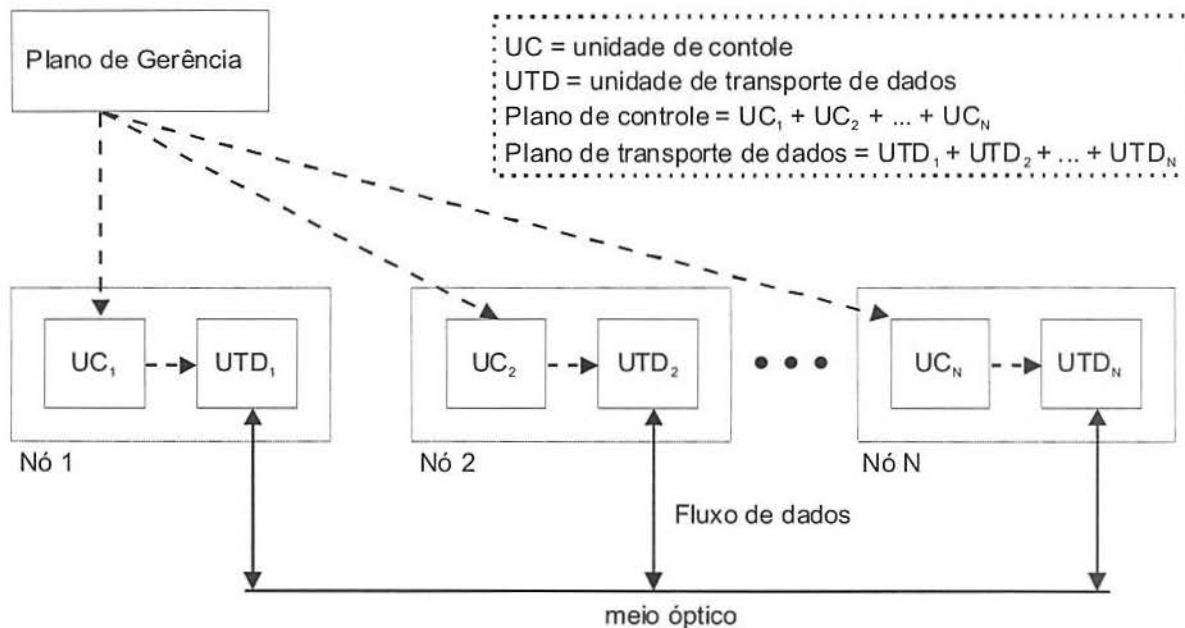


Figura 3-2 Inter-relacionamento entre os planos funcionais

Por fim, a comunicação entre os clientes e a rede ocorre da seguinte forma: caso deseje transmitir dados o cliente deve realizar uma solicitação para a rede indicando as necessidades e características do tráfego. Para tal pode ser utilizada uma sinalização (exemplo: GMPLS - *Generalized Multi-Protocol Label Switching* [15]) que solicita ao plano de gerência a configuração da rede para atender a sua demanda, como mostrado na figura 3-3.

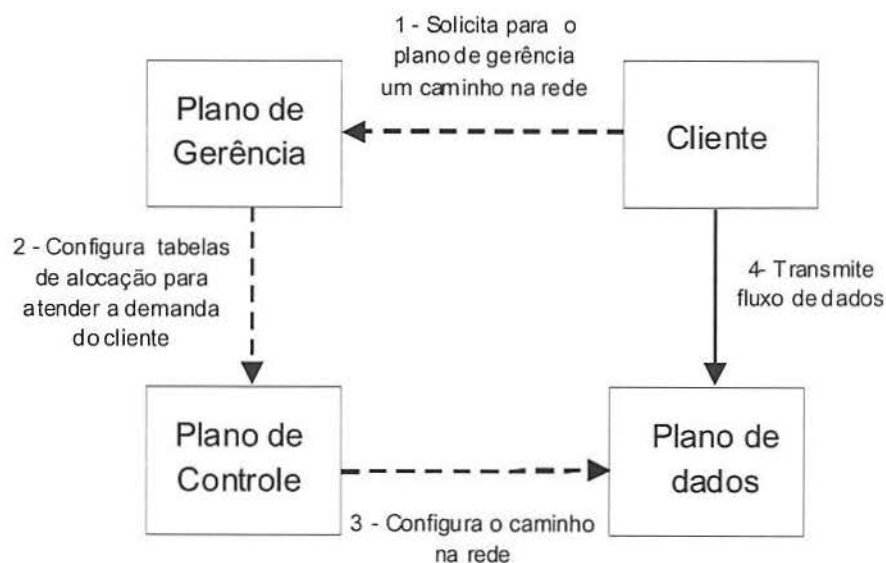


Figura 3-3 Solicitação do cliente e configuração da rede

3.3. Serviço de transporte

Seguindo o modelo das redes de dados de nova geração, a RCPO deve prover diferentes tratamentos no transporte de dados, de acordo com a classe e qualidade de serviço (QoS – *quality of service*). Uma importante característica da arquitetura da RCPO é a grande flexibilidade e facilidade que ela apresenta com relação à diferenciação e o tratamento do tráfego, permitindo a implementação de três diferentes tipos de classe de serviço (CoS – *class of service*): melhor esforço, reservada e transparente.

O serviço de transporte de melhor esforço consiste no compartilhamento de um dado número de *slots* entre os nós da rede tendo todos os nós o mesmo direito de acesso a estes *slots*, sendo assim considerado justo para todo o tipo de demanda. Esta classe de serviço não garante banda pré-definida para nenhum fluxo de dados e nem tempo de latência máxima. Entretanto, é possível realizar a priorização do tráfego através de diferentes tipos de qualidade de serviço. Para tal, as filas de transmissão são divididas em sub-filas, tendo cada uma delas um diferente valor de QoS. A diferenciação no tratamento das sub-filas ocorre através do ajuste de tempo de montagem da rajada e na prioridade de transmissão, controladas por parâmetros definidos pelo plano de controle da rede. A figura 3-4 ilustra o conceito. Neste exemplo existem quatro sub-filas de transmissão para cada fila de transmissão, sendo associada uma fila de transmissão para cada nó destino. A sub-fila com valor de QoS mais prioritário possui tempo de montagem de rajada mais baixo que as demais. Desta forma, mesmo que os dados recebidos não sejam suficientes para preencher 100% da capacidade do *slot*, é solicitada a transmissão assim que o tempo de montagem for vencido. Isso obviamente gera uma ineficiência na utilização do canal, porém diminuiu o tempo de latência de forma a tentar atingir os requisitos que um dado serviço demanda.

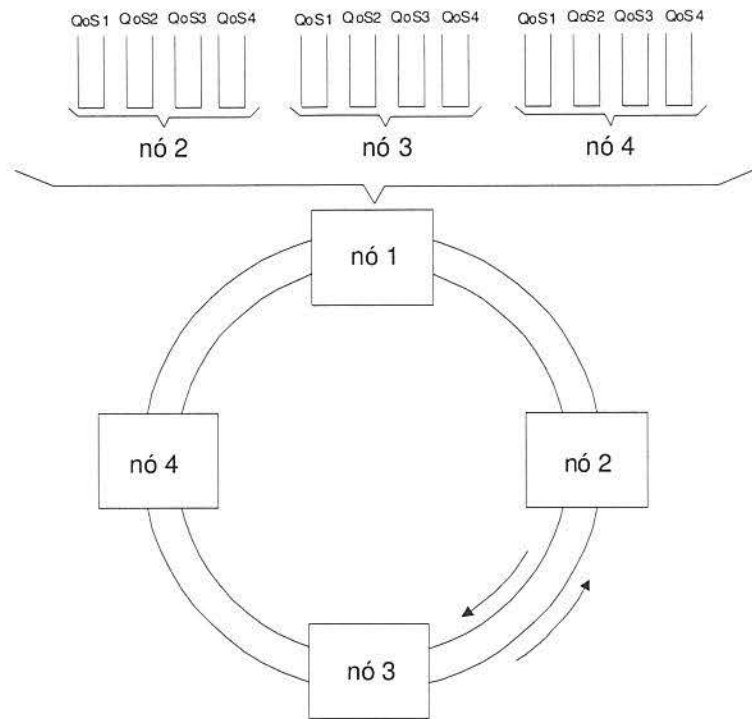


Figura 3-4 Classe de serviço de melhor esforço

Além do tempo de montagem mais baixa uma rajada mais prioritária deve ter preferência na transmissão quando outras rajadas concorrem por um mesmo *slot*.

Esta classe de serviço é definida como de melhor esforço pois não há como garantir o tempo máximo de espera pela transmissão, nem banda mínima, sendo estes valores diretamente relacionados com o tráfego total carregado pela rede versus a sua capacidade.

A princípio, havendo apenas a classe de serviço de melhor esforço, todos os nós poderiam utilizar qualquer *slot* disponível para as transmissões, independentemente do nó de destino. Entretanto esta abordagem trás dois problemas. O primeiro problema é a reordenação de pacotes, no exemplo caso do nó 1 desejar transmitir dados para o nó 4 ele poderá utilizar *slots* tanto do canal 1 quando do canal 2. Porém neste caso o canal 2 é o caminho mais curto entre os nós 1 e 4, enquanto o canal 1 é caminho mais longo. Assim é possível que um pacote transmitido pelo canal 1 chegue depois do pacote transmitido pelo canal 2 mesmo que este tenha sido transmitido primeiro. A reordenação é um grave problema principalmente para tráfegos do tipo TCP (*Transmission Control Protocol*) que são orientados a conexão e devem ser evitados. Assim através da tabela de alocação de recurso é possível forçar todo o tráfego do nó 1 para o nó 2 para os *slots* do canal 2.

O segundo problema é a má utilização dos *slots* na rede. Utilizando o mesmo exemplo é importante notar que caso o nó 1 utilizar *slots* do canal 1 para transmitir dados para o nó 4 estes *slots* ficarão ocupados por toda a rotação. No caso da utilização do canal 2 os *slots* ficam imediatamente livres assim que passarem pelo nó 4, podendo ser utilizados para os demais tráfegos. Igualmente este problema é resolvido através da manipulação da tabela de alocações.

A segunda classe de serviço suportada pela RCPO é a reservada. Com esta classe é possível garantir o tempo de acesso e a banda de forma pré-determinada. A implementação ocorre, como o próprio nome indica, com a reserva de um certo número de *slots* para um fluxo. Esta reserva ocorre com a utilização de tabelas na unidade de plano de controle. A Figura 3-5 ilustra um exemplo de uma rede de quatro nós e um canal de dados circulando em um anel de fibra, sendo o canal dividido em quatro *slots* de tempo.

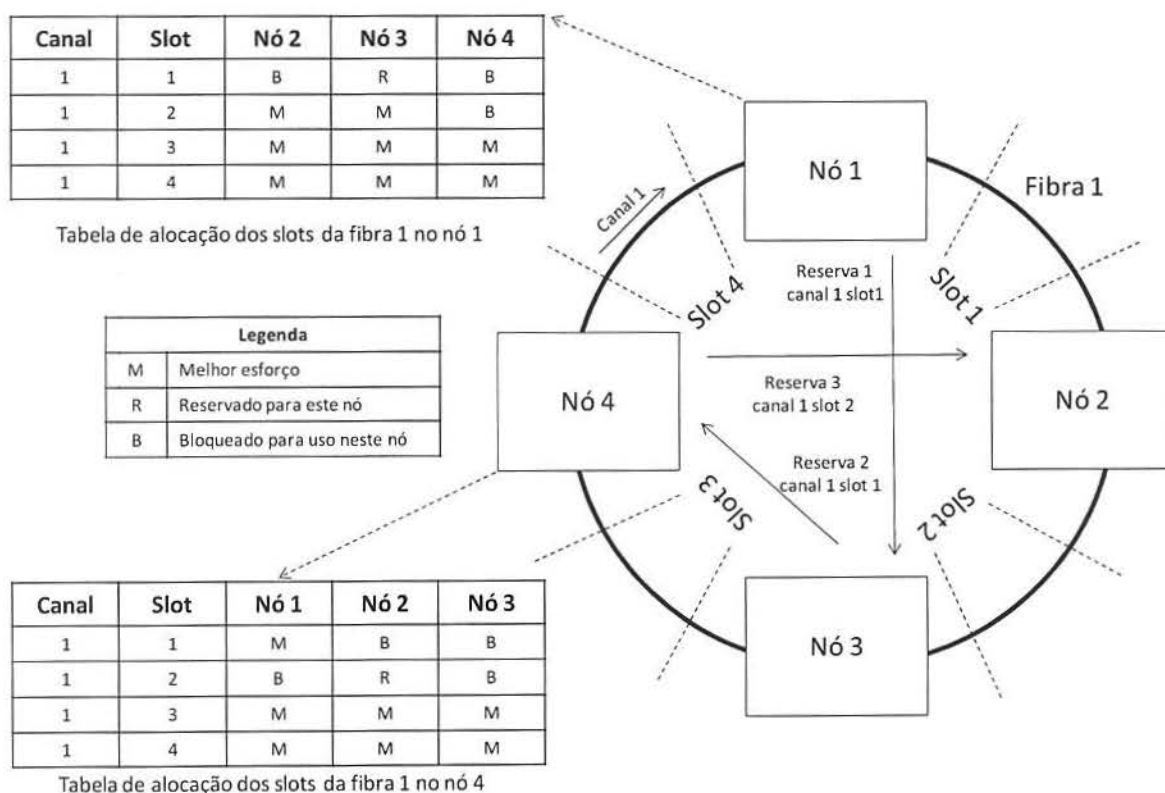


Figura 3-5 Classe de serviço reservada

A tabela de alocação dos *slots* é implementada na unidade de plano de controle dos nós e controla o uso dos *slots* pelo nó, assim mesmo que um canal em um *slot* esteja disponível o nó necessita antes consultar esta tabela para verificar se pode ou não utilizá-lo. A tabela de cada nó é controlada pelo plano de gerência, que possuindo a visão completa da rede e tendo todas as requisições dos clientes pode realizar este controle de forma a buscar a melhor configuração possível.

No exemplo existem três reservas, a primeira do nó 1 ao nó 3 utilizando o *slot* 1, a segunda do nó 3 ao nó 4, também utilizando o *slot* 1, e a terceira, do nó 4 ao nó 2 utilizando o *slot* 2. As reservas causam bloqueio do canal no *slot* para as demais transmissões, seja do próprio nó ou dos demais nós que viriam a utilizar este recurso nos mesmos trechos em que é usado para a reserva. Entretanto não invalida a sua utilização em trechos diferentes. No exemplo, o *slot* 1 do canal 1 é utilizado em duas reservas, do nó 1 para o 3, e do 3 para o 4, uma vez que estas reservas não se sobrepõem. Já no caso da terceira reserva, do nó 4 ao nó 2, é necessário utilizar-se outro *slot* uma vez que o *slot* 1 já não poderia atender esta solicitação devido ao conflito entre o trecho do nó 1 ao nó 2.

Através da tabela de alocação dos *slots* também é possível controlar a distribuição do tráfego na rede, de forma a otimizar a utilização canais, principalmente quando existe recursos redundantes na rede. No exemplo da figura 3-6 existem dois anéis de fibra cada um com um canal, sendo os canais contra-propagantes.

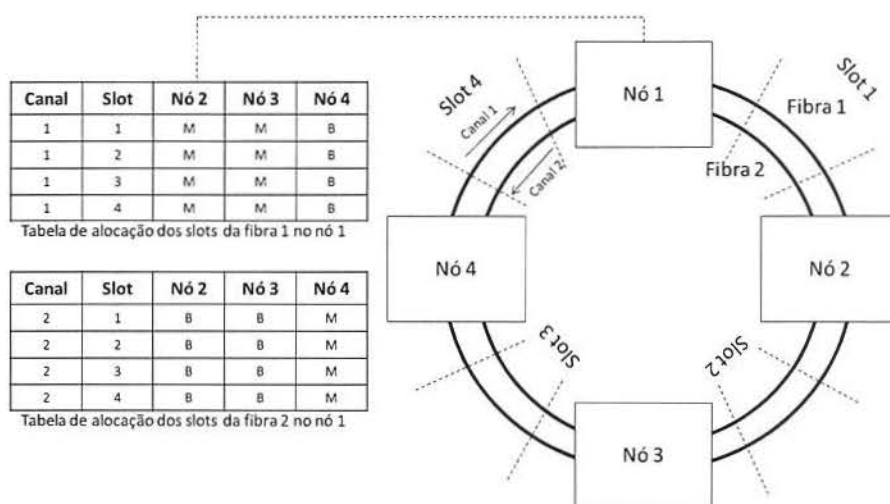


Figura 3-6 Utilização da tabela de alocação de *slots* para a distribuição do tráfego

Nenhuma das classes de serviço descritas, de melhor esforço e reservada, garantem a entrega dos dados, pois, como o controle de acesso ao meio evita qualquer tipo de colisão e por se tratar de meio de transmissão óptico (com baixa probabilidade de erro), supõe-se que ocorrerão poucas falhas na transmissão, não sendo eficiente a utilização de qualquer mecanismo de retransmissão na camada 2, que geraria um *overhead* desnecessário. Assim, a confirmação do recebimento dos dados fica a encargo das camadas superiores.

A terceira classe de serviço suportada pela RCPO é a transparente. Sua realização é possível dada a uma peculiaridade da rede: toda a sinalização do plano controle ocorre através de um único canal que é dedicado para este fim. Desta forma, quando não ocorre transmissão em nenhum *slot* de um canal em um segmento da rede não existe nenhum sinal neste comprimento de onda no segmento, permitindo o transporte de qualquer tipo de informação e em qualquer taxa, alcançando a mesma transparência dos sistemas WDM.

Esta característica, típica das redes de chaveamento de comprimento de onda, também é utilizada na RCPO, realizando neste caso a comutação do comprimento de onda baseado na solicitação do cliente sem qualquer necessidade de analisar os campos de endereçamento dos quadros transmitidos.

A figura 3-7 ilustra um exemplo do serviço de transporte de dados com classe de serviço transparente. Neste caso os dados devem chegar de um cliente já convertidos no domínio óptico, para assim serem comutados a qualquer outro nó da rede. Desta maneira o cliente 1 pode, por exemplo, transmitir seus dados diretamente para o cliente 2 ou para o cliente 3 ou ainda para ambos, caso se trate de um fluxo *multicast*. A transmissão ocorre sem qualquer alteração dos dados, da mesma forma como se tivessem sido transmitidos em uma rede óptica passiva.

Como a transmissão do cliente 1 para os clientes 2 e 3 utiliza os seguimentos de rede entre o nó 1 e o nó 2, e entre o nó 2 e o nó 3, e ainda supondo que a mesma ocorra utilizando o comprimento de onda do canal 1, as tabelas que descrevem a alocação dos *slots* nos nós 1 e 2 devem marcar com B todos os *slots* do canal 1 para todos os possíveis destinos. Para o nó 3 os *slots* devem ser bloqueados apenas para o destino no nó 1 e 2, porém a transmissão de dados do nó 3 para o nó 4 pode ocorrer livremente e em todos os *slots* deste canal, criando assim a possibilidade de reuso do canal, mesmo quando este está reservado em parte do anel.

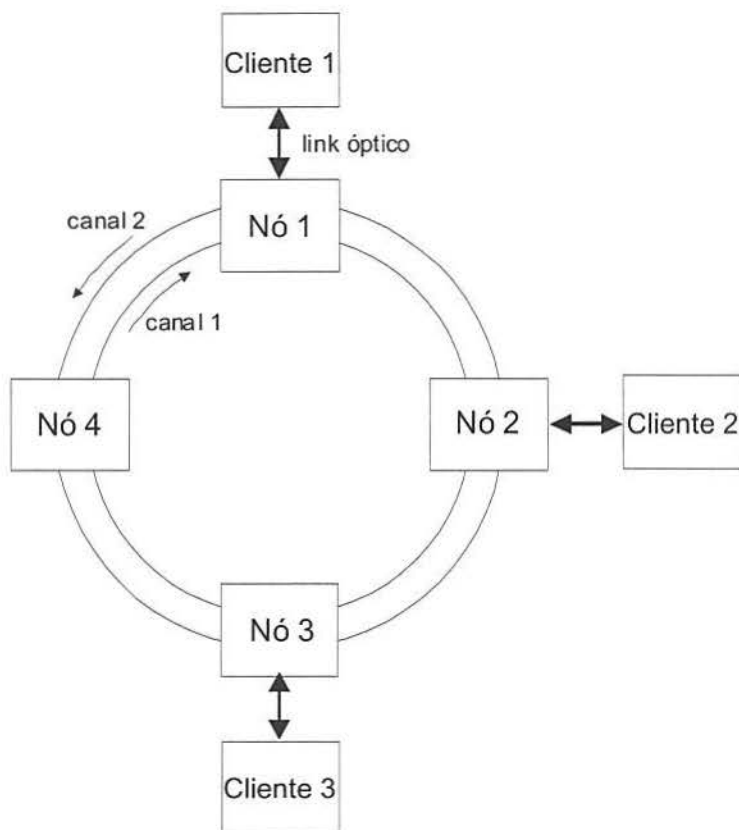


Figura 3-7 Classe de serviço transparente

A tabela 3-2 apresenta um resumo das três classes de serviço suportadas pela RCPO, comparando as suas características, os tráfegos e as aplicações alvo.

Visando a necessidade do aumento da largura de banda entre dois nós da rede é possível associar, para as classes de serviço de melhor esforço e reservada, mais de uma fila de transmissão para um mesmo nó destino, criando-se assim um grupo de fila.

Neste caso, quanto mais filas forem associadas ao grupo maior será a vazão para este destino. Esta associação de filas a um grupo pode ocorrer de forma incremental, aumentando a banda de transmissão para o nó de forma escalável.

A utilização de grupos de filas e a forma de distribuição de tráfego entre as filas de um grupo fica a critério do projeto de implementação, uma vez que a sua existência não é fundamental para a operação da rede.

Cós	Características	Tráfegos alvo	Aplicações alvo
De melhor esforço	Comprimento de onda e acesso compartilhado, com baixa prioridade na transmissão e suscetível a latência variável	Rajadas, com baixa tolerância a perdas, mas tolerante a atrasos	e-mail, FTP, <i>web</i> , <i>peer-to-peer</i>
Reservada	Comprimento de onda compartilhado mas com acesso garantido	Rajada ou contínuo, com baixa tolerância a perdas e atrasos	Vídeo, áudio, voz, multimídia, teleconferência, HDTV, computação destituída, controle, gerência
Transparente	Segmento de comprimento de onda de uso exclusivo	A encargo do protocolo de transporte cliente	ATM, SDH, Frame relay

Tabela 3-2 Classes de serviço de transporte da rede

3.4. Controle de acesso ao meio

Vários protocolos MAC (*medium access control*) foram estudados para esta arquitetura de rede (exemplo: [16] e [17]), sendo alguns de baixo desempenho e de fácil implementação e outros de alto desempenho e de difícil implementação.

O protocolo adotado para a realização do controle de acesso ao meio da classe de serviço de melhor esforço situa-se entre os dois extremos e baseia-se na combinação de versões modificadas dos protocolos PAT (*Packet Aggregate Transport*) [18] e SAT [19].

O protocolo PAT responde pelo controle de acesso ao meio propriamente dito independentemente do que seja um acesso justo ou não. Seu foco é vazão, eficiência e escalabilidade. Este protocolo utiliza a mesma abordagem da RCPO, onde a capacidade total do anel é dividida em *slots* de tempo com duração fixa, podendo cada *slot* transportar um único pacote óptico que se traduz em um ou mais pacotes elétricos. É importante notar que uma arquitetura de rede OPS purista define um pacote óptico como um meio capaz de transportar um único pacote elétrico. Entretanto, para atender este requisito seriam necessários pacotes ópticos de duração na ordem de algumas dezenas de nanosegundos.

