

UNIVERSIDADE FEDERAL DO PARÁ
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

PROPOSTA DE ARQUITETURA PARA O ESTABELECIMENTO DINÂMICO
DE CONEXÕES DETERMINÍSTICAS E BUSCA DE RECURSOS
MULTIDOMÍNIO EM REDES GRID OBS BASEADAS NO GMPLS

ANDERSON MARQUES NUNES

DM - 23 / 2011

UFPA – CT - PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM – PARÁ – BRASIL
2011

UNIVERSIDADE FEDERAL DO PARÁ
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ANDERSON MARQUES NUNES

PROPOSTA DE ARQUITETURA PARA O ESTABELECIMENTO
DINÂMICO DE CONEXÕES DETERMINÍSTICAS E BUSCA DE
RECURSOS MULTIDOMÍNIO EM REDES GRID OBS BASEADAS NO
GMPLS

Dissertação submetida à Banca
Examinadora do Programa de
Pós-Graduação em Engenharia
Elétrica da UFPA para a
obtenção do Grau de Mestre em
Engenharia Elétrica com Ênfase
em Computação Aplicada

UFPA – CT - PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM – PARÁ – BRASIL
2011

UNIVERSIDADE FEDERAL DO PARÁ
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

PROPOSTA DE ARQUITETURA PARA O ESTABELECIMENTO
DINÂMICO DE CONEXÕES DETERMINÍSTICAS E BUSCA DE
RECURSOS MULTIDOMÍNIO EM REDES GRID OBS BASEADAS NO
GMPLS

AUTOR : ANDERSON MARQUES NUNES

DISSERTAÇÃO APRESENTADA AO PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DA PARÁ COMO
REQUISITO PARCIAL PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA
ELÉTRICA.

APROVADA EM: 17/06/2011.

BANCA EXAMINADORA:

Prof. Dr. Eduardo Coelho Cerqueira
(ORIENTADOR - UFPA)

Prof. Dr. Antônio Jorge Gomes Abelém
(MEMBRO - UFPA)

Prof. Dr. João Crisóstomo Weil Albuquerque Costa
(MEMBRO - UFPA)

VISTO:

Prof. Dr. Marcus Vinicius Alves Nunes
(COORDENADOR DO PPGE/CT/UFPA)

Aos meus pais, que desde sempre estiveram ao meu lado apoiando todas as decisões que eu tomei com a certeza de estar certo, apesar de em algumas não estar. Amo vocês.

AGRADECIMENTOS

Deus, Você tem um parágrafo exclusivo de agradecimento na vida de todos, inclusive na minha pelo Senhor cuidar não só de nós, mas também de quem amamos. A Ti devemos tudo!

Ao meus amados pais Sonia e Antonio, que estiveram ao meu lado nos bons e principalmente nos piores momentos, firmes e fortes, dispostos a não largarem a minha mão nunca! Aos meus queridos irmãos Rafael e Fabiane, que me ajudaram a deixar um pouco de lado as dificuldades com as suas alegrias e por me lembrar que nem só de trabalho viverei. Aos meus sogros Arlene e Moreira por nesse período referente ao mestrado terem me aceitado como um filho para o resto da vida. A Meg, Mabel e Hanna, minhas cadelinhas amadas, que independente do que eu diga ou faça sempre estarão do meu lado esperando um afago.

Aos meus amigos do coração Andrei e Marina, pessoas que definem perfeitamente o significado da palavra amizade.

Ao meu eterno coordenador/amigo Dionne, um exemplar engenheiro que me incentivou desde a graduação, assim como ao meu co-orientador Abelém, por ter acreditado no momento em que fui procura-lo e ter estado ao meu lado durante este longo trabalho e é claro, ao meu orientador Eduardo, que abraçou o meu trabalho como se fosse dele.

As agências de fomento CNPQ, RNP e FAPESPA pelo suporte financeiro.

Aos amigos Henrique e Thiêgo, colegas que com o tempo se tornaram grandes amigos. Aos grandes companheiros de trabalho Rafael Lopes, Billy, Rômulo, Airton, Raphael, Salvatti, Vagner e em especial ao Davison, Rafael Esteves, Fernando e Mariana pessoas que me ajudaram a conquistar essa vitória.

A mulher da minha vida Caroline, uma das maiores incentivadoras da minha busca pelo conhecimento, agradeço a você pelo carinho, companheirismo, dedicação, amor e simplesmente por acreditar em mim mais do que eu mesmo! Te amo!!

“Nobody is gonna hit as hard as life.
But it ain’t about how hard you hit.
It’s about how hard you can get it
and keep moving forward. How
much you can take and keep
moving forward. That’s how
winning is done!”
(Rocky Balboa)

SUMÁRIO

Lista de Tabelas	x
Lista de Figuras	xi
Lista de Siglas	xii
Resumo.....	xv
Abstract	xvii
Capítulo 1	1
Introdução	1
1.1 – Redes de Próxima Geração.....	1
1.2 – Motivação	5
1.3 – Objetivos.....	8
1.4 – Organização do trabalho.....	9
Capítulo 2	11
Desafios em Redes para Grades Computacionais.....	11
2.1 – Grades Computacionais.....	11
2.1.1 – Introdução	11
2.1.2 – Propriedades de uma Grade	13
2.1.3 – Padronização de Grades	15
2.2 – Redes Ópticas em Grades.....	16
2.2.1 – Introdução	16
2.2.2 – Requisitos de Rede para Grades.....	21
2.3 – Grids OBS.....	23
2.3.1 – Comutação em Rajadas Ópticas.....	23
2.3.2 – Redes OBS em Grades	26
2.4 – Considerações Finais.....	36
Capítulo 3	38
Trabalhos Relacionados.....	38
3.2 Redes OBS em Grades.....	48
3.3 Qualidade de Serviço em Redes OBS	50
3.4 Roteamento <i>Anycast</i>	56
3.5 Considerações Finais	58
Capítulo 4	60
Arquitetura para o Oferecimento de Conexões Determinísticas Multidomínio em Redes GOBS Baseadas no GMPLS.....	60
4.1 – Qualidade de Serviço em Redes OBS Baseadas no GMPLS.....	61
4.2 – Conexões Determinísticas e Seleção de Recursos	63
4.3- Servidor GOBS	65
4.4 – Servidor GOBS Raiz	66
4.5 – Seleção de Rotas.....	67
4.6 – Monitoramento de Recursos	70
4.7 – Funcionamento da Proposta.....	70
4.8 – Modificações na Sinalização GMPLS.....	73
4.9 – Considerações Finais.....	74
Capítulo 5	76
Simulação e Análise dos Resultados	76

5.1 - Adaptações Necessárias no Simulador.....	76
5.2 - Cenário da Simulação.....	81
5.3 - Resultados Obtidos	84
5.4 Considerações Finais.....	93
Capítulo 6	95
6.3 Contribuições	95
6.2 - Conclusões	96
6.3 - Trabalhos Futuros.....	98
Referências Bibliográficas.....	99

Lista de Tabelas

TABELA 1: COMPARAÇÃO ENTRE OS PARADIGMAS DE COMUTAÇÃO ÓPTICA [6].	25
TABELA 2: GMPLS X REQUISITOS DE APLICAÇÕES [11].	39
TABELA 3: EXEMPLO DE ENTRADA NO SERVIDOR GOBS.	66
TABELA 4: EXEMPLO DE ENTRADA NO SERVIDOR GOBS RAIZ.	67
TABELA 5: DISPOSIÇÃO DOS NÓS NA TOPOLOGIA.	82
TABELA 6: NÚMERO DE COMPRIMENTOS DE ONDA POR CLASSE.....	83
TABELA 7: PROBABILIDADE DE BLOQUEIO MÁXIMA PERMITIDA POR CLASSE.....	84

Lista de Figuras

FIGURA 1: EXEMPLO DE APLICAÇÃO DE ASTROFÍSICA USANDO GRADES [2].	2
FIGURA 2: COMUTAÇÃO DE RAJADAS ÓPTICAS[60].	25
FIGURA 3: GUNI E GRNI EM UMA REDE OBS[33].	28
FIGURA 4: REDES HETEROGÊNEAS COM GMPLS [43].	34
FIGURA 5: ARQUITETURA OBS SOBRE GMPLS [9].	36
FIGURA 6: ARQUITETURA PROPOSTA EM [48].	41
FIGURA 7: PSI TROCADO ENTRE DOMÍNIOS ÓPTICOS[49].	43
FIGURA 8: MODELO DE CONVERSORES DE COMPRIMENTO DE ONDA PROPOSTO EM [51].	44
FIGURA 9: PLANO DE CONTROLE E SUA INTERAÇÃO COM O MIDDLEWARE DA GRADE E O PLANO DE GERENCIAMENTO [2].	45
FIGURA 10: ARQUITETURA [52].	47
FIGURA 11: ARQUITETURA GOBS[46].	48
FIGURA 12 SINALIZAÇÃO NO OBS PROGRAMÁVEL [33].	49
FIGURA 13 VISÃO GERAL DA ARQUITETURA DO DQMA [57].	53
FIGURA 14: SERVIDOR GOBS[59].	55
FIGURA 15: ALGORITMO DE SELEÇÃO DE ROTAS.	68
FIGURA 16: FLUXOGRAMA DE SELEÇÃO DE ROTAS.	69
FIGURA 17: ARQUITETURA PARA O APROVISIONAMENTO DE CONEXÕES DETERMINÍSTICAS.	71
FIGURA 18: INTEGRAÇÃO ENTRE OS PLANOS NA ARQUITETURA PROPOSTA.	72
FIGURA 19:ESTRUTURA DO LOBS <i>NODE</i> [57].	78
FIGURA 20: AGENTE GRIDJOB.	80
FIGURA 21: INTERAÇÃO ENTRE OS COMPONENTES DESENVOLVIDOS PARA O NS-2.	80
FIGURA 22: TOPOLOGIA UTILIZADA NAS SIMULAÇÕES.	81
FIGURA 23: MODELO DE SIMULAÇÃO 1 COM OS SEUS CIRCUITOS INTERDOMÍNIO ESTÁTICO EM DESTAQUE (VERDE).	82
FIGURA 24: NÚMERO DE REQUISIÇÕES POR NÓ.	85
FIGURA 25: PROBABILIDADE DE BLOQUEIO X CLASSES DE SERVIÇO (NÓ 3).	86
FIGURA 26: PROBABILIDADE DE BLOQUEIO X CLASSES DE SERVIÇO (NÓ 13).	86
FIGURA 27: PROBABILIDADE DE BLOQUEIO X CLASSES DE SERVIÇO (NÓ 28).	86
FIGURA 28: PROBABILIDADE DE BLOQUEIO X CLASSE DE SERVIÇOS (NÓ 41).	87
FIGURA 29: PROBABILIDADE DE BLOQUEIO X CLASSE DE SERVIÇOS (NÓ 57).	87
FIGURA 30: PROCESSAMENTO MÉDIO UTILIZADO.	88
FIGURA 31: ESPAÇO DE ARMAZENAMENTO MÉDIO UTILIZADO.	89
FIGURA 32: ATRASO FIM-A-FIM.	90
FIGURA 33: NÚMERO DE REQUISIÇÕES DESCARTADAS NOS NÓS DE COMPUTAÇÃO.	91
FIGURA 34: PROBABILIDADE DE BLOQUEIO (STATIC X DYN - CLASSE 0).	92
FIGURA 35: PROBABILIDADE DE BLOQUEIO (STATIC X DYN - CLASSE 1).	92
FIGURA 36: PROBABILIDADE DE BLOQUEIO (STATIC X DYN - CLASSE 2).	93

Lista de Siglas

AAA	<i>Authentication, Authorization, Accounting</i>
AF	<i>Assured Forwarding</i>
API	<i>Application Programming Interface</i>
AS	<i>Autonomous System</i>
ASTB	<i>Application Specific Topology Builder</i>
BCP	<i>Burst Control Packet</i>
BE	<i>Best Effort</i>
CAC	<i>Call Admission Control</i>
CR-LDP	<i>Constraint Based Label Distribution Protocol</i>
CWDM	<i>Coarse Wavelength Division Multiplexing</i>
DDRP	<i>Domínio-to-Domain Routing Protocol</i>
DQMA	<i>Dynamic QoS Management Agent</i>
DRAGON	<i>Dynamic Resource Allocation via GMPLS Networks</i>
DWDM	<i>Dense Wavelength Division Multiplexing</i>
EF	<i>Expedited Forwarding</i>
ESA	<i>End System Agent</i>
FDL	<i>Fiber Delay Line</i>
FLOP	<i>FLoating point Operations Per Second</i>
FSC	<i>Fiber Switching Capable</i>
GENI	<i>Global Environment for Network Innovations</i>
GFLOP	<i>Giga FLOP</i>
GLIF	<i>Global Lambda Integrated Facility</i>
GMPLS	<i>Generalized Multiprotocol Label Switching</i>
GOBS	<i>Grid OBS</i>
GRNI	<i>Grid Resource to Network Interface</i>
GUNI	<i>Grid User to Network Interface</i>
IETF	<i>Internet Engineering Task Force</i>
IP	<i>Internet Protocol</i>
IPU	<i>Input Processing Unit</i>

IDRA	<i>Inter-Domain Routing Agents</i>
IDRP	<i>Inter-Domain Routing Protocol</i>
ITU	<i>International Telecommunication Union</i>
JET	<i>Just Enough Time</i>
JIT	<i>Just In Time</i>
L2SC	<i>Layer 2 Switch Capable</i>
LAUC-VF	<i>Latest Available Unused Channel – Void Filling</i>
LHC	<i>Large Hadron Collider</i>
LMP	<i>Link Management Protocol</i>
LOBS	<i>Labeled Optical Burst Switching</i>
LSC	<i>Lambda Switching Capable</i>
LSP	<i>Label Switched Path</i>
MEMS	<i>Micro-Electrical-Mechanical System</i>
MPLS	<i>Multiprotocol Label Switching</i>
NARB	<i>Network Aware Resource Broker</i>
OBS	<i>Optical Burst Switching</i>
OBGP+	<i>Optical Border Gateway Protocol</i>
OCS	<i>Optical Circuit Switching</i>
OEO	<i>Optical-Electronic-Optical</i>
OGF	<i>Open Grid Forum</i>
OGSA	<i>Open Grid Services Architecture</i>
OOO	<i>Optical-Optical-Optical</i>
OPS	<i>Optical Packet Switching</i>
OPU	<i>Output Processing Unit</i>
OSPF-TE	<i>Open Shortest Path First – Traffic Engineering</i>
OWNS	<i>Optical WDM Network Simulator</i>
OXC	<i>Optical Crossconnect</i>
PCE	<i>Path Computation Element</i>
PHB	<i>Per Hop Behavior</i>
PSC	<i>Packet Switch Capable</i>
PXC	<i>Photonic Crossconnect</i>
QoS	<i>Quality of Service</i>
RDT	<i>Route Discovery Threshold</i>

