

Arquitetura e Protótipo de Rede de Pacotes em Anel com Transmissão e Comutação no Domínio Óptico

¹Marcos R. Salvador, ¹Eduardo Mobilon, ¹Vinicius G. de Oliveira, ¹Rodrigo Bernardo, ¹Luis R. Monte, ¹Mateus M. Pereira, ¹Leonardo Pezzolo, ¹Maria F. M. B. B. F. de Oliveira, ¹Antonio D. Coral, ¹Alberto Paradisi, ²Furio Damiani, ²Peter Jürgen Tatsch, ²Nelson L. S. da Fonseca, ²Marcelo M. Uesono

¹Fundação CPqD Centro de Pesquisa e Desenvolvimento em Telecomunicações
13086-902 – Campinas – SP – Brasil

²Universidade Estadual de Campinas (Unicamp)
13084-971 – Campinas, SP, Brasil

{marcosrs, mobilon, vgarcia, bernardo, lmonte, mpereira, pezzolo, falcão, acoral, paradisi}@cpqd.com.br
{furio, peter}@dsif.fee.unicamp.br
{fonseca, marcelo.uesono}@ic.unicamp.br

***Abstract.** This paper presents a novel network technology for the next technological transition in telecommunications. Using burst assembly and optical packet switching techniques and optical networking technologies available in the market, the network offers high performance, efficiency and flexibility at low cost. The paper describes the architecture of the network as well as the modules being prototyped for proof of concept. Performance results obtained from computational simulation experiments are shown and discussed.*

***Resumo.** Este artigo apresenta uma tecnologia de rede inovadora voltada para a próxima transição tecnológica das telecomunicações. Usando técnicas de montagem de rajadas e de comutação de pacotes no domínio óptico e tecnologias maduras disponíveis no mercado, a rede oferece desempenho, eficiência e flexibilidade com baixo custo. Este artigo descreve a arquitetura desta rede assim como os módulos desta arquitetura que estão sendo prototipados para prova de conceito. Resultados de desempenho obtidos em simulação computacional são apresentados e discutidos.*

1. Introdução

O reconhecimento da necessidade de novas arquiteturas e tecnologias de rede tem levado muitas universidades e muitos centros de pesquisa e desenvolvimento de excelência ao redor do mundo a investirem na busca de novas arquiteturas e tecnologias de rede para a próxima transição tecnológica das telecomunicações. Dentre as arquiteturas de rede em estudo, destacam-se as que maximizam o uso das tecnologias ópticas para reduzir o chamado gargalo eletrônico. Essas arquiteturas representam uma mudança do 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 rede óptica em estudo, porém de uma forma geral elas podem ser classificadas de acordo com as técnicas de comutação óptica que utilizam. São elas: comutação de comprimento de onda (*wavelength routing*), que equivale à comutação de circuito, comutação de pacote (*optical packet switching - OPS*) e comutação de rajada (*optical burst switching - OBS*). Todas elas apresentam vantagens e desvantagens, porém as mais inovadoras têm se concentrado nas duas últimas em vista da maior flexibilidade que elas oferecem.

Este artigo apresenta uma tecnologia de rede óptica inovadora para a próxima transição tecnológica das telecomunicações que desde 2004 está em desenvolvimento nos laboratórios do CPqD em colaboração com grupos universitários. Denominada BlasteRing (*Burst and Lambda Services for a Transparent and Ethernet services Ring*), esta arquitetura de rede em anel combina técnicas de montagem de rajadas com comutação óptica de pacotes e apoia-se em tecnologias ópticas de rede disponíveis comercialmente a preços não proibitivos. Especificamente, este artigo descreve a arquitetura da rede concebida e os protótipos dos módulos de hardware que compõem esta rede. O artigo também descreve um dos protocolos de controle de acesso ao meio (*medium access control - MAC*) desenvolvidos neste trabalho e analisa resultados de desempenho obtidos em um simulador computacional que modela este protocolo sobre a arquitetura da rede proposta bem como o desempenho dos módulos de hardware prototipados.

O restante deste artigo está organizado da seguinte maneira: na seção 2 descreve-se a arquitetura da rede. Na seção 3 descreve-se a arquitetura do nó da rede e os protótipos desenvolvidos a partir dos módulos que compõem esta arquitetura. Na seção 4 descreve-se o protocolo MAC e alguns aspectos da sua implementação. Na seção 5 apresenta-se e discute-se os resultados de desempenho da rede obtidos em simulações computacionais. Na seção 6 finaliza-se o artigo com alguns comentários.

2. A Rede

A arquitetura da rede foi concebida para aplicações metropolitanas (embora não esteja limitada a essas aplicações), onde os requisitos dos usuários e das aplicações fazem-se bastante presentes e, por esta razão, onde provavelmente estará o próximo gargalo das redes de telecomunicações.

Em virtude do foco em aplicações metropolitanas, a arquitetura da rede baseia-se na topologia de anel, bastante comum nessas aplicações, o que além de contribuir para a simplicidade e baixo custo de operação da rede também facilita a aceitação da tecnologia proposta em virtude do reaproveitamento das infra-estruturas existentes. A recente padronização da tecnologia *Resilient Packet Ring (RPR)* [802.17-2004 2004] pelo IEEE e a oferta de produtos baseados nesta tecnologia bem como os trabalhos de evolução do SDH (SDH-NG) confirmam o interesse nas vantagens da topologia em anel e das tecnologias nela baseadas.

Para atender aos requisitos das aplicações metropolitanas, a arquitetura da rede suporta serviços Ethernet e serviços transparentes do tipo um-para-um e um-para-muitos sob demanda. Especificamente, a arquitetura oferece os serviços Ethernet de melhor esforço, Ethernet reservado e Transparente reservado. O foco nos serviços Ethernet e o suporte a serviços transparentes conferem à arquitetura a flexibilidade necessária para

atender a tecnologia (Ethernet) na qual o transporte tende a convergir e outras, sejam elas legadas (ex.: SDH) ou ainda inexistentes.