Como isto dificulta muito a implementação, dada à necessidade de se processar eletricamente o cabeçalho do pacote óptico, o protocolo PAT propõe a abordagem de agregação de pacotes elétricos. Entretanto, com o PAT original a solicitação da transmissão do pacote eletrônico ocorre tão cedo este chegue a uma fila vazia, agregando os demais pacotes à fila enquanto a transmissão não seja liberada. Porém, mesmo que não cheguem outros pacotes elétricos, ou mesmo que alguns outros poucos cheguem mas não atinjam a capacidade de transmissão do pacote óptico, este é alocado da mesma forma, gerando assim uma sub-utilização do recurso.

Por esse motivo a modificação do protocolo PAT proposto para a RCPO, chamado de PAT+, utiliza o conceito de montagem de rajada das redes OBS, visando uma melhor utilização dos *slots*. Nesta abordagem uma solicitação de transmissão ocorre dada a necessidade de se transmitir uma rajada já formada. Isso resulta em melhor desempenho na transmissão e na possibilidade de se utilizar pacotes óticos de maior duração sem impactar demasiadamente na eficiência da utilização dos canais, o que simplifica e diminui o custo da implementação.

Entretanto o PAT original também é aplicado na RCPO para a classe de serviço reservada, neste caso como o canal no *slot* já é reservado para apenas um nó transmissor não é necessário realizar a agregação da rajada a fim de se otimizar o uso da rede. Desta forma tão logo o canal fique disponível para transmissão dos dados os pacotes devem ser transmitidos independentemente do seu número e tamanho.

Para evitar que alguns nós da rede tenham mais oportunidades de acesso que outros é utilizada uma versão modificada do protocolo SAT. No protocolo original um sinal de controle denominado SAT (de *satisfied*), implementado por um bit no cabeçalho dos *slots*, circula pelos nós do anel atribuindo quotas de transmissão (em número de pacotes óticos) pré-determinadas. No caso do nó receber o campo SAT com valor 1, possuir algum dado para transmitir, e ainda não tiver utilizado o seu número predeterminado de *slots*, este irá reter o SAT (atualizando em 0 o campo no cabeçalho que o trouxe) até que não haja mais dados para transmitir ou que tenha sido utilizada a sua cota de *slots*.

A não liberação do SAT impede que os outros nós tenham suas quotas renovadas, fazendo com que em algum momento parem de transmitir, permitindo que o nó que possui o SAT

tenha oportunidade para utilizar sua quota. Uma vez satisfeito o nó libera o SAT atualizando em 1 o bit apropriado do primeiro cabeçalho que aparecer.

O protocolo SAT alterado, denominado de SAT+, segue o mesmo princípio do PAT+ substituindo a abordagem de pacotes por rajadas. Desta maneira um nó é considerado satisfeito mesmo quando existam dados para serem transmitidos, caso estes não tenham resultado no fechamento de uma rajada. Existe apenas um sinal SAT por anel havendo uma instância do protocolo para cada anel.

O protocolo SAT+ pode ser aplicado apenas para o transporte de dados que se enquadram na classe de serviço de melhor esforço, uma vez que não há competição pelo recurso nas demais classes. A figura 3-8 apresenta o fluxograma do protocolo SAT+.

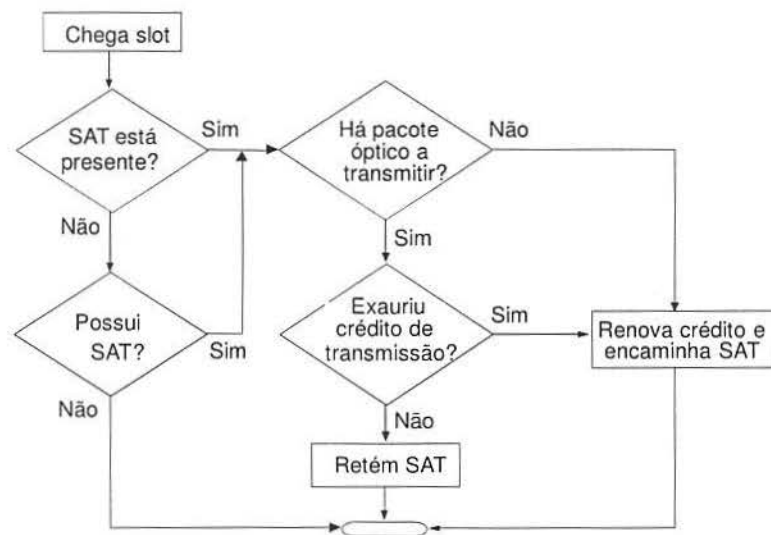


Figura 3-8 Fluxograma do protocolo SAT+

Uma vez descrito como ocorre a montagem da rajada e a transmissão do pacote óptico, é importante definir em maiores detalhes a relação entre os *slots* e o cabeçalho. Na RCPO os *slots* de cada canal são sincronizados de forma a chegarem juntos em um nó e possuem um cabeçalho compartilhado que é convertido para o domínio eletrônico e processado no nó da rede. Através do processamento do cabeçalho são determinadas as ações sobre os *slots*, divididas em quatro casos: o pacote pode estar livre, podendo ser utilizado para a transmissão; o pacote pode ter como destino o próprio nó, sendo realizada assim uma

operação de *drop*; o pacote pode ser endereçado a um nó à frente, sendo realizada uma operação de *forward*; ou ainda o pacote pode conter um endereço de *multicast*, neste caso são realizadas as duas operações anteriores de forma simultânea. A Figura 3-9 ilustra a relação entre cabeçalhos e *slots*. O cabeçalho corrente sempre traz a informação dos *slots* anteriores, isso dá ao nó o tempo de um *slot* para executar o processamento das informações.

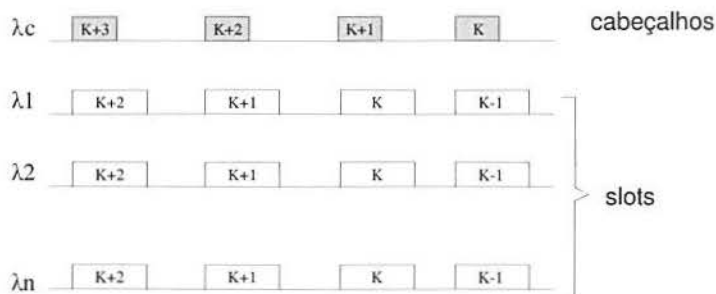


Figura 3-9 Cabeçalhos e *slots*

Por fim a figura 3-10 resume o controle de acesso ao meio da RCPO através do fluxograma do processamento do cabeçalho e da transmissão dos pacotes ópticos no anel para a classe de serviço de melhor esforço, já levando em conta a aplicação em conjunto com o protocolo SAT+ e a utilização da tabela de alocação descrita na seção 3.3.

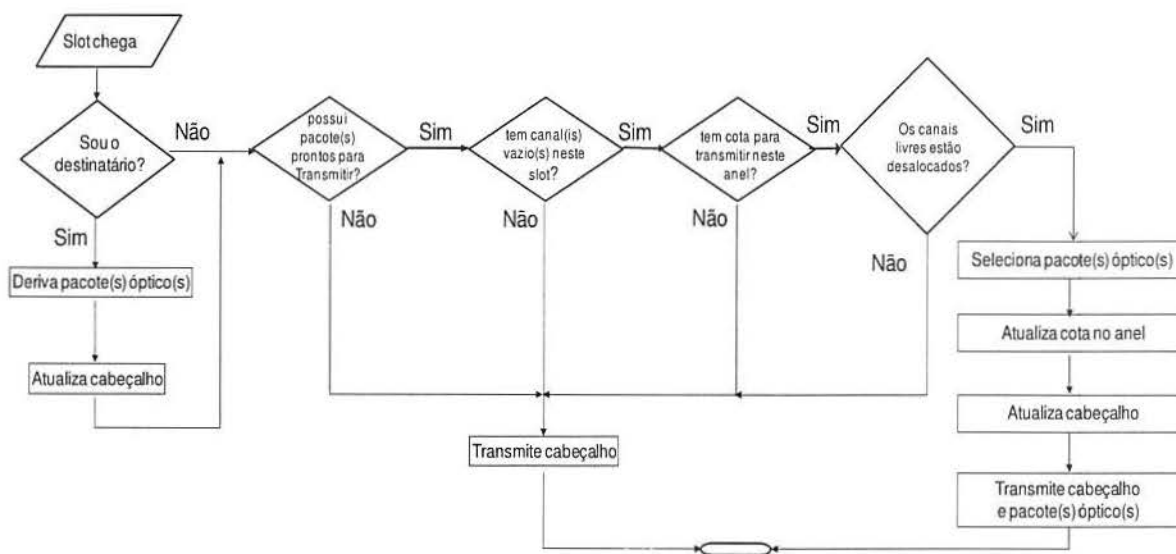


Figura 3-10 Fluxograma do processamento do cabeçalho e a transmissão dos pacotes ópticos para a classe de serviço de melhor esforço

3.4.1. Formato do cabeçalho

A Figura 3-11 apresenta a estrutura básica do cabeçalho utilizada na RCPO.

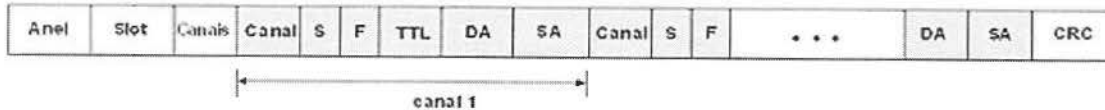


Figura 3-11 Estrutura do cabeçalho

Anel: indica por qual anel o cabeçalho está circulando.

Slot: indica qual é o *slot* de tempo deste canal.

Canais: indica o número de canais descritos no cabeçalho.

Canal: indica qual é o canal que os próximos campos: S, F, TTL, DA e SA se referem.

S (*Status*): campo que indica se o *slot* está vazio (0) ou ocupado (1). Um canal vazio contém os campos S, DA e SA preenchidos com zeros.

F (*Fairness*): campo que indica se o sinal SAT+ está presente (1) ou ausente (0).

TTL (*Time-to-live*): campo que indica o número máximo de *hops* (nós) que a rajada contida no *slot* do canal pode atravessar no anel. Cada nó da rede decrementa este campo toda vez que encaminha o *slot* para o próximo nó. No entanto, se o valor for zero, então o nó descarta a rajada contida no *slot* e marca o *slot* do canal como vazio.

DA (*Destination Address*): campo que indica o destinatário daquele *slot*. Se DA for igual ao endereço do nó que está processando o *slot*, a rajada é recebida e o *slot* é marcado como vazio, podendo ser reusado pelo próprio nó para a transmissão de uma nova rajada ou por nós subseqüentes. Também podem ser usados endereços de *multicast* e *broadcast* uma vez que a rede possui esse suporte, que é realizado no domínio óptico.

SA (*Source Address*): campo que indica o remetente daquele *slot*. Se SA for igual ao endereço do nó que está processando o *slot*, a rajada é descartada e o *slot* é marcado como vazio, podendo ser reusado pelo próprio nó para a transmissão de uma nova rajada ou por nós subseqüentes.

CRC (*Cyclic Redundancy Check*): calculado sobre o conteúdo de todos os outros campos, permite a cada nó determinar se o cabeçalho está íntegro. Ao receber o cabeçalho todo nó calcula um CRC sobre o conteúdo do cabeçalho (excluindo o campo CRC) e o compara com o CRC contido no cabeçalho. Se forem diferentes, a rajada é descartada (se houver

uma) e o *slot* é marcado como vazio. O campo F também é marcado com zero independentemente do seu valor anterior.

Os tamanhos dos campos estão reduzidos visando o processamento do cabeçalho do *slot* e a execução das ações de forma rápida. Estes campos podem ser facilmente aumentados para tamanhos mais comumente adotados na indústria com o uso de um processador de cabeçalho de maior desempenho.

3.4.2. Inserção e sincronização dos cabeçalhos

Para o bom funcionamento da rede os cabeçalhos que circulam pelo canal de controle devem ser síncronos, com intervalo fixo entre eles. O número de cabeçalhos em um anel pode variar de acordo com o tamanho total do anel e o tempo de processamento máximo requerido pelo nó, sendo que todos os nós devem ter o mesmo tempo de processamento.

Não existe a necessidade de haver o mesmo número de *slots* em anéis diferentes pois controle de cada anel deve ser independente. Desta forma o tamanho de cada anel pode variar. O número de cabeçalhos, e por consequência de *slots*, em um anel é calculado da através da descoberta do tempo de rotação do cabeçalho no anel. Para isto um nó é eleito para inserir um pacote que circulará no anel, cronometrando-se o seu tempo. Ao receber este pacote um nó da rede deve retê-lo pelo mesmo tempo fixo que reterá os cabeçalhos durante a operação da rede. Assim, considerando um anel de 100km em uma rede com 10 nós e tempo de retenção de 20us, o tempo de rotação do anel é de 700us segundos dado que a luz leva 500us para atravessar os 100Km de fibra e o tempo de retenção de 10 nós é de 200us.

O menor tempo de *slot* deve ser o tempo que o nó toma para processar o cabeçalho. Desta forma o número máximo de *slots* que o anel deste exemplo poderia ter é 35, dada a divisão de 700 por 20.

Entretanto é importante haver uma margem entre *slots* para que o projeto de implementação seja possível. No caso, adicionando-se uma margem de 20%, o número de *slots* no anel cairá para 28, e o tempo entre *slots* será de 25us. Assim, o nó responsável por calcular o tempo de rotação deve também inserir os cabeçalhos de forma a espaçá-los com 25us.

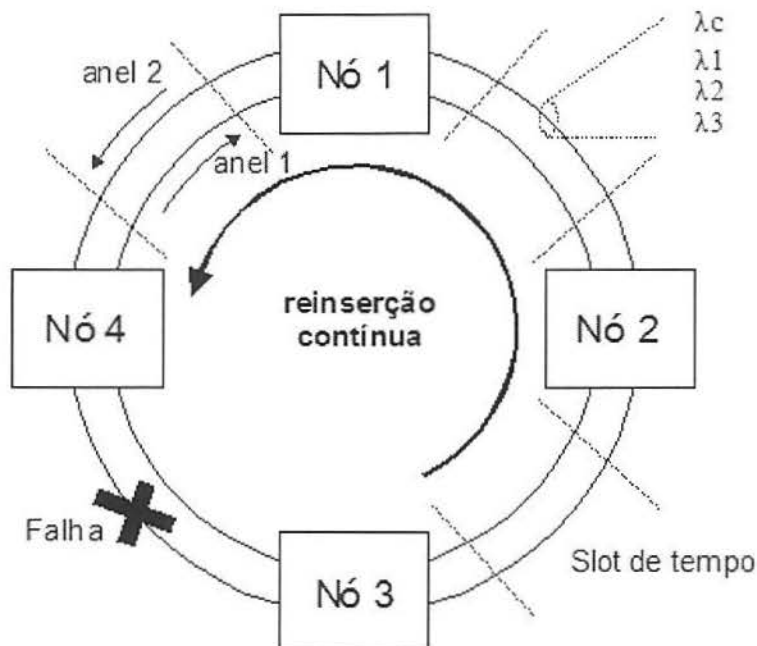


Figura 3-13 Mecanismo de reinserção contínua

A detecção da queda de fibra é baseada na falha da recepção dos cabeçalhos, uma vez que existe um canal dedicado para o transporte destes por anel. No exemplo, quando o nó 3 deixa de receber os cabeçalhos no anel 2 este entrará em modo de reinserção contínua, de forma que os demais não percebam a falha. Assim a comunicação continua a ocorrer mesmo que de forma limitada. Evidentemente que os dados enviados pelos nós 1 e 2 ao nó 3, e todos os dados enviados pelo nó 4 neste anel serão perdidos até que a gerência da rede intervenha como será mostrado no serviço de proteção de caminho, abordado na seção 3.5.1.

O mesmo processo também é válido mesmo para uma segunda falha neste anel, como mostrado na figura 3-14. Agora, com uma nova falha entre o nó 2 e nó 1, o nó 1 também entra no estado de reinserção contínua para o anel 2, reinserindo os cabeçalhos que seguem para nó 4.

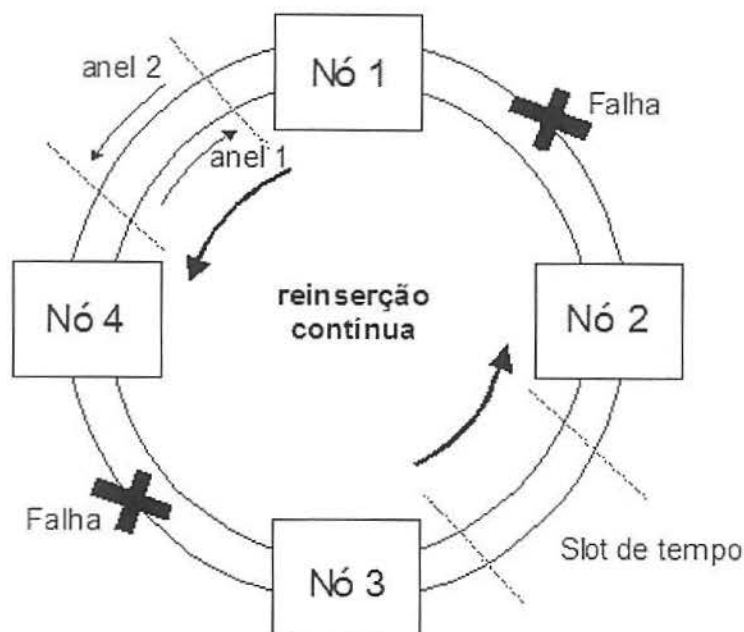


Figura 3-14 Tolerância a múltiplas falhas de fibra

Entretanto, erros na recepção dos cabeçalhos podem ocorrer resultando em sua perda mesmo que não haja nenhum problema nos *links* ópticos. Neste caso é da responsabilidade de qualquer nó da rede reinserir o cabeçalho sem entrar em modo de reinserção contínua. O modo de reinserção contínua é apenas acionado quando o nó deixa de receber na sequência todos os cabeçalhos do anel, o que caracteriza uma queda de fibra.

É importante destacar que não existe a necessidade de sincronia entre os cabeçalhos de anéis diferentes, sendo o controle de sincronização realizado individualmente por anel, igualmente a tolerância a falhas de um anel não depende da existência de um segundo anel na rede.

A tolerância a múltiplas falhas de fibra torna a RCPO também tolerante a falha de nós. Neste caso é importante observar que a queda de um nó pode ser traduzida como uma sequência de quedas de fibra. No exemplo da figura 3-15 o nó 4 sofreu uma falha e não é mais operacional. Neste caso, o nó 3 identificará a falha da conexão entre ele e o nó 4 no anel 2 e o nó 1 a falha da conexão entre ele e o nó 4 no anel 1. Assim, o nó 3 entra em modo de reinserção contínua no anel 2 e o nó 1 no anel 1.

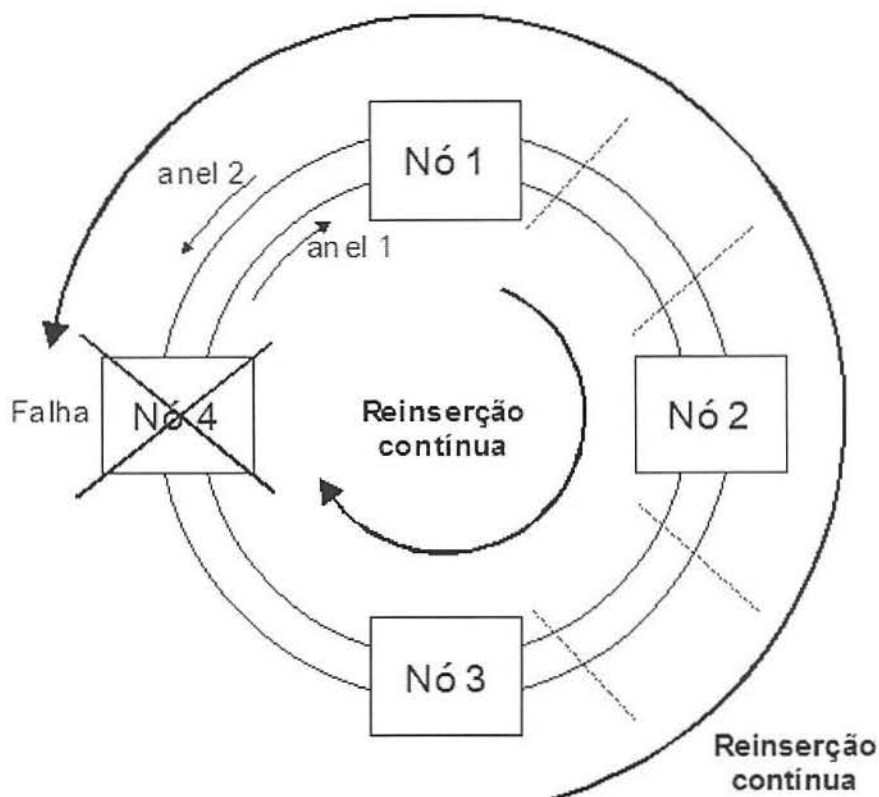


Figura 3-15 Tolerância a falha de nó

3.5.1. Serviço de proteção de caminho

Através da tolerância a falhas é possível prover na RCPO o serviço de proteção de caminho. Este serviço é altamente associado às transmissões ópticas devido a capacidade das redes SDH em realizar a proteção em menos de 50 ms, tornando-se um referencial de qualidade.

Na RCPO para que a proteção de caminho seja possível é fundamental a intervenção do plano de gerência e a utilização da tabela de alocação de *slots* apresentada na seção 3.3.

Para exemplificar o serviço de proteção de caminho a figura 3-16 ilustra uma rede de 4 nós, dois anéis cada um com um canal de dados sendo estes contra-propagantes e com 4 *slots* por canal. A figura também apresenta as tabelas de alocação de *slots* em ambos os anéis para cada nó da rede. As tabelas estão ordenadas de forma a prover a melhor distribuição de tráfego entre os canais disponíveis de forma a buscar a menor distância entre dois nós e de forma que dois canais não podem ser utilizados para o transporte entre dois nós, evitando assim a ocorrência da reordenação de pacotes em um mesmo fluxo de dados.

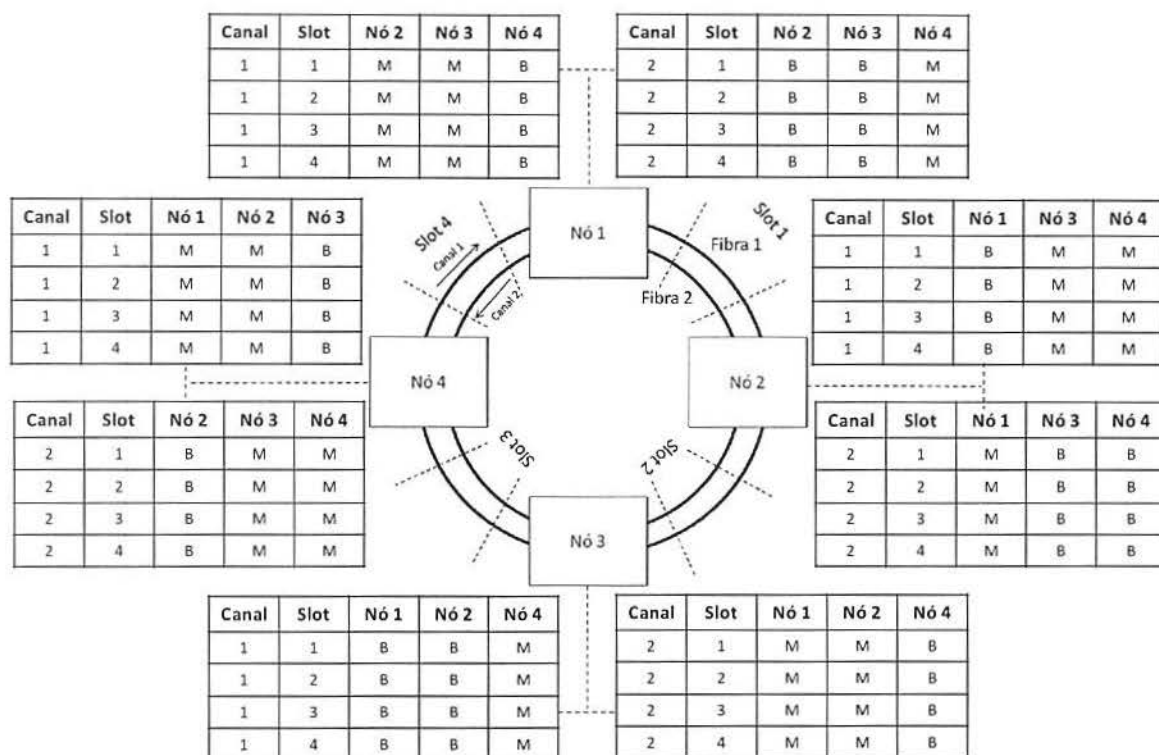


Figura 3-16 Distribuição de tráfego a fim de prover proteção de caminho

Desta forma as mesmas tabelas podem ser reordenadas a fim de evitar que os nós continuem a trafegar dados pelo trecho do anel em falha. Assim, a figura 3-17 mostra o mesmo cenário agora com um evento de falha entre o trecho do anel 1 entre os nós 3 e 4. As células em cinza são aquelas que foram alteradas pela gerência a fim de evitar que os nós continuem a trafegar dados que se perderiam em função da falha.