RSVP-TE	<i>ReSource ReserVation Protocol – Traffic Engineering</i>
SCU	<i>Switching Control Unit</i>
SOA	<i>Semiconductor Optical Amplifier</i>
SOA	<i>Service Oriented Architecture</i>
SONET	<i>Synchronous Optical Network</i>
SU	<i>Switching Unit</i>
TDM	<i>Time Division Multiplexing</i>
TE	<i>Traffic Engineering</i>
TED	<i>Traffic Engeneering Database</i>
TTM	<i>Transit Tunnel Manager</i>
VLBI	<i>Very Long Baseline Interferometry</i>
VLSR	<i>Virtual Label Switch Router</i>
VNTM	<i>Vitrual Network Topology Manager</i>
WDM	<i>Wavelength Division Multiplexing</i>
WE	<i>Wait Unit</i>

Resumo

A capacidade de processamento das instituições de pesquisa vem crescendo significativamente à medida que processadores e estações de trabalho cada vez mais poderosos vão surgindo no mercado. Considerando a melhoria de desempenho na área de redes de computadores e visando suprir a demanda por processamento cada vez maior, surgiu a ideia de utilizar computadores independentes conectados em rede como plataforma para execução de aplicações paralelas, originando assim a área de computação em grade. Em uma rede que se encontra sob um mesmo domínio administrativo, é comum que exista o compartilhamento de recursos como discos, impressoras, etc. Mas quando a rede ultrapassa um domínio administrativo, este compartilhamento se torna muito limitado. A finalidade das grades de computação é permitir compartilhamento de recursos mesmo que estes estejam espalhados por diversos domínios administrativos.

Esta dissertação propõe uma arquitetura para o estabelecimento dinâmico de conexões multidomínio que faz uso da comutação de rajadas ópticas (OBS – *Optical Burst Switching*) utilizando um plano de controle GMPLS (*Generalized Multiprotocol Label Switching*). A arquitetura baseia-se no armazenamento de informações sobre recursos de grade de sistemas autônomos (AS – *Autonomous Systems*) distintos em um componente chamado Servidor GOBS Raiz (Grid OBS) e na utilização do roteamento explícito para reservar os recursos ao longo de uma rota que satisfaça as restrições de desempenho de uma aplicação. A validação da proposta é feita através de simulações que mostram que a arquitetura é capaz de garantir níveis de desempenho diferenciados de acordo com a classe da aplicação e proporciona uma melhor utilização dos recursos de rede e de computação.

PALAVRAS CHAVE: Redes Ópticas de Próxima Geração, Sistemas Autônomos, Qualidade de Serviço, Computação em Grade.

Abstract

The processing capacity of research institutions has been growing significantly as processors and workstations increasingly powerful emerging in the market. Considering the improvement of performance in the area computer networks and seeking to fill the demand for ever greater processing, the idea to use networked computers as independent platform for running parallel applications, thus creating the area grid computing. In a network that is under the same administrative domain, is common for there to share resources such as disks, printers, etc. But when the network goes beyond an administrative domain, this sharing becomes very limited. The purpose of computing grids is to allow sharing resources even if they are scattered in different areas administration.

This thesis proposes an architecture for dynamic establishment of connections multidomain which makes use of optical burst switching (OBS) using a GMPLS control plane (Generalized Multiprotocol Label Switching). The architecture is based on storing information about grid resources of autonomous systems in a separate component called GOBS Root Server (Grid OBS) and the use of explicit routing to reserve resources along a route that satisfies the performance constraints of an application. The validation of the proposal is made through simulations that show that the architecture is capable of levels of performance differed according to the application class and provides a better utilization of network resources and computing.

KEYWORDS: Next-Generation Optical Networks, Autonomous System, Quality of Service, Grid Computing.

Capítulo 1

Introdução

Neste capítulo serão discutidas as questões relacionadas aos desafios impostos pelo modelo de computação em grade para as redes de próxima geração. A Seção 1.1 discute o papel das redes de computadores nesse novo cenário de computação distribuída em larga escala, as Seções 1.2 e 1.3 apresentam respectivamente a motivação e os objetivos do trabalho. A Seção 1.4 apresenta a organização da dissertação.

1.1 – Redes de Próxima Geração

Aplicações científicas de alto desempenho continuam em destaque no cenário científico atual, como física de alta energia, astronomia, bioinformática, telemedicina, visualização remota, entre outras. *E-science*, nada mais é do que a caracterização desta nova realidade de pesquisa onde ocorre colaboração entre diversos grupos de pesquisa em localidades distintas, inclusive em outros continentes.

Estas aplicações tem como principais características o grande volume de informações manipulado, que pode chegar à ordem de *terabytes* (10^{12} bytes) e *petabytes* (10^{15} bytes) de dados, o alto poder de computação exigido que muitas vezes chega a *teraflops* (10^{12} operações de ponto flutuante por segundo) e a utilização de recursos heterogêneos, com diferentes arquiteturas de hardware e sistemas operacionais.

Com a sua definição primária, a computação em grade (*grid computing*)[1] é um modelo que propõe o uso de recursos computacionais de várias máquinas situadas em localizações distintas no globo, com o intuito de suprir as necessidades de aplicações que exigem grande poder de computação. As grades de uma forma geral, são compostas por além de computadores, por grandes repositórios de dados, equipamentos científicos controlados remotamente, dispositivos de visualização, sensores, entre outros, que oferecem possibilidades inéditas de cooperação entre cientistas. A Figura 1 ilustra um exemplo de aplicação de *e-science* (e.g., astrofísica) que usa grades.

Vale ressaltar que neste caso de aplicação em grade, faz-se necessário a simulação de eventos raros no universos, por exemplo, a simulação de buracos negros. Essas simulações geralmente requerem centenas de *terabytes* de memória, onde conseqüentemente geram petabytes de informação e sua execução costuma levar dias, mesmo fazendo o uso de milhares de processadores.

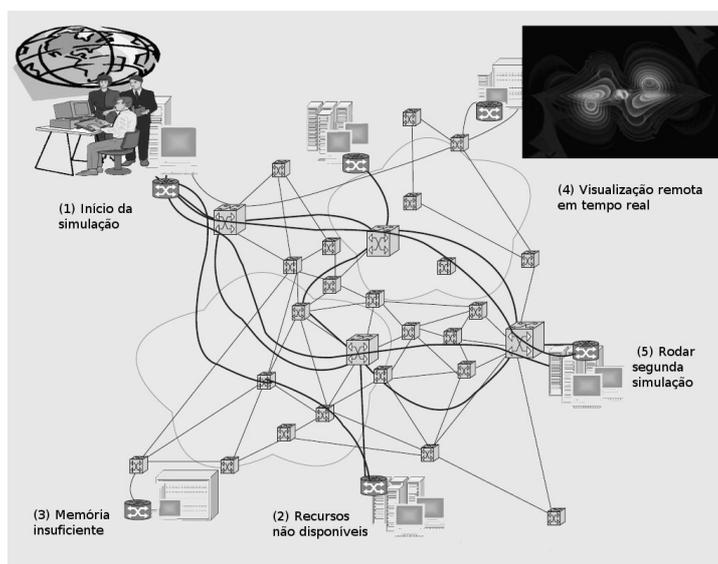


Figura 1: Exemplo de aplicação de astrofísica usando grades [2].

Neste exemplo, um pesquisador submete uma simulação que será executada em algum ponto de uma grade (1). Em determinados momentos os recursos computacionais requeridos poderão não estar disponíveis (2). O evento (3) acontece quando não há memória suficiente para possibilitar o armazenamento remoto dos dados gerados. Os resultados da simulação são então apresentados em dispositivos de visualização remota em tempo real (4). Com base nos resultados obtidos, o pesquisador pode solicitar novas simulações (5).

Devido ao alto custo de implantação de novos recursos, tais como memória, armazenamento e processamento, a computação em grade mantém-se como uma ótima opção, uma vez que se utilizam recursos já existentes, aproveitando a capacidade disponível dos mesmos para processar a demanda de uma aplicação. Com isso, organizações que não poderiam arcar com os custos de aquisição de novos equipamentos dedicados para conduzir seus experimentos possuem uma alternativa mais barata para viabilizar suas pesquisas.

As redes de computadores são fundamentais para a viabilidade das grades e, por isso, grandes esforços de pesquisa estão sendo conduzidos na tentativa de se propor modelos de comunicação mais adequados para atender os requisitos deste método de computação. Atualmente, os recursos de rede já são vistos no contexto de grades como parte de sua infraestrutura de recursos. Antes, a infraestrutura de rede que dava suporte a computação em grade era considerada um componente importante, porém não totalmente integrada aos demais recursos que constituíam a grade, mas sim, funcionando como um serviço externo. A rede é um recurso chave, da mesma forma que as unidades de processamento, a memória, os dispositivos de entrada e saída, entre outros, para assim proporcionar novas possibilidades de integração e combinação de serviços [3].

Além disso, em virtude do grande volume de dados manipulado, as grades necessitam de uma infraestrutura de comunicação robusta que seja adaptável às particularidades desse modelo. Os avanços nos meios de transmissão, principalmente no que se refere às tecnologias ópticas, foram feitos na tentativa de atender as exigências das novas aplicações dentre as quais se podem destacar as aplicações de grades que possuem requisitos estritos de desempenho.

Outro aspecto importante diz respeito ao controle da infraestrutura da rede que dá suporte à grade. A tendência é fazer com que os recursos de rede sejam alocados para as aplicações de forma dinâmica e transparente sem que haja a preocupação com a intervenção humana e fazendo com que essa alocação esteja em conformidade com os requisitos das aplicações da grade, ou seja, que os recursos selecionados possam garantir o funcionamento adequado das aplicações.

A definição de planos de controle e gerenciamento para essas redes permitirá que as aplicações possam solicitar e liberar recursos automaticamente, sem a intervenção de um administrador, e, além disso, possibilitará a aplicação de políticas de qualidade de serviço, de autenticação, autorização e contabilização, bem como mecanismos de descoberta de recursos e de segurança. Por essa razão, o desenvolvimento de propostas de plano de controle para a computação em grade vem ganhando bastante força na última década e impulsionado a área de controle e gerenciamento de redes.

1.2 – Motivação

O aperfeiçoamento nas tecnologias de transmissão era necessário devido aplicações de rede, como voz sobre IP, videoconferência, jogos on-line, streaming de vídeo, entre outras, que se caracterizam por serem baseadas em conteúdo multimídia e serem suscetíveis a determinadas propriedades da rede, como atraso, variação de atraso e perdas, baseando-se nesse fator, as redes ópticas tomaram o principal foco como meio de transmissão por serem um dos únicos capazes de suprir todas estas necessidades.

Dentre esses avanços podem-se destacar a multiplexação por divisão de comprimento de onda (WDM – *Wavelength Division Multiplexing*) [4], que permite que sinais ópticos distintos sejam transmitidos em uma única fibra óptica utilizando múltiplos comprimentos de onda (λ s) e proporcionando taxas de transmissão a ordem de terabits por segundo. Em outra frente, foram conduzidas pesquisas no intuito de contornar as limitações existentes nos sistemas de comutação atuais, como a necessidade de conversão do sinal óptico para o domínio elétrico e vice-versa, o que limitava a taxa máxima de transmissão com que o dispositivos de comutação poderiam trabalhar. A comutação puramente óptica veio a superar essa dificuldade, culminando em tecnologias que permitem que o sinal trafegue totalmente no domínio óptico, sem conversão. Isto possibilita taxas de transmissão próximas do limiar teórico das fibras ópticas (~100 Tbps) [5].

As abordagens de comutação puramente óptica são conhecidas como, a comutação de circuitos ou comutação de λ s, a comutação de pacotes ópticos e a comutação de rajadas ópticas. A comutação de λ s é semelhante à comutação de circuitos eletrônica, onde há a criação de um caminho de luz dedicado (*lightpath*) que

fica ativo durante toda a transmissão. Na computação de pacotes ópticos as informações de controle (cabeçalho) são processadas eletronicamente e os dados são encaminhados no domínio óptico. A principal dificuldade dessa alternativa é a ausência de estruturas de armazenamento temporário que sejam adequadas para os sinais ópticos, uma vez que os cabeçalhos necessitam de um tempo maior para serem processados, devido às limitações de velocidade dos dispositivos de comutação atuais, o que pode acarretar em problemas de sincronização [6].

A comutação de rajadas ópticas (OBS – *Optical Burst Switching*) [7] surgiu como uma alternativa híbrida entre a comutação de lambdas e a comutação de pacotes. A principal característica desse tipo de comutação é adotar a separação entre dados de controle e dados de aplicação, reservar os recursos somente pelo período de tempo necessário para a transmissão e por possuir um nível de granularidade de informações variável, com rajadas possuindo tamanhos intermediários entre um pacote e um circuito, o que facilita a comutação sem que haja desperdício de recursos.

Muitas aplicações científicas manipulam grandes quantidades de informação, sendo que a estrutura da Internet atual baseada na comutação de pacotes não se mostra adequada para esse tipo de aplicação, que normalmente necessitam de certas garantias (e.g., uma determinada taxa mínima de transmissão, ausência de perdas, atraso mínimo) para o seu funcionamento adequado. Por essa razão, as redes ópticas de próxima geração baseadas na comutação puramente óptica são naturalmente as melhores opções para constituir a chamada grade óptica (*Optical Grid*) [8].