O mecanismo de proteção usa o roteamento alternativo no nó origem para recuperação rápida da falha e, ao contrário de SDH, não impede o uso de um anel de proteção enquanto este não for necessário na ocorrência de uma falha no anel de trabalho correspondente.

A Figura 1 ilustra a arquitetura da rede, que consiste de N nós conectados por F pares de anéis de fibra óptica do mesmo tamanho, cada um com L comprimentos de onda (ou lambdas) habilitados por meio da tecnologia de multiplexação por divisão de comprimento de onda (*wavelength division multiplexing - WDM*). Desses L comprimentos de onda, L-1 são usados para o transporte de dados (isto é, pacotes ópticos) e um é usado para transporte de pacotes de controle e dos cabeçalhos dos pacotes ópticos. A transmissão do cabeçalho do pacote em comprimento de onda separado é necessária para permitir o processamento (eletrônico) que resultará na decisão acerca do que fazer com o pacote propriamente dito -o pacote óptico está indisponível para processamento.

Para proteção e para flexibilidade de roteamento, as transmissões nos comprimentos de onda de um anel de fibra seguem uma direção, ao passo que as transmissões nos comprimentos de onda do outro anel de fibra seguem na direção contrária.

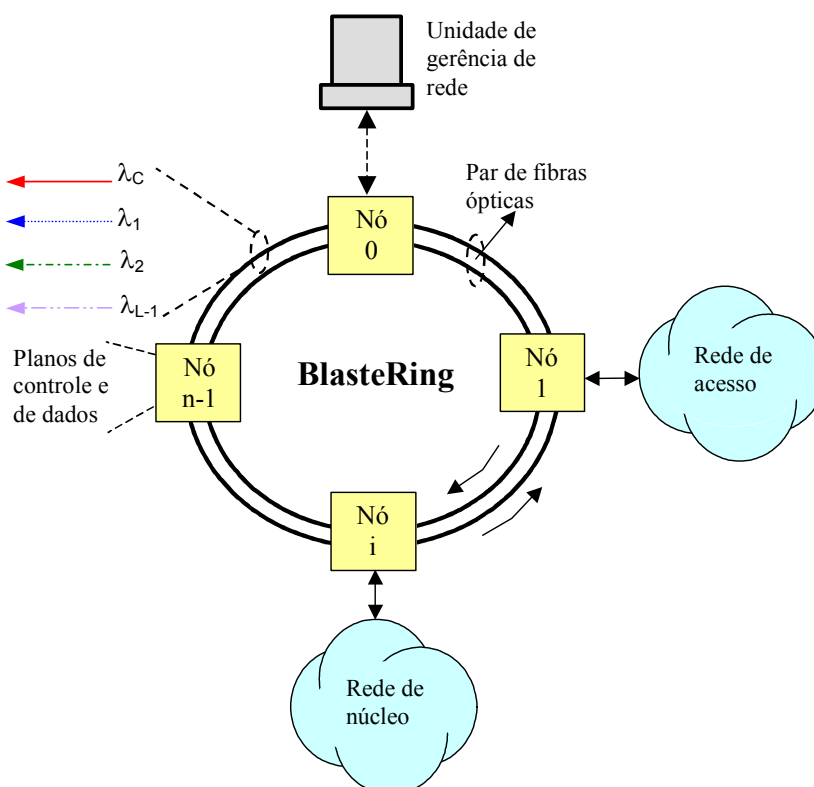


Figura 1 - Rede Óptica de Pacotes

O protótipo da rede em desenvolvimento consiste de quatro nós e dois anéis de fibra de 100km de extensão, cada um contendo dois comprimentos de onda: um para o

transporte de informações de controle e outro para o transporte de dados. Na primeira versão somente o plano de dados será implementado, pois é a parte fundamental para provar o conceito.

3. O Nó da Rede

A Figura 2 ilustra a arquitetura do nó da rede. Esta arquitetura consiste de uma unidade de plano de controle (UPC), uma unidade de montagem de rajada e encaminhamento (UMRE), que é o coração do nó da rede, e uma matriz de comutação óptica de pacotes (MCOP). A UPC é responsável pela operação automática da rede, o que inclui a manutenção de tabelas diversas (ex.: roteamento, recursos) e o atendimento às solicitações de serviços dos usuários. Esta unidade disponibiliza uma interface aos usuários (*user-network interface* – UNI) para que estes possam solicitar os serviços da rede dinamicamente. Esta unidade será implementada na próxima versão do protótipo da rede usando linguagem C em ambiente Linux PC e seguirá o padrão *Generalized MultiProtocol Label Switching* (GMPLS) [Mannie 2004].

Do lado do emissor, a UMRE é responsável pela montagem das rajadas a partir dos quadros Ethernet provenientes dos clientes e pela transmissão dessas rajadas, como pacotes ópticos, aos seus destinatários. Do lado receptor, a UMRE é responsável por extrair e segregar os quadros Ethernet da rajada recebida (como pacote óptico) do anel e por transmitir esses quadros aos seus clientes.