Assim, para que a gerência possa atuar na rede e realizar a proteção de caminho o nó que entrou no modo de reinserção contínua deve notificá-la imediatamente. Igualmente, ao voltar a receber os cabeçalhos este deve novamente avisar a gerência que a falha foi corrigida para que as tabelas sejam ajustadas para a operação normal.

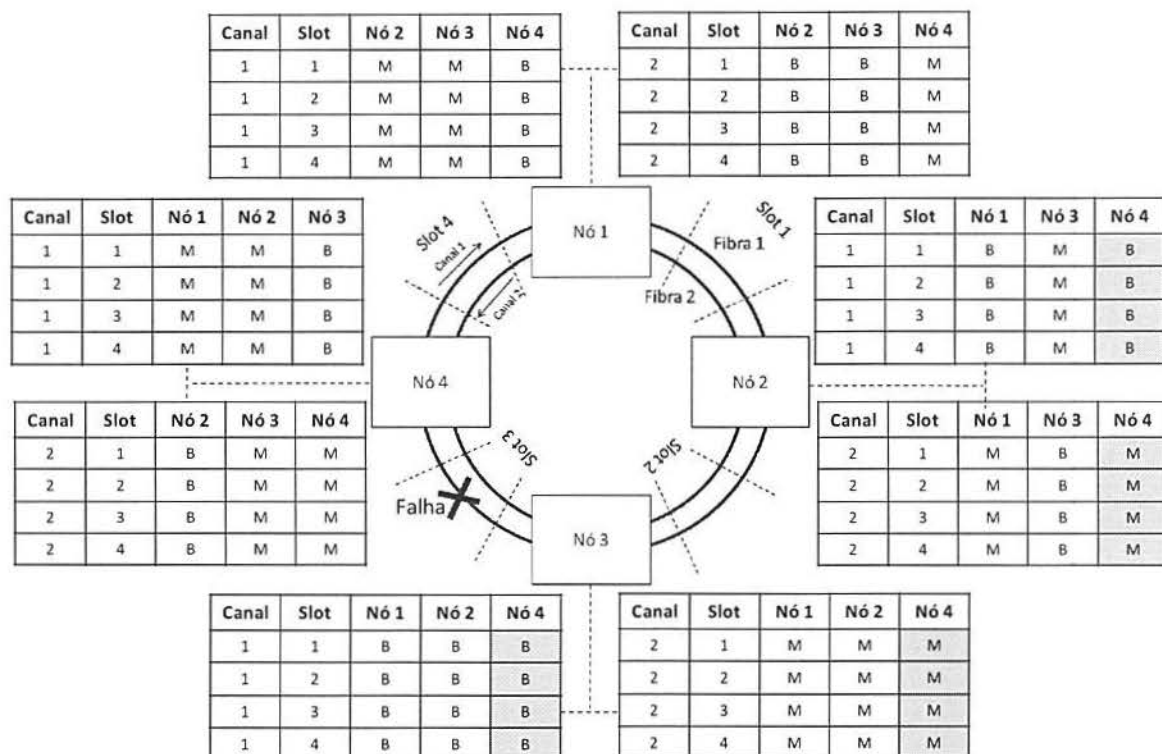


Figura 3-17 Serviço de proteção de caminho

Desta forma a comunicação dos nós com o plano de gerência vale especial atenção. Dado que a proteção ocorre através da atuação da gerência deve ser levando em conta no projeto de implementação como um nó enviará informações para a gerência, e como esta atuará nos nós. A especificação arquitetônica descrita na seção 3.2 permite que a comunicação com a gerência ocorra usando-se os recursos da própria rede ou utilizando-se uma DCN separada. No caso da primeira abordagem deve-se construir uma rede com redundância suficiente para atender o número máximo de falhas que se pretende suportar. Por exemplo, se for especificado que o serviço de proteção deve tolerar até duas quedas de fibra simultâneas a rede deve se constituir de pelo menos três anéis para que no pior caso a comunicação com a gerência seja possível para todos os nós. Igualmente, se for suportada a queda de um nó, deve ser possível comunicar-se com a gerência através de pelo menos dois nós, caso a falha ocorra em um destes.

3.6. Resultados simulados de desempenho

O objetivo dos experimentos de simulação foi demonstrar a viabilidade da utilização dos protocolos utilizados no plano de controle da RCPO para a classe de serviço de melhor esforço, para assim definir uma arquitetura de rede que apresente um bom desempenho, tanto no cenário com tráfego simétrico como assimétrico. Desta forma, a especificação da rede deve criar um modelo de nó que possibilite a implementação de tais protocolos.

Outros experimentos com mais comprimentos de ondas de dados, seja com uma única fibra óptica ou com duas transmitindo em direções opostas (com roteamento pelo menor caminho), são importantes para a comprovação de que esta arquitetura não é apenas passível de ser implementada, mas apresenta vantagens significativas em termos de desempenho e utilização de recursos em comparação às redes atuais.

Entretanto, foge do escopo do trabalho de especificação realizar simulações mais complexas, pois isso tomaria mais tempo e atrasaria o projeto de prototipagem, foco do projeto.

Porem, este limitante de tempo não implica que novas propostas de protocolos e métodos de controle de acesso ao meio não possam ser feitos. De fato, como será visto no capítulo 5, o projeto de implementação buscou a construção de um protótipo que não apenas atenda os requisitos especificados, mas que também permita melhoramentos ao longo do tempo com alterações que, através de simulações computacionais mais elaboradas, comprovadamente apresentem vantagens.

Os resultados de desempenho da rede apresentados nesta seção foram obtidos em um simulador computacional de eventos discretos desenvolvido na linguagem Java. Seguindo as recomendações do trabalho *On Credibility of Simulation Studies of Telecommunication Networks* [20] o simulador usa um intervalo de confiança como critério de parada e o gerador de números pseudo-aleatórios Mersenne Twister [21], contido na biblioteca Colt-lib, desenvolvida pela CERN [Organização Européia para Estudos Nucleares].

Os experimentos consideraram uma rede com 16 nós ligados por um anel de 100km de extensão com uma única fibra óptica e um único comprimento de onda para o transporte de dados (em uma única direção).

A taxa de transmissão e o tamanho dos *slots* adotados nos experimentos são de 10Gb/s e de 64KB, respectivamente.

Foram considerados dois cenários de tráfego: simétrico e assimétrico. No primeiro todos os nós geram a mesma carga de tráfego para todos os outros nós da rede. No segundo o nó 16 gera a mesma carga de tráfego para os outros nós, enquanto que os nós de 1 a 15 geram 50% do tráfego para o nó 16 e o restante para os outros nós de forma simétrica. Nos dois cenários a chegada de pacotes segue o modelo de *Poisson* e os tamanhos dos pacotes a distribuição apresentada no trabalho *The nature of the beast: recent traffic measurements from an Internet backbone* [22]. Foram assumidos quatro tamanhos de rajada: 64B (neste caso um único pacote em uma fila já constitui uma rajada), 16KB, 32KB e 64KB.

A Tabela 3-3 apresenta os parâmetros de entrada adotados nas simulações. A figura 3-18 apresenta a vazão total da rede em função da carga total injetada na rede no cenário simétrico. Vazão total consiste da soma da vazão de cada nó, determinada pela razão entre o total de bytes de dados transmitidos e a capacidade (em bytes) disponibilizada para aquele nó durante o experimento.

As curvas mostram que o tamanho de rajada não influenciou a vazão da rede e que em todos os casos foram obtidas vazões bastante altas. Nota-se que as vazões estão ligeiramente abaixo da carga injetada. Isto ocorre por causa do tamanho dos *slots*, do baixo número resultante de *slots* disponíveis na rede e do tamanho das rajadas, o que dificulta o acesso e gera um acúmulo de pacotes nas filas nos nós da rede. Este acúmulo tem reflexos nos atrasos médios enfrentados pelos pacotes, conforme mostra a Tabela 3-4, e também tem reflexos na taxa de descarte de pacotes, embora a 150% de carga ainda não ocorra uma perda de pacotes sequer.

O atraso médio consiste da média dos atrasos enfrentados por todos os pacotes do momento em que entram na fila até o momento em que são completamente transmitidos.

Parâmetro	Valor
Numero de anéis	1 (transmissão unidirecional)
Tamanho do anel (km)	100
Numero de canais de dados por fibra	1
Numero de nós	16
Taxa de transmissão por canal	10Gb/s
Tamanho do slot (bytes)	64K (resulta em 9 slots)
Tamanho de rajada (bytes)	64, 16K, 32K, 64K
Tamanho de pacote (bytes)	64 (60%), 596 (15%), 700 (5%), 800 (5%), 1100 (7%), 1500 (8%)
Quota de transmissão ($k=1$)	Numero de slots
Tráfego	Chegada: <i>poisson</i> Distribuição: simétrica; assimétrica
Intervalo de confiança (%)	90
Tamanho de cada fila de pacotes (bytes)	128K

Tabela 3-3 Parâmetros de entrada

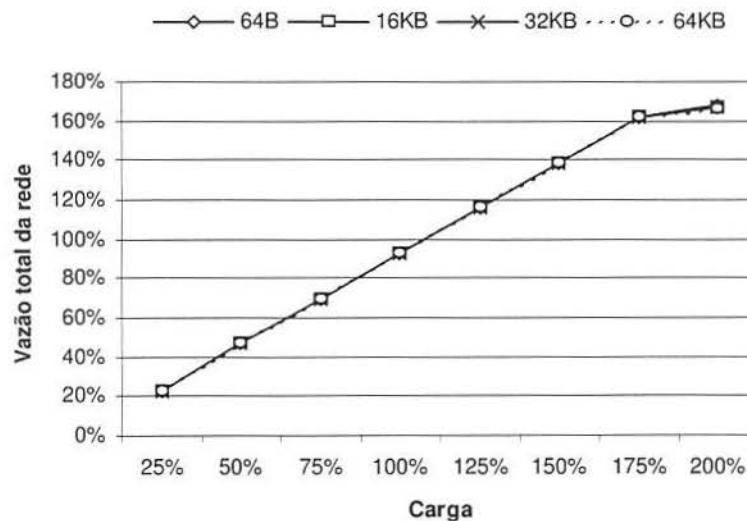


Figura 3-18 Vazão total da rede versus tráfego no cenário simétrico

Percebe-se pelos números contidos na Tabela 3-4 que à baixas cargas as rajadas menores levam a atrasos mais baixos, o que é compreensível uma vez que, apesar de os *slots* trafegarem sub-utilizados, ainda não ocorre uma competição tão grande por eles. A situação começa a se inverter a partir de 100% de carga, quando a competição por *slots* torna-se acirrada. Nesta condição a espera pela montagem de rajadas maiores e o conseqüente uso mais eficiente dos *slots* mostra-se vantajoso. Outro fator importante é que a simulação não

considera a montagem de rajada por tempo máximo, como previsto na arquitetura da rede. Neste caso espera-se que a latência diminua, porém também é esperado que a vazão seja menor dado a subutilização do *slot*.

Tamanho de rajada	Carga		
	50%	100%	150%
64B	3,8ms	3,8ms	4,2ms
16KB	3,7ms	3,8ms	4ms
32KB	7ms	3,6ms	4ms
64KB	13ms	6,5ms	4,5ms

Tabela 3-4 Atraso médio dos pacotes

A Figura 3-19 apresenta o grau de justiça da rede em função da carga total injetada na rede no cenário simétrico. O grau de justiça neste cenário é determinado pela razão entre a vazão do nó com maior vazão e a vazão do nó com menor vazão. Quanto mais próximo de 1 melhor é a justiça de acesso da rede. É interessante notar neste gráfico que o grau de justiça aumenta à medida que a carga também aumenta. Isto ocorre porque com o aumento da transmissão nos nós estes exaurirão as suas cotas de transmissão mais rapidamente, liberando o SAT em menor tempo. Quanto maior for a circulação do SAT na rede mais eficiente é a atuação deste protocolo.

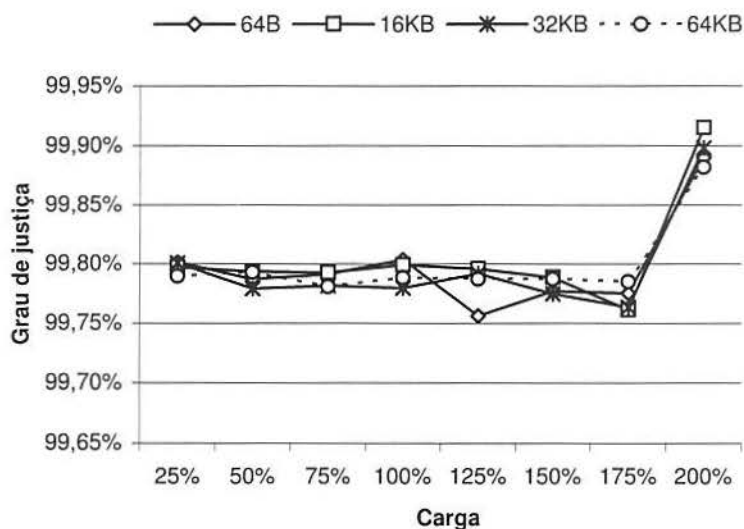


Figura 3-19 Grau de justiça no acesso versus carga no cenário simétrico

De forma geral as curvas mostram que, apesar das pequenas variações, o tamanho da rajada não influenciou de forma considerável o grau de justiça da rede que em todos os casos foi próximo do ótimo.

A Figura 3-20 apresenta a vazão total da rede em função da carga total injetada na rede no cenário assimétrico. As curvas mostram a influência do tamanho da rajada na vazão total da rede. Quanto maior a rajada maior é a vazão alcançada. Isto ocorre porque no cenário assimétrico a competição pelos *slots* é bastante acirrada mesmo em baixas cargas. Portanto, quanto maior a rajada transmitida em um *slot* melhor é a eficiência e, conseqüentemente, melhor são os parâmetros de desempenho. Como a quota de transmissão é dada em número de *slots*, quanto mais cheio estiver o *slot* melhor é o desempenho.

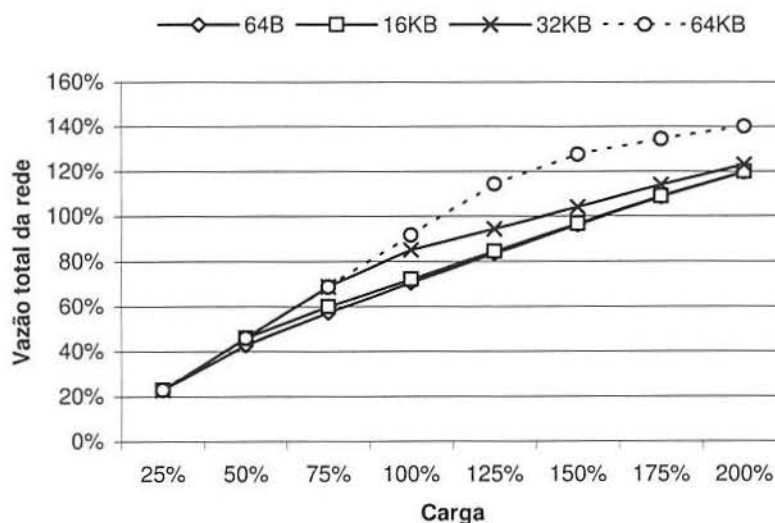


Figura 3-20 Vazão total da rede versus tráfego no cenário assimétrico

Assim como no cenário simétrico e pelo mesmo motivo, a vazão não aumenta na mesma proporção que a carga. Entretanto o acúmulo de pacotes nas filas ocorre já com baixas cargas, sendo que a taxa de descarte de pacotes é maior para rajadas de tamanhos menores. Por exemplo, para o tamanho de rajada de 64B, descartes de pacotes já ocorrem em algumas filas de alguns nós a uma carga de apenas 25%, sendo que para o tamanho de rajada de 64KB isto ocorre somente quando a carga está próxima de 75%. Portanto, no cenário assimétrico é ainda mais evidente a necessidade do uso eficiente dos *slots*.

A Figura 3-21 apresenta o grau de justiça da rede em função da carga total injetada na rede no cenário assimétrico. Apesar da diferença na distribuição do tráfego, o total de tráfego gerado em cada nó da rede é o mesmo e, portanto, aplica-se o mesmo cálculo de grau de justiça do cenário simétrico.

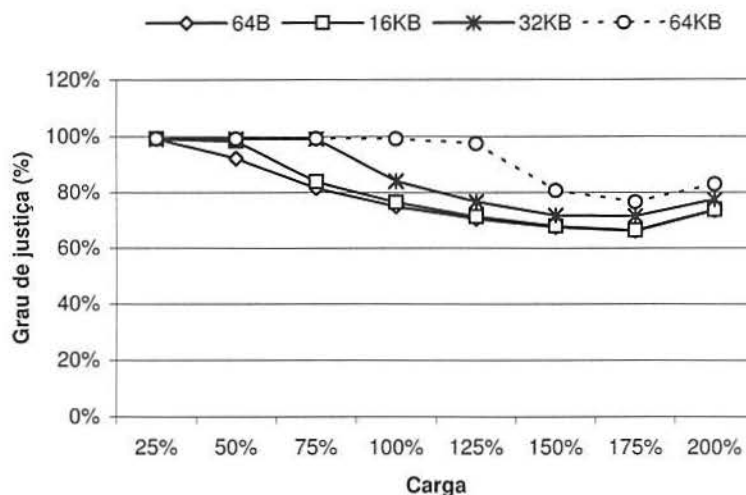


Figura 3-21 Grau de justiça no acesso versus carga no cenário assimétrico

Nota-se que quanto menor o tamanho da rajada maior é a degradação do grau de justiça em função da carga da rede. Isto ocorre em virtude do uso mais eficiente dos escassos recursos de rede (*slot*) com rajadas de tamanho maior e por causa da limitação de SAT+ em lidar com distribuições assimétricas de tráfego.

3.7. Interconexão

Para o transporte de dados com classes de serviço de melhor esforço ou reservada a rede se interconecta com outras redes e clientes através de interfaces Ethernet, e só entende os formatos de quadro definido nos padrões IEEE 802.3 [23] e IEEE 802.1q [24]. Os formatos destes quadros são mostrados na figura 3-22.

A classificação do tráfego em classes de serviço de melhor esforço ou reservada é realizada através dos campos de endereçamento e VLAN ID, tanto usados em conjunto quanto separadamente. Já a determinação do tipo de qualidade de serviço é feita através do campo de prioridade. Desta forma, os quadros Ethernet IEEE 802.3 não podem ser diferenciados

em qualidade de serviço, sendo encaminhados para a mesma sub-fila dentro de uma fila de transmissão.

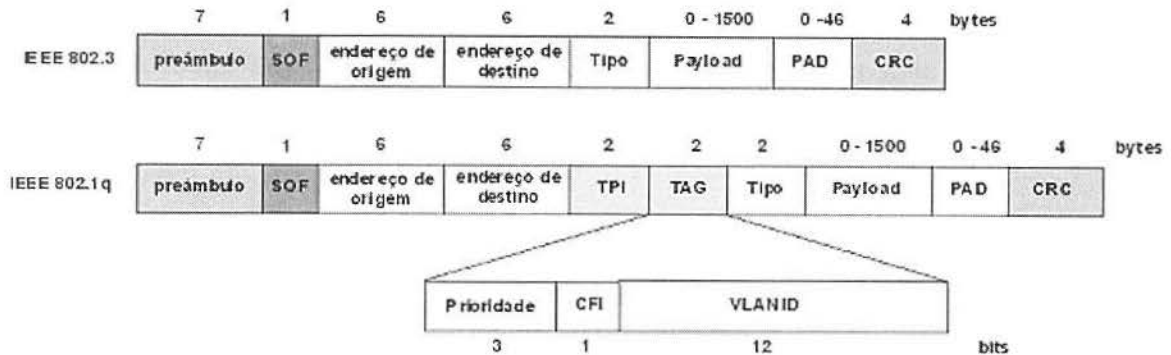


Figura 3-22 Formato dos quadros Ethernet IEEE 802.3 e IEEE 802.1q/p

Para o transporte de dados com classe de serviço transparente, qualquer tipo de sinal óptico com qualquer modulação e em qualquer taxa de transmissão pode ser suportado, apenas respeitando-se o comprimento de onda do laser, que deve estar dentro da janela de 1.550 nm.

A interconexão entre os nós da rede ocorre por interfaces proprietárias. A utilização destas é necessária para atender as características de transmissão e sincronização inerentes às redes de comutação de pacotes ópticos.

Porém os clientes interligados nas interfaces Ethernet não têm ciência das suas particularidades, sendo privados de entender as complexidades da rede de pacotes ópticos. Para eles a RCPO é uma nuvem, um meio de comunicação qualquer, sendo a comutação realizada da mesma forma definida no padrão IEEE 802.1d [25]. Assim, os clientes têm a mesma visão lógica uns dos outros que teriam se estivessem ligados a um único comutador Ethernet.

3.8. Discussão

Criar uma arquitetura de rede de comutação de pacotes ópticos é um trabalho árduo que demanda o profundo conhecimento de diversos conceitos, indo desde os fundamentos básicos da camada física até os diferentes tipos de tráfego gerados pelas aplicações.

Dentro deste trabalho existem questões onde é necessário optar entre desempenho e simplicidade. Por exemplo, uma rede OPS com acesso ao meio por *slots* atingiria sua

eficiência máxima se um pacote óptico contivesse apenas um único pacote elétrico, havendo suficientes pacotes ópticos dentro do *slot* para ocupá-lo totalmente. Entretanto, tal especificação não é possível de ser implementada com as tecnologias disponíveis na taxa de gigabit ou superior, dada a necessidade de se processar cabeçalhos e encaminhar pacotes ópticos na ordem de poucos nanossegundos.

Em contrapartida a RCPO propõe um modelo inovador, onde a técnica de montagem de rajadas das redes OBS pode ser utilizada para agregar diversos pacotes elétricos em um único pacote óptico. Assim, a solicitação de transmissão deste pacote óptico acontece apenas quando a rajada composta de diversos pacotes elétricos atinge o seu limite. Esta abordagem aumenta a utilização do meio óptico, porém traz uma considerável desvantagem: o aumento da latência na transmissão.

Para contornar esse problema a RCPO prevê uma parametrização no critério do tempo de vencimento de uma rajada. Assim, tráfegos que não toleram uma alta latência podem ser encaminhados para filas com mais baixo tempo de montagem de rajadas. Evidentemente estas filas tendem a produzir rajadas de menor tamanho, gerando ineficiência na utilização do *slot*, e por conseqüência, no canal óptico. Assim, o dimensionamento entre o tamanho de rajadas e tempo de latência em função das demandas de tráfego e as limitações físicas do nó da rede é a chave para se obter bons resultados de desempenho, e deve ser analisado para cada cenário em particular.

Entretanto, no caso dos requisitos de um dado tráfego, como a latência ou a banda de transmissão, serem muito restritos, uma melhor solução é a utilização de outro diferencial da arquitetura da RCPO - as diferentes classes de serviço.