A comutação OBS apresenta características que a tornam atrativa também para a computação em grade comparada ao que se utiliza atualmente nas grades baseadas em redes ópticas, normalmente comutação de lambdas. Nas redes OBS a informação pode possuir tamanho variável, que permite diferentes perfis de tráfego. As tarefas

(jobs) da grade podem ser mapeadas diretamente para rajadas ópticas. Além disso, o fato de haver separação entre dados de controle e de aplicação abre possibilidades para a implementação de mecanismos de descoberta de recursos, segurança, entre outros[9].

No que diz respeito ao plano de controle, a arquitetura GMPLS (*Generalized Multiprotocol Label Switching*) [10] é considerada uma alternativa para as redes puramente ópticas em virtude da facilidade de integração com a tecnologia WDM, quando os rótulos são associados com os comprimentos de onda da fibra e em virtude das funcionalidades de engenharia de tráfego e gerenciamento de enlaces, além de proporcionar flexibilidade de operação. O GMPLS é uma opção para a computação em grade, pois permite o provisionamento dinâmico de rotas, o que faz sentido em um ambiente heterogêneo e denso como a grade, tornando a operação da rede mais eficiente.

O aumento da capacidade de transmissão e a utilização de esquemas de comutação mais eficientes nem sempre garantem que uma determinada aplicação tenha seus requisitos de desempenho completamente satisfeitos, pois, se a quantidade de usuários e aplicações que podem solicitar os serviços de uma grade forem extremamente elevados é provável que comprometerá o funcionamento da mesma, o que na maioria das vezes implica em uma utilização intensa dos recursos disponíveis.

Oferecer conexões que apresentem poucas perdas, baixo atraso e baixa variação de atraso é fundamental para as aplicações em grade computacionais. Sendo assim, oferecer qualidade de serviço e controle de recursos para as aplicações de grades torna-se desafiador, ao ponto de abrir novas possibilidades de pesquisa na área de redes ópticas de próxima geração.

1.3 – Objetivos

O objetivo principal desta dissertação é propor um modelo para o estabelecimento automático de conexões interdomínio com garantias de desempenho dentro do contexto de uma rede que trabalha com o paradigma de comutação de rajadas ópticas em um ambiente de grade com plano de controle inspirado no GMPLS. Este modelo permite selecionar dinamicamente rotas que possam atender os requisitos de desempenho de uma aplicação e a partir daí, utilizando protocolos com funcionalidades de Engenharia de Tráfego, realizará a reserva efetiva dos recursos necessários da rede.

Nesta dissertação é proposto um componente responsável por gerenciar a disponibilidade de informações referentes aos recursos de rede e de computação, controle de requisições e encaminhamento de novas rotas interdomínio para sistemas autônomos (AS – *Autonomous Systems*) chamado de Servidor GOBS (*Grid OBS*) Raiz. Também são utilizados agentes de monitoração para atualizar as estatísticas de utilização dos recursos, assim como um servidor GOBS comum que auxiliará no cálculo de rotas dentro do seu AS. Com isso, pretende-se dar tratamento diferenciado para classes distintas de aplicações aproveitando a existência de recursos subutilizados na rede para evitar bloqueios excessivos, bem como para proporcionar uma melhor utilização dos recursos de computação garantindo os níveis de serviço definidos previamente.

Existem discussões na literatura sobre a natureza das funcionalidades de um plano de controle para a computação em grade [2][11]. Algumas funções são mais adequadas para serem realizadas por uma entidade de gerenciamento centralizada e outras se beneficiam de um plano de controle distribuído [11]. A proposta desta

dissertação procura estabelecer um balanceamento entre essas características, através da centralização das informações dos recursos, o que permite um cálculo preciso de rotas, e a distribuição de funções de sinalização e monitoração de recursos, o que permite rapidez na reserva dos recursos e na atualização das informações associadas a estes.

Este trabalho busca estender o paradigma de comunicação *anycast* para levar fortemente em consideração os recursos de rede, ou seja, os nós intermediários (núcleo) ao longo da rota para garantir os níveis de serviço desejados pela aplicação.

Outro objetivo desta dissertação está relacionado à proposta de alternativas de sinalização mais ágeis comparadas às tradicionais, com o intuito de minimizar o atraso na reserva de caminhos e aproveitar as características vantajosas da comutação de rajadas ópticas, como a não-obrigatoriedade de confirmação e para aproveitar de maneira mais realista o cenário em grade, é proposto um modo de gerenciamento de proativo que aplica políticas que se antecipam a na ocorrência de problemas na rede.

1.4 – Organização do trabalho

O restante desta dissertação está organizada da seguinte forma: o Capítulo 2 aborda questões relacionadas aos novos requisitos impostos pelo modelo de grades às redes ópticas e destaca os principais desafios que devem ser resolvidos para adequá-las a esse cenário.

No Capítulo 3 são apresentados os principais trabalhos relacionados à proposta dessa dissertação que serviram de base para o desenvolvimento da mesma.

A arquitetura proposta para o provisionamento automático de conexões determinísticas é descrita no Capítulo 4, onde são detalhados o Servidor GOBS Raiz, o

mecanismo de seleção de rotas, a arquitetura de monitoração e os modos de gerenciamento de qualidade de serviço.

O Capítulo 5 consiste na análise da proposta onde são descritos os componentes de simulação desenvolvidos, os cenários de simulação utilizados na análise e os resultados obtidos com as respectivas discussões.

Finalmente, o Capítulo 6 encerra o trabalho com uma avaliação geral da proposta e sugere alternativas de trabalhos futuros.

Capítulo 2

Desafios em Redes para Grades

Computacionais

Este capítulo aborda as principais questões referentes à adequação das redes de computadores ao modelo de computação em grade, uma vez que as redes atuais não contemplam totalmente os requisitos impostos pela computação em grade, como roteamento interdomínio, estabelecimento rápido e dinâmico de recursos, autenticação, autorização e contabilização, entre outros.

A Seção 2.1 apresenta os conceitos fundamentais de grades, lista as principais características que tornam as grades diferentes de outros sistemas de computação e apresenta algumas aplicações que demandam o uso de grades. A Seção 2.2 discute o papel das redes ópticas no contexto de grades. A Seção 2.3 aborda o uso da comutação de rajadas ópticas em grades, as iniciativas em andamento para padronizar essa integração e o plano de controle, com ênfase na solução GMPLS. Por fim, a Seção 2.4 apresenta as considerações finais do capítulo.

2.1 – Grades Computacionais

2.1.1 – Introdução

Devido o alto custo implementado nos recursos locais tornou-se inviável a aquisição de grande parques tecnológicos para até mesmo para grandes centros de pesquisa que necessitam de alto poder de computação. Visando este problema que

ocorre com todos, a computação em grade surgiu como alternativa para viabilizar a resolução de problema. O conceito de grade é uma analogia ao sistema de geração e distribuição de energia elétrica, que oferece facilidade de uso e está disponível para uma grande quantidade de usuários de forma transparente. A ideia é que a grade possa ser acessada por usuários que podem estar localizados em diferentes locais, sem que se conheçam os detalhes da implementação e do funcionamento dos serviços oferecidos.

De acordo com [12], uma grade é baseada no “compartilhamento de recursos coordenado e na resolução de problemas em organizações virtuais, multi-institucionais e dinâmicas”. Uma organização virtual é definida como um conjunto de indivíduos ou instituições que compartilham recursos de forma coordenada. São exemplos de organizações virtuais provedores de aplicações, provedores de serviços de armazenamento de dados, iniciativas de pesquisa envolvendo diversos grupos, consórcios empresariais, etc.

A computação em grade teve um crescimento significativo em um momento onde a demanda por aplicações científicas sofisticadas impulsionou o desenvolvimento de soluções que fossem capazes de contornar as limitações de processamento e armazenamento dos sistemas de computação a nível local. Por essa razão, as grades surgiram como uma alternativa de custo relativamente mais baixo para viabilizar diversas pesquisas científicas, e, além disso, para permitir um nível maior de colaboração entre grupos de pesquisa que podem se encontrar espalhados geograficamente.

A ideia fundamental das grades é o compartilhamento de recursos de forma coordenada. Isto implica que deve haver regras que regem esse compartilhamento.

Por exemplo, que recursos estão disponíveis, quem pode acessar esses recursos e de que forma ocorrerá o acesso [13].

Grades também se caracterizam de uma forma geral pelo baixo custo comparado a alternativas locais como clusters e supercomputadores paralelos, uma vez que se beneficiam da disponibilidade de uma grande quantidade de recursos já existentes. Outra característica é a heterogeneidade, pois a grade é composta de uma variedade de sistemas cada um com suas particularidades de hardware e software. Por essa razão, as grades são ambientes que pressupõem um alto grau de abstração para facilitar o compartilhamento e a integração desses recursos.

2.1.2 – Propriedades de uma Grade

Uma grade possui características que a diferenciam de outros sistemas de computação. Algumas dessas características são descritas a seguir [3]:

- **Abstração:** a visão da grade como um ambiente virtualizado permite a definição de uma variedade de serviços customizáveis sem que se precise conhecer detalhes dos componentes que implementam esses serviços, tais como sua localização e funcionamento. Isto significa que, do ponto de vista do usuário, a grade nada mais é que um aglomerado de recursos que podem ser solicitados e integrados para servir a um propósito específico.
- **Flexibilidade:** a flexibilidade está relacionada com a capacidade que a grade tem de se organizar para atender determinadas solicitações. Por exemplo, pode-se agendar a utilização de um serviço para um período de tempo futuro, realizar mudanças na alocação de recursos em tempo real, e assim por diante, de forma semelhante ao que é feito em sistemas de

workflow.

- Controle descentralizado: grades se caracterizam por proporcionar funcionalidades que não dependem da existência de uma entidade centralizada de gerenciamento. Os recursos são controlados de forma distribuída, pois estão normalmente localizados em domínios administrativos distintos.
- Pervasividade: as grades podem englobar uma infinidade de sistemas e se espalhar em áreas geográficas bastante extensas, inclusive rompendo barreiras continentais.
- Compartilhamento de recursos: é a principal motivação para justificar a existência de uma grade, onde se deseja que uma grande quantidade de recursos distribuídos seja disponibilizada para as diversas organizações virtuais que fazem parte da grade. Os recursos são compartilhados pelos membros de uma organização virtual, mas podem ser utilizados por mais de uma organização virtual ao mesmo tempo [14].
- Escalabilidade: grades são altamente escaláveis, uma vez que podem possuir diferentes extensões, sendo possíveis implementações tanto a nível local como a nível continental.
- Determinismo: em grades é desejável que certas aplicações possam ter garantias de níveis de serviço e de recursos disponíveis para que funcionem adequadamente.
- Segurança: a questão da segurança é fundamental em grades, principalmente para ambientes de grande extensão, que englobam entidades (reais ou virtuais) distintas, que por sua vez possuem restrições de acesso a determinados recursos e informações.

Foster [10] sugere uma lista de propriedades que devem ser satisfeitas para caracterizar uma grade. De acordo com essa lista uma grade é um sistema que:

- Coordena recursos que não estão sujeitos a um controle centralizado, ou seja, permite a participação de usuários e recursos situados em domínios de gerenciamento distintos.
- Utiliza protocolos e interfaces de propósito geral, abertos e padronizados, facilitando o desenvolvimento de novos serviços.
- Oferece vários tipos não-triviais de qualidade de serviço, o que permite atender aos diferentes perfis de usuários que fazem parte da grade.

2.1.3 – Padronização de Grades

Para popularizar o uso de grades e facilitar o desenvolvimento de aplicações e serviços para esses ambientes é necessário a criação de padrões bem definidos. A principal instituição responsável por regulamentar e padronizar as grades é o *Open Grid Forum* (OGF) [15]. O OGF é a fusão do *Global Grid Forum* (GGF) e da *Enterprise Grid Alliance* (EGA).

Dentre os diversos padrões estabelecidos pelo OGF pode-se destacar o OGSA (*Open Grid Services Architecture*) [16] que define uma arquitetura aberta e padronizada para o desenvolvimento de aplicações de grade. O OGSA especifica uma arquitetura orientada a serviços (SOA – *Service Oriented Architecture*) para grades baseada em Web Services definindo serviços como: Gerência de Execução, Dados, Gerência de Recursos, Auto-Gerenciamento e Informações [74].

2.2 – Redes Ópticas em Grades

2.2.1 – Introdução

As redes ópticas se constituem na base da infraestrutura de comunicação de grades que estão sendo construídas atualmente, em virtude principalmente das altas taxas de transmissão que são possíveis com essa tecnologia, o que é fundamental para as grades, uma vez que certas aplicações científicas geram uma grande quantidade de informações, que muitas vezes devem ser movidas de/para múltiplos centros de pesquisa, laboratórios, etc. Portanto, para agilizar a transmissão desses dados gerados é necessária uma rede de alta velocidade operando com taxas de transmissão que variam de dezenas de gigabits por segundo até terabits por segundo.

As grades que utilizam redes ópticas como um recurso chave, ou seja, como entidades de primeira classe, da mesma forma que os recursos de computação, são denominadas grades ópticas [8]. As grades ópticas visam atender a necessidade de aplicações que demandam a transmissão de grandes conjuntos de dados em longas distâncias. Alguns exemplos de aplicações e iniciativas que fazem o uso de grades e que utilizam redes ópticas são listadas a seguir:

- LHC (Large Hadron Collider): o LHC [17] é o maior acelerador de partículas do mundo, onde a quantidade de dados gerada pode chegar a dezenas de petabytes por ano, sendo que essas informações são distribuídas para sites remotos ao redor do globo para análise.