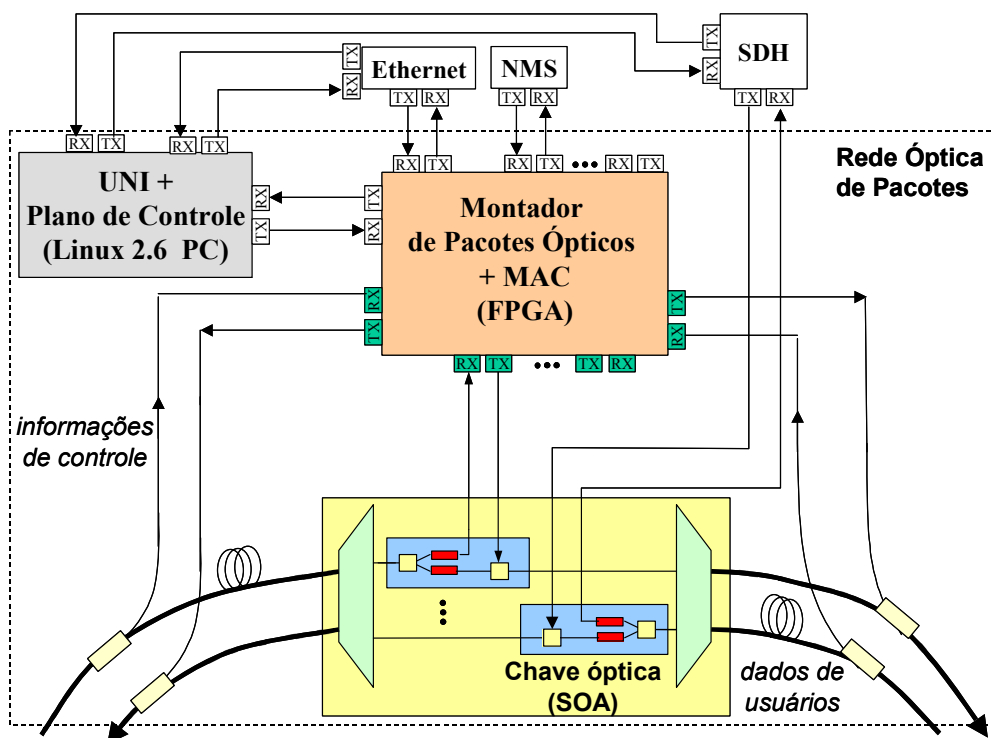


Figura 2 - Arquitetura do Nó da rede

A UMRE contém um conjunto de filas virtuais de saída (*virtual output queuing* - VOQ) criadas, em tempo de inicialização, para serviços Ethernet de melhor esforço e também um número adicional de filas criadas sobre demanda do plano de controle para serviços Ethernet reservado; o tráfego dos serviços transparentes não passa pela UMRE.

Pacotes provenientes de redes clientes são armazenados nessas filas de acordo com a classe de equivalência de encaminhamento (*forwarding equivalence class* - FEC) a que pertencem, sendo que pacotes que pertencem à mesma FEC são armazenados na mesma fila. Dois pacotes pertencem à mesma FEC quando apresentam certas características comuns (ex.: mesmo nó destino, mesmo nó destino e mesma prioridade).

Periodicamente, o montador de rajadas inspeciona todas as filas procurando por aquelas que atendem a pelo menos um dos critérios de montagem: i) tamanho da fila igual ou maior que o tamanho definido de rajada; ii) tempo de permanência em fila do primeiro pacote da fila (*head of the line* – HOL) igual ou maior que o período de montagem de rajada. O resultado deste processo é uma lista de requisições de transmissão, sendo uma para cada rajada montada. O módulo MAC interno ao UMRE, responsável por controlar o acesso aos anéis ópticos, então decide quais rajadas podem ser transmitidas e qual rota e anel de fibra e comprimento de onda cada transmissão deve usar. Esta decisão leva em conta as determinações do plano de controle armazenadas nas diversas tabelas mantidas pelo nó.

O hardware eletrônico que implementa a UMRE consiste de uma placa principal e de duas placas filhas (veja Figura 3 e Figura 4). A placa principal dispõe de um componente de lógica programável (*field programmable gate array* - FPGA) de alta densidade que permite a implementação de todo um circuito digital composto por máquinas de estado e lógica combinacional. Esse circuito digital é extremamente complexo para ser desenvolvido como um bloco único e, portanto, foi dividido em diversos blocos funcionais que estão sendo desenvolvidos de forma paralela, integrados e então programados no dispositivo.

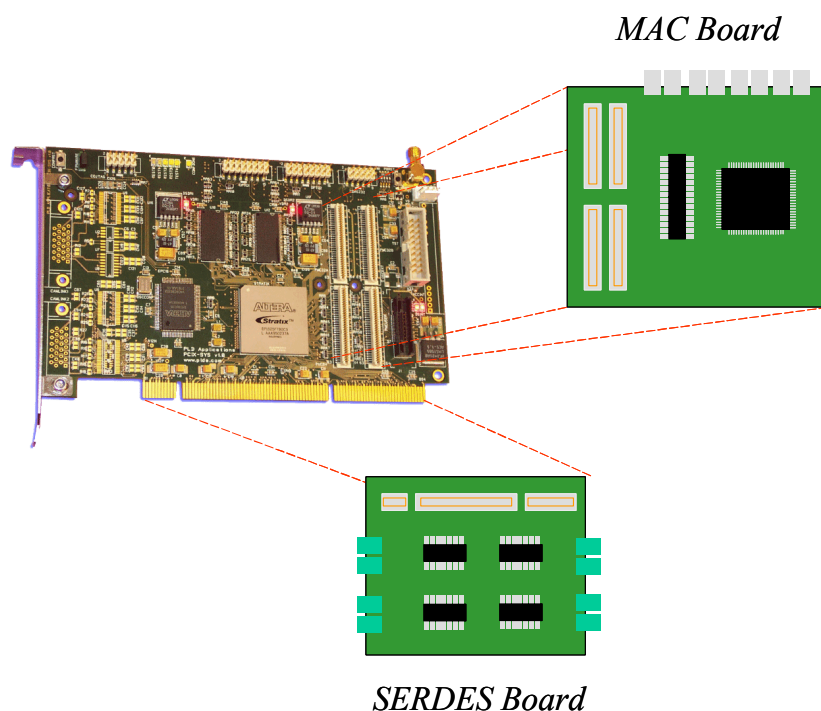


Figura 3 - Esquemático do hardware eletrônico que implementa a UMRE

A primeira placa filha, denominada MAC Ethernet, é responsável pelo controle de tráfego Ethernet cliente e, no projeto atual, disponibiliza quatro interfaces ópticas

Gigabit Ethernet em padrão SFP (*Small Form Pluggable*). Esta placa baseia-se em um projeto de alta complexidade com desafios tecnológicos concentrados no desenho da placa de circuito impresso, repleto de linhas de transmissão de impedância controlada e comprimento equalizado para sinais com taxas de bit de até 1,25 gigabit por segundo.