É possível, por exemplo, determinar o tempo de latência máximo e a banda mínima para um determinado tráfego, utilizando para isso a classe de serviço reservada, podendo-se atingir valores tão rígidos quando necessários através da simples adição de mais *slots* dedicados a esse fluxo.

Além desta classe, a RCPO ainda permite que qualquer tipo de protocolo com qualquer modulação e taxa circule pela rede através da classe de serviço transparente, usando os conceitos das redes OCS.

Outro considerável diferencial da RCPO é a sua tolerância a falhas de múltiplas conexões através do mecanismo de reinserção contínua. Com isto é possível realizar a proteção de

caminho através da atuação da gerência, mantendo a simplicidade do nó. Para os nós não existe o conceito de falhas, para estes existem apenas recursos livres ou bloqueados, ficando a complexidade de decisão para o plano de gerência. Esta abordagem apesar de simplificar a implementação dos elementos de rede trás um significativo problema: o tempo de atuação para a correção de erros. Tipicamente a atuação via plano de gerência é mais lenta que via plano de controle. Assim sendo, é importante levar isto em conta no projeto de implementação, buscando-se meios para que o plano de gerência possa atuar o mais rápido possível na rede. Também é possível que novas propostas surjam a fim de se implementar funcionalidades no plano de controle para contornar este problema.

Em suma, a RCPO apresenta uma arquitetura de rede moderna, com todas as funcionalidades necessárias para atender de forma eficiente os novos serviços de telecomunicações. Entretanto todas estas funcionalidades de nada valeriam se a RCPO fosse uma rede de difícil implementação e operação. É necessário garantir todos os serviços descritos neste capítulo sem perder o foco da viabilidade da prototipagem.

O próximo capítulo descreve a arquitetura do nó da RCPO, que é o último passo antes do projeto de implementação. A arquitetura do nó define quão complexa, cara e difícil é a construção da rede.

Capítulo 4

Arquitetura do Nó

O capítulo anterior apresentou a especificação arquitetônica da RCPO com todos os seus requisitos e funcionalidades. Porém, para atingir tais requisitos é necessário especificar uma arquitetura de nó condizente com a rede proposta, sem perder o foco na viabilidade de implementação, objetivo final do projeto.

Assim, este capítulo especifica a arquitetura do nó da RCPO que é dividido em dois blocos funcionais: unidade de controle e unidade de transporte de dados, como mostrados na figura 4-1. O conjunto das unidades de controle e transporte de dados, implementadas em todos os nós da rede, resultam, respectivamente, nos planos de controle e transporte de dados da RCPO.

Ao contrário desses planos a gerência da rede é centralizada, podendo se comunicar utilizando a unidade de transporte de dados ou utilizando uma DCN. No caso de ser ligada diretamente esta pode utilizar uma ou mais unidades caso seja necessária a redundância de acesso, desta forma a gerência pode atingir através do plano de dados todas as demais unidades do plano de controle. Entretanto, a gerência não se comunica diretamente com o plano de transporte de dados, o utilizando apenas para transporte das informações. A atuação da gerência neste plano é indireta, ocorrendo por meio do plano de controle. A arquitetura do nó apresentado neste capítulo utiliza a abordagem da comunicação com a gerência via plano de dados uma vez que esta é mais complexa e requer maior atenção.

Com relação aos canais, o comprimento de onda reservado para o fluxo dos cabeçalhos em cada um dos anéis é removido na entrada do nó através de um filtro, entrando diretamente na unidade de controle. Os demais comprimentos de onda seguem para a unidade de transporte de dados, passando antes por um *delay* de fibra. Assim, os comprimentos de

onda de dados serão atrasados em relação ao comprimento de onda de controle, tendo a unidade de plano de controle um maior tempo para processar o cabeçalho.

Em paralelo ao processamento do cabeçalho a unidade de transporte de dados recebe e classifica os quadros Ethernet enviados pelos clientes, montando a rajada eletrônica e solicitando a unidade de plano de controle a transmissão dos pacotes ópticos já montados.

Por sua vez, a unidade de controle analisa o cabeçalho, verificando se existem dados destinados a este nó e se há disponibilidade para a transmissão de seus próprios dados. Feito isso, a unidade de transporte de dados é propriamente configurada e o cabeçalho já processado é re-inserido à frente.

As próximas seções desse capítulo destinam-se a detalhar estas duas unidades que compõem o nó da rede.

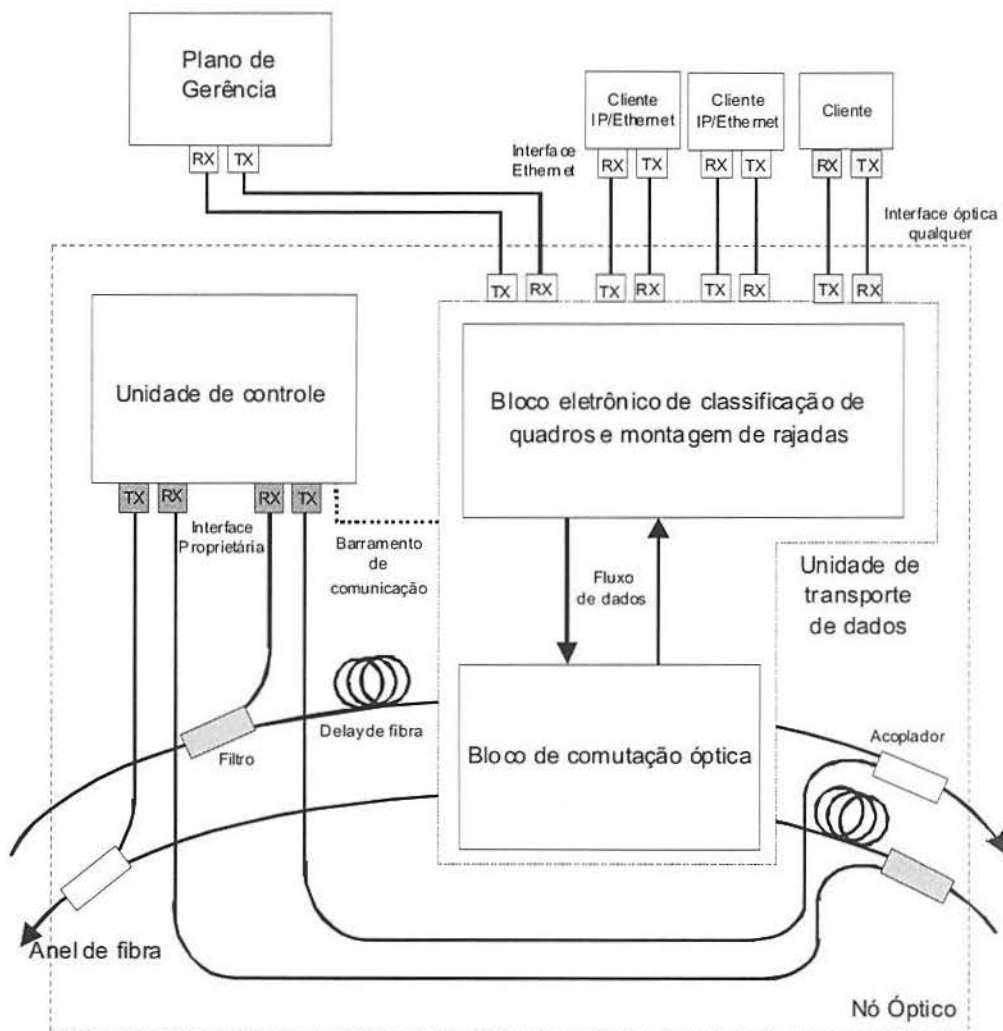


Figura 4-1 Visão arquitetônica do nó

4.1. Unidade de controle

Como explicado na seção 3.4 a unidade de controle realiza fundamentalmente a função de controle de acesso ao meio através do processamento do cabeçalho dos *slots*. Para tal esta é dividida em blocos com funções específicas, como mostrado na figura 4-2.

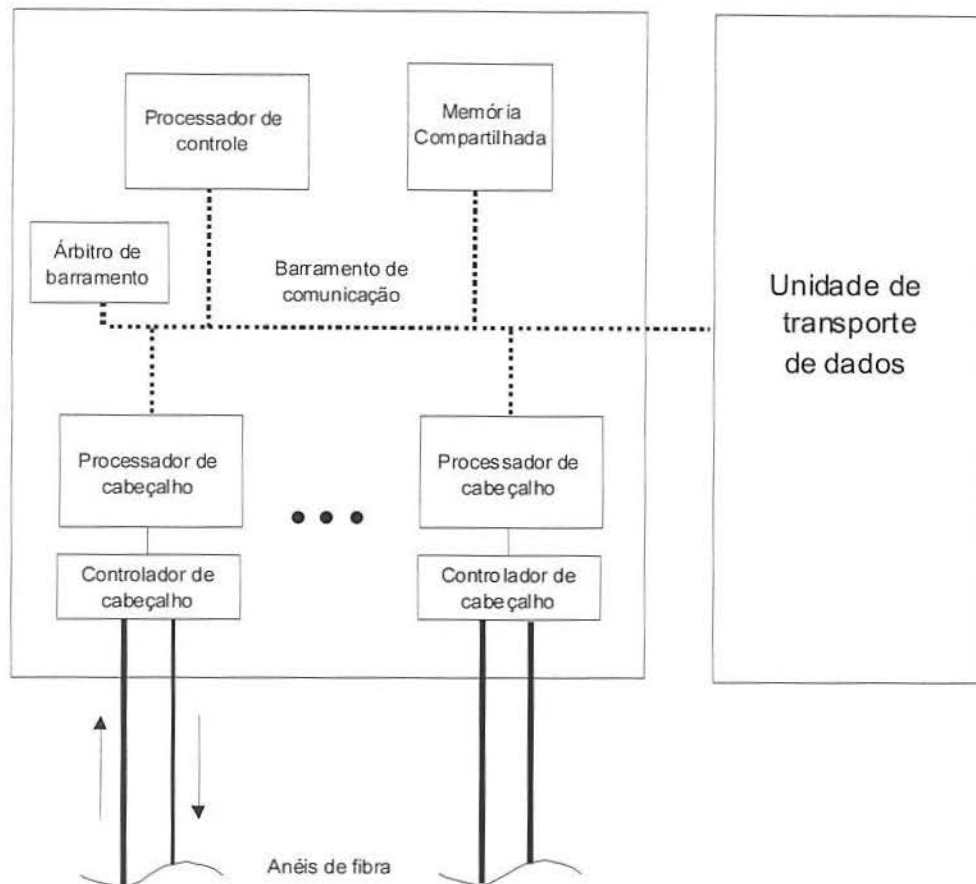


Figura 4-2 Unidade de controle

O Processador de controle recebe comandos do plano de gerência via o barramento de dados através do plano de transporte de dados. Através dos comandos da gerência o processador de controle organiza as tabelas de alocação dos *slots*, como explicado na seção 3.4. Estas tabelas são armazenadas na memória compartilhada e serão utilizadas pelos processadores de cabeçalhos para as tomadas de decisão na alocação do *slot*. O processador de controle deve também alertar o plano de gerência sobre as falhas de *link* detectadas pelos processadores de cabeçalho, como mostrado na seção 3.4.

O processador de controle deve ainda realizar o mapeamento das filas de dados (controladas pela unidade de transporte de dados) com os nós da rede, mantendo uma tabela que traduz o destino que cada fila alcança.

O processador de cabeçalho atua em conjunto com um controlador de cabeçalho dedicado ao seu uso, sempre existindo um par de processador e controlador por anel da rede. O controlador é um bloco necessário para garantir os rígidos requisitos de sincronização dos cabeçalhos. Nos processadores de cabeçalho é executado o protocolo SAT+. a reinserção de cabeçalhos também é executada nestes processadores, de forma individual para cada anel de fibra na rede.

Como o mesmo barramento é utilizado para toda a comunicação do nó, tanto entre a unidade de gerência e controle, a unidade de controle e transporte de dados, e internamente na própria unidade de controle, torna-se necessária a utilização de um bloco chamado árbitro de barramento que controlará todo o acesso entre os dispositivos pertencentes a esse barramento, evitando assim colisões resultantes das tentativas de dois ou mais periféricos tentarem acessar um mesmo recurso.

4.2. Unidade de transporte de dados

A unidade de transporte de dados se divide em dois blocos macros: o bloco eletrônico de classificação de quadros e montagem de rajadas e o bloco de comutação óptica. Esses dois blocos em conjunto realizam todo o transporte de dados carregados pelo nó da RCPO.

4.2.1. Bloco eletrônico

O bloco eletrônico de classificação de quadros e montagem de rajadas, ou simplesmente bloco eletrônico, tem a função de classificar o tráfego dos clientes ligados a RCPO, separando-o em filas e sub-filas de transmissão de acordo com o nó destino, a classe e a qualidade de serviço. Além desta também são funções deste bloco: a montagem de rajadas de acordo com os critérios de tamanho e tempo definidos pelo plano de controle, a solicitação para o plano de controle da transmissão das rajadas prontas e a transmissão da rajada. A figura 4-3 apresenta os sub-blocos funcionais que compõem o bloco eletrônico.

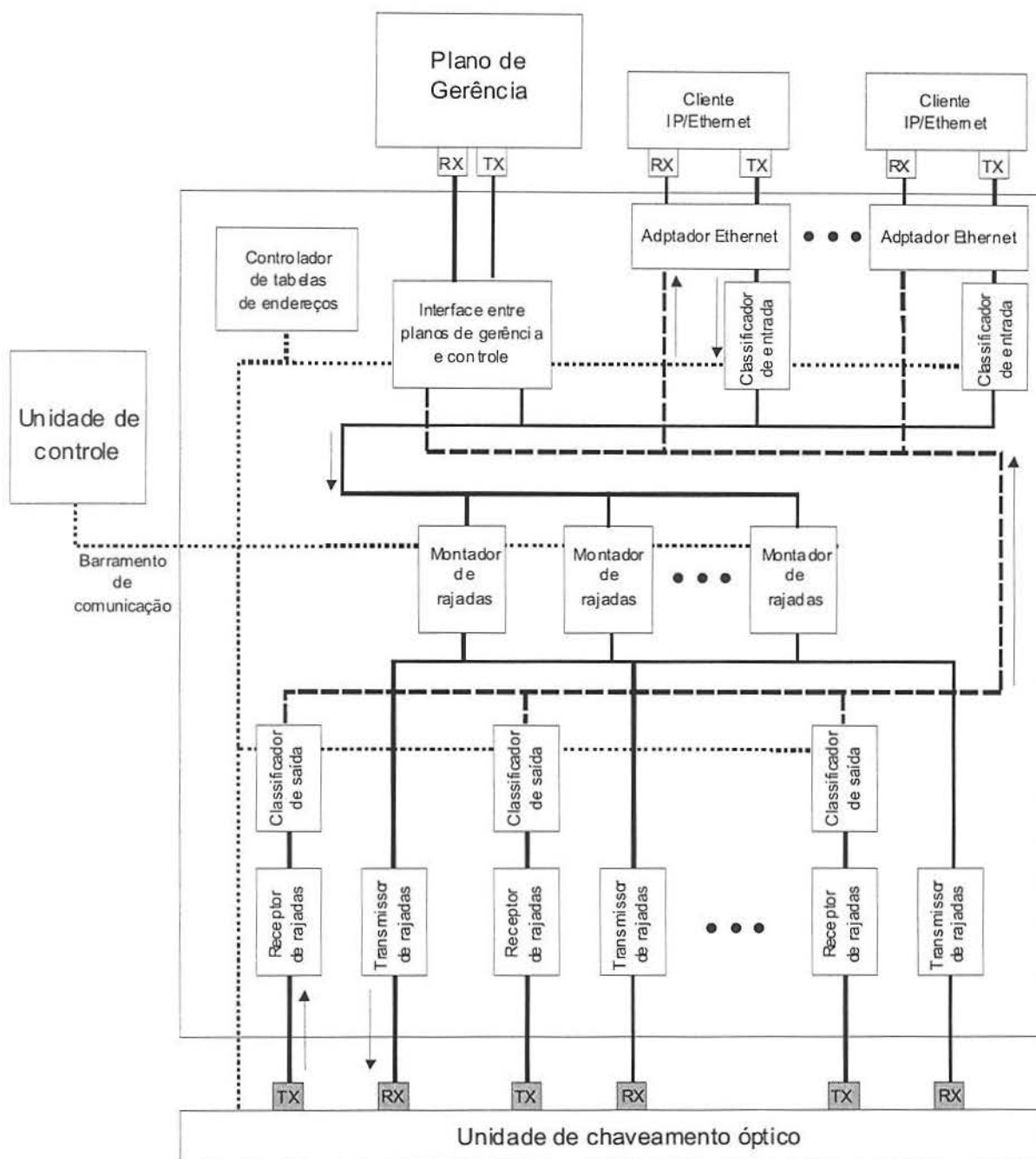


Figura 4-3 Bloco eletrônico de classificação de quadros e montagem de rajadas

Os quadros dos clientes que entram no nó passam inicialmente pelo sub-bloco Adaptador Ethernet. Como o atual padrão Ethernet não prevê um campo apropriado que contenha o

tamanho do quadro, que é variável, é necessário acrescentar este campo ao quadro para facilitar o seu encaminhamento pelos próximos sub-blocos¹.

O Adaptador Ethernet deve então realizar toda a comunicação inerente ao protocolo Ethernet e acrescentar o campo extra contendo a informação do tamanho do quadro.

Após a adaptação o quadro segue para o Classificador de entrada, que consulta o Controlador de tabelas de endereço a fim de decidir para qual fila de transmissão o quadro deve ser comutado. Para esta tomada de decisão diversos campos do quadro podem ser utilizados como o endereço MAC de destino, o número da VLAN, os parâmetros de QoS, de acordo com o grau de sofisticação da prototipagem do nó.

Para realizar este encaminhamento de quadros diversas abordagens podem ser seguidas, como a auto-aprendizagem de endereços, definido pelo padrão IEEE 802.1d, ou mesmo a criação de tabelas estática de encaminhamento. A implementação da política de encaminhamento de quadros é realizada pelo sub-bloco Controlador de tabelas de endereços e a sua abordagem fica a critério do projeto de implementação.

Já propriamente classificado o quadro segue para o Montador de rajadas. Este sub-bloco implementa os protocolos PAT, para a classe de serviço reservado, e PAT+, para a classe de serviço de melhor esforço, montando a rajada eletrônica baseando-se nos critérios de tamanho, além de também poder fechar a rajada de acordo com o parâmetro de tempo máximo. Quando uma fila atinge um destes dois critérios de montagem, tamanho e tempo, o sub-bloco solicita a unidade de plano de controle a sua transmissão. Assim, quando houver algum *slot* disponível, a unidade de controle enviará a liberação da transmissão ao montador de rajadas, que passará todos os quadros que compõe esta rajada ao sub-bloco Transmissor de rajadas².

Deve existir um sub-bloco de Transmissão de rajadas para cada canal de dados suportado pelo nó. Este sub-bloco deve preparar a rajada eletrônica de forma a constituir o pacote óptico determinado no item 3.4.3.

1- A determinação do tamanho do quadro ocorre no nível elétrico, através de sinais na interface MII.

2- Uma fila de transmissão do montador pode conter diversas rajadas prontas para a transmissão, assim, o fechamento de uma rajada não deve impedir a utilização da fila para a montagem de outra rajada.

Para a recepção de pacotes ópticos o processo inicia-se na chegada de um cabeçalho que contenha a informação de que um *slot* carrega dados destinados a este nó. Após a atuação da unidade de controle no bloco de comutação óptica, os dados chegarão ao Receptor de rajadas que removerá a formatação realizada pelo Transmissor de rajadas, passando os quadros ethernet ao Classificador de saída, que também consultará o Controlador de tabelas de endereços para tomar a decisão de para qual cliente o quadro deverá ser encaminhado. Antes de chegar ao cliente o quadro passa novamente por um Adaptador Ethernet que removerá o campo de tamanho, tornando-o novamente em um quadro Ethernet padrão. Além de classificar, comutar e transmitir o tráfego dos clientes, o plano de transporte de dados é utilizado ainda para o fluxo de informações entre a unidade de gerência e as unidades de controle. Para tal, a unidade de gerência se conecta a uma unidade de transporte de dados através do sub-bloco Interface entre planos de gerência e controle. Este sub-bloco se interliga com a unidade de controle do nó em que reside através do barramento de comunicação e com as demais unidades de controle da rede através do plano de transporte de dados, sendo estes dados encaminhados a uma fila de transmissão que atinja o nó desejado, da mesma forma como ocorre com os dados dos clientes.

4.2.2. Bloco de comutação óptica

O segundo bloco que compõe a unidade de transporte de dados é o bloco de comutação óptica. Ele é formado por elementos ópticos que em conjunto dão uma grande flexibilidade nas operações de chaveamento de comprimentos de onda que, quando executadas em um período de tempo bem determinado, permitem a operação sobre os pacotes ópticos que circulam pelo anel.

A figura 4-4 ilustra a arquitetura do bloco. O principal elemento é o comutador óptico, apresentado em maiores detalhes na figura 4-5. O comutador é composto por três chaves ópticas simples e dois acopladores ópticos. Por se tratar de uma rede de topologia em anel este comutador deve permitir basicamente a execução de três operações: adição, remoção e passagem.

Quando um *slot* contém um pacote destinado ao nó, a unidade de controle deve preparar o comutador óptico do canal deste *slot* para a operação de remoção. Nesta operação fecha-se a chave 1, permitindo que os dados fluam para dentro do nó, e abre-se a chave 3, impedindo

que os dados continuem a circular pelo anel. No caso contrário, quando o nó transmite um pacote óptico, fecha-se a chave 2 e abre-se a chave 3. É importante notar que ambas as operações podem ser realizadas simultaneamente; quando um nó remove um pacote do anel

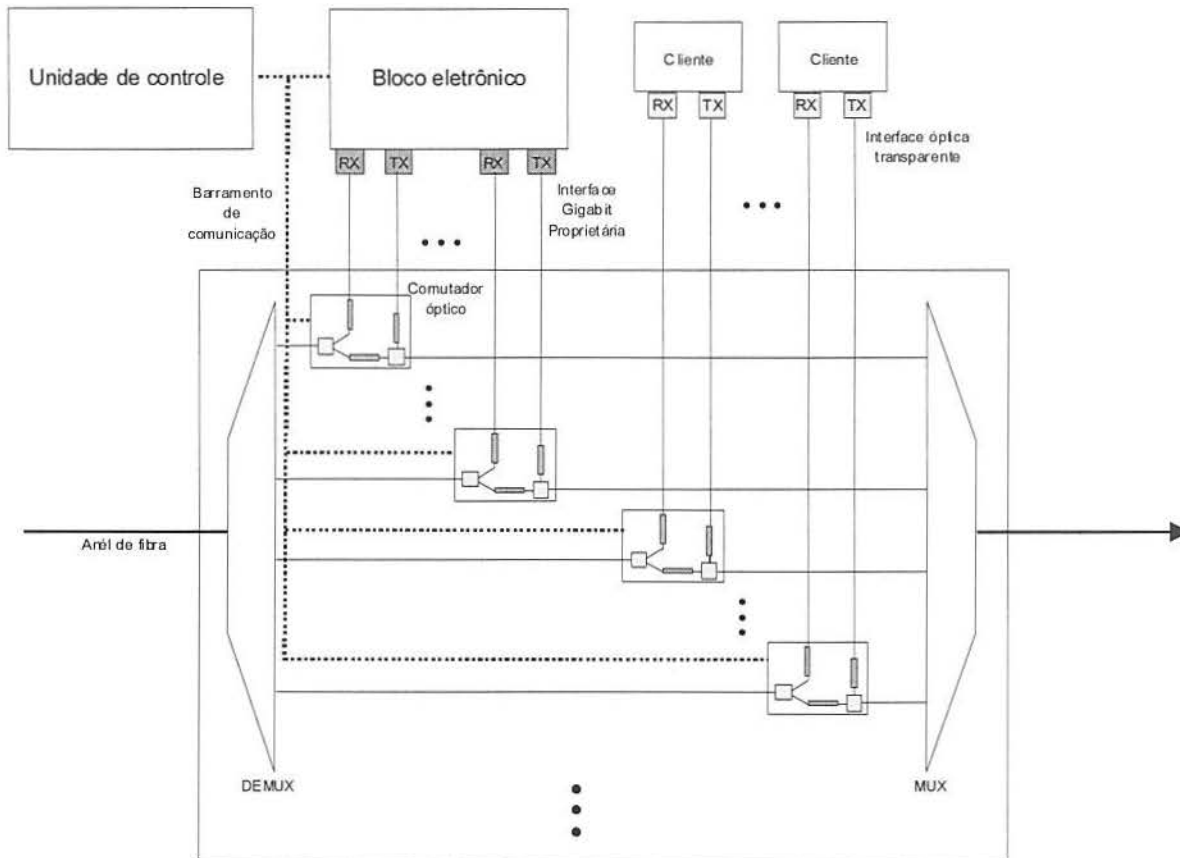


Figura 4-4 Bloco de comutação óptica

o mesmo *slot* já estará livre para a transmissão, podendo ser reusado imediatamente pelo nó. No caso do *slot* não ser destinado ao nó ele deverá passar por este sem que ocorra remoção ou adição. Desta forma o comutador deve realizar a operação de passagem abrindo as chaves 1 e 2 e fechando a chave 3.