- VLBI (Very Long Baseline Interferometry): Dentro do projeto EVN (European VLBI Network) [18], informações coletadas a partir de rádio telescópios distribuídos pela Europa são processadas em uma grade local.
- Edição de conteúdo multimídia: atualmente a manipulação de vídeo em alta definição demanda capacidades computacionais que não podem ser satisfeitas localmente. Grades muitas vezes são utilizadas para proporcionar facilidades de edição e visualização remota de conteúdo multimídia.
- OptIPuter [19]: É uma iniciativa que visa interligar recursos computacionais através de redes ópticas utilizando o protocolo IP. O objetivo é permitir que cientistas que manipulam grandes quantidades de informação possam visualizar, analisar e correlacionar esses dados de forma integrada.
- Mineração de dados: consiste na análise de grandes conjuntos de dados para identificar padrões, relações, anomalias que possam ser de interesse. Devido ao aumento crescente dessas massas de dados, se faz necessário muitas vezes que estes sejam distribuídos e analisados paralelamente em diversas localidades.
- CineGrid [20]: CineGrid é uma iniciativa que propõe a distribuição de conteúdo multimídia em altíssima definição. Entre os objetivos estão: a distribuição segura de conteúdo proprietário da indústria cinematográfica, videoconferência em alta definição (HDTV) e a educação remota através de aulas ministradas por de meio de vídeo de alta qualidade.

Avanços nas tecnologias de redes ópticas foram realizados nos últimos anos com o objetivo de ampliar a capacidade de transmissão e proporcionar largura de banda suficiente para o suporte a novos serviços. Um desses aperfeiçoamentos é

multiplexação por divisão de comprimento de onda (WDM) e a sua variante em larga escala (DWDM – *Dense WDM*), que permite que sinais distintos possam ser transmitidos simultaneamente na mesma fibra em diferentes comprimentos de onda cada um operando na casa de gigabits por segundo [21].

Além disso, as redes ópticas de próxima geração deverão ser baseadas na comutação puramente óptica, que elimina a necessidade de conversão do domínio eletrônico para o domínio óptico e vice-versa (OEO) dentro dos dispositivos de comutação. Por meio de tecnologias como MEMS (*Micro-Electrical Mechanical Systems*) e SOA (*Semiconductor Optical Amplifiers*), pretende-se desenvolver dispositivos de comutação capazes de realizar o encaminhamento de sinais totalmente no domínio óptico (OOO), sem a necessidade de conversão, o que permitiria taxas de transmissão mais altas que as obtidas atualmente devido às limitações de velocidade de inerentes aos dispositivos eletrônicos de comutação [22].

Os dispositivos baseados em MEMS possuem velocidades de comutação mais baixas se comparadas às obtidas em comutadores baseados em SOAs [75]. Por essa razão, MEMS são mais adequados quando não se exigem comutadores muito rápidos, como os geralmente utilizados na comutação de circuitos ópticos e na comutação de rajadas ópticas. Já SOAs são necessários para viabilizar a comutação de pacotes ópticos, pois as velocidades de comutação requeridas neste caso são bastante elevadas.

O desenvolvimento de novos planos de controle para essas redes, como os baseados na arquitetura GMPLS definida pela IETF, permite que os recursos de rede sejam acessados de forma direta e dinâmica pelos usuários e aplicações, o que facilita a visão da rede como uma coleção de recursos que podem ser reservados de acordo com uma demanda específica.

Esses e outros avanços contribuíram para que as redes ópticas sejam consideradas como blocos essenciais para a construção de uma infraestrutura de comunicação que atenda as exigências da computação em grade. Esforços estão sendo realizados para identificar e explorar as vantagens de se utilizar redes ópticas em grades e propondo formas de se abstrair os recursos das redes ópticas, tornando-os entidades de grade que se caracterizam por serem gerenciáveis, dinâmicas e capazes de se adequar às demandas complexas das aplicações de grade e proporcionando, assim, uma melhor utilização dos recursos da rede [8]. Nesse sentido, se faz necessária a integração do plano de controle com o middleware da grade, para proporcionar alocação de recursos de rede de forma transparente, semelhante ao que ocorre com os recursos de computação. O grupo de trabalho NSI (*Network Service Interface*) do OGF (*Open Grid Forum*) atua no desenvolvimento e na padronização de soluções para criar uma interface entre o middleware da grade e a infraestrutura de rede que a suporta [23].

Várias iniciativas práticas utilizando redes ópticas para interligar recursos de computação de alta capacidade vêm sendo desenvolvidas ao redor do mundo. Alguns desses projetos são listados a seguir:

O projeto EnLIGHTened [24] consiste no uso de redes de alta velocidade para conectar recursos de computação de alto desempenho e instrumentos científicos distribuídos geograficamente de forma dinâmica e coordenada. O projeto procura estudar formas de integrar o plano de controle com o middleware da grade em cenários com uma grande quantidade de recursos heterogêneos que por sua vez são requisitados com bastante frequência. Algumas questões investigadas estão relacionadas com monitoração e descoberta de recursos, cumprimento de acordos de

serviço, garantias de desempenho, extensões para protocolos de sinalização e controle, entre outras.

Os projetos DRAGON (*Dynamic Resource Allocation via GMPLS Optical Networks*) [25] e PHOSPHORUS [26] têm como objetivo investigar questões relacionadas ao estabelecimento de conexões que podem atravessar vários domínios e que serão disponibilizadas para aplicações avançadas, como as grades. O DRAGON propõe um plano de controle baseado no GMPLS para garantir, além do estabelecimento dinâmico de caminhos em domínios distintos, o cumprimento de políticas de AAA (Autenticação, Autorização e Aferição) e de qualidade de serviço. O PHOSPHORUS pretende servir como base de testes para o oferecimento de serviços que englobem: múltiplas redes de pesquisa, planos de controle e gerenciamento, equipamentos.

A iniciativa GENI (*Global Environment for Network Innovations*) [27] propõe uma infraestrutura de rede para servir de base para experimentos de ciência e engenharia em redes. Nessa estrutura os usuários poderão solicitar uma quantidade específica de recursos para um determinado experimento cuja realização não seria possível localmente. Se o experimento crescer de tal forma que a quantidade de recursos alocados não for mais suficiente, o usuário poderá solicitar mais recursos para o seu experimento caso seja possível. O GENI tem como intuito de servir como um laboratório para o estudo de novas arquiteturas e soluções de rede.

Na Europa, o projeto FIRE (*Future Internet Research and Experimentation*) [28] pretende ser um ambiente que permite a experimentação de novas arquiteturas de rede na prática em condições semelhantes às encontradas em um ambiente de rede de produção. O FIRE pode ser entendido como um conjunto de *testbeds* que compartilham uma infraestrutura de rede. No Japão, a iniciativa [29] propõe a

construção de uma rede de nova geração que visa contornar as limitações existentes na Internet atual no que diz respeito ao desenvolvimento de uma nova arquitetura de rede que atenda requisitos tecnológicos e sociais.

2.2.2 – Requisitos de Rede para Grades

Em virtude das exigências impostas pelas aplicações de grades, surgem novos requisitos que as redes de computadores devem satisfazer para se adequar às características desse modelo de computação se caracteriza por permitir o controle e o gerenciamento descentralizado dos recursos que fazem parte da mesma, sem que seja necessária a existência de uma entidade central de gerenciamento. Em redes isso significa que os recursos podem ser alocados e controlados de forma dinâmica e independente, sendo que a aplicação tem papel fundamental na reserva dos recursos interagindo com o middleware da grade e com o plano de controle da rede. Alguns dos principais requisitos de rede que surgem no contexto de grades são listados a seguir [3]:

- Roteamento interdomínio: grades normalmente são constituídas de recursos de computação heterogêneos que podem estar situados em domínios diferentes, passando por redes que pertencem a organizações distintas. Nesse contexto algumas questões que devem ser levantadas estão principalmente relacionadas à confidencialidade das informações que trafegarão nessas redes, já que informações privadas podem passar por redes de terceiros; ao provisionamento automático de recursos sem a intervenção humana, sendo que isso deve ser feito da forma mais transparente possível para a aplicação; e ao oferecimento de garantias de qualidade de serviço fim-a-fim, onde cada domínio deve se responsabilizar por proporcionar o desempenho requerido.

- **Agendamento Futuro de Conexões:** muitas aplicações de grade e *e-science* necessitam solicitar recursos para um tempo futuro. Essas situações são mais comuns quando se deseja reservar uma grande quantidade de recursos por um período de tempo definido. Por exemplo, a transmissão de informações coletadas por instrumentos científicos remotos exige que os enlaces ao longo do caminho até o destino onde essas informações serão processadas sejam reservados previamente devido à grande quantidade de dados envolvida nesse tipo de comunicação. Por essa razão, se fazem necessárias extensões aos protocolos de sinalização e controle para que estes sejam capazes de reservar recursos para um tempo futuro com uma duração específica.
- **Estabelecimento Rápido de Conexões:** certas classes de aplicações (e.g., multimídia) são sensíveis a propriedades da rede como atraso e variação de atraso. Portanto, se o tempo necessário para o estabelecimento de um caminho for muito grande comparado ao tempo de transmissão e propagação da informação em si então esse atraso pode comprometer o desempenho da aplicação de uma maneira geral. O uso de planos de controle distribuídos como o GMPLS permite que os caminhos sejam construídos de forma mais ágil, já que esse estabelecimento não depende de uma consulta a um elemento centralizado.
- **Suporte a diferentes perfis de utilização:** as grades são normalmente utilizadas por uma grande variedade de usuários e aplicações que possuem diferentes perfis de tráfego e de utilização dos recursos da grade, incluindo aí os recursos de rede. Algumas aplicações científicas necessitam de uma grande quantidade de recursos por um longo período, por exemplo, os dados gerados pelo acelerador

de partículas LHC deverão ser da ordem de *petabytes*, enquanto usuários e aplicações comuns manipulam uma quantidade menor de informações.

- **Determinismo:** Certas aplicações de grade necessitam de garantias de desempenho para o seu funcionamento adequado. Essas garantias podem envolver um limite máximo para as perdas, latência, disponibilidade de uma determinada quantidade de recursos para um período específico de tempo, entre outras. Determinismo está associado com a ideia de que os parâmetros de desempenho estabelecidos pelo usuário/aplicação devem ser totalmente satisfeitos, ou seja, proporcionando qualidade de serviço de forma absoluta.

2.3 – Grids OBS

2.3.1 – Comutação em Rajadas Ópticas

A comutação puramente óptica surgiu com o objetivo de eliminar a necessidade de se converter o sinal do domínio óptico para o domínio eletrônico e vice-versa. Essa limitação resulta em uma diminuição das taxas de transmissão que podem ser oferecidas, uma vez que os equipamentos de comutação eletrônicos não são capazes de processar informação em taxas muito elevadas quando comparadas à capacidade disponível nas fibras ópticas. Em virtude dessa limitação, e as redes baseadas nesse tipo de comutação ficam restritas a taxas de transmissão na ordem de dezenas de *gigabits* por segundo, e, além disso, essa conversão implica em custos adicionais aos equipamentos de comutação. Na comutação puramente óptica¹, os dados em si são transmitidos no domínio totalmente óptico, de forma transparente, o que permite que se atinjam taxas de transmissão mais elevadas que as atuais.

¹ O termo comutação puramente óptica, a partir deste ponto será caracterizado pelo termo comutação óptica.

Foram propostas três alternativas para proporcionar a comutação óptica que diferem basicamente no nível de granularidade da informação e na complexidade/custo de implementação. Essas três formas de se implementar a comutação óptica são descritas a seguir [4]:

- Comutação de circuitos ópticos (*OCS – Optical Circuit Switching*): a comutação de circuitos ópticos, ou comutação de lambdas, se caracteriza por sua semelhança com a comutação de circuitos eletrônica, onde há o estabelecimento prévio de um caminho de luz (*lightpath*) dedicado para a transmissão. A vantagem dessa técnica é a baixa complexidade de implementação, mas, por outro lado, os recursos da rede são subutilizados.
- Comutação de pacotes ópticos (*OPS – Optical Packet Switching*): a comutação de pacotes ópticos procura se aproximar ao máximo da comutação de pacotes eletrônica, uma vez que a comutação de pacotes proporciona multiplexação estatística, resultando em uma alta utilização da rede. No entanto, a comutação de pacotes no contexto óptico apresenta muitos desafios. O principal está relacionado à ausência de uma estrutura de armazenamento temporário a nível óptico que seja simples e barata, e esse *buffer* óptico é necessário uma vez que as informações de controle do pacote (cabeçalho) ainda precisam ser processadas de forma eletrônica. Outra questão se refere à sincronização entre os cabeçalhos e os dados, o que pode acarretar em diminuição nas taxas de transmissão.
- Comutação de rajadas ópticas (*OBS – Optical Burst Switching*): a comutação de rajadas ópticas se propõe a contornar as limitações dos outros paradigmas, como a baixa utilização dos recursos e a complexidade de implementação. Na comutação OBS, os pacotes são agregados em unidades chamadas rajadas antes de serem transmitidos no domínio óptico. Existe uma sinalização preliminar (*BCP – Burst Control Packet*) que é responsável por reservar os recursos necessários ao longo da rota pelo tempo necessário para a transmissão da rajada.

As vantagens da comutação de rajadas ópticas são: a alocação sob demanda dos recursos permite uma utilização mais eficiente, já que os recursos são reservados

somente durante o tempo necessário para a transmissão da rajada; na maioria das implementações de OBS não há a necessidade de confirmação acerca da reserva dos recursos, o que diminui a latência; o processamento eletrônico dos pacotes de controle abre possibilidades para a implementação de esquemas de qualidade de serviço, segurança, entre outros e a granularidade intermediária de uma rajada permite a utilização de dispositivos de comutação menos complexos e, conseqüentemente, mais baratos. A Figura 2 ilustra o funcionamento básico do OBS. A Tabela 1 mostra uma comparação entre os paradigmas de comutação óptica.

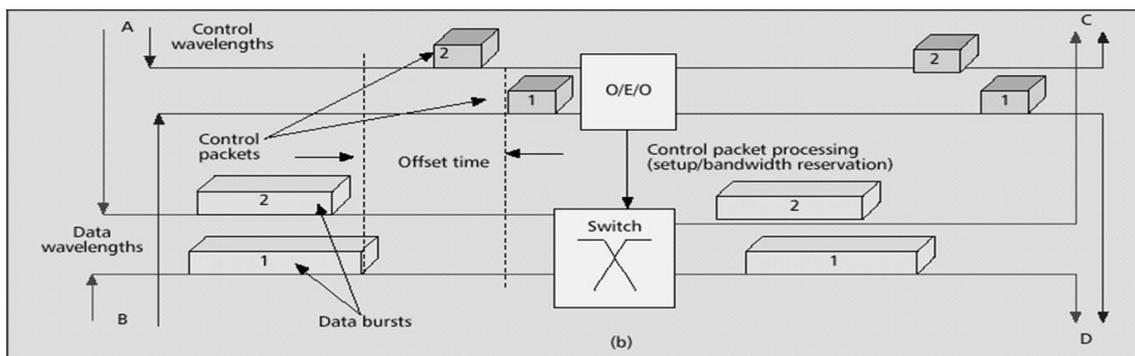


Figura 2: Comutação de Rajadas Ópticas[60].