A segunda placa filha é responsável pela conversão serial-paralelo (*Serializer-Deserializer - SERDES*) necessária na transmissão e na recepção de pacotes ópticos. Esta placa disponibiliza quatro interfaces proprietárias de transmissão e recepção, cada uma operando em taxas de até 1,25 gigabit por segundo. O desenvolvimento das interfaces proprietárias fez-se necessário pela incapacidade que a interface Gigabit Ethernet tem de operar no modo rajada nos tempos exigidos pelo projeto. Esta placa exhibe um projeto de média complexidade.

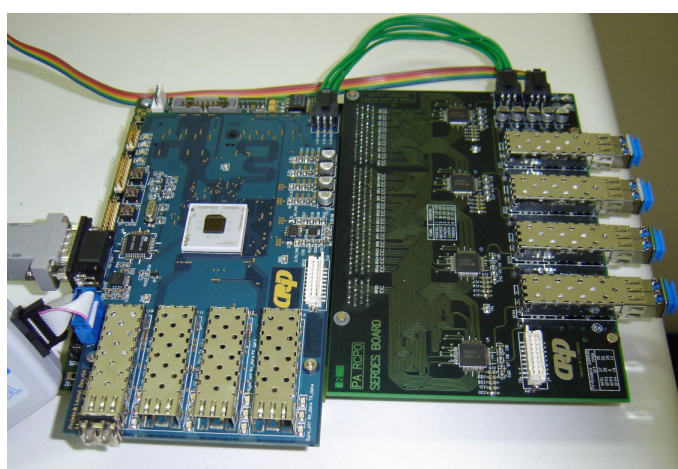


Figura 4 – Foto do protótipo de hardware da UMRE

A MCOP consiste de um número de unidades de comutação óptica (UCO), onde o número exato depende da configuração de transmissão e recepção do nó. Cada UCO consiste de um acoplador que divide o sinal de entrada em dois braços, cada um com uma chave óptica do tipo liga/desliga baseada em amplificador óptico a semiconductor (*semiconductor optical amplifier - SOA*).

A Figura 5 apresenta o esquemático da UCO e a Figura 6 uma foto de um protótipo da UCO, já com sua placa de eletrônica de controle correspondente. Cada SOA necessita de um controle de temperatura e um circuito de excitação (*driver*) de alta velocidade para fornecer a corrente necessária para o dispositivo amplificar (estado ON) ou não (estado OFF) o sinal óptico. Esta placa é a responsável por essas funções.

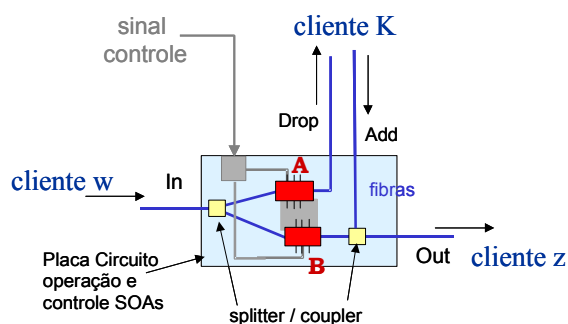


Figura 5 - Esquemático da UCO

A UCO pode ser configurada em um dos quatro estados:

- Deriva: este estado ocorre quando o nó detecta o pacote óptico destinado a ele e é configurado ligando-se a chave óptica do braço superior e desligando-se a do braço inferior;
- Insere: este estado ocorre quando o nó resolve transmitir o pacote óptico e é configurado desligando-se as chaves ópticas dos braços superior e inferior;
- Encaminha: este estado é usado quando o nó detecta que o pacote óptico é destinado a outro nó e é configurado desligando-se a chave óptica do braço superior e ligando-se a do braço inferior;
- Deriva e encaminha: este estado é usado quando o nó detecta que ele é um dos destinatários do pacote óptico e é configurado ligando-se as chaves ópticas dos braços superior e inferior.

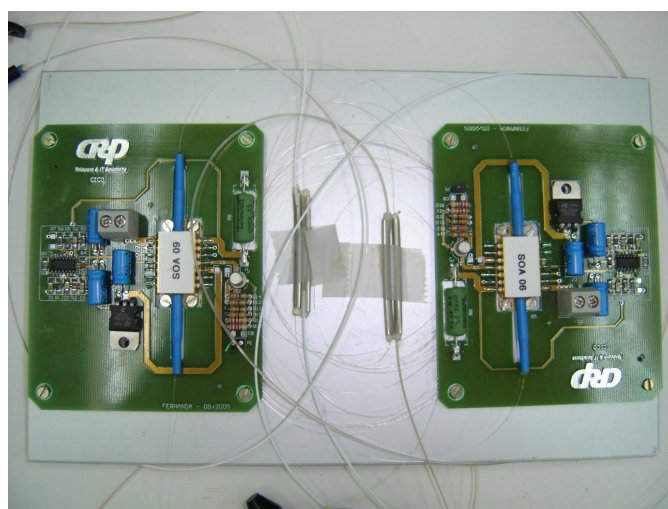


Figura 6 - Foto do Protótipo da UCO

O modo de funcionamento básico do nó da rede, considerando-se apenas um anel de fibra e um comprimento de onda de dados para clareza, é o descrito a seguir. Na entrada do nó, o cabeçalho do pacote óptico é enviado a UMRE. O pacote óptico correspondente segue para uma linha de atraso baseada em fibra (*fiber delay line* - FDL) que irá retê-lo tal que quando ele chegue a MCOP o módulo MAC já tenha processado o cabeçalho, tomado as decisões apropriadas e configurado a MCOP adequadamente. Saindo da FDL o pacote óptico é demultiplexado e enviado a UCO correspondente. Neste ponto, o pacote óptico é derivado, encaminhado ou derivado e encaminhado. Neste ponto também pode ser inserido um novo pacote óptico no anel. O tempo de processamento do cabeçalho é de 1 microsegundo, o que no nosso projeto resulta em uma FDL com 200 metros de comprimento.