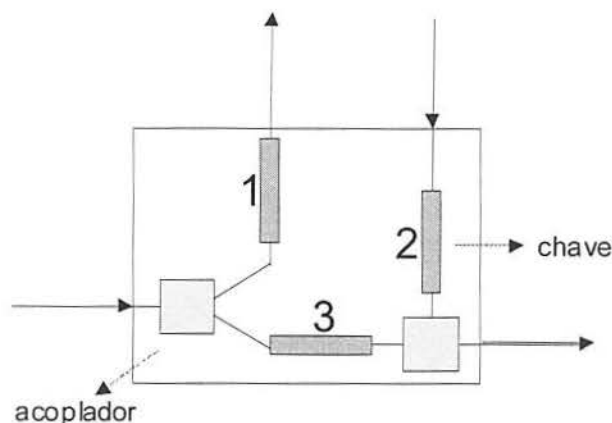


Figura 4-5 Comutador óptico

Outra característica importante deste comutador é a capacidade de realizar as operações de remoção e passagem simultaneamente, dando a rede a funcionalidade de *multicast* no domínio óptico. Tal funcionalidade otimiza o uso dos recursos da rede além de ser fundamental para a implementação de novos serviços multimídia que estão ganhando força no mercado de telecomunicações, como o IPTV [26].

O bloco de comutação óptica deve conter um comutador para cada canal de dados suportado no nó. Também deve existir um par de demultiplexador e multiplexador óptico para cada anel de fibra da rede.

4.3. Estados dos anéis no nó

Cada anel pode estar em um de três estados no nó da RCPO: estado de inicialização, operação e proteção, como mostrado na figura 4-6. Ao ser ligado o processador de cabeçalho considera que o anel de sua responsabilidade está no estado de inicialização, podendo tomar uma de duas possíveis atitudes: caso o nó seja o responsável por inserir os cabeçalhos neste anel este assim o fará (respeitando o espaçamento definido entre os cabeçalhos) até que receba de volta todos os cabeçalhos transmitidos na ordem correta; caso o nó não seja o responsável ele apenas esperará pela recepção de todos os cabeçalhos do anel, repassando para frente todos os cabeçalhos que receber (respeitando o tempo de processamento do cabeçalho, que deve ser fixo e igual em todos os nós da rede). Em qualquer um dos dois casos o nó que receber todos os cabeçalhos do anel na sequência correta saíra do estado de inicialização entrando no estado de operação. A determinação do nó que deve inserir os cabeçalhos, o número de cabeçalhos, o espaçamento entre

cabeçalhos e o tempo de processamento são pré-determinados, ficando a critério do projeto de implementação a forma de realizá-lo.



Figura 4-6 Estados do anel no nó da rede

Passada a fase de inicialização o anel entrará no estado de operação, onde todos os canais de dados podem ser utilizados para carregar o tráfego dos usuários. Caso o nó deixe de receber algum cabeçalho o anel entrará no estado de proteção neste nó. Na volta do recebimento correto dos cabeçalhos o anel novamente entrará no estado de operação.

Em condições normais o anel nunca deverá voltar ao estado de inicialização, a não ser que o processador de cabeçalho responsável por este anel no nó seja reinicializado.

4.4. Discussão

A especificação arquitetônica do nó da rede é o último passo antes do projeto de implementação. Esta especificação tem a difícil tarefa de manter dois compromissos que tendem a divergir em diversos aspectos. São eles: o compromisso de realizar as funcionalidades descritas na especificação da arquitetura da rede e o compromisso com a viabilidade da implementação, criando uma arquitetura de nó simples o bastante para ser implementado com os componentes hoje disponíveis no mercado, porém com as funcionalidades inovadoras propostas pela RCPO.

Alguns pontos da arquitetura do nó merecem atenção especial. Por exemplo, é necessária a realização de uma eficiente comunicação entre as unidades de controle e transporte de dados, o que envolve diversos blocos e sub-blocos funcionais que necessitam trocar dados e comandos entre si. Toda esta comunicação é representada na especificação arquitetônica do nó por um barramento, chamado de barramento de comunicação. Este é um ponto que tende a apresentar dificuldade na implementação dada a sua complexidade, podendo ser decisivo na escolha da tecnologia utilizada.

Além da comunicação entre as unidades, outros aspectos relevantes são encontrados internamente em cada unidade. Na unidade de controle existem dois desafios bem definidos: realizar o processamento do cabeçalho da forma rápida (na ordem de microssegundo) e manter a sincronização entre o cabeçalho e os *slots*.

No primeiro caso busca-se diminuir o tempo de processamento de cabeçalho utilizando-se um processador por anel, criando-se um sistema multi-processado. Em contrapartida este processamento distribuído traz um aumento da complexidade de operação, principalmente no que diz respeito ao acesso dos recursos comuns como a memória compartilhada.

No segundo caso o bloco Controlador de cabeçalhos visa manter a exata sincronização entre cabeçalhos. Para a implementação deste bloco deve ser empregada uma tecnologia que consiga atender estes rígidos parâmetros de tempo, o que também deve ser um ponto de decisão no trabalho de prototipagem.

Na unidade de transporte de dados vários outros pontos são relevantes, como a implementação do Controlador de tabelas de endereços. A descrição arquitetônica do nó deixa esta questão para o projeto de implementação. De fato, diversas abordagens podem ser assumidas, desde a criação de endereçamentos estáticos, o que tornaria a implementação mais simples porém pouco atrativa do ponto de vista de operação, até a implementação da auto-aprendizagem de endereços, que, ao contrário, facilitaria a operação da rede, mas aumentaria a complexidade do bloco eletrônico. Esta questão está altamente relacionada com o grau de sofisticação do protótipo que se pretende alcançar.

Do ponto de vista da transmissão dos pacotes ópticos no anel, algumas questões também surgem, como a necessidade de uma rápida sincronização no domínio eletrônico para o recebimento de pacotes. Como a interface de recepção não permanece conectada ao longo da operação da rede, sendo utilizada apenas quando um pacote efetivamente atinge o nó através desta, é necessário utilizar mecanismos de rápida recuperação de relógio, o que inviabiliza o uso de tecnologias de transmissão já existentes para o transporte de dados no anel como a Ethernet que necessita de tempos na ordem de milissegundos para realizar a sincronização entre duas interfaces. Na especificação as interfaces do anel da rede são descritas como proprietárias, uma vez que deverão ser desenvolvidas para atender os requisitos de tempo que o protótipo proposto necessita alcançar.

Ainda com relação a transmissão dos pacotes ópticos, o tempo de atuação nas chaves do comutador óptico é outro aspecto importante. No item 3.6 as simulações de desempenho foram construídas sobre um cenário onde os *slots* possuem um tamanho de 64KB a uma taxa de transmissão de 10 Gbps resultando em um tempo de aproximadamente 50 microssegundos. Como os cabeçalhos seguem um *slot* a frente do *slot* que está carregando a informação, o processamento de cabeçalho deve ser realizado em um tempo de duração nesta mesma ordem. Entretanto, a atuação nas chaves é apenas uma das tarefas necessárias no processamento do *slot*. Desta forma, para atender um cenário similar ao proposto nas simulações, objetivando-se assim alcançar um igual desempenho, é necessário criar uma matriz de comutação que responda aos comandos da unidade de controle em poucos microssegundos. Esta requisição de tempo exclui diversas tecnologias de chaves ópticas como as chaves mecânicas, eletro-ópticas e termo-ópticas, que respondem de forma muito mais lenta. Assim, cabe ao projeto de implementação buscar tecnologias de chaves compatíveis com os tempos de operação da rede.

Como visto, diversos são os pontos que merecem especial atenção no desenvolvimento do nó da rede, tornando o projeto de implementação um desafio. Entretanto, mesmo atendendo todos os requisitos da rede, a arquitetura do nó da RCPO não apresenta nenhuma parte impossível de ser realizada com as tecnologias já disponíveis atualmente, como será visto no próximo capítulo.

Capítulo 5

Projeto de implementação

Este capítulo apresenta o projeto de implementação adotado, iniciando-se com a análise das tecnologias candidatas a serem empregadas no desenvolvimento. A partir desta análise são apresentados os componentes eleitos para o desenvolvimento do protótipo e, em seguida, a descrição do projeto de implementação. Após a especificação do projeto de implementação é apresentado o estado atual do protótipo construído e os testes até agora realizados.

Ao final, é realizada a discussão sobre os assuntos abordados ao longo do capítulo, destacando-se os pontos favoráveis e os possíveis problemas encontrados na solução adotada.

5.1. Análise das tecnologias candidatas

Como abordado na seção de discussão do capítulo anterior a implementação do nó da RCPO apresenta vários e diversificados desafios. Do ponto de vista de prototipagem tais desafios podem ser separados em dois mundos distintos: o eletrônico e o óptico.

Os próximos itens apresentam o estudo das tecnologias analisadas para o emprego no projeto de prototipagem do nó da RCPO e a escolha dos componentes adotados.

5.1.1. Componentes eletrônicos

Do ponto de vista de processamento de dados no domínio elétrico, a RCPO apresenta dois principais desafios: o processamento de cabeçalhos e o transporte de dados dos clientes.

Na seção 4.1 foram apresentados os blocos Processador de cabeçalho e Controlador de cabeçalho como os responsáveis pelo tratamento deste. Como visto anteriormente, processar o cabeçalho significa analisar os seus campos a fim de verificar se existem *slots* carregando informações para o nó, ou se existem *slots* vazios para a transmissão de seus próprios dados. Feita esta verificação diversas consultas a tabelas e a atuação em outros blocos devem ocorrer de acordo com os procedimentos descritos nos capítulos 3 e 4.

Assim, para o processamento do cabeçalho, deve-se buscar uma tecnologia que permita a implementação do protocolo SAT+ de forma simples além de ser capaz de executar as tarefas acima descritas na ordem de microssegundos.

A escolha natural para esta função, e amplamente utilizada em comutadores comerciais como *switches* e roteadores, são os processadores de propósito geral. Mais comumente conhecidos como microprocessadores. Estes permitem a fácil implementação de protocolos através do uso linguagens de programação de médio nível como o ANSI-C [27] e apresentam performance compatível com os tempos desejados. Dado o requisito de tempo de microssegundos deve-se escolher uma arquitetura de processador de 32 bits ou superior, e com frequência de *clock* na ordem de centenas de Mega-hertz.

Existe uma grande variedade de arquiteturas de microprocessadores com tais características e disponíveis comercialmente a baixo custo, não sendo assim este um ponto crítico na prototipagem do nó da RCPO.

Para o aumento da flexibilidade na programação os microprocessadores podem ainda operar munidos de um sistema operacional. O uso de sistemas operacionais facilita a programação devido aos serviços que este provê, como: o interfaceamento entre os *softwares* de aplicação e o *hardware*, o escalonamento de processos e o controle de acesso a memória. Em contrapartida um sistema operacional introduz características não desejáveis, como o aumento da latência na execução de tarefas e a perda da determinação do tempo de tratamento de eventos. Uma variação de sistema operacional conhecido como sistema operacional de tempo real (RTOS – *real time operating system*) visa à diminuição desses efeitos e é comumente utilizado em aplicações similares.

Para a execução da sincronização dos cabeçalhos uma diferente abordagem deve ser realizada. O desafio neste caso está em garantir o mesmo tempo de retenção do cabeçalho no nó para que a rede opere corretamente.

Entretanto o processamento do cabeçalho pode sofrer grandes variações de tempo de acordo com diversas variáveis como o número de *slots* destinados ao nó e o número de transmissões solicitadas pela unidade de transporte de dados. Em qualquer caso, quando o cabeçalho chega ao nó, o Controlador deve iniciar um temporizador antes de passar o cabeçalho para o Processador que, independentemente da variação, deverá devolver o cabeçalho já processado em um tempo menor que o máximo estipulado. Ao o receber de volta o Controlador manterá o cabeçalho no nó até que vença a temporização, o re-inserindo na rede imediatamente após este evento.

Para garantir que esta contagem seja fixa é necessária à utilização de um temporizador em conjunto com uma lógica que execute a comunicação com o Processador, a recepção e a re-inserção do cabeçalho na rede. Dado o requisito de tempo determinístico o controlador deve ser implementado em *hardware* evitando-se assim o uso de *software* para esta tarefa. Para tal algumas tecnologias podem ser utilizadas como: o uso de componentes digitais discretos, componentes de lógica programável ou componentes dedicados conhecido como ASICs (*Application Specific Integrated circuits*).

Por causa da complexidade e custo o projeto de um ASICs torna-se inviável para a aplicação na RCPO. Em contrapartida o emprego de componentes discretos como contadores e portas lógicas é de baixo custo e de fácil utilização. Porém uma vez realizado impossibilita a alteração do projeto de forma simples.

A prototipagem com componentes de lógica programável como FPGAs (*Field Programmable Gate Array*) e CPLDs (*Complex Programmable Logic Device*) apresenta-se como solução intermediária. Com boa flexibilidade de reprogramação, capacidade em termos de quantidade de elementos lógicos, velocidade de operação e custo aceitável o desenvolvimento com tais componentes torna-se ainda mais atrativo devido a possibilidade do uso de linguagens de descrição de hardware, como o VHDL (*VHSIC Hardware Description Language*) [28] que facilitam o desenvolvimento de elementos mais complexos diminuindo o tempo de projeto.

Para o transporte dos dados dos clientes, diversos pontos devem ser analisados para a escolha dos componentes capazes de atender os requisitos. Um dos primeiros aspectos é o interfaceamento entre os clientes e o nó. Como visto ele ocorre através de interfaces *Ethernet* que, por sua vez, é uma tecnologia muito difundida e de fácil acesso. Assim para a realização deste interfaceamento diversos componentes integrados estão disponíveis comercialmente fornecendo conectividade nas taxas de 10/100 Mbps, e 1/10 Gbps. Em geral é utilizando um conjunto de dois componentes integrados que realizam, respectivamente, a adaptação elétrica - conhecido como *chip* PHY (da palavra *physical*), e o controle de acesso ao meio, chamado de *chip* MAC.

Uma vez recebido os quadros dos clientes é necessário realizar a classificação, a montagem da fila e a transmissão dos pacotes ópticos. Estas operações apresentam o maior desafio do ponto de vista de implementação eletrônica no nó da RCPO pois, além de serem tarefas complexas, necessitam de desempenho. Tendo este desafio em mente algumas tecnologias se destacam como possíveis soluções.

Os *Network Processors* (NPs), por exemplo, são componentes integrados desenvolvidos especialmente para o processamento de pacotes. Em sua maioria são compostos por um processador de propósito geral que opera em conjunto com uma série de processadores específicos para o tratamento de pacotes. Coordenados por *software* que roda no microprocessador, os processadores de pacotes executam os procedimentos desejados.

Apesar de serem flexíveis os NPs apresentam limitações e podem não ser apropriados caso seja necessária a realização de operações muito específicas, como por exemplo a adição de campos especiais no quadro *Ethernet*.

Outro componente integrado que pode ser proposto neste caso é o *Switch On a Chip* (SOC), muito popular em equipamentos comutadores comerciais. Os SOCs possuem altíssimo desempenho e capacidade de comutação na ordem de dezenas de gigabits por segundo. Entretanto apresentam baixa flexibilidade não permitindo, por exemplo, a alteração no processo de encaminhamento de pacotes. Para que a utilização do SOC seja possível na RCPO é necessária à adição de componentes que implementem as funcionalidades específicas de uma rede de pacotes ópticos, como a agregação de pacotes em uma fila de transmissão, que operem em conjunto com o SOC.

Uma terceira possibilidade é a utilização de componentes de lógica programável também para esta função. Tais componentes apresentam desempenho e capacidade compatíveis com as necessidades e permitem total flexibilidade para a implementação de qualquer funcionalidade desejada. Entretanto projetar uma matriz de comutação pode ser uma atividade árdua e que eventualmente apresentará problemas de escalabilidade no acréscimo de interfaces de clientes e de anel.

A escolha entre NP, SOC ou lógica programável deve levar em conta principalmente as questões de simplicidade de implementação, funcionalidades desejadas, escalabilidade e desempenho, sendo necessário priorizar um ou dois desses pontos em relação aos demais visto que nenhuma tecnologia possui todas estas características.

Por fim o último desafio no que diz respeito ao transporte de dados é a transmissão do pacote óptico no anel. Dada a característica de rede de comutação óptica de pacotes um transmissor apenas se conecta a um receptor quando efetivamente existirem dados para serem transmitidos. Isso impede que a sincronização entre transmissor e receptor seja permanente, como ocorre com as interfaces de diversas tecnologias como o ATM [29] ou o próprio *Ethernet*. Nestes exemplos ocorre constantemente a troca de quadros especiais, chamados de quadros de *idle*, que mantém a sincronização entre as interfaces mesmo quando não existem dados para serem transmitidos.

No caso de redes de comutação de pacotes ópticas a sincronização deve ocorrer sobre demanda, ou seja, imediatamente antes da recepção dos dados dos usuários. Entretanto, dado o fato de o *slot* possuir duração de dezenas de microssegundos, a sincronização deve ocorrer em bem menos tempo (na ordem de poucos microssegundos), caso contrário, o *overhead* do pacote óptico será alto.

Para que isso ocorra as interfaces de transmissão no anel devem ser de implementação proprietária uma vez que nenhuma interface padrão possui tais características. Esta interface proprietária deve ser composta por um componente SERDES (*serializer/deserializer*) de rápida recuperação de relógio. Existem componentes integrados SERDES disponíveis comercialmente que atendem os requisitos acima mencionados na taxa de transmissão de gigabit e 10 gigabit por segundo. Desta forma a implementação da interface de transmissão do anel não se apresenta como um ponto proibitivo na construção do nó da RCPO.

5.1.2. Componentes ópticos

Ao contrário do processamento eletrônico, que abrange diversas atividades, apenas uma tarefa deve ser realizada no domínio óptico: a comutação dos pacotes ópticos.

Porém as opções de tecnologias ópticas disponíveis são enormemente inferiores quando comparadas à eletrônica. Esta escassez se acentua ainda mais dada à necessidade de se comutar pacotes com duração de poucas dezenas de microssegundos.

A seção A.6 apresenta as tecnologias de chaves ópticas existentes atualmente identificando as suas características como: o número de portas, perdas, *crosstalk*, e, principalmente, o tempo de comutação.

Nota-se que apenas dois tipos de chaves ópticas atendem o requisito de tempo da RCPO, sendo elas as chaves de Niobato de Lítio (LiNbO_3), e as chaves baseadas em SOAs (*Semiconductor Optical Amplifier*).

As chaves de Niobato de Lítio são em geral usadas como moduladores externos e possuem grande dependência da polarização da luz. Apesar de serem as chaves ópticas mais rápidas, com atuação na ordem de pico segundos, possuem maior custo em comparação com as chaves SOA. De fato o tempo de atuação desse tipo de chave está bem abaixo do necessário não se justificando em termos custo e a complexidade com relação ao SOA.

Assim o SOA operando em modo de chaveamento é a opção mais razoável para o desenvolvimento do comutador ópticos descrito na seção 4.2.2. Para maiores informações sobre o princípio de funcionamento do SOA consulte a seção A.5.2.

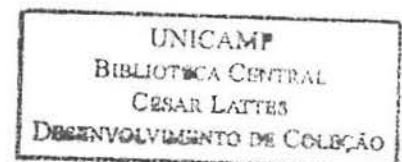
5.1.3. Tecnologias adotadas

Uma vez analisadas as tecnologias candidatas são apresentados agora os componentes escolhidos para a composição do nó da rede.

Dada a grande flexibilidade dos componentes de lógica programável a escolha para a implementação do bloco eletrônico da unidade de transporte de dados foi um componente FPGA de alta densidade, modelo Stratix, do fabricante Altera [30]. Esse componente possui internamente elementos de lógica programável, blocos de memória RAM, multiplicadores e PLLs (*Phase Locked Loop*) em quantidade necessária para a prototipagem dos sub-blocos descritos no capítulo 4.

Esse FPGA também é utilizado para a implementação da unidade de controle. Para tal é adotado um processador de 32 bits do tipo *softcore*. Um *softcore* é um componente eletrônico digital geralmente desenvolvido em linguagem de descrição de hardware que pode ser utilizado em diversas aplicações, sem que o desenvolvedor necessite conhecer as complexidades internas do componente. Um *softcore* possui diversas flexibilidades em comparação os componentes integrados como a possibilidade de ser removido ou alterado mesmo depois da finalização do projeto de *hardware* dada às próprias características de reconfiguração dos componentes de lógica programável.

Existem inúmeros *softcore* disponíveis hoje no mercado, realizando as mais diversificadas funções. Os *softcores* em geral são fornecidos gratuitamente pelos fabricantes de componentes de lógica programável ou comercializados por empresas especializadas neste tipo de desenvolvimento. O *softcore* de processador utilizado é chamado de NIOS-II [31] e é fornecido pelo fabricante Altera. Este possui uma arquitetura de 32 bits e pode atingir performance superior a 100 MIPS (*millions instructions per second*). Para o desenvolvimento do *software* que roda neste processador é utilizada a linguagem C através do uso de ferramenta de desenvolvimento e compilador também fornecidos pelo fabricante. Para a interconexão dos clientes *Ethernet* foi escolhido o componente integrado IXF1104 do fabricante Intel [32]. Este componente integra quatro dispositivos MACs com capacidade de transmissão de 1 Gbps. A figura 5-1 ilustra a interconexão do componente. Este componente é ligado ao FPGA via o barramento SPI-3 (*System Packet Interface Level 3*). Um microcontrolador de 8 bits é ligado ao IXF1104 para a configuração dos registradores no processo de inicialização do componente.



No caso dos clientes se conectarem por interfaces ópticas na taxa de 1 Gbps não é necessário a utilização de componentes PHY, sendo o IXF1104 ligado diretamente a *transceivers* SFP (*small form pluggable*), que farão a conversão eletro-óptica.

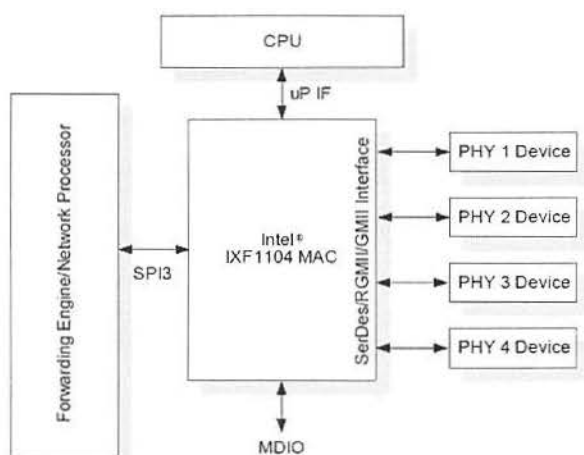


Figura 5-1 IXF 1104 e interconexões

Para a transmissão e recepção dos pacotes ópticos e dos cabeçalhos, o componente integrado TLK2201A do fabricante *Texas Instruments* [33] foi adotado. Este componente SERDES provê a rápida sincronização necessária para esse tipo de comunicação realizando a recuperação de relógio com poucas dezenas de bits de sincronização. Além desta característica a taxa de transmissão é de gigabit por segundo de forma a manter a compatibilidade com o tráfego gerado pelos clientes. É necessária a utilização de um componente por canal no anel óptico que também é conectado a um *transceiver* SFP a fim de realizar a conversão eletro-óptica.