A figura ilustra o processo de sinalização em redes OBS conhecido como JET (*Just Enough Time*) [30] que é o mais representativo das características do OBS. Nesse esquema, os pacotes de controle (numerados 1 e 2) são enviados em um canal de controle e são processados eletronicamente com o objetivo de configurar os dispositivos ao longo da rota para a chegada futura da rajada de dados. Após o tempo de ajuste (*offset time*) as rajada são enviadas no domínio óptico em comprimentos de onda distintos através de encaminhamento direto (*cut-through*) sem a necessidade de conversão para o domínio eletrônico.

Tabela 1: Comparação entre os paradigmas de comutação óptica [6].

Paradigma de comutação	Utilização de banda	Latência de Sinalização	Velocidade de comutação requerida	Overhead de processamento e sincronização	Adaptação ao tráfego
<i>Comutação de circuitos ópticos</i>	Baixa	Alta	Baixa	Baixo	Baixa
<i>Comutação de pacotes ópticos</i>	Alta	Baixa	Alta	Alto	Alta
<i>Comutação de rajadas ópticas</i>	Alta	Baixa	Média	Baixo	Alta

É possível observar que a comutação de rajadas ópticas reúne os aspectos positivos dos demais paradigmas, como a alta utilização dos recursos disponíveis, a baixa latência de sinalização devido à ausência de confirmação, um baixo *overhead* de processamento e sincronização, uma velocidade de comutação intermediária, pois o tempo que a rajada necessita para atravessar o comutador é normalmente maior do que na comutação de pacotes, uma vez que as rajadas possuem granularidade superior a de um simples pacote. Além disso, existe o tempo de ajuste que permite que os dispositivos de comutação tenham mais tempo para serem configurados e receber a rajada. Como as rajadas podem ser de diferentes tamanhos, a comutação OBS pode atender diferentes tipos de aplicação.

2.3.2 – Redes OBS em Grades

2.3.2.1 – Introdução

Conforme explicitado na seção anterior, dentro do contexto de redes ópticas de nova geração existem basicamente três abordagens de comutação óptica: a comutação de lambdas (circuitos), a comutação de pacotes ópticos e a comutação de rajadas ópticas. A comutação de rajadas ópticas (OBS) procura minimizar as limitações dos outros paradigmas de comutação, como a baixa utilização dos recursos e a complexidade de implementação [6].

As redes ópticas usadas em grades baseadas na comutação de lambdas, ou circuitos, que são as mais utilizadas atualmente, apresentam algumas desvantagens dentro do contexto de grades [31]:

- A construção de um circuito dedicado para cada requisição oriunda de uma aplicação de alto desempenho pode não se mostrar eficiente em todos os casos, principalmente quando não há tráfego contínuo, o que

pode ocasionar o desperdício de recursos;

- Grades apresentam uma grande variedade de usuários e aplicações que implicam na imprevisibilidade nas requisições de serviços. Uma arquitetura inflexível como a baseada em comutação de lambdas pode limitar a quantidade de usuários que poderiam ser atendidos simultaneamente;
- Em algumas situações, as demandas das aplicações de grade têm uma duração pequena comparada o tempo de estabelecimento de um recurso dedicado, exceto quando se trata *terabytes* e *petabytes* de informação.

A comutação OBS vem sendo considerada uma séria candidata para atender a computação em grade por uma série de razões, dentre as quais se destacam [33]:

- As requisições (tarefas) de uma aplicação podem ser diretamente mapeadas para rajadas ópticas. A granularidade variável das informações na rede OBS permite diferentes perfis de tráfego;
- A separação entre dados de controle (BCP) e dados de aplicação (rajadas) proporciona transmissões de dados de forma totalmente óptica, sem a necessidade de se converter o sinal do domínio óptico para o domínio eletrônico e vice-versa;

O processamento eletrônico dos pacotes de controle permite a adição de novas funcionalidades no contexto de grades, como descoberta inteligente de recursos e segurança.

2.3.2.2 – Arquitetura de uma Rede Grid OBS

Em uma rede grid OBS, podem-se destacar dois tipos de nós: os nós de núcleo (*core nodes*) e os nós de borda (*edge nodes*). Dentro do contexto de grades, existem diferentes perfis de utilização dos recursos disponíveis que resultam em requisições de tamanho e durações variáveis. Isto implica que poderá haver grande variação no tamanho das rajadas que serão transmitidas pela rede e a mesma deve se adaptar a essas variações, sendo que para rajadas pequenas (de 100 a 1000 *bytes*) a velocidade de comutação exigida é extremamente alta. Portanto, para os nós de núcleo, o desafio é construir dispositivos que sejam capazes de fazer a comutação em velocidades muito altas, quando necessário. Logo, o ideal é que na rede se tenha comutadores de alta velocidade, baseados em SOAs para as rajadas menores em conjunto com comutadores de baixa velocidade, baseados em MEMS para as rajadas de maior granularidade.

Os nós de borda em uma rede grid OBS são responsáveis pelas funções de classificação de tarefas; agregação de tráfego e montagem de rajadas; bem como a definição de interfaces entre a rede e os usuários (*GUNI – Grid User to Network Interface*) e entre a rede e os recursos da grade (*GRNI – Grid Resource to Network Interface*) [33], que são ilustradas na Figura 3.

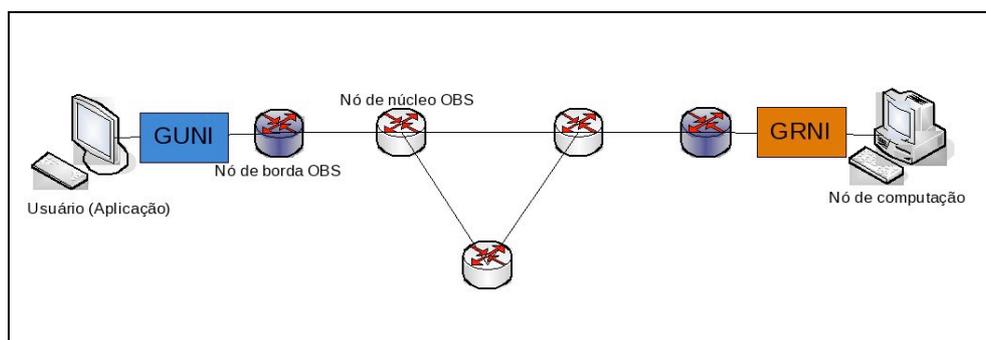


Figura 3: GUNI e GRNI em uma rede OBS[33].

A classificação de tarefas é responsável por auxiliar na implementação de serviços diferenciados em grades (*GridDiffServ*) e também no escalonamento de tarefas

e na descoberta de recursos [33]. A agregação de tráfego e montagem de rajadas varia de acordo com o tipo de mecanismo para determinar o processo de montagem, por exemplo, um tempo máximo de montagem ou o tamanho máximo de rajadas.

A GUNI se constitui de um conjunto de funcionalidades que permitem que o usuário interaja diretamente com os serviços oferecidos pela grade. Por exemplo, usuário pode solicitar caminhos para uma determinada tarefa que obedeçam a certas restrições, como banda, atraso, etc. Para isso, são necessários aperfeiçoamentos ao plano de controle das redes para proporcionar essas facilidades.

Algumas funções da GUNI são: alocação flexível de banda, solicitação de acordos de serviço já existentes, estabelecimento de caminhos de forma automática, classificação e agregação de tráfego.

Para garantir que a rede tenha conhecimento acerca da disponibilidade dos recursos de computação (*e.g.*, processamento e armazenamento) é necessário que exista uma interface entre a rede e os demais componentes da grade. Esse papel é feito pela GRNI, cujas funções principais são: envio das tarefas para os sistemas finais, suporte a esquemas de reserva antecipada de recursos e propagação do estado dos recursos de computação disponíveis.

2.3.2.3 – Sinalização e Controle

Em redes OBS existe uma sinalização representada pelo BCP que é responsável por reservar os recursos ao longo da rota e configurar os dispositivos de comutação para que a rajada possa ser encaminhada totalmente no domínio óptico. O BCP é processado de forma eletrônica e, após um determinado tempo, denominado tempo de ajuste (*offset time*), a rajada é enviada. As alternativas de sinalização para redes OBS mais conhecidas são o JET (*Just Enough Time*) [30] e o JIT (*Just In Time*) [32]. Essas

abordagens diferem pelo fato de que no JET a reserva pode ser atrasada para um tempo futuro e a liberação dos recursos é implícita sem a necessidade de nenhum tipo de mensagem, já no JIT a reserva é imediata e a liberação deve ser feita de forma explícita através de uma mensagem que sinalize a liberação.

Existem várias maneiras de manipular esse tempo de ajuste, sendo normalmente relacionadas com a implementação de esquemas de qualidade de serviço. Um tempo de ajuste pequeno diminui o atraso de uma maneira geral, porém gera muita informação de controle e exige dispositivos de comutação mais rápidos. Por outro lado, um tempo de ajuste mais elevado pode reduzir o bloqueio e a quantidade de sinalização, no entanto, a latência é maior.

Outra questão importante diz respeito ao escalonamento de rajadas. Escalonamento de rajadas consiste em determinar qual o comprimento de onda mais adequado para receber uma rajada, procurando evitar ao máximo o bloqueio e minimizar a latência. Alternativas como o LAUC-VF [34] procuram reduzir o espaço (*gap*) entre as rajadas utilizando uma técnica chamada *void filling*, alocando novas rajadas nos espaços vazios quando for possível.

A definição de um plano de controle para gerenciar de forma mais ágil e eficiente os recursos nas redes ópticas vem sendo bastante investigada nos últimos anos e diversas soluções foram propostas para reduzir a intervenção humana na gerência e reduzir os custos de operação da rede. Um plano de controle caracteriza uma série de funcionalidades, e seus respectivos protocolos, relacionadas à operação de uma infraestrutura de rede. Entre as principais funções de um plano de controle estão o roteamento, que define o caminho por onde as informações serão encaminhadas, e a sinalização que é responsável pela reserva efetiva dos recursos ao longo da rota definida pelo protocolo de roteamento.

De acordo com GLIF [35], um plano de controle define uma “infraestrutura e uma inteligência distribuída que controla o estabelecimento e a manutenção de conexões na rede, incluindo protocolos e mecanismos para disseminar essa informação; e algoritmos para construir caminhos ótimos entre dois pontos”. Portanto, o plano de controle é parte fundamental em redes de próxima geração e na computação em grade essa importância é acentuada, uma vez que os usuários e aplicações deverão ter acesso direto aos recursos e aos serviços da grade, e para isso é necessário que a rede tenha inteligência suficiente para proporcionar o controle dinâmico de conexões de forma automática e transparente.

Diversas propostas de plano de controle vêm sendo propostas para as redes ópticas de próxima geração, sendo a maioria baseada na arquitetura GMPLS (*Generalized Multiprotocol Label Switching*) [10] definida pelo IETF. O GMPLS estende a comutação baseada em rótulos [36] para incluir vários tipos de interface, como:

- Interfaces PSC (*Packet Switch Capable*): são interfaces que reconhecem pacotes e encaminha dados baseado no cabeçalho do pacote ou no valor do rótulo MPLS.
- Interface L2SC (*Layer-2 Switch Capable*): são interfaces que reconhecem frames/células, comutando dados com base no cabeçalho do frame/célula. São exemplos de dispositivos que suportam interfaces desse tipo são comutadores Ethernet e comutadores ATM.
- Interfaces TDM (*Time Division Multiplexing Capable*): esse tipo de interface comuta dados com base em intervalos de tempo fixos que se repetem em ciclos, a exemplo do que ocorre nas redes SONET/SDH.

- Interfaces LSC (*Lambda Switch Capable*): são interfaces baseadas no comprimento de onda que transporta os dados, normalmente utilizadas em OXCs e PXCs, dispositivos capazes de operar no nível de comprimento de onda.
- Interfaces FSC (*Fiber Switch Capable*): são interfaces que comutam dados com base na posição física dos dados, por exemplo, uma porta física de um comutador óptico OXC.

O GMPLS é baseado em três funcionalidades principais: roteamento, sinalização e gerenciamento de enlaces. Roteamento nesse contexto está mais associado à descoberta de recursos na rede através do gerenciamento e na divulgação de informações sobre a topologia [37], sendo que o protocolo utilizado para esse fim é o OSPF-TE [38]. A sinalização é responsável por reservar os recursos para o estabelecimento de conexões. Os protocolos de sinalização mais utilizados são o RSVP-TE [39] e o CR-LDP [40], sendo que o CR-LDP vem caindo em desuso. O gerenciamento de enlaces permite que cada nó da rede determine a identidade dos seus vizinhos e os detalhes da conectividade entre esses nós, essas informações são necessárias para a descoberta de topologias e para a sinalização. O LMP [41] é o protocolo que realiza o gerenciamento de enlaces na arquitetura GMPLS.

A utilização de um plano de controle distribuído como o GMPLS em grades é adequada para situações onde se deseja um estabelecimento rápido de conexões e facilidades de descoberta de recursos e engenharia de tráfego. O GMPLS oferece essas funcionalidades por meio de seus protocolos: OSPF-TE, para roteamento, RSVP-TE, para sinalização e LMP, para gerência de enlaces, além de ser facilmente integrado às redes ópticas de nova geração, onde os rótulos (*labels*) são associados aos comprimentos de onda na fibra. Em suma, o plano de controle GMPLS oferece

inteligência ao núcleo da rede, o que é de fundamental importância para as grades, onde se deseja um controle melhorado de todos os recursos, incluindo dos recursos de rede.