Na saída do nó o pacote óptico encaminhado ou inserido é multiplexado, juntamente com seu cabeçalho correspondente, e deixa o nó em direção ao seguinte.

Cabe ressaltar que a arquitetura da rede foi definida de forma modular para permitir que partes dela tivessem utilidade técnica ou comercial mesmo que separadamente ou que partes dela pudessem ser substituídas por outras de maior

interesse. Por exemplo, a MCOP permite que a UPC seja substituída por outra de tal forma que, conforme sua natureza, a rede possa operar no modo comutação de circuito, comutação de pacote, comutação de rajada ou até como uma combinação destas.

4. Protocolo MAC

Vários protocolos MAC foram desenvolvidos e estudados neste trabalho (ex.: [Salvador, Uesono e da Fonseca 2005][Uesono, Salvador e da Fonseca 2005]), sendo alguns de baixo desempenho e de fácil implementação e outros de alto desempenho e de difícil implementação. Um protocolo que situa-se entre os dois extremos e que está sendo implementado neste trabalho baseia-se na combinação de versões modificadas dos protocolos PAT (*Packet Aggregate Transport*) [Salvador 2003] e SAT [Cidon e Ofek 1993].

O primeiro protocolo 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. Neste protocolo, a capacidade total do anel é dividida em fatias de tempo (*time slots*) com duração fixa. Cada slot pode transportar um único pacote óptico, onde um pacote óptico pode transportar um ou mais pacotes de dados (quadros Ethernet) agrupados em uma rajada. Cada slot possui um cabeçalho que é convertido para o domínio eletrônico e processado em cada nó da rede para que este possa determinar se o slot está vazio ou, se estiver ocupado, em que lambda está o pacote correspondente e o que fazer com aquele pacote. A Figura 7 ilustra a organização entre cabeçalhos e slots.

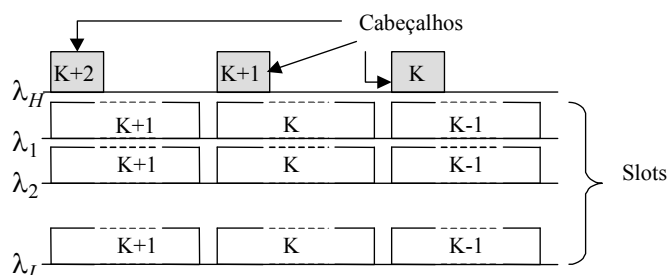


Figura 7 – Slots e cabeçalhos

A Figura 8 apresenta a organização dos pacotes em uma rajada. Os campos *Preamble* e “K” *words* são necessários para permitir a sincronização do receptor com rajadas de tamanho variado. O campo *Payload* subdivide-se em várias segmentos de tamanho fixo, cada um contendo um ou mais quadros Ethernet. Todo segmento contém um campo com *bytes* de preenchimento (*bytes of padding - BP*), um campo definindo o tamanho do campo BP (*padding size - PS*) e um último campo de verificação de integridade do quadro (*cyclic redundancy check - CRC*).

A transmissão de uma rajada é possível somente mediante a chegada de um slot vazio. Diferentemente de PAT, onde um slot pode ser alocado para transmitir um pacote óptico que pode conter um único pacote de dados independentemente do seu tamanho, neste protocolo a alocação de um slot vazio pode ocorrer somente se houver uma rajada pronta para transmissão. Em função de decisões de engenharia de tráfego ou de roteamento tomadas pela UPC, uma rajada pronta para transmissão só pode ser enviada

em um slot vazio se o slot e o lambda em que a rajada está para ser transmitida forem aceitáveis do ponto de vista de engenharia de tráfego ou de roteamento¹.

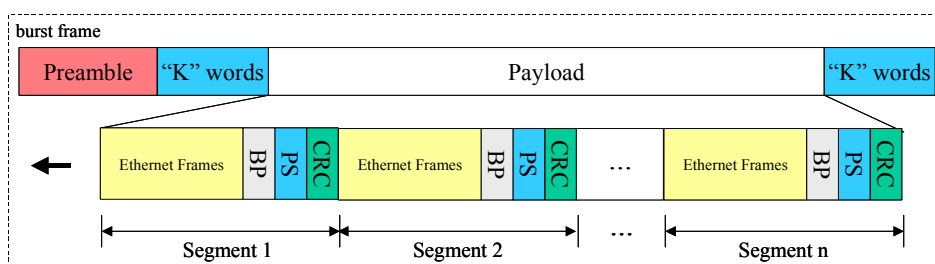


Figura 8 – Estrutura de uma rajada

O pacote óptico contido em um slot é removido ao chegar no destino de forma que aquele slot possa ser reusado por aquele nó ou por outros subseqüentes. A desmontagem da rajada ocorre conforme explicado anteriormente.