Por fim a matriz de comutação óptica, descrita na seção 4.2.2, é implementada utilizando-se SOAs como chaves ópticas. Os SOAs além de possuírem capacidade de chaveamento na ordem de nanossegundos realizam a amplificação do sinal óptico compensando as perdas causadas pelos demais componentes ópticos do nó como: *delays* de fibra, *splitters* e *mux*.

A figura 5-2 mostra o encapsulamento de um componente SOA. Apesar de existirem SOAs disponíveis comercialmente os utilizados na RCPO foram encapsulados nos laboratórios do CPqD.

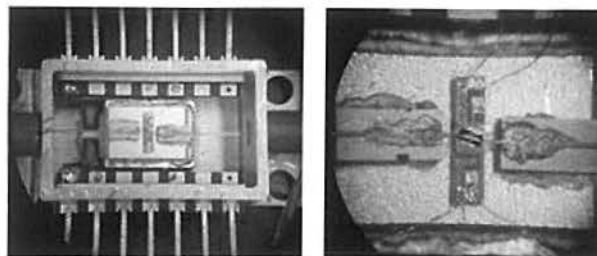


Figura 5-2 Encapsulamento do SOA

5.2. Projeto de implementação

Após o estudo das tecnologias e a escolha dos componentes inicia-se o processo de prototipagem do nó da rede. Em função da capacidade dos componentes adotados definiu-se que o nó prototipado possui quatro interfaces Ethernet ópticas operando na taxa de 1Gbps, destas interfaces três são utilizada pelos clientes e uma unidade de gerência. O nó pode possuir até dois anéis através de suas quatro interfaces de rede, sendo utilizadas duas interfaces por anel, uma para dados e a outra para os cabeçalhos.

5.2.1. Placa base e placas filhas

Um FPGA é utilizado para a implementação de parte da unidade de transporte de dados e da unidade de controle. Ao invés de se desenvolver uma placa de circuito impresso com um FPGA optou-se por utilizar uma das muitas placas disponíveis no mercado. A placa escolhida foi um dos modelos comercializados pela empresa PLDA[34] dado o FPGA Stratix de alta capacidade e performance existente nela. Esse componente possui 32 mil elementos lógicos e permite sinais digitais de até 125 MHz. Além disto, a placa possui um banco de 32 Mega bytes de memória RAM e uma grande quantidade de pinos disponíveis. A figura 5-3 exibe a placa adotada.

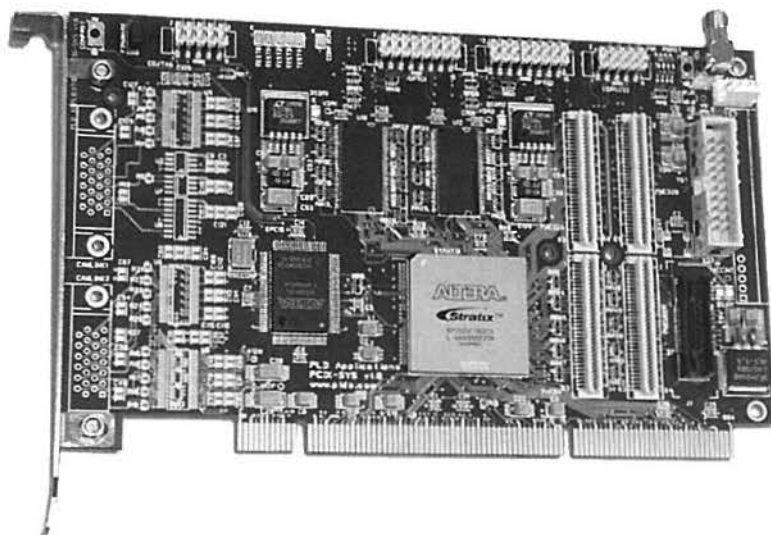


Figura 5-3 Placa do fabricante PLD

Além desta placa, chamada de placa base, foi necessário o desenvolvimento de duas placas filha que se conectam diretamente na placa base, e uma terceira placa para o controle da comutação óptica.

Os clientes *Ethernet* da rede se ligarão ao nó através do *chip* IXF1104. Assim tornou-se necessário o desenvolvimento de uma placa com esse componente. Esta primeira placa filha, chamada de placa MAC, se liga à placa base através do conector mezanino localizado na parte superior. A figura 5-4 ilustra a placa MAC.

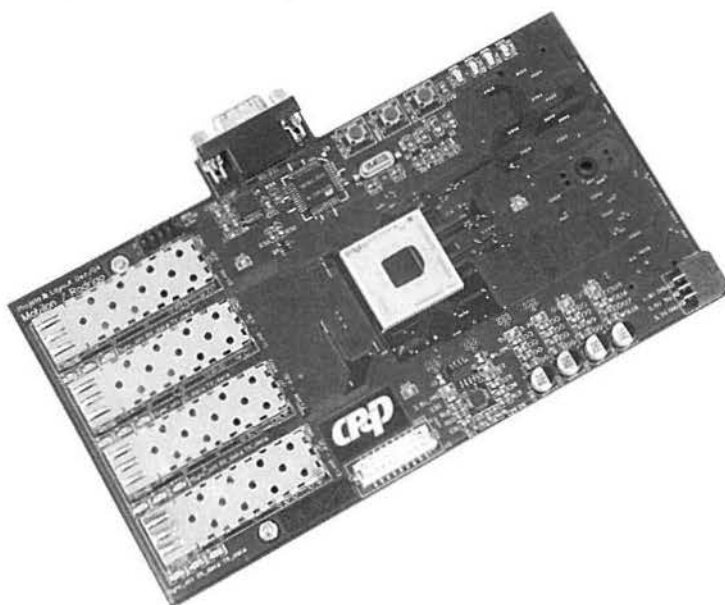


figura 5-4 Placa MAC

Para a transmissão no anel óptico é utilizado o *chip* TLK2201A. A placa desenvolvida, chamada de placa Serdes, realiza o acesso ao anel e é constituída de quatro desses componentes, permitindo assim a transmissão e recepção em quatro diferentes canais. A figura 5-5 ilustra a placa.

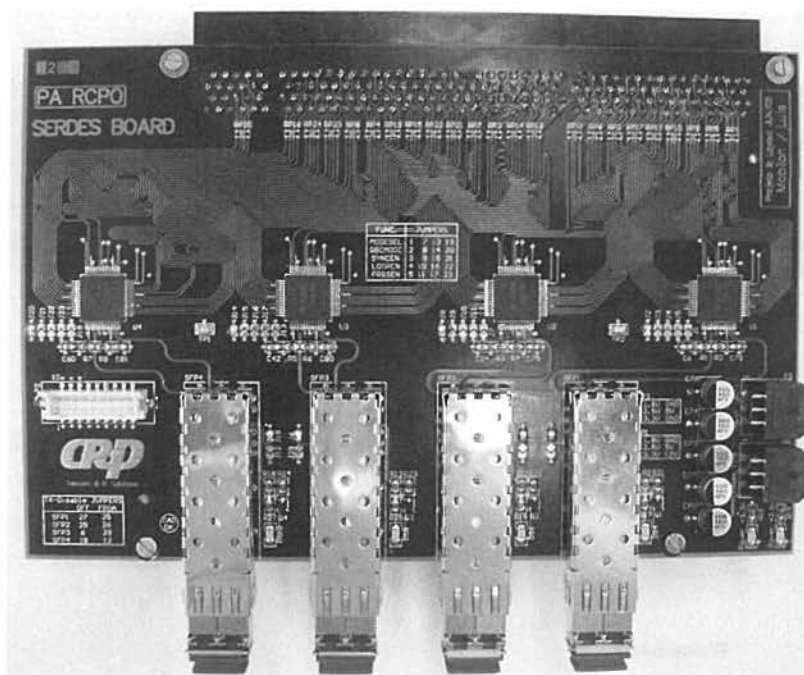


Figura 5-5 Placa Serdes

A placa Serdes se conecta a placa base através do conector PCI (*Peripheral Component Interconnect*), os comprimentos de onda utilizados no anel são selecionados através dos SFPs utilizados.

A última placa filha necessária para a criação do nó da rede é a placa SOA. Esta placa deve acomodar um componente SOA de forma a realizar as funções necessárias para o funcionamento correto do mesmo como o controle de temperatura e corrente de polarização. Para a criação da matriz de comutação óptica diversas placas SOA devem ser agrupadas de forma a atingir a configuração desejada. A figura 5-6 ilustra uma placa SOA. O controle desta ocorre através dos pinos dos conectores localizados na parte superior da placa base.

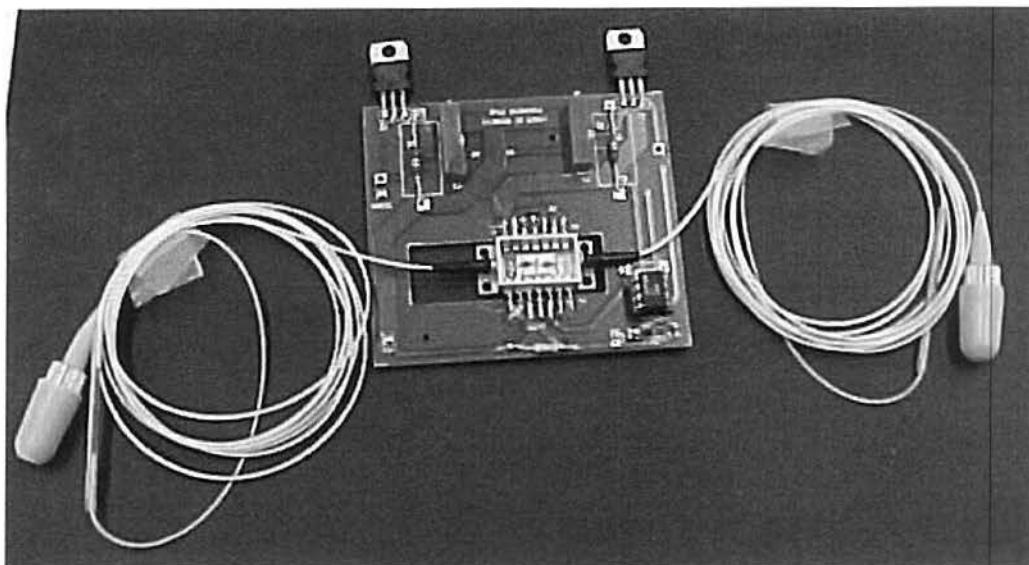


Figura 5-6 Placa SOA

Por fim a figura 5-7 ilustra como ocorre a ligação entre as placas filhas e a placa base. Para maiores detalhes sobre a construção das placas que constituem o nó da RCPO consulte[35].

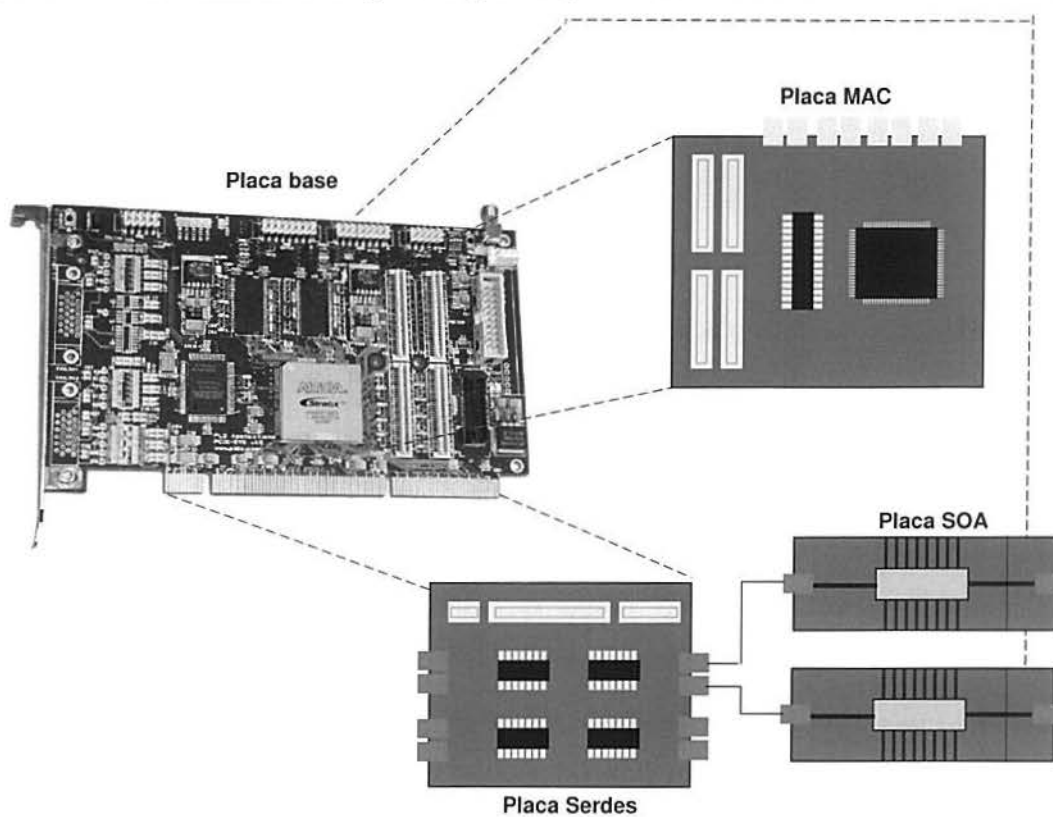


Figura 5-7 Ligação entre a placa base e as placas filhas

5.2.2. Blocos digitais

Uma vez abordado o desenvolvimento das placas que constituem o nó da rede o último passo da descrição do projeto de prototipagem constitui-se da explanação dos blocos digitais que compõe a unidade de controle e a unidade de transporte de dados. Estes blocos são implementados no FPGA da placa base em linguagem VHDL. A figura 5-8 ilustra os blocos.

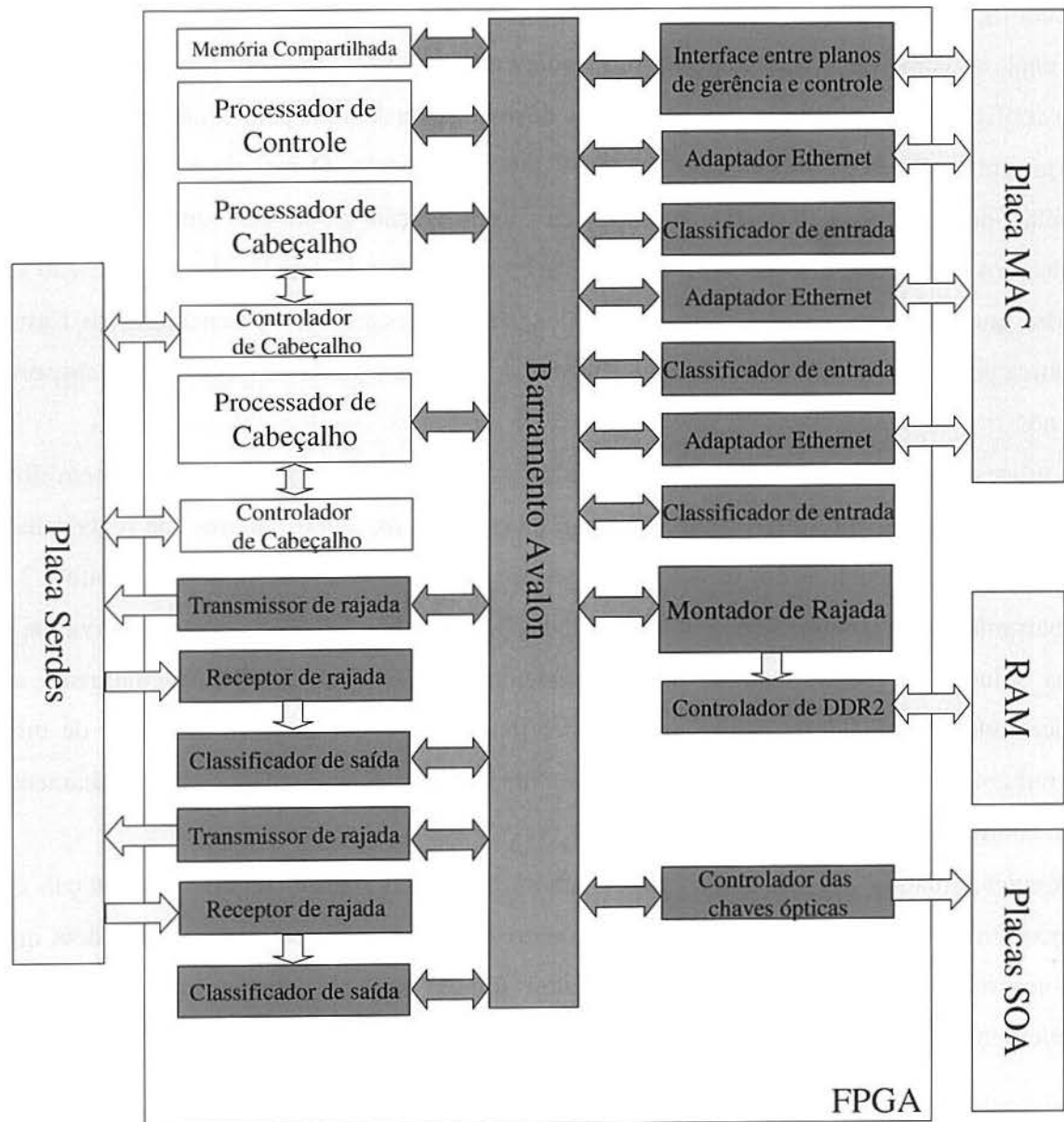


Figura 5-8 Blocos digitais

Unidade de controle e barramento de comunicação

Os blocos em azul da figura 5-8 compõem a unidade de controle do nó e seguem o comportamento descrito na seção 4.1. A unidade é composta por três processadores NIOS-II. O primeiro processador, chamado de processador de controle, realiza as funções de configuração inicial, preenchimento de tabelas de endereços e comunicação com o plano de gerência.

As configurações iniciais dizem respeito aos parâmetros da rede como: o tempo de rotação do anel, tamanho do *slot* e tempo de processamento de cabeçalho. Todos esses parâmetros são configurados de forma estática na rede e devem ser fornecidos pelo usuário diretamente na programação deste processador ou via plano de gerência. O mesmo acontece com as tabelas de endereçamento. Por motivos de simplificação as tabelas que mapeiam os endereços MAC dos clientes com as filas de transmissão e as tabelas de classificação de saída, que mapeiam os endereços MAC dos clientes locais, são preenchidas de forma estática pelo processador em cada classificador de entrada e saída na unidade de transporte do nó.

Os processadores de cabeçalhos são dedicados para cada anel da rede, assim existem dois processadores no protótipo que, em conjunto com os controladores de cabeçalho, processam os cabeçalhos dos dois anéis de acordo com os critérios definidos no capítulo 3.

O barramento de comunicação entre os blocos digitais, chamado de barramento Avalon, é uma solução de barramento de dados fornecida pelo fabricante Altera. Igualmente ao processador NIOS-II esta é uma solução do tipo *softcore* que permite a criação de uma barramento de multi-conexões onde diversos blocos podem trocar dados simultaneamente sem condições de concorrência.

Entretanto quando dois ou mais blocos tentam acessar o mesmo recurso a condição de concorrência é inevitável. Para tal o barramento Avalon possui árbitros que impedem que os acessos colidam. Desta forma o Avalon é uma solução eficaz que simplifica a implementação da comunicação das unidades de controle e transporte de dados.

Unidade de transporte de dados

Os blocos em verde da figura 5-8 compõem a unidade de transporte de dados do nó. Esse bloco eletrônico prototipado é um *subset* do bloco proposto no item 4.2.1. Por motivos de

simplificação não são implementadas algumas das funcionalidades descritas como, por exemplo, o controlador de tabelas de endereços.

A única forma de classificação implementada nesta versão do protótipo é a por endereços MAC. Outras formas de classificação como *vlan*s e *bits* de prioridade não são suportados pelo protótipo. Por consequência o montador de rajadas não suporta a separação do tráfego em sub-filas, não implementando a diferenciação de qualidade de serviço tanto para a classe de serviço de melhor esforço quanto para a classe de serviço reservada.

Para a agregação do tráfego em rajadas as filas controladas pelo montador são armazenadas no banco de memória RAM DDR-2 disponível na placa base. Para a realização do acesso à este banco de memória um controlador de RAM DDR-2 é utilizado. Este bloco também é um *softcore* fornecido pelo fabricante Altera.

Os blocos adaptador *Ethernet*, transmissor e receptor de rajadas, e comunicação entre planos de controle e gerência realizam as mesmas funcionalidades descritas na seção 4.2.1.

Por fim o controlador das chaves ópticas executa os comandos enviados pela unidade de controle nas chaves ópticas de acordo com a operação desejada.

Foge do escopo desse trabalho detalhar a construção interna desses blocos e os códigos de descrição de hardware desenvolvidos, para maiores detalhes sobre estes consulte [36].

5.3. Atual estado do nó prototipado

A fim de se realizar a prova de conceito foram prototipados três nós de rede. Como visto um nó constitui-se de uma placa base contendo a FPGA onde é implementado o plano de dados, uma placa MAC que realiza o interfaceamento dos clientes, uma placa Serdes para a interligação com a rede e placas SOA, no caso duas, para o chaveamento óptico. Por motivos de simplificação apenas um comprimento de onda é utilizado para o transporte de dados e apenas um cliente é ligado a cada nó. A figura 5-9 ilustra o nó prototipado.

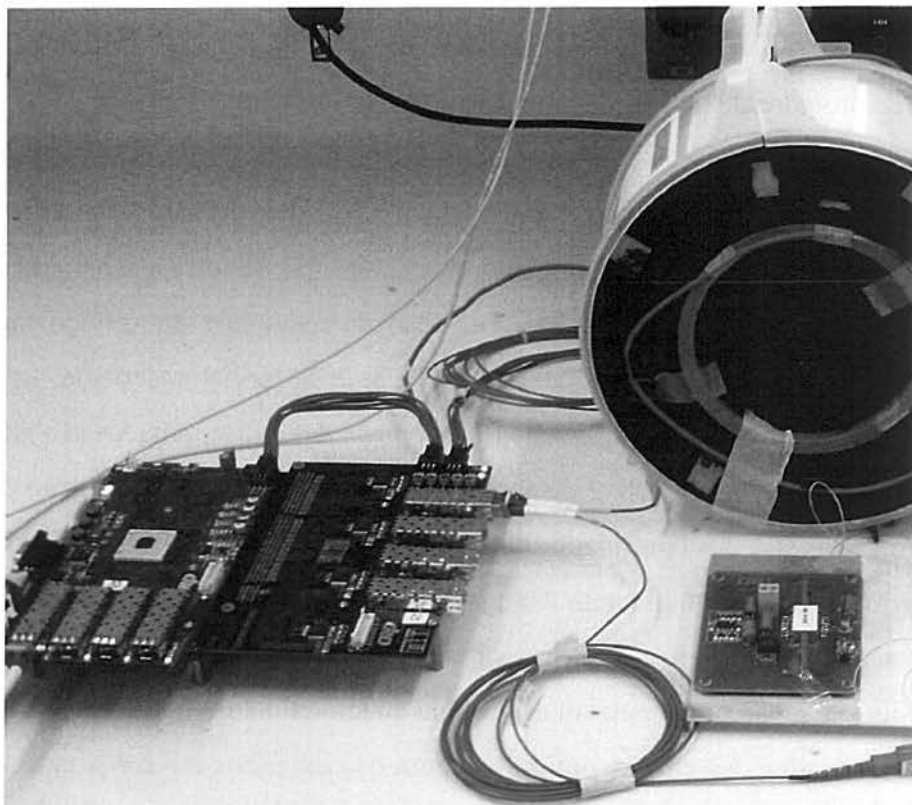


Figura 5-9 Nó prototipado

No atual status do projeto nem todos os blocos digitais descritos na seção 5.2.2 foram implementados. A figura 5-10 mostra quais blocos estão em funcionamento atualmente no nó da rede.