Ao contrário do MPLS tradicional, o GMPLS dá suporte a LSPs (*Label Switched Paths*) bidirecionais e a novas formas de utilização e representação do rótulo como o rótulo geral (*generalized label*), rótulo comutado por banda de ondas (*waveband switching label*) e o rótulo sugerido (*suggested label*). O rótulo geral pode representar um pacote, um comprimento de onda de fibra, uma fatia de tempo ou uma porta física. Rótulos comutados por banda de ondas definem um conjunto de comprimentos de onda (*lambdas*) contíguos que podem ser comutados da mesma forma. O rótulo sugerido é usado quando se quer agilizar a configuração dos equipamentos internos da rede. Nesse caso, sugere-se um rótulo para o dispositivo que pode aceitá-lo ou não. Esse processo é bem mais rápido do que a distribuição de rótulos tradicional, que se assemelha ao estabelecimento de um circuito [10][42].

É importante ressaltar que o GMPLS permite a cooperação entre redes heterogêneas de forma simplificada e transparente, onde é possível estabelecer caminhos rotulados que são compostos por diversos domínios de comutação. Por exemplo, pode-se estabelecer uma rota que tenha restrições de largura de banda e atraso para fluxos que iniciam em uma rede IP, são transportados por SONET, e encaminhados por meio de um comprimento de onda específico de uma fibra. A Figura 4 ilustra esse conceito de LSPs heterogêneos.

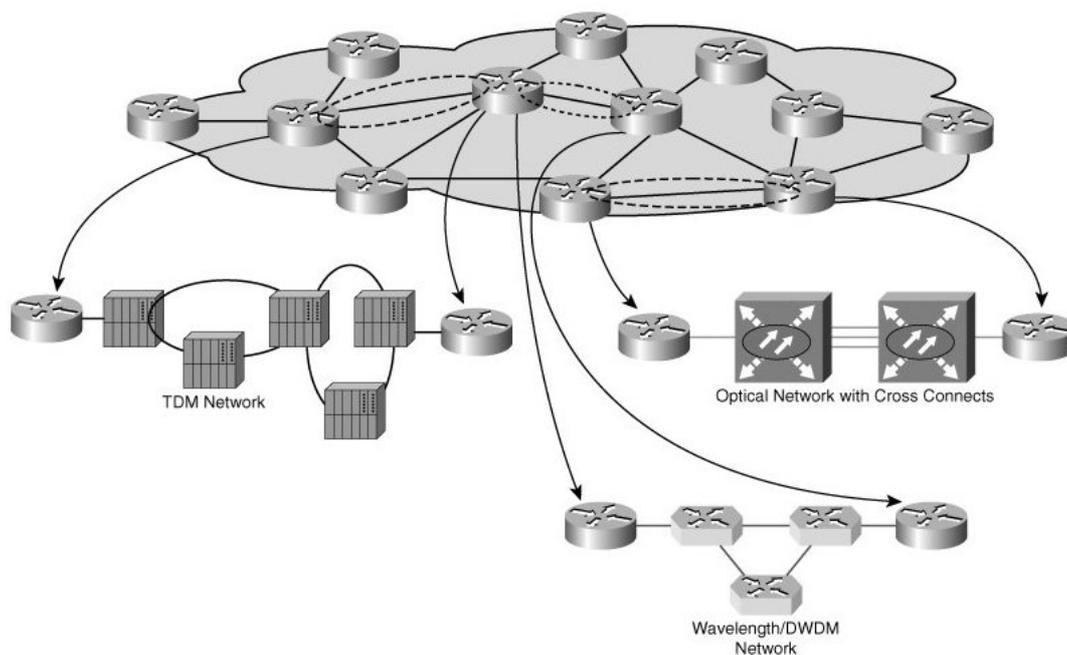


Figura 4: Redes Heterogêneas com GMPLS [43].

O GMPLS também oferece suporte a uma variedade de esquemas de proteção e restauração de enlaces. Caso ocorra uma falha em um enlace, este poderá ser recuperado utilizando um dos seguintes mecanismos: Proteção de LSP (*1+1 LSP Protection*, *1:1 LSP Protection*), Proteção de enlace (*1+1 Link Protection*, *1:1 Link Protection*, *1:N Link Protection*), Restauração de LSP (*End-to-end LSP Restoration with Re-provisioning*), Otimizado (*Enhanced*) e Tráfego Extra (*Extra Traffic*) [42] [44].

O esquema de proteção de LSP 1+1 faz uso de um LSP redundante pelo qual trafega uma cópia do fluxo de dados do LSP principal. Na ocorrência de uma falha o LSP redundante assume o papel do principal. Na abordagem proteção de LSP 1:1 cria-se um caminho dedicado alternativo, simultaneamente ao estabelecimento do LSP principal, por onde os dados serão enviados se o LSP principal estiver inoperante.

Na proteção de enlace 1+1 um enlace auxiliar recebe uma cópia dos dados que trafegam no enlace que está sendo protegido e o nó de saída decide de qual deve receber o tráfego. Na técnica de proteção de enlace 1:1, um enlace alternativo previamente

estabelecido é utilizado em caso de falhas. Na proteção de enlace 1:N existe 1 enlace de reserva compartilhado entre N enlaces ativos.

O esquema otimizado define que outros mecanismos, como anéis de proteção, devem ser utilizados para assegurar um grau de confiabilidade mais elevado do que o esquema 1+1. No mecanismo de Tráfego Extra, um LSP pode ser estabelecido a partir de recursos que estão protegendo outros tráfegos, sendo que na ocorrência de falhas este LSP pode sofrer preempção, isto é, pode dar lugar ao tráfego dos enlaces que falharam. No esquema de restauração de LSP, uma nova rota é calculada somente após a ocorrência da falha sem nenhuma garantia acerca da existência de recursos para o estabelecimento dessa nova rota [45].

Existem discussões a respeito da coexistência da sinalização nativa das redes OBS (JET/JIT) com um plano de controle como o GMPLS. Para redes menores e conexões de curta duração a sinalização nativa OBS é mais adequada. Já o GMPLS é mais atrativo para facilitar a operação de redes de grande extensão (nacionais e continentais), com conexões de longa duração e alta utilização dos recursos, onde a automatização de funções como o estabelecimento, liberação e restauração de caminhos é fundamental para tornar o gerenciamento da rede mais eficiente.

Portanto, conciliar a existência dessas duas abordagens é importante para garantir que os diversos tipos de usuários que fazem uso dos serviços da grade possam ser atendidos satisfatoriamente. Assim, o OGF propõe uma arquitetura sobreposta para combinar essas duas dimensões de controle [9]. A visão geral da arquitetura é apresentada na Figura 5.

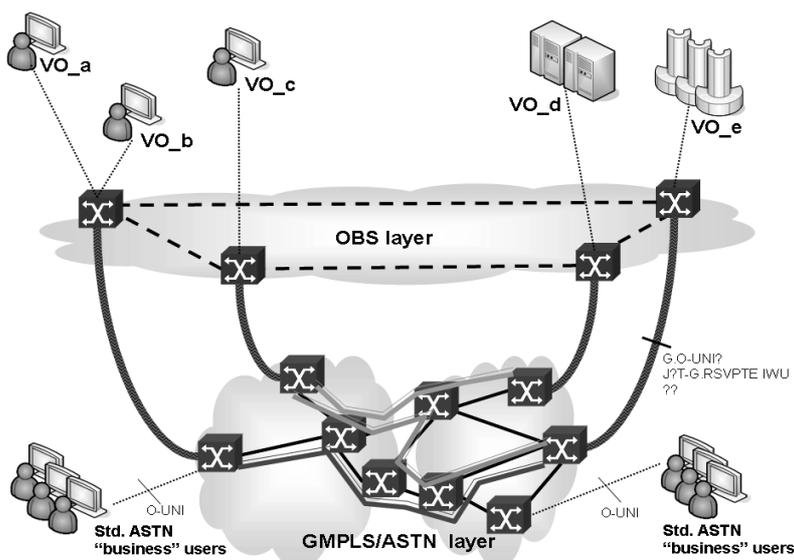


Figura 5: Arquitetura OBS sobre GMPLS [9].

A arquitetura proposta pelo OGF é composta por nós OBS de borda e de núcleo; bem como por nós híbridos, capazes de entender tanto a sinalização OBS e os protocolos da arquitetura GMPLS. Por essa razão, pode-se considerar que essa arquitetura é um modelo de rede híbrida, que pode trabalhar com diferentes tipos de comutação e esquemas de sinalização. Em grades, essa característica é fundamental uma vez que as demandas dos usuários e aplicações podem mudar com bastante frequência.

2.4 – Considerações Finais

Este capítulo apresentou conceitos básicos de computação em grade, desafios de rede que aparecem no contexto de grades e apresentou motivos para que a comutação de rajadas ópticas seja considerada uma excelente alternativa para as novas gerações de grades. A ideia agora é ver a rede como uma entidade de primeira classe, da mesma forma que os processadores, a memória, entre outros recursos, para assim proporcionar serviços de comunicação avançados para os usuários/aplicações da grade.

Novos requisitos surgem para tornar as redes mais adequadas ao cenário de grades, como roteamento interdomínio, agendamento futuro, determinismo, entre outros.

As redes de próxima geração baseadas na comutação de rajadas ópticas possui características que as tornam vantajosas para serem usadas em grades, como a utilização mais eficiente dos recursos, a possibilidade de diferentes perfis de aplicação coexistirem na mesma rede e o oferecimento de funcionalidades avançadas de controle para essas aplicações. Outro desafio que surge nesse contexto diz respeito a tornar a operação da rede mais dinâmica e independente da intervenção humana, o que é fundamental para grades que precisam lidar com uma grande quantidade de requisições simultâneas .

A utilização de planos de controle distribuídos como o GMPLS vem sendo considerada uma opção viável para auxiliar nesse gerenciamento automático da rede.

Capítulo 3

Trabalhos Relacionados

Este capítulo apresenta os principais trabalhos encontrados na literatura relacionados à proposta desta dissertação. As propostas listadas estão associadas a um ou mais aspectos considerados neste trabalho, como a definição de um plano de controle para grades, a utilização de novas tecnologias como a comutação de rajadas ópticas na computação em grade, questões de determinismo (qualidade de serviço garantida), roteamento *anycast*, entre outros.

Os quatro organismos de normalização competentes são a *International Telecommunications Union* (ITU), a *Internet Engineering Task Force* (IETF), *Optical Internetworking Forum* (OIF) e *Open Grid Forum* (OGF) analisaram alguns dos temas relacionados com o domínio de várias redes ópticas. Em 2002, a OIF propôs o *Domínio-to-Domain Routing Protocol* (DDRP). As desvantagens da DDRP é que ela representa uma grande mudança no sistema de roteamento e não é apropriado para a proteção de caminho. O IETF propôs o framework *Generalized Multi-Protocol Label Switching* (GMPLS) que estende os recursos de *Multi-Protocol Label Switching* (MPLS) para a provisão de conexões comutadas por circuito através de abstrações rótulo para comprimentos de onda, timeslots, etc. A *International Telecommunications Union-Telecommunication Standardization Sector* (ITU-T) tenha especificado uma ampla rede de comutação automática óptica quadro (ASON).

Na literatura, há apenas poucos trabalhos que lidam com redes ópticas multidomínio, há ainda menos trabalhos que estudam a possibilidade de integração com a computação em grade.

Em [11], é feita uma discussão acerca das diferentes necessidades que as aplicações de grade possuem no que diz respeito ao controle e ao gerenciamento. Uma constatação é a de que existem aplicações cujos requisitos são atendidos de forma mais adequada por uma entidade de gerenciamento centralizado, outras que se beneficiam melhor de um plano de controle distribuído, sendo que a tendência é considerar uma alternativa híbrida que agregue funcionalidades de ambas as abordagens. São apresentadas algumas iniciativas de protótipos reais como a Ultra Science Network (USN), o DRAGON e a rede CHEETAH. Os autores focam no plano de controle GMPLS e fazem algumas considerações a respeito das funcionalidades que o GMPLS oferece, e de que forma as mesmas contribuem para satisfazer os requisitos da computação em grade. A Tabela 2 ilustra um resumo das principais funcionalidades do GMPLS e sua relação com algumas características de aplicações de grade.

Tabela 2: GMPLS x Requisitos de Aplicações [11].

Funcionalidades do Plano de Controle GMPLS	Tipo de aplicação		
	<i>Banda requerida (pequenas e médias quantidades)</i>		<i>Banda requerida (grande quantidade)</i>
	Escalonamento imediato	Escalonamento futuro	Escalonamento futuro
Roteamento distribuído	Alto	Baixo	Baixo
Estabelecimento de conexões em tempo quase real	Alto	Baixo	Baixo
Auto descoberta	Alto	Alto	Alto
Restauração rápida	Baixo	Alto	Alto

Os valores “alto” e “baixo” indicam o grau de importância que uma determinada funcionalidade tem para cada tipo de aplicação. Aplicações que

requerem estabelecimento imediato de conexões de baixas e médias granularidades se beneficiam do roteamento distribuído e do estabelecimento rápido de conexões, no entanto como essas conexões geralmente têm curta duração, um mecanismo de restauração rápida não é tão fundamental. Por outro lado, em situações onde se deseja alocar uma grande quantidade de banda, normalmente feitas para um intervalo de tempo futuro, uma rápida restauração em caso de falhas é fundamental, já o estabelecimento em tempo quase real e o roteamento distribuído não fazem tanta diferença já que as conexões não precisam ser configuradas imediatamente.

A arquitetura de referência para rotas de computação multicamada utilizando a estrutura do *Framework Path Computation Element* PCE é apresentada em [47], onde o *Virtual Network Topology Manager* (VNTM), que nada mais é do que um Plano de controle virtual é introduzido na camada inferior com um PCE criar caminhos nas camadas de menor capacidade, quando suficientes não está disponível nas camadas mais altas.