Para evitar que alguns nós da rede tenham mais oportunidades de acesso que outros, também está sendo implementado uma versão modificada do protocolo SAT. Neste protocolo, um sinal de controle denominado "SAT" (de *satisfied*), implementado por um bit no cabeçalho dos slots (1-SAT presente; 0-SAT ausente), circula pelos nós do anel atribuindo quotas de transmissão (em slots) pré-determinadas. Se o nó que recebe o SAT possuir pelo menos uma rajada para transmitir e ainda não tiver enviado 1 rajadas, então este nó irá reter o SAT (atualizando em 0 o bit apropriado do cabeçalho do slot que trouxe o SAT) até que não haja mais rajadas para transmitir ou que tenham sido transmitidas 1 rajadas. A não liberação do SAT impede que os outros nós tenham suas quotas renovadas, fazendo com que em algum momento parem de transmitir, desta forma permitindo que o nó que possui o SAT tenha oportunidades para utilizar sua quota. Uma vez satisfeito, o nó libera o SAT (atualizando em 1 o bit apropriado do cabeçalho do primeiro slot que aparecer).

A Figura 9 apresenta a estrutura básica do cabeçalho do slot já na versão em prototipagem. Com exceção dos campos S e F, todos os outros campos são o comumente usados nas tecnologias comerciais:

- S (*Status*): campo que indica se o slot está vazio (0) ou ocupado (1). Um slot vazio contém os campos S, TTL, DA, SA e Reservado 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 pode atravessar no anel. Cada nó da rede decrementa este campo em um 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 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, então a rajada é recebida

¹ Esta funcionalidade não está sendo implementada na versão atual do protótipo.

e o slot é marcado como vazio, podendo ser reusado pelo próprio nó para a transmissão de uma nova rajada ou por nós subseqüentes;

- SA (*Source Address*): campo que indica o remetente daquele slot. Se SA for igual ao endereço do nó que está processando o slot, então 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;
- Reservado: reservado para uso futuro. Este campo não é processado;
- 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, então 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. Cabe ao nó mestre detectar a ausência do SAT e reinseri-lo na rede. Antes de enviar o slot para o próximo nó, o nó calcula o CRC sobre o novo cabeçalho e atualiza o campo CRC do cabeçalho com este novo CRC.

Os tamanhos dos campos estão reduzidos visando o processamento do cabeçalho do slot e a execução das ações correspondentes em 1 microsegundo. Esses campos podem ser facilmente aumentados para os tamanhos adotados na indústria com o uso de um processador de cabeçalho de maior desempenho.

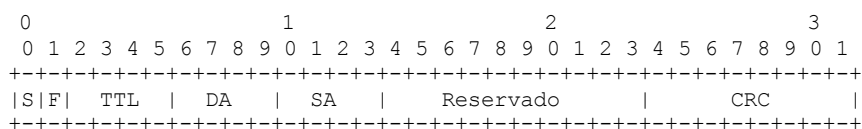


Figura 9 – Estrutura básica do cabeçalho do slot

Os dois protocolos estão sendo implementados em C (sem sistema operacional) no modo de processador embarcado disponível na FPGA (denominado NIOS-II), o que traz alto desempenho e simplifica enormemente o desenvolvimento dos protocolos.

Ao nó mestre (primário ou secundário) cabe a responsabilidade de inserir os cabeçalhos na rede no momento da inicialização ou, na condição de erro, reinserir um cabeçalho ou o sinal SAT ausente. A ausência de um cabeçalho ou do sinal SAT é notada pelo nó mestre em no máximo uma rotação no anel. A reinserção desses no anel ocorre na próxima rotação. Portanto, em um anel com 100km e 16 nós o tempo de detecção de erro é igual a 516 microsegundos e o de recuperação a 1.3 milisegundos. Aplicando-se uma certa margem de tolerância pode-se dobrar esses valores que ainda assim os tempos de proteção serão bastante baixos.

Na versão atual do protótipo está prevista somente a proteção de um enlace qualquer via detecção de erro baseada no processamento dos cabeçalhos dos slots –por questões de complexidade de implementação. Na ocorrência de falha em um enlace, detectada pelos nós da rede pela não chegada dos cabeçalhos correspondentes, cada nó passa a transmitir no outro anel. Portanto, em condições normais de operação os dois

anéis são usados para transmissão de dados, porém na condição de falha apenas o anel sem problemas pode ser usado.

Uma versão mais sofisticada de mecanismo de proteção, baseada no mecanismo de proteção de RPR, está em estudo e provavelmente será implementada na próxima versão da rede para permitir proteção contra vários tipos de falhas, sempre usando os dois anéis.

5. Resultados de Desempenho

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 de [Pawlikowski, Joshua Jeong e Ruth Lee 2002], o simulador usa intervalo de confiança como critério de parada e o gerador de números pseudo-aleatórios Mersenne Twister [Matsumoto e Nishimura 1998], contido na biblioteca Colt-lib, desenvolvida pela Organização Européia para Estudos Nucleares (CERN).

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 (numa única direção). Resultados de experimentos realizados com mais comprimentos de ondas de dados, seja com uma única fibra óptica ou com duas fibras ópticas transmitindo em direções opostas (com roteamento pelo menor caminho), não apresentados neste trabalho, confirmaram que o desempenho da rede aumenta linearmente com o aumento do número de comprimentos de ondas de dados.

A taxa de transmissão e o tamanho dos slots adotados nos experimentos é de 10Gb/s e 64KB de tamanho equivalente. O tamanho do slot é suficiente para compensar o tempo de processamento de cabeçalho, que equivalente a 10Kb por slot, resultando numa ineficiência de transporte de 13,5%.

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 em [Claffy, Miller e Thompson 1998]. 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 1 apresenta os parâmetros de entrada adotados nas simulações.

A Figura 10 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, que é 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 e do baixo número resultante de slots disponíveis na rede e do tamanho das rajadas, o que dificulta o acesso e gera um certo acúmulo de pacotes nas filas nos nós da rede. Este

acúmulo tem reflexos nos atrasos médios enfrentados pelos pacotes, conforme mostra a Tabela 2 - também tem reflexos na taxa de descarte de pacotes, embora a 150% de carga ainda não ocorra uma perda de pacotes sequer. 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.