Apesar do protótipo do nó não estar totalmente concluído, os elementos existentes possibilitam a transmissão de dados entre os nós através do link óptico, como será visto na próxima seção.

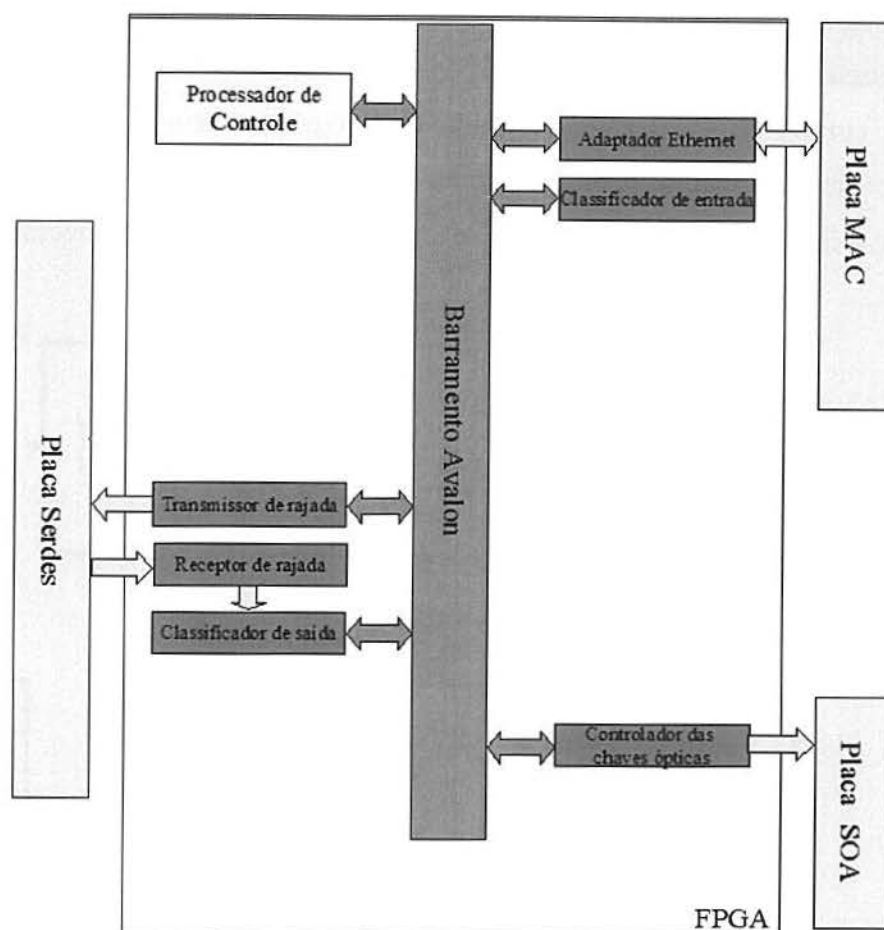


Figura 5-10 Blocos digitais implementados

5.3.1. Diagrama em blocos do nó prototipado

A figura 5-11 apresenta o diagrama em blocos do nó prototipado. Além dos blocos digitais mostrados na figura 5-10, que constituem um subconjunto da proposta apresentada na seção 4.2.1, a matriz de comutação óptica também não foi construída em total conformidade ao descrito na seção 4.2.2. Na proposta da figura 4-5 são utilizados três chaves SOAs. Entretanto na prototipagem optou-se por utilizar apenas duas chaves. Como consequência todas as rajadas, mesmo as não destinadas ao nó, serão recebidas por este. Neste caso o bloco classificador de saída possui a função de encaminhar ou não o pacote de acordo com o endereçamento. Assim quando tratar-se de um pacote destinado a um cliente de outro nó este bloco o descartará. Com este recurso no hardware eletrônico uma chave óptica é economizada.

Em contrapartida a chave SOA ligada a transmissão do nó deve ser sempre utilizada. Neste caso a necessidade se dá em função do módulo SFP sempre transmitir um sinal óptico conhecido como CW (Carrier Wave) mesmo quando não existem dados saindo do nó para a rede. Este sinal contínuo impediria o uso deste trecho do anel por outro nó, assim sendo é necessária a chave óptica para impedir a sua transmissão quando não é necessário.

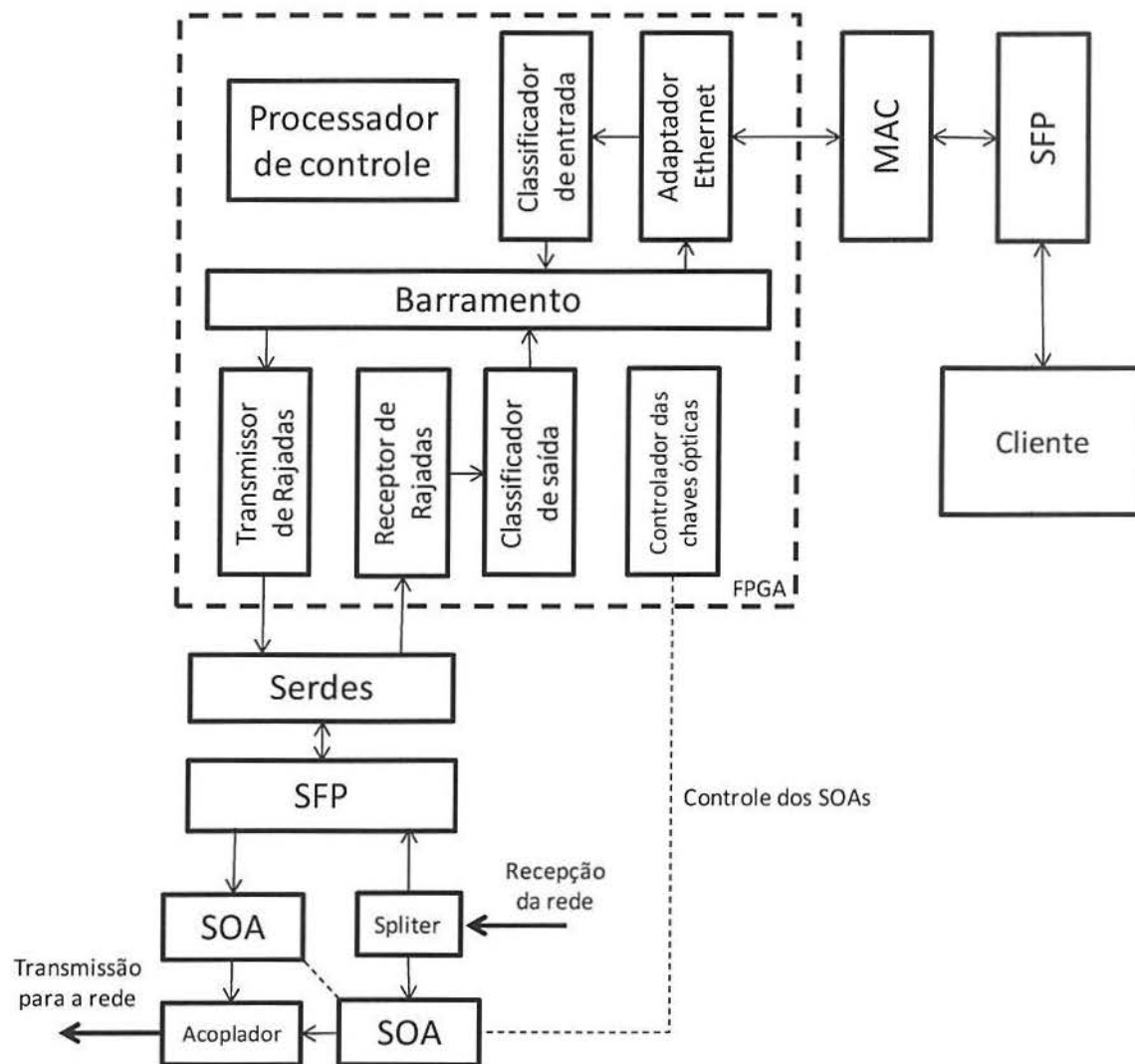


Figura 5-11 Diagrama em blocos do nó prototipado

5.4. Cenários de teste e resultados

A fim de se verificar o funcionamento dos nós prototipados e a validade de alguns dos conceitos propostos neste trabalho foram elaborados dois cenários de teste. O primeiro

constitui-se de dois nós ligados ponto a ponto e o segundo uma rede constituída de três nós. Ambos os cenários são apresentados em detalhes nas próximas seções.

5.4.1. Transmissão entre dois nós ligados ponto a ponto

O primeiro cenário de teste constitui-se de dois nós ligados ponto a ponto. O objetivo deste cenário é validar a operação do hardware construído, garantindo que o nó é capaz de transmitir e receber corretamente os dados que trafegam na rede óptica. A figura 5-12 ilustra o cenário.

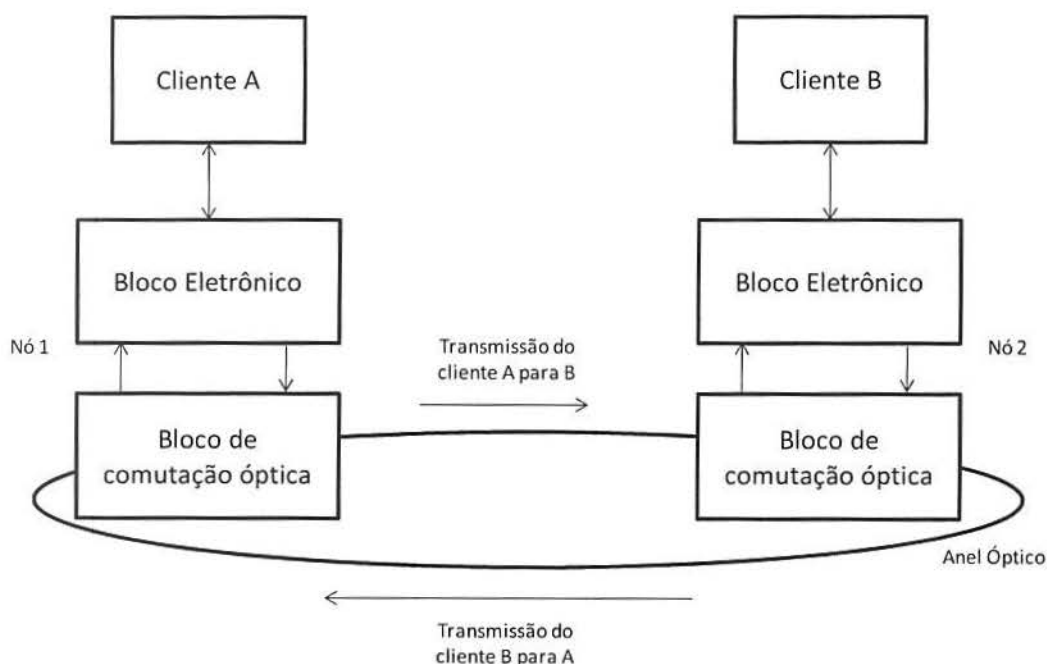


Figura 5-12 Cenário de teste com dois nós ligados ponto a ponto

Neste cenário há dois fluxos de dados, o primeiro é a transmissão do cliente A para o cliente B e o segundo a transmissão do cliente B para o cliente A.

Para tornar as transmissões possíveis é necessário que o processador de controle do nó 1 realize algumas configurações. Primeiramente é preciso associar na tabela de endereço do bloco digital classificador de entrada o endereço MAC do cliente B ao endereço interno do bloco digital transmissor de rajadas. Apesar de na atual implementação haver um único bloco transmissor de rajadas este é endereçável uma vez que se espera continuar o desenvolvimento do protótipo a fim de se atingir total conformidade com o nó descrito no

capítulo 4, neste caso haverá mais de um bloco deste tipo e o endereçamento torna-se necessário.

De fato o classificador deveria enviar o quadro ethernet a um montador de rajadas como especificado na seção 4.2.1 que por sua vez agruparia os quadros em rajadas seguindo os critérios definidos na seção 3.3. Entretanto não havendo montador de rajada na atual implementação o classificador de entrada foi desenvolvido para enviar todo o quadro que ingresse no nó diretamente ao montador de rajadas.

Como consequência será montada uma rajada com todos os campos definidos na seção 3.4.3 para cada quadro ethernet. Isto gera ineficiência na utilização do anel óptico devido ao *overhead* de cabeçalho. Entretanto isto é feito uma vez que o bloco transmissor de rajadas foi desenvolvido visando a especificação completa do nó, que contará com o montador de rajadas. Assim sendo quando houver este bloco o transmissor de rajadas já estará totalmente apto a operar conforme a especificação.

O próximo passo é configurar as chaves SOA. Para isto é necessário fechar a chave ligada ao transmissor do nó e abrir a chave ligada ao receptor. Desta forma o nó está apto a transmitir e receber dados na rede óptica.

Por fim o último passo é configurar a tabela de endereçamento do bloco classificador de saída associando o endereço MAC do cliente A ao bloco adaptador ethernet ligado a este cliente. Mais uma vez é realizado o endereçamento mesmo havendo apenas um bloco adaptador ethernet, visando-se assim a especificação completada do nó.

Estes mesmos passos realizados no nó 1 devem ser executados no nó 2 apenas mudando-se o endereçamento para que neste nó o cliente B seja o transmissor e o cliente A o receptor.

Cada cliente ligado a um nó constitui-se de um micro-computador rodando sistema operacional GNU/Linux. Foi utilizado o aplicativo VideoLAN em cada cliente para a transmissão e recepção de dois *streamings* de vídeo com áudio entre os clientes, o primeiro partindo do cliente A para o cliente B e o segundo do cliente B para o cliente A.

Resultados

Constatou-se que cada *streaming* chegou a taxa de 10 Mbps de pico não sendo observada nenhuma falha visualmente perceptível na recepção. Sendo os *streamings* transmitidos com protocolo de transporte UDP (User Datagram Protocol), ou seja, sem suporte a

retransmissão no caso de falha no recebimento, assume-se que os nós estão operando sem perder pacotes, ou com perdas muito baixas. Desta forma o resultado do teste é considerado válido uma vez que todos os elementos do hardware do nó prototipado estão operando como esperado.

5.4.2. Transmissão entre três nós ligados em anel

O primeiro cenário de teste teve o objetivo de validar o nó prototipado garantindo que cada elemento opere como o esperado. Passada esta etapa o próximo cenário visa construir uma rede com três nós como mostrado na figura 5-13.

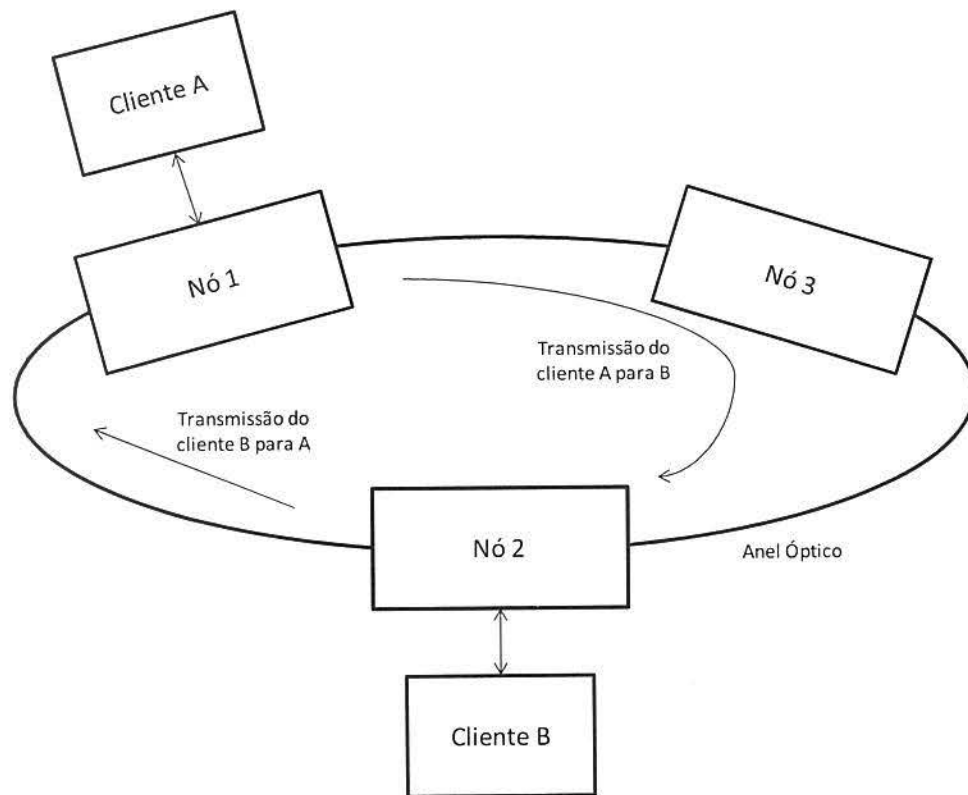


Figura 5-13 Cenário de teste com três nós ligados em anel

Neste cenário tanto o nó 1 quando o nó 2 possuem a mesma configuração descrita no cenário anterior, porém agora um terceiro nó é acrescentado. O objetivo é comprovar a funcionalidade de *bypass* que o nó da rede deve possuir. Desta forma os dados endereçados

a clientes ligados a outros nós que passem por este nó não necessitam de conversão optoeletrônica para atingirem o seu destino.

Para tal as chaves SOA do nó 3 devem estar configuradas de forma que a chave ligada a recepção esteja fechada e a chave ligada a transmissão esteja aberta. Assim os dados que são recebidos através da rede são simplesmente re-inseridos no anel sem modificação alguma.

Resultados

Os mesmos *streamings* utilizados no primeiro teste foram reutilizados neste novo cenário. Igualmente não foi observada nenhuma interrupção visualmente perceptível nas recepções. Assim o resultado do teste é considerado válido uma vez que a função de *bypass* se comportou da forma esperada.

5.5. Discussão

A prototipagem do nó é a prova de conceito para o modelo proposto ao longo desse trabalho. De fato a RCPO se apresenta como solução viável do ponto de vista de implementação apesar das muitas inovações que apresenta.

No início do capítulo diversas tecnologias, tanto eletrônicas quanto ópticas, foram analisadas a fim de se escolher qual melhor atenderia os propósitos da RCPO. Esta questão se mostrou particularmente difícil na escolha da tecnologia a ser adotada para o desenvolvimento do bloco eletrônico da unidade de transporte de dados. Como visto as três tecnologias candidatas: *Network Processor*, *Switch On a Chip* e lógica programável apresentam vantagens e desvantagens significativas.

A tecnologia de lógica programável, na forma de um componente FPGA, foi a escolha para esse projeto de implementação dada a sua principal característica – a grande flexibilidade na criação dos blocos necessários.

Entretanto esta mesma vantagem traz empecilhos no desenvolvimento. Nesta solução todas as funcionalidades necessárias para a implementação de um comutador de pacotes, desde as mais básicas, devem ser desenvolvidas do início, resultando em uma tarefa árdua. Além disto, tarefas comuns dos comutadores como a classificação de pacotes e a auto

aprendizagem de endereços que são realizadas de forma muito eficiente e sofisticada nos SOCs e NPs apresentam grande limitação no protótipo construído.

Da mesma forma a escalabilidade não é um ponto forte do protótipo sendo muito difícil a adição de novas interfaces de clientes ou novos canais no anel óptico uma vez que as placas de interfaceamento construídas não prevêm esta possibilidade.

Outro problema que se revelou ao longo do desenvolvimento foi a limitação no transporte de sinais no FPGA. Os sinais das interfaces gigabit, tanto na interligação com os clientes quando com os canais do anel óptico, trafegam pelo FPGA com uma frequência de barramento de 125 MHz. Esta frequência é o valor limite do componente adotado que apresentou grandes dificuldades no roteamento dos elementos lógicos por causa desta limitação, o que consumiu boa parte do tempo de projeto. Entretanto esse problema não se demonstra crítico uma vez que os novos componentes FPGA disponíveis no mercado suportam frequências internas na ordem de 500 MHz e ainda estão evoluindo neste aspecto. Por fim, os testes realizados indicam que as tecnologias escolhidas atendem os requisitos da rede, mesmo os nós não estando ainda totalmente concluídos.

Capítulo 6

Conclusões e Trabalhos Futuros

Este trabalho se propôs a descrever a arquitetura e prototipagem de uma rede de chaveamento pacotes ópticos em anel que também utilizasse conceitos de chaveamento de rajadas e comprimento de ondas.

Como visto no capítulo 2 existem diversos trabalhos de pesquisa ao redor do mundo com interessantes propostas de arquiteturas semelhantes, porém existe uma forte relação entre funcionalidades e complexidade, sendo as arquiteturas com mais funcionalidades aquelas que apresentam maior complexidade na implementação. Assim sendo no capítulo 3 foi apresentada a arquitetura da RCPO com as suas diversas funcionalidades, como: a divisão temporal do anel por *slots*, o mecanismo de proteção contra falha de caminho, a diferenciação por QoS do tráfego, classe de serviço reservada e *multicast* no domínio óptico. Estas características colocam a arquitetura da RCPO como uma das mais completas, além disto o recurso de montagem de rajadas e a utilização da classe de serviço transparente adicionam a arquitetura da RCPO características inerentes das redes OBS e OCS, a tornando uma arquitetura totalmente inovadora que mescla o conceito das redes de comutação de pacotes ópticos com as redes de comutação em rajadas e comutação em comprimento de onda.

Assim a primeira conclusão deste trabalho é que a arquitetura da RCPO é completa em termos de funcionalidades, estando preparada para atender os novos serviços de rede, além de apresentar um novo conceito unindo em uma única rede os três principais modelos de comutação óptica hoje existentes: OCS, OBS e OPS.

Os resultados de simulação, também apresentados no capítulo 3, levam a segunda conclusão do trabalho - os protocolos PAT+ e SAT+ propiciam uma eficiente utilização do anel óptico para diversos tamanhos de pacotes tanto para o cenário simétrico quanto

assimétrico, além de garantir a justiça no acesso ao anel entre os nós. Assim, a arquitetura da RCPO pode ser considerada não apenas completa e inovadora, como também eficiente no transporte de dados e justa para o acesso ao anel.

Entretanto utilizando apenas a especificação arquitetônica da rede não seria possível determinar o quão complexa é a sua implementação, e se de fato todas as funcionalidades propostas eram viáveis. Assim tornou-se necessário definir a arquitetura do nó da rede. Este trabalho foi realizado no capítulo 4. O primeiro passo foi definir e dividir dentro do nó as unidades de controle e de transporte de dados. Tendo-as definido a próxima ação foi descrever em detalhes os elementos que compõe cada unidade a fim de se atender todas as funcionalidades que cada unidade deve possuir. Como resultado chega-se a terceira conclusão deste trabalho – a RCPO possui uma arquitetura de nó bem definida que atende uma grande gama dos requisitos de rede e que aponta para as dificuldades na implementação, porém garantindo que o nó é totalmente viável com as tecnologias comercialmente disponíveis.