Já [48], é proposto uma arquitetura com *framework* de particionamento adaptativo, que pode utilizar de uma maneira efetiva os recursos de rede disponíveis e ao mesmo tempo estabilizar a quantidade de anúncio de rotas disponíveis juntamente com os seus Sistemas Autônomos (AS). Este método de divisão dos recursos de rede é baseado na reserva futura de recursos para tráfego interdomínio, sendo que neste é utilizado políticas de modificação da partilha de recursos com o objetivo de manter a capacidade de transporte disponível entre limites especificados. A figura abaixo descreve a arquitetura proposta.

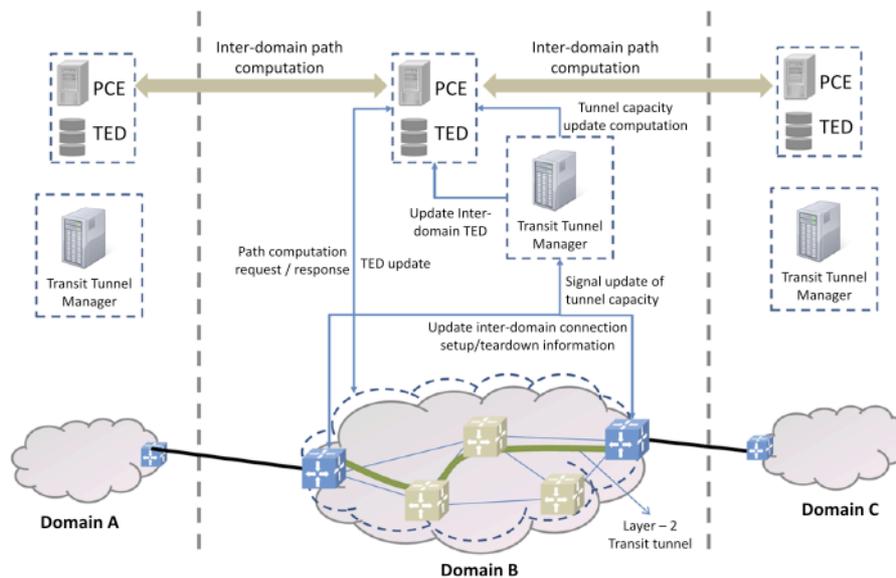


Figura 6: Arquitetura proposta em [48].

Os autores desenvolveram um plano de controle baseado no PCE que é composto dos seguintes componentes principais:

- PCE (*Path Computation Element*): Componente responsável pelo estabelecimento de conexões. Realiza o roteamento interdomínio entre os PCEs dos domínios distintos com base no TED.
- TED (*Traffic Engineering Database*): É o gerador do cálculo dos caminhos intra/interdomínio e a base de dados para a tomada de decisão do PCE.
- TTM (*Transit Tunnel Manager*): É introduzido na arquitetura para governar a inicialização e a operação dos túneis de trânsito dentro de um domínio.

O TED é utilizado pelo PCE para calcular caminhos. Este é dividido em TED intradomínio e TED-interdomínio. O PCE usa o TED-intradomínio para calcular caminhos dentro de um domínio, enquanto PCE de diferentes domínios falam uns com os outros e usam o TED-interdomínio para calcular trajetórias de trânsito entre os nós de borda para pedidos de interdomínio. Os anúncios de rotas disponíveis entre

os PCEs podem ser inicializados nos nós de fronteira como pode ser visto normalmente em caso de BGP ou também pode ser inicializado pelo PCE, que tem uma visão completa da topologia de trânsito do domínio. O TTM é introduzido na arquitetura para governar a inicialização e a operação dos túneis de trânsito dentro de um domínio. Como visto na figura 6, o TTM recolhe informações sobre o estado dos túneis de trânsito dos nós de fronteira correspondente, e é responsável pelo desencadeamento das atualizações de capacidade determinadas pelas políticas correspondentes. Os túneis de trânsito, são estabelecidos na forma de túneis na camada 2 entre os nós de borda com a capacidade reservada e este mesmo nó é responsável pela admissão de conexões interdomínio nestes túneis.

Em [49], foi proposto o OBGp+ (*Optical Border Gateway Protocol Plus*), que é uma extensão de um protocolo de vetor de percurso que calcula e divulga o *Path State Information* (PSI) entre domínios ópticos. O PSI proposto é altamente condensado na forma de um único valor inteiro. A fim de evitar o aumento do número de mensagens de roteamento associado com a atualização da PSI, foi desenvolvido uma carga de atualizações em mensagens de *Keepalive* trocadas entre OBGp+ e vizinhos. Os autores mostraram que, integrando apenas simples e altamente agregados PSI e o OBGp (na forma do OBGp +), é possível melhorar drasticamente a sua performance, sem aumentar o número ou a frequência de atualizações de roteamento trocadas entre os correntes.

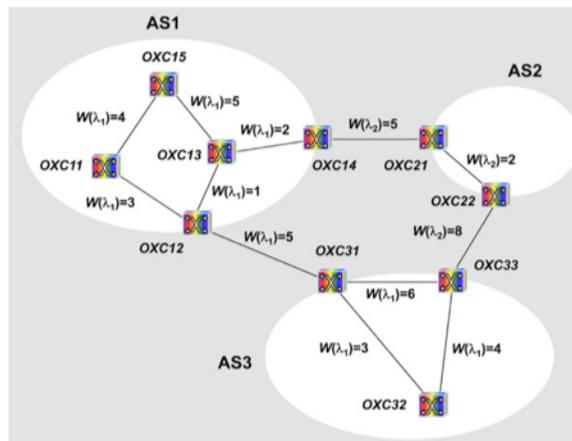


Figura 7: PSI trocado entre domínios ópticos[49].

Em [50], um novo modelo de controle distribuído percurso é proposto, que se baseia na implantação de *Inter-Domain Routing Agents* (IDRAs). Estes autores propõem o protocolo de roteamento executado entre os agentes, como um protocolo de roteamento baseado em IDRAs chamado *Inter-Domain Routing Protocol* (IDRP). O IDRP é capaz de reduzir significativamente a taxa de bloqueio comparado com o de OBGP. No entanto, faz-se necessário o uso de conversores de comprimento de onda físicos para reduzir a probabilidade de bloqueio ocorrer e estes conversores possuem um custo demasiadamente elevado.

Foi proposto em [51] a utilização de conversores de comprimento de onda em um ambiente multidomínio. Para tal foram realizadas modificações importantes nos já existentes protocolos de roteamento interdomínio em redes ópticas, dentre eles o IDRP, OBGP e o OBGP+. Foram identificados dois tipos de canais com comprimento de onda de desocupados em qualquer comutador de borda dado: conversor e não-conversor de canais. Um canal conversor é composto de diferentes tipos de comprimentos de onda em ambos os lados do comutador, exigindo, portanto, a conversão de comprimento de onda se este for utilizado para o estabelecimento de um circuito. Um canal não-conversor, por outro lado, é composto do mesmo comprimento de onda em ambos os lados do comutador e não necessita de conversão

de comprimento de onda [51]. Essas modificações foram integradas nestes protocolos sem alterar os seus aspectos algorítmicos para reduzir significativamente a sua taxa de bloqueio. Um exemplo mostrando na figura 8 onde comutadores de borda de dois domínios ligados por uma conexão interdomínio. Este exemplo mostra o número de comprimentos de onda e conversores de comprimento de onda disponíveis nos comutadores de borda entre os dois domínios.

O revés desta proposta é aplicação em um cenário óptico, pois a mesma só é possível em um cenário com comutação de circuitos ópticos e para esta dissertação foi apresentado que a comutação de rajadas seria uma melhor solução para a comutação em grades.

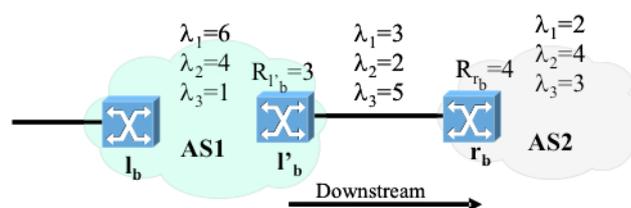


Figura 8: Modelo de conversores de comprimento de onda proposto em [51].

Segundo [2], em um ambiente de grade, as interações com as aplicações e o controle da infraestrutura de rede são feitas por três entidades fundamentais: o middleware da grade, o plano de controle e o plano de gerenciamento, conforme ilustrado na Figura 9. O middleware da grade é responsável, entre outras funções, por prover uma API (*Application Programming Interface*) para que as aplicações possam solicitar os serviços oferecidos pela grade.

O plano de gerenciamento engloba funções de propósito geral como gerenciamento de segurança e desempenho, configuração, que precisam de informações detalhadas dos recursos da rede e são normalmente baseadas no modelo

cliente/servidor. O plano de controle é responsável principalmente pelas funções de roteamento e sinalização que são implementadas de forma descentralizada por questões de escalabilidade e para reduzir o custo de operação e aumentar a eficiência no estabelecimento de conexões dinâmicas, o que é fundamental dentro do contexto de grades. O GMPLS é discutido como uma opção atrativa nesse cenário, embora existam questões em aberto, como roteamento interdomínio, restrições de camada física, coordenação com outros recursos da grade e escalonamento de recursos para um tempo futuro.

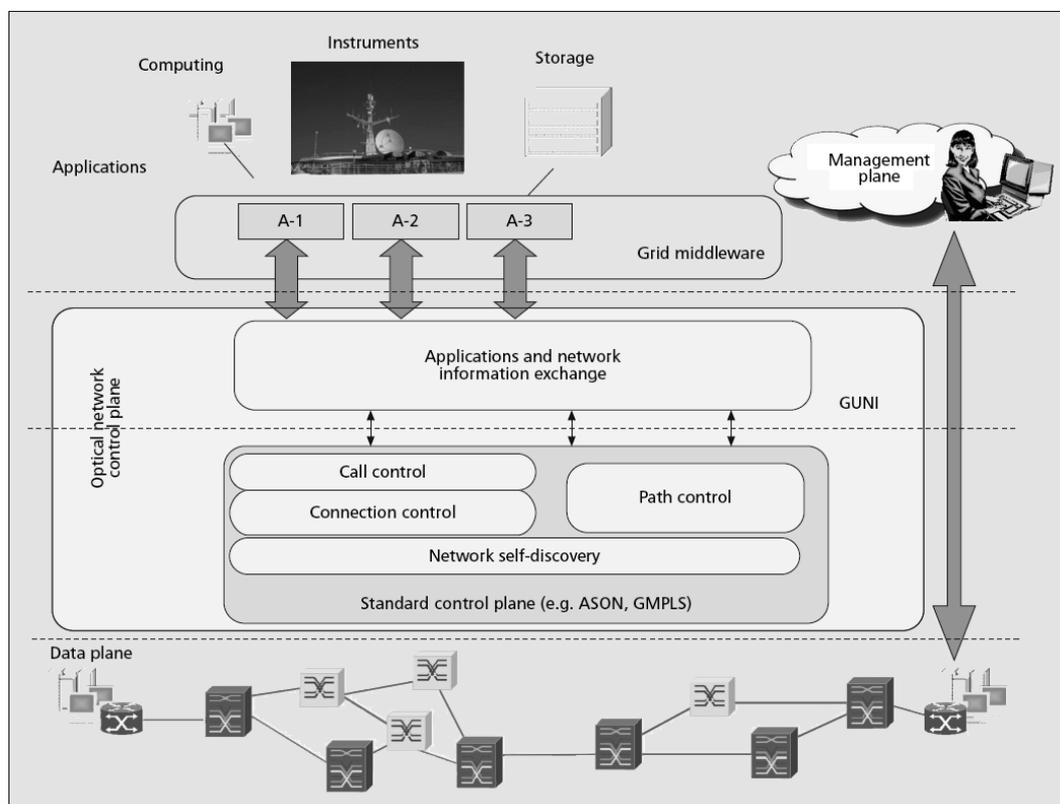


Figura 9: Plano de controle e sua interação com o middleware da grade e o plano de gerenciamento [2].

A iniciativa DRAGON (*Dynamic Resource Allocation via GMPLS Optical Networks*) é apresentada em [52]. Especificamente, o DRAGON procura desenvolver

tecnologias que permitem o provisionamento dinâmico de recursos de rede, que podem estar situados em domínios administrativos distintos e que envolvem tecnologias de rede variadas.

O projeto DRAGON desenvolveu um plano de controle baseado em GMPLS que é composto dos seguintes componentes principais:

- NARB (*Network Aware Resource Broker*): Componente responsável por gerenciar um domínio. Possui as funções de: roteamento interdomínio, abstração de topologia, gerenciamento de LSPs e aplicação de políticas.
- VLSR (*Virtual Label Switched Router*): Permite que os dispositivos não-GMPLS possam fazer parte do estabelecimento de caminhos fim-a-fim.
- ESA (*End System Agent*): Software que fica nos limites do plano de dados, ou seja, na borda da rede.
- ASTB (*Application Specific Topology Builder*): Permite que uma aplicação faça requisições de LSPs específicos.

A Figura 10 ilustra a visão geral da arquitetura DRAGON, com dois domínios.