Tabela 1 – Parâmetros de entrada

Parâmetro	Valor
Nr. de fibras ópticas no anel	1 (transmissão unidirecional)
Tamanho do anel (km)	100
Nr. de lambdas de dados por fibra	1
Nr. de nós	16
Taxa de transmissão por lambda	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 ($k=l$)	Nr. de slots
Tráfego	Chegada: <i>poisson</i> Distribuição: simétrica; assimétrica
Intervalo de confiança (%)	90
Tamanho de cada fila de pacotes (bytes)	128K

Percebe-se pelos números contidos na Tabela 2 que a 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.

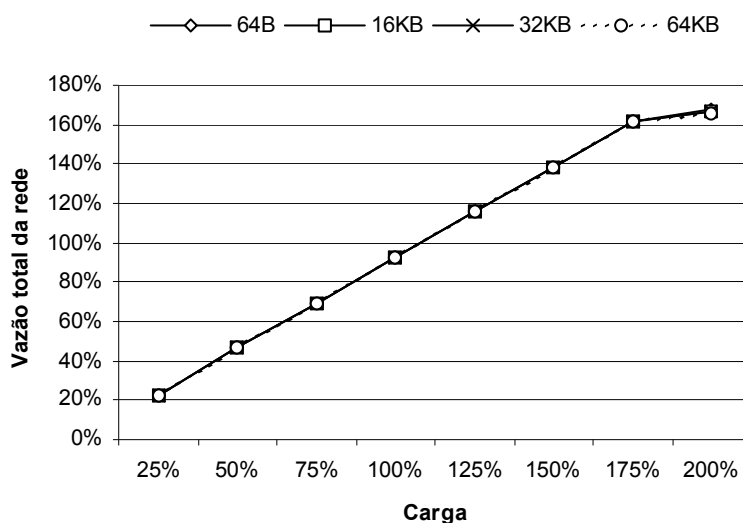


Figura 10 – Vazão total da rede (cenário simétrico)

Tabela 2 – Atraso médio dos pacotes (cenário simétrico)

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

A Figura 11 apresenta o grau de justiça da rede em função da carga total injetada na rede no cenário simétrico. 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.

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.

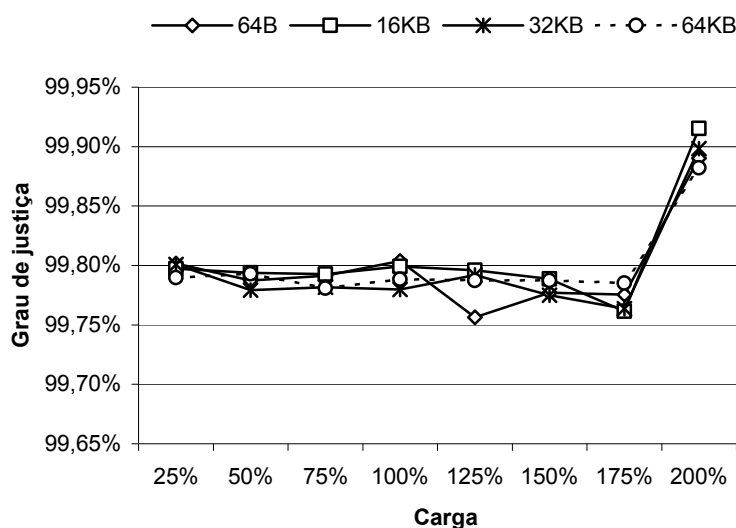


Figura 11 – Grau de justiça de acesso da rede (cenário simétrico)

A Figura 12 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.

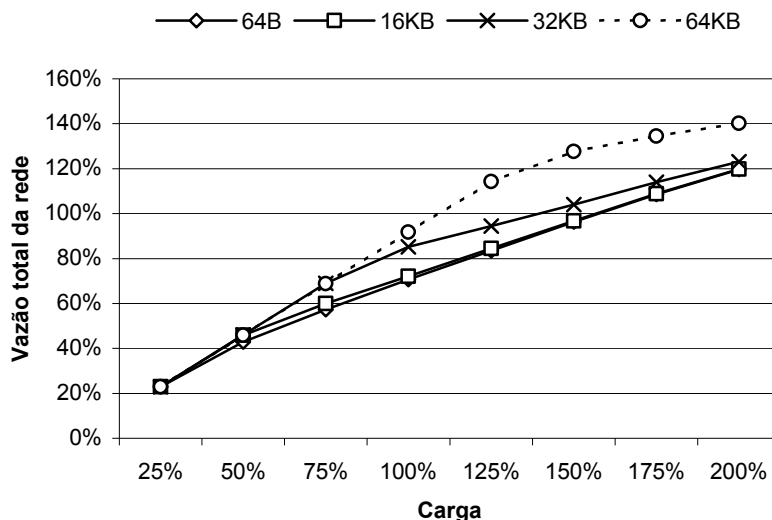


Figura 12 – Vazão total da rede (cenário assimétrico)

Assim como no cenário simétrico, e pelo mesmo motivo, a vazão não aumenta na mesma proporção que a carga. O acúmulo de pacotes nas filas ocorre já a 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 13 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.

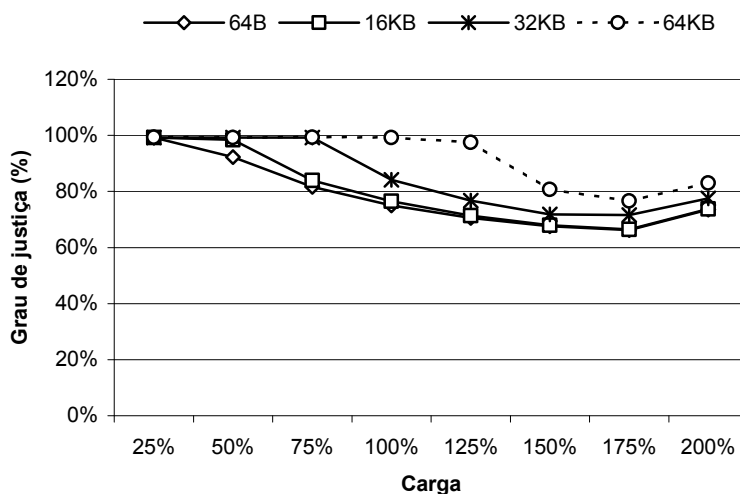


Figura 13 – Grau de justiça de acesso da rede (cenário assimétrico)

As curvas apresentadas na Figura 13 são um reflexo das curvas apresentadas na Figura 12 e das explicações daquelas curvas. 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+ (e outros derivados de SAT) em lidar com distribuições assimétricas de tráfego [Salvador 2003].