Terminado o trabalho de especificação iniciou-se o projeto de implementação. O objetivo deste foi analisar as possíveis tecnologias de componentes, tanto ópticos quando eletrônicos, que possuem as características necessárias para atender os requisitos da rede. Além desta questão também foi objetivo do projeto de implementação construir um protótipo que seja o mais flexível possível para assim facilitar o trabalho de prototipagem. Por exemplo, a decisão de se implementar em software os protocolos do plano de controle através de um processador de propósito geral permite que o desenvolvimento deste seja mais fácil, permitindo correções e modificações que se façam necessárias sem a necessidade de alteração no hardware. Em contrapartida a parte da unidade de controle que demanda rígidos tempos de sincronização, o que impede a sua implementação em software, foi realizada utilizando hardware digital em lógica programável. A própria decisão da utilização de elementos de lógica programável foi também baseada na idéia de alcançar alta flexibilidade, permitindo que os blocos funcionais possam ser implementados de forma incremental, além de permitir correções e melhoramentos ao longo do desenvolvimento.

Assim uma vez alcançado um protótipo que permitisse o tráfego de dados na rede foram iniciados os testes preliminares na RCPO. O primeiro teste visou garantir o funcionamento de todos os elementos do nó até então desenvolvidos. Já o objetivo do segundo teste foi a

construção de um primeiro cenário de rede utilizando três nós ligados em anel. Os resultados positivos destes dois testes levaram a quarta e última conclusão deste trabalho – as tecnologias escolhidas para a prototipagem atendem os requisitos da rede, além de apresentarem boa flexibilidade para permitir que a implementação seja incremental e para que correções e melhorias possam ser feitas ao longo do desenvolvimento.

Os trabalhos futuros ainda são diversos. Primeiramente é necessário concluir a prototipagem do nó, principalmente para que seja possível a realização da montagem de rajadas com todos os critérios propostos, e o controle de *slots* através do fluxo dos cabeçalhos trocados entre as unidades de controle.

Tendo em vista o protótipo completo outro importante trabalho é o desenvolvimento de cenários de testes que objetivam validar todas as funcionalidades da rede a fim de garantir que os resultados da simulação de desempenho se concretizem na rede desenvolvida.

Por fim também é necessário iniciar o trabalho da especificação e implementação do plano de gerência da RCPO, que visa apresentar uma visão unificada de todos os nós e permita realizar o controle integrado dos planos de controle e de dados.

Referências Bibliográficas

- [1] Scarabucci, R.R et al: Testbeds and Research Infrastructures for the Development of Networks and Communities, 2005. Tridentcom 2005. First International Conference on 23-25 Feb. 2005 Page(s):242 - 251 Digital Object Identifier 10.1109/TRIDNT.2005.29
- [2] Optical Networks Magazine, USA: SPIE, v. 3, n. 1, jan./fev. 2002. 92p.
- [3] Qiao, C.; Yoo, M: Optical burst switching (OBS) – a new paradigm for an optical Internet. Journal of High Speed Networks, Vol. 8, No. 1, pp 69-84, Maio 1999.
- [4] Optical Networks Magazine, USA: SPIE, v. 3, n. 6, Nov/Dez. 2002. 96p.
- [5] Martins, B. M.: Analise de tráfego, capacidade e proteção em redes de pacotes ópticos com chaveamento fotonico. Universidade Estadual de Campinas . Faculdade de Engenharia Elétrica e de Computação, 2007.
- [6] B. Mukherjee: WDM-based local lightwave networks part I: Single-hop systems. IEEE Network magazine, Vol. 6, No 3, pp. 12-27, Maio 1992.
- [7] I. Chlamtac, V. Elek, A. Fumagalli, L. G. Kazovsky, P. T. Poggiolini: A Contention/Collision Free WDM Ring Network for Multi Gigabit Packet Switched Communication. Journal of High Speed Networks, Vol. 4, No 2, pp 201-219, 1995.
- [8] P. Poggiolini, S. Benedetto, Theory of Subcarrier Encoding of Packet Headers in Quasi-All-Optical Broadband WDM Networks, Journal of Light Technology, Vol 12, No 10, pp 1869-1881, Out 1994.
- [9] C. S. Chen, B. S. Park, J. D. Shin, J. M. Jeong: A broadband ring network: multichannel optical slotted ring. Elsevier Computer Networks and ISD Systems, Vol. 27, No. 9, pp 1387-1398, 1995.

- [10] M. A. Summerfield: MAWSON: A Metropolitan Area Wavelength Switched Optical Network. Procs. Of 3rd Asia Pacific Conference (APCC'97), Vol 1, pp. 327-331, Sydney, Australia, Dez 7-10, 1997.
- [11] K. V. Shrikhande et al.: HORNET: A Packet-Over-WDM Multiple Access Metropolitan Area Ring Network. IEEE Journal on Selected Areas in Communication, Vol. 18, No. 10, pp 2004-2016, Out 2000.
- [12] Carena et al.: RINGO: A demonstrator of WDM optical packet network on a ring topology. Proc. 6th IFIP Working Conference on Optical Network Design and Modeling (ONDM), Torino, Itália, Fev 2002.
- [13] W. J. Goralski: Sonet/SDH. McGraw-Hill/OsborneMedia; 3rd edition. 687p, Oct. 4, 2002.
- [14] D. Dey, T. Koonen, M. R. Salvador: Network Architecture of a Packet-Switched WDM LAN/MAN. Proc. 5th Annual Symposium of the IEEE/LEOS Benelux Chapter, pp 256-259, Delft, Holanda, Out 30, 2000.
- [15] A. Farrel, I. Bryskin: GMPLS: Architecture and Applications. Morgan Kaufmann; illustrated edition edition, p. 400, Jan. 3, 2006.
- [16] M. R. Salvador, M. M. Uesono, N. L. Fonseca: Protocolos Fairness para Controle de Acesso ao Meio em Redes de Comutação de Pacotes com Topologia. Anel. Cadernos de Tecnologia (CPqD), Brasil, v. 1, n. 1, pp 143-154, 2005.
- [17] M. M. Uesono, N. L. Fonseca, M. R. Salvador: Protocolos de Justiça de Acesso em Redes Ópticas de Pacotes em Anel. In: XXIII Simpósio Brasileiro de Redes de Computadores, 2005. Anais do XXIII Simpósio Brasileiro de Redes de Computadores, 2005. pp 1-14.
- [18] M. R. Salvador: MAC Protocols for optical packet-switched WDM rings. Enschede, Holanda: Twente University Press, 2003. 192p. ISBN 90-365-1862-8.
- [19] I. Cidon, Y. Ofek: MetaRing – a full-duplex ring with fairness and spatial reuse. IEEE Transactions on Communications, Vol. 41, No 1, pp. 110-120, Jan 1993.
- [20] K. Pawlikowski, H. D. J. Jeong, J. S. R. Lee: On Credibility of Simulation Studies of Telecommunication Networks. IEEE Communication Magazine, Vol. 40, No. 1, pp 132-139, Jan 2002.

- [21] M. Matsumoto, T. Nishimura: Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. ACM Transactions and Modeling and Computer Simulation, Vol. 8, No. 1, pp 3-30, Jan.1998
- [22] K. Claffy, G. Miller, K. Thompson: The nature of the beast: recent traffic measurements from an Internet backbone. Internet Society (ISOC) Internet Summit (Inet), Genebra, Suíça, Jul. 21-24 ,1998.
- [23] IEEE 802.3-2002 IEEE Standard for Information technology – Telecommunication and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA-CD) Access Method and Physical Layer Specification.
- [24] IEEE 802.1Q-1998 IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks.
- [25] IEEE 802.1D-2004 Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges.
- [26] H. J. Gunn: The basics of IPTV. International Engineering Consortium; illustrated edition, p. 182, Maio 10, 2007.
- [27] R. C. Mayer: Linguagem C ANSI. McGraw-hill, p. 217, 1989.
- [28] S. Yalamanchili: VHDL: a starter's guide. Upper Saddle River: Pearson Education, p.241, 2005.
- [29] J. Edmonds: ATM network planning and implementation. International Thomson Computer, p.277, 1997.
- [30] Altera Corporation: <http://www.altera.com>.
- [31] Altera Corporation: Nios II Processor Reference Handbook, http://www.altera.com/literature/hb/nios2/n2cpu_nii5v1.pdf.
- [32] Intel Corporation: <http://www.intel.com>.
- [33] Texas Instruments: <http://www.ti.com>.
- [34] PLDA: <http://www.plda.com>.
- [35] E. Mobilon, et al.: Hardware Design and Prototype of an Optical Packet-Switched Ring Network Node. MWSCAS - The 49th IEEE International Midwest Symposium on Circuits and Systems, 2006, San Juan. IEEE, 2006.

- [36] E. Mobilon et al.: Hardware Architecture for Optical Packet and Burst Switching Applications. ITS - International Telecommunications Symposium, 2006, Fortaleza. IEEE, 2006.
- [37] C. H. Henry, G. E. Blonder, and R. F. Kazarinov. "Glass waveguide on silicon for hybrid optical packaging. J. Lightwave Technol, vol 7, pp. 1530-1939, 1989.
- [38] Koga, M.; Matsrnmoto, T. High-isolation polarization-insensitive optical circulator for advanced optical communication systems. Lightwave Technology, Journal of Volume 10, Issue 9, 1992 pp. 1210 – 1217.
- [39] R. Ramaswami, K. N. Sivarajan: Optical Networks – A Pratical Perspective. Morgan Kaufmann, Inc. USA, p. 632, 1998.
- [40] Rocha, M.L.; Kashyap, R.; Souza, R.F.; Paradisi, A.; Barros, M.R.X.; Coral, C.: System impact of the physical length of unapodized chirped fiber Bragg gratings on dispersion compensation. Microwave Theory and Techniques, IEEE Transactions on Volume 50, Issue 1, Part 1, Jan. 2002 pp. 88 – 93.
- [41] J. B. Rosolem; R. F. Silva: Comparação e Implementação de Analisadores de Canais Ópticos para Sistemas DWDM. Momag, Belo Horizonte, 2006.
- [42] B. Mukherjee: Optical WDM Networks. Springer; Jan. 2006.
- [43] M. J.F. Digonnet: Rare-Earth-Doped Fiber Lasers and Amplifiers. CRC; Mai. 2001.
- [44] M. J. Connelly: Semiconductor Optical Amplifiers. Springer; Jan 2002.

Apêndice A

Glossário de Componentes Ópticos

Os sistemas de comunicação óptica modernos utilizam uma série de componentes como acopladores, amplificadores, chaves, filtros e multiplexadores. Estes mesmos componentes são imprescindíveis para a criação das redes ópticas, e o entendimento destas está ligado ao conhecimento das funcionalidades dos elementos que as compõe. A seguir apresenta-se um glossário com a definição e uma pequena introdução sobre os principais elementos necessários para a criação das redes ópticas.

Acopladores

Um acoplador direcional é um componente utilizado para combinar diferentes sinais ou para dividir a potência de um mesmo sinal entre dois caminhos diferentes. Uma possível construção de um acoplador é dada pela fusão de duas fibras. Outra possibilidade é através de guias de ondas utilizando a tecnologia de óptica integrada [37]. A figura A-1 apresenta um acoplador 2x2 que utiliza uma porcentagem da potência do sinal de entrada na porta 1

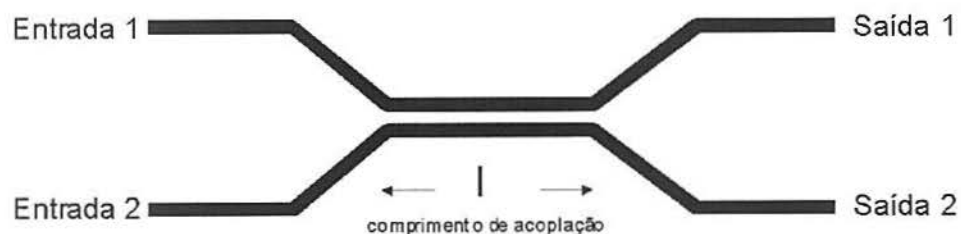


Figura A-1 Acoplador óptico

Isoladores e Circuladores

Isoladores e circuladores [38] são exemplos de componentes passivos não recíprocos, ou seja, componente cujo efeito físico muda de acordo com o sentido de propagação da luz no componente. Isoladores permitem a passagem da luz em apenas um sentido, bloqueando a passagem desta no sentido oposto. Circuladores igualmente permitem a passagem da luz em apenas um sentido, porém os circuladores também permitem a remoção e adição dos sinais com baixa perda através de diversas portas, como mostrado na figura A-2.

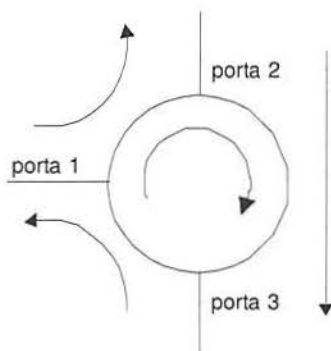


Figura A-2 Circulador óptico

Filtros

Filtros ópticos são elementos capazes de filtrar um dado comprimento de onda ou uma faixa espectral de um sinal WDM. São abordados a seguir dois modelos clássicos de filtros ópticos: o filtro com grade de Bragg e o filtro de Fabry-Perot. Atualmente existem muitos outros tipos de filtros, mais complexos e de melhor desempenho, como o TFF (*Thin-Film Filter*) e o AWG (*Arrayed Waveguide Grating*), porém foge do escopo abordá-los. Para informações mais detalhadas sobre filtros ópticos consulte [39].

Filtro com grade de Bragg

Grades de fibra óptica [40] são componentes interessantes para diversas aplicações como filtragem, remoção/adição de canais e compensação de distorção em sistemas de transmissão. Para todas as aplicações acima mencionadas a grade executa basicamente a mesma operação: a reflexão de uma faixa espectral.

De maneira geral, qualquer perturbação periódica no meio de transmissão resulta em uma grade de Bragg. A figura A-3 ilustra o conceito.

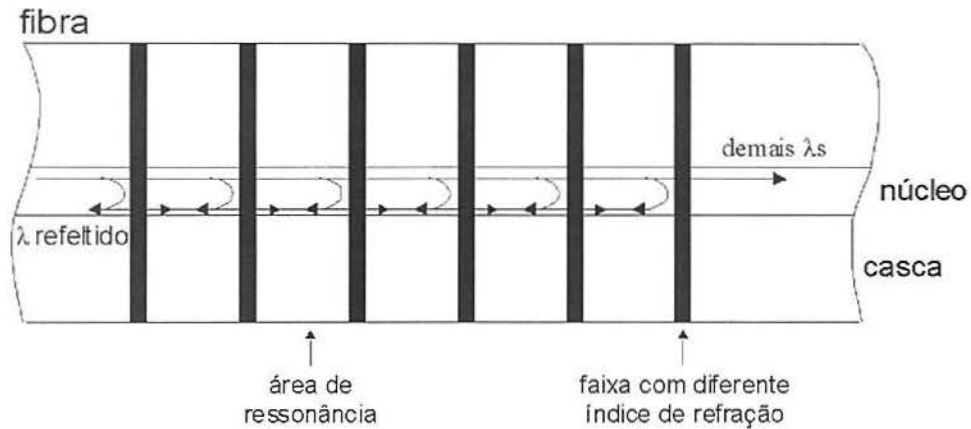


Figura A-3 Grade de Bragg

Filtro de Fabry-Perot

Um filtro de Fabry-Perot consiste de uma cavidade formada de dois espelhos de alta refletividade postos em paralelo, como mostrado na figura A-4. Este tipo de filtro também pode ser chamado de interferômetro de Fabry-Perot ou etalon.

Uma interessante aplicação do filtro de Fabry-Perot para a construção de um OSA (*Optical Spectrum Analyzer*), pode ser encontrada em [41].

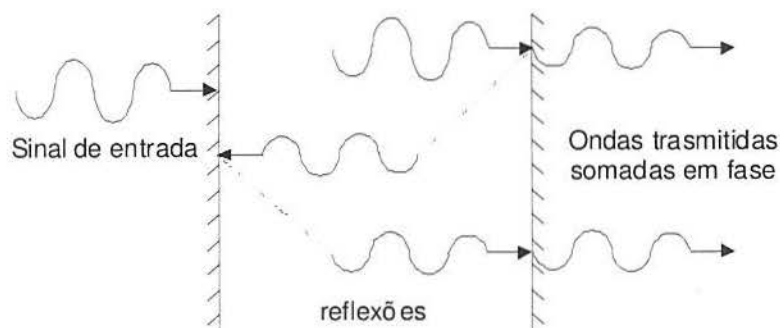


Figura A-4 Filtro de Fabry-Perot

Multiplexadores, Demultiplexadores, OADMs e roteadores de comprimentos de onda

Um multiplexador tem a função de direcionar diferentes comprimentos de ondas que entram em portas distintas, a uma mesma saída. Já um demultiplexador por sua vez separará os canais que entram por uma única porta para as demais portas de saída [42].

Com o arranjo de multiplexadores e demultiplexadores é possível criar outros componentes como o roteador de comprimento de onda. Este elemento traz uma grande flexibilidade para as redes ópticas e permite a comutação de canais no domínio óptico, sem a necessidade da conversão elétrica. A figura A-5 mostra um roteador de comprimentos de onda estático. Como o uso de chaves ópticas é possível criar um elemento que realize a comutação de forma dinâmica, de acordo com a necessidade momentânea da rede. As chaves ópticas são apresentadas na sequência.

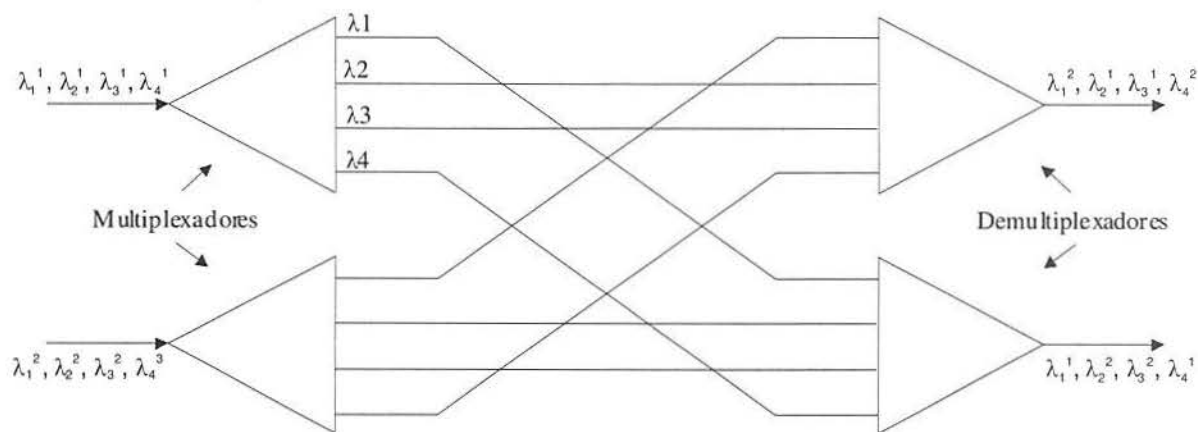


Figura A-5 Roteador de comprimentos de onda

Além da comutação dos canais, outras duas operações fundamentais são a inserção e remoção de canais ao longo da rede (*add and drop*). Para tal existe o componente OADM, um exemplo de OADM é mostrado na figura A-6, e utiliza um circulador em conjunto com uma grade de Bragg e um acoplador.

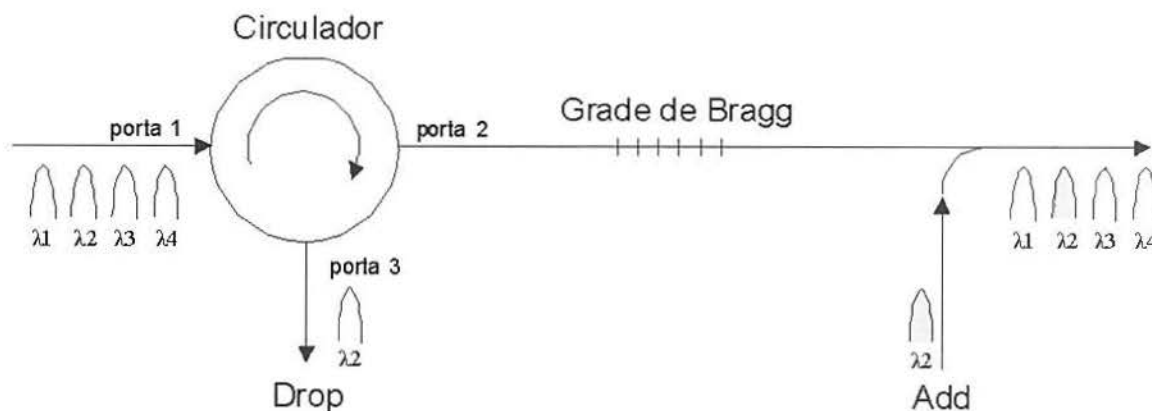


Figura A-6 Add and Drop

Amplificadores ópticos

Antes do surgimento dos amplificadores ópticos a única solução para este problema era a utilização de regeneradores eletro-ópticos. Os amplificadores ópticos possuem diversas vantagens com relação aos regeneradores. Por exemplo, eles não são dependentes da taxa de transmissão ou do tipo de modulação usado na transmissão. Enquanto a mudança nestes dois quesitos implicaria na troca de todos os regeneradores, os amplificadores não sofreriam alterações. O regenerador tem o seu custo diretamente associado à taxa de transmissão utilizada, tornando-se mais caro à medida que esta aumenta. Já os amplificadores não possuem esta relação, tendo capacidade de transmissão muito mais elevada.

A principal vantagem do amplificador está na sua utilização nos sistemas WDM. Um único amplificador óptico é suficiente para amplificar todos os comprimentos de onda transmitidos por uma fibra.

Em contra partida um amplificador óptico também possui algumas desvantagens. Com os amplificadores é possível apenas realizar a amplificação do sinal e não a correção na forma e no tempo do bit, ao contrário dos regeneradores que possuem esta capacidade. Outra característica negativa é a adição de ruído ASE (*Amplified Spontaneous Emission*). O ruído ASE é similar ao ruído branco na transmissão eletrônica, sendo composto de componentes em todo o espectro de frequência. A sua existência é prejudicial à recepção do sinal.

EDFA

O amplificador a fibra dopada a érbio EDFA (*Erbium-Doped Fiber Amplifier*) [43] consiste de um dado comprimento de fibra dopada com érbio (Er^{3+}). Esta fibra recebe o sinal de um laser de bombeio assim como o sinal que se deseja amplificar. Em geral o laser de bombeio possui comprimento de onda de 980 nm ou 1480 nm enquanto que o sinal transmitido está na faixa de 1550 nm. Através do princípio da emissão estimulada novos fótons coerentes com o sinal de entrada serão gerados, resultando na amplificação.

SOA

O amplificador óptico a semicondutor SOA (*Semiconductor Optical Amplifier*) [44] não se mostrou tão eficiente para o uso na amplificação de sinais como o EDFA. No entanto este

dispositivo apresenta outras características que o tornam interessante para o uso em redes ópticas, principalmente exercendo a função de chave óptica e conversor de comprimento de onda. A atuação do SOA como chave pode atingir tempos na ordem de 1ns, enquanto no EDFA este tempo é de 10ms.

Chaves ópticas

Chaves ópticas são elementos amplamente utilizados em redes ópticas. A escolha de uma chave óptica para uma certa aplicação está principalmente relacionada ao seu tempo de chaveamento. A tabela A-1 apresenta os tempos de chaveamento requeridos para as principais necessidades de um sistema óptico.

Aplicação	Tempo de chaveamento requerido
Aprovisionamento	10 - 100ms
Proteção	1 – 10ms
Comutação de pacotes ópticos	100ns
Modulação externa	10ps

Tabela A-1 Tempo de atuação das chaves em função da aplicação

Em função do tempo de chaveamento e de outras características como número de portas e perda deve-se selecionar o tipo de chave óptica que atenda uma determinada demanda. A tabela A-2 apresenta as características típicas de algumas tecnologias de chaves ópticas.

Tipo	Portas	Perda (dB)	Crosstalk (dB)	Tempo de chaveamento
Mecânica/ Termo-óptica	8 x 8	3	55	10 ms
Eletro-óptica	8 x 8	10	30	2 ms
LiNbO ₃	4 x 4	8	35	10 ps
SOA	4 x 4	0	60	1 ns

Tabela A-2 Características das chaves ópticas