6. Conclusões

Este trabalho apresentou os resultados de um projeto de pesquisa e desenvolvimento de três anos que desde 2004 está em andamento nos laboratórios do CPqD e que visa o desenvolvimento de uma tecnologia inovadora de rede em anel com transmissão e comutação de pacotes no domínio óptico. Este projeto diferencia-se da maioria dos demais encontrados na literatura nacional e internacional (ex.: Hornet [Shrikhande et al 2000], Ringo [Carena 2002], Flamingo [Dey, van Bochove, Koonen, Geuzebroek e Salvador 2001], Omega [Rocha et al 2002]) na sua busca por um protótipo laboratorial para prova de conceito e no uso de tecnologias e componentes ópticos e eletrônicos disponíveis no mercado a preços não proibitivos.

A arquitetura da rede foi projetada para oferecer serviços do tipo Ethernet assim como transparentes (para tecnologias legadas como SDH), sejam dos tipos um-para-um ou um-para-muitos, sob demanda, com grande flexibilidade. Visando escalabilidade e baixo custo, a rede pode trabalhar apenas com fibras ópticas ou também com WDM, de preferência com rajadas de tamanho grande, pois quanto mais pacotes forem comutados com o processamento de um único cabeçalho melhor é a escalabilidade da rede. O custo dos SOAs utilizados nas chaves ópticas sugere que quanto menor for o número de lambdas e quanto maior for a taxa de transmissão por lambda, menor será o custo de instalação da rede. Isto graças ao compartilhamento dos lambdas que a comutação óptica de pacotes oferece.

O projeto encontra-se em estágio final de prototipagem. Quatro protótipos completos do nó da rede, com exceção do software do protocolo MAC, foram montados e passaram com sucesso nos testes, incluindo ligação ponto-a-ponto entre pares para transporte de vídeo real. O software do protocolo MAC descrito neste trabalho está em desenvolvimento e até que esteja pronto não será possível o teste dos protótipos ligados em rede. Embora pareça que há pouco a ser feito, a ligação desses nós em rede pode apresentar surpresas, principalmente na camada de transmissão, em virtude da dinâmica da rede e dos seus efeitos indesejados que mesmo apesar dos esforços comumente ocorrem.

Agradecimentos

Este projeto está sendo financiado com recursos do Fundo de Desenvolvimento das Telecomunicações (Funttel) e conta com a colaboração de Marbey Mosso e Henrique Portela da Silva, da PUC-Rio, e de Felipe Rudge Barbosa, da Unicamp.

Referencias

802.17-2004 (2004) IEEE Standard for Information Technology -Telecommunications and information exchange between systems - Local and metropolitan area networks -

Specific requirements - Part 17: Resilient packet ring (RPR) access method & physical layer specifications.

- Carena et al. (2002) "RINGO: A demonstrator of WDM optical packet network on a ring topology", Anais do 6^o. IFIP Working Conference on Optical Network Design and Modelling (ONDM), Torino, Itália.
- Cidon, I. e Ofek, Y. (1993) "Metaring – a full-duplex ring with fairness and spatial reuse", IEEE Transactions on Communications, Vol. 41, No. 1, pp. 110-120.
- Claffy, K., Miller, G. e Thompson, K. (1998) "The nature of the beast: recent traffic measurements from an Internet backbone", Internet Society (ISOC) Internet Summit.
- Dey, D., van Bochove, A., Koonen, A.M.J., Geuzebroek, D. e Salvador, M. R. (2001) "FLAMINGO: A Packet-switched IP-over-WDM All-optical MAN", In Procs. of 27th European Conference on Optical Communications (ECOC'2001), pp. 480-4812, Amsterdã, Holanda.
- Mannie, E. (2004). Generalized Multi-Protocol Label Switching (GMPLS) Architecture. IETF RFC 3945.
- Matsumoto, M. e Nishimura, T. (1998) "Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator", ACM Transactions on Modeling and Computer Simulation, Vol. 8, No. 1, pp.3-30.
- Pawlikowski, K., Joshua Jeong, H.-D. e Ruth Lee, J.-S. (2002) "On Credibility of Simulation Studies of Telecommunication Networks", IEEE Communications Magazine, Vol. 40, No. 1, pp. 132-139.
- Salvador, M. R. (2003), MAC Protocols for Optical Packet-Switched WDM Rings, Twente University Press, ISBN 90-365-1862-8.
- Salvador, M. R., Uesono, M. M. e da Fonseca, N. L. S. (2005) "A Local Fairness Protocol for Optical Packet-switched WDM Ring Networks", Anais do 40^o. IEEE International Conference on Communications (ICC), Seul, Coréia do Sul.
- Shrikhande, K. V. et al. (2000) "HORNET: A Packet-Over-WDM Multiple Access Metropolitan Area Ring Network", In IEEE Journal on Selected Areas in Communications, Vol. 18, No. 10, pp. 2004-2016.
- Uesono, M. M., Salvador, M. R. e da Fonseca, N. L. S. (2005) "A Packet Ring Fairness Protocol and its impact on TCP fairness", Anais do 48^o. IEEE Global Communications (Globecom), St. Louis, EUA.