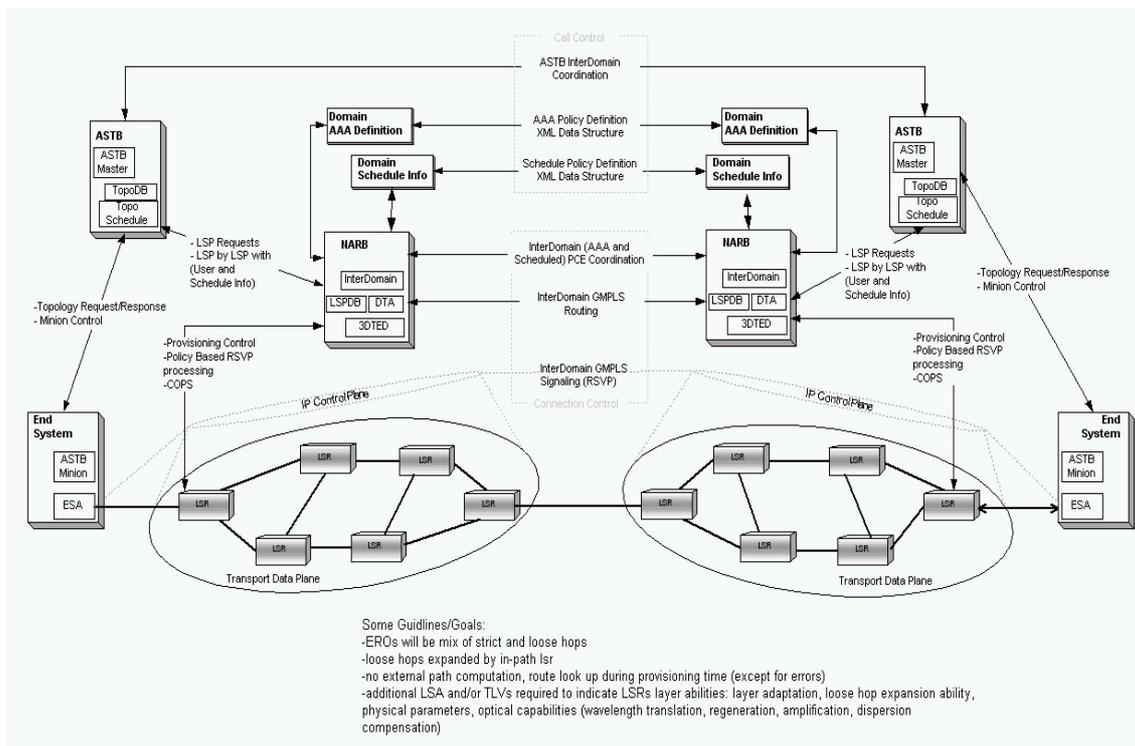


Figura 10: Arquitetura [52].

É possível observar a interação entre os principais elementos que compõem a arquitetura DRAGON. O NARB define três dimensões básicas para os parâmetros associados aos recursos de rede: dimensão AAA (Autenticação, Autorização e Aferição), dimensão de escalonamento e a dimensão GMPLS que são as restrições padrão de engenharia de tráfego do GMPLS. Os sistemas finais, através dos ESAs (*End System Agents*) interagem com o plano de controle, e seus respectivos protocolos de roteamento/sinalização (e.g., OSPF-TE e RSVP-TE) para proporcionar o estabelecimento de caminhos fim-a-fim (entre ESAs) sob demanda. Quando a topologia é complexa, ou quando a aplicação deseja um conjunto específico de LSPs para si, o módulo ASTB é acionado, sendo que o ASTB repassa a requisição para o NARB para que este determine se é possível ou não atender a demanda solicitada.

3.2 Redes OBS em Grades

Em [46] é proposta uma arquitetura para utilizar a comutação OBS em grades computacionais. A arquitetura proposta é baseada na comunicação *anycast*, onde não há a determinação prévia do destino para o qual os dados serão encaminhados e cada tarefa é mapeada para uma rajada óptica com o intuito de simplificar a implementação da arquitetura. A Figura 11 ilustra essa arquitetura.

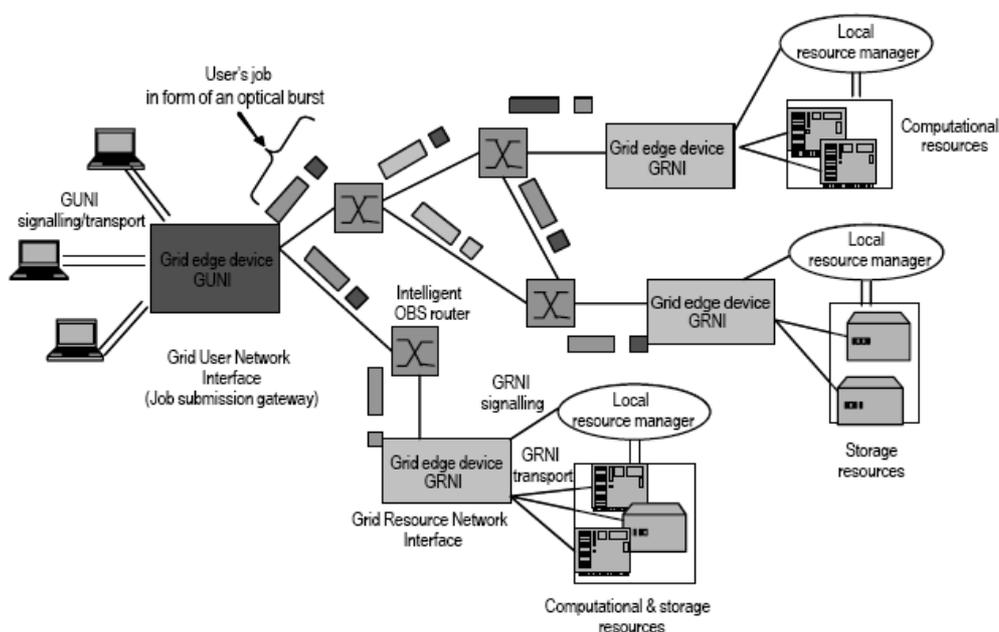


Figura 11: Arquitetura GOBS[46].

Neste modelo, quando uma rajada chega a um nó intermediário, ele determina para onde a mesma deve ser encaminhada com base nas informações fornecidas anteriormente pelo pacote de controle (BCP). A resposta das requisições deve ser enviada para a origem utilizando a comunicação ponto-a-ponto (*unicast*).

Em [33] é proposta uma arquitetura OBS “programável”, onde existem elementos ativos na rede, responsáveis por auxiliar na descoberta de recursos. Nesse caso, a sinalização é realizada em duas vias, isto é, em um primeiro passo existe uma sinalização para descobrir a rota a qual será utilizada para transmitir a rajada e, posteriormente, existe outra sinalização para reservar os recursos ao longo da rota descoberta. A sinalização de descoberta é enviada através de comunicação multicast para grupos de nós de destino, e um nó destino que seja capaz de processar a requisição informa a origem desta capacidade. A origem seleciona, dentre os vários destinos obtidos, aquele que melhor se adequa à requisição. A Figura 12 ilustra esse modelo.

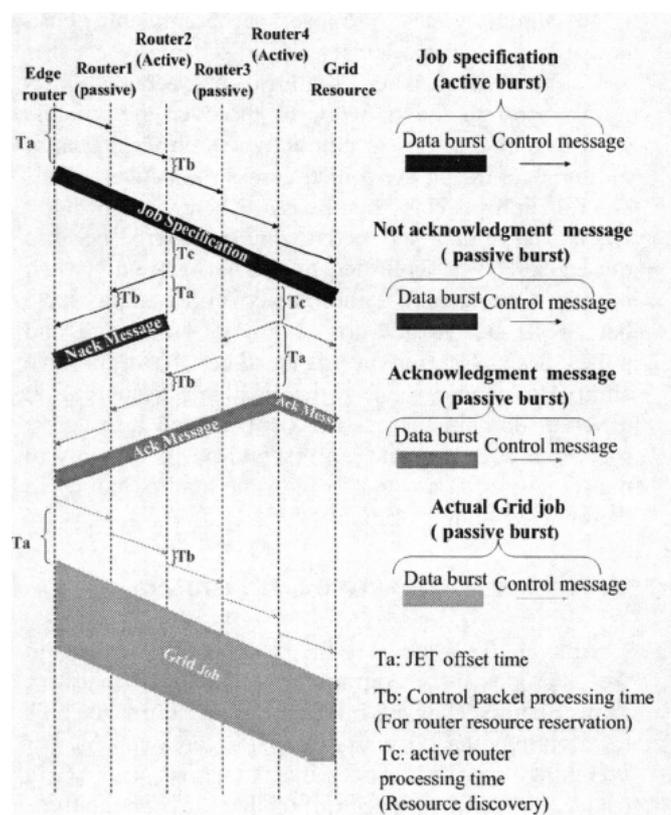


Figura 12 Sinalização no OBS programável [33].

O problema desta abordagem é o retardo excessivo para se obter uma rota, já que a sinalização é feita em duas etapas, a primeira para descobrir a rota e a segunda para sinalizar os recursos ao longo da rota selecionada.

3.3 Qualidade de Serviço em Redes OBS

Em geral, pode-se prover qualidade de serviço em redes OBS pela diferenciação de parâmetros da rajada em função da classe de serviço associada à mesma. Tipicamente essas diferenciações incluem: diferenciação no tempo de ajuste, diferenciação por políticas, diferenciação na construção da rajada e diferenciação no escalonamento da rajada [53]. Existem dois modelos básicos para qualidade de serviço (QoS) em redes OBS: QoS relativa e QoS absoluta [54].

No modelo relativo o desempenho de cada classe não é definido quantitativamente a partir de níveis de QoS determinados, mas sim em função do desempenho da classe em relação às demais, ou seja, tem-se a garantia de que uma classe de maior prioridade terá melhor desempenho que uma classe de menor prioridade, mas não há certeza acerca do cumprimento de certas restrições de QoS em virtude da influência causada pela proporção do tráfego das diferentes classes. Já o modelo absoluto trabalha através do estabelecimento de níveis determinados de QoS que devem ser obedecidos. Por exemplo, pode-se definir um valor máximo de probabilidade de bloqueio para certa classe de rajadas de alta prioridade. No modelo relativo, não há a definição desse valor máximo, define-se apenas que o valor da probabilidade de bloqueio seja menor que o experimentado por rajadas de classes menos prioritárias. Na QoS absoluta o valor da probabilidade de bloqueio não deve ser maior que o definido para a classe.

Em [30] é proposto uma adaptação ao protocolo de reserva de recursos JET. Um tempo de ajuste (*offset*) diferente é associado a cada classe de serviço. A idéia básica é aumentar o valor do offset das rajadas das classes de maior prioridade. Com isso, os nós possuem um maior intervalo de tempo para alocar os recursos necessários e, assim, a probabilidade de bloqueio de uma rajada das classes de maior prioridade tende a ser reduzida. Porém, como as rajadas de maior prioridade, em sua maioria, contêm pacotes de mídias com restrições de tempo, como voz e vídeo, o aumento no valor do offset provoca um aumento não desejável no atraso fim-a-fim.

Os mecanismos de controle de admissão estático e dinâmico de rajadas são propostos por [54]. O princípio dos dois mecanismos é o mesmo, ou seja, reservar um determinado número de comprimento de onda em um enlace para cada classe de serviço. Ambos se baseiam na ocupação do enlace para admitir as rajadas de uma determinada classe de serviço e, desta forma, diferenciar a probabilidade de bloqueio experimentada por cada classe. Os dois mecanismos usam o protocolo JET e a implementação é feita em todos os nós da rede OBS, onde cada rajada ocupa um comprimento de onda durante a sua transmissão e cada nó OBS tem capacidade total de conversão de comprimentos de onda. A diferença fundamental entre os dois mecanismos é que o estático define quantos e quais comprimentos de onda são alocados para cada classe, já o dinâmico define somente a quantidade reservada de comprimentos de onda para cada classe de serviço e não quais serão exatamente os comprimentos de onda associadas a cada uma das classes de serviço.

A utilização da arquitetura de serviços diferenciados (*Diffserv*) para oferecer qualidade de serviço em redes OBS é proposta em [55]. Os pacotes de controle de rajada são processados eletronicamente para prover tratamento diferenciado às rajadas correspondentes através de diferentes PHBs (*Per Hop Behaviors*) referentes

aos serviços suportados: EF (*Expedited Forwarding*), AF (*Assured Forwarding*) e BE (*Best Effort*). A definição desses PHBs tem impacto no processo de montagem de rajadas que varia de acordo com a classe de serviço das mesmas.

Foi proposto por [53] a diferenciação nos serviços experimentados por cada classe, através da variação dos parâmetros de construção das rajadas, onde o temporizador, o limite para o tamanho mínimo e o limite para o tamanho máximo das rajadas, podem ser ajustados para reduzir o atraso e a probabilidade de bloqueio das rajadas de uma dada classe de serviço. O objetivo é determinar como os pacotes serão agregados para garantir os requisitos de QoS de cada um.

Em [56], é apresentada uma proposta de qualidade de serviço dinâmica utilizando controle de admissão em cada nó da rede OBS. Os autores propõem um novo modelo de protocolo de admissão para rajadas. Este modelo utiliza uma arquitetura de rede OBS onde cada nó é composto de um unidade de comutação (SU – *Switching Unit*), unidade de espera (WU – *Wait Unit*), formada por fibras de retardo (FDL – *Fiber Delay Lines*), unidade de controle de comutação (SCU – *Switching Control Unit*) (responsável por fazer reserva de recursos e contenção), unidade processamento de entrada (IPU – *Input Processing Unit*) e unidade de processamento de saída (OPU – *Output Processing Unit*). Quando se deseja enviar uma rajada, o nó de borda OBS envia o BCP com informações de qualidade de serviço e estatísticas de atraso. Os nós intermediários recebem o BCP, verificam os valores de QoS e fazem a estimativas para as métricas de atraso e bloqueio. Caso ocorra bloqueio, a unidade de controle envia a rajada para uma FDL. Se a FDL estiver ocupada a rajada é bloqueada. Se após a rajada sair da FDL o bloqueio ainda estiver ocorrendo, a rajada é bloqueada. Caso a estimativa do atraso seja maior que a definida em métrica de QoS, a rajada também será bloqueada

Em [57], é apresentada uma proposta para utilizar a engenharia de tráfego do GMPLS desviando fluxos de rajadas para rotas alternativas e, assim, abrir possibilidades para o oferecimento de QoS do tipo absoluta. No entanto, a abordagem empregada para guiar as decisões de engenharia de tráfego é baseada em análises prévias do comportamento das classes de serviço na rede. Ou seja, a mudança de rotas não é feita de forma dinâmica, o que pode comprometer a qualidade de serviço em cenários de utilização variável da rede, como uma grade.

O trabalho em [58] procura utilizar a engenharia de tráfego para rerrotear de forma dinâmica os fluxos prejudicados quando o nível de qualidade de serviço começa a ficar abaixo dos limites definidos, ou seja, quando ocorre uma quebra de contexto. É proposto um agente denominado DQMA (*Dynamic QoS Management Agent*) para monitorar os fluxos no núcleo da rede OBS. O DQMA é responsável por coletar e armazenar informações sobre os fluxos de cada classe de serviço. Dessa forma, é possível verificar quando o nível de serviço de um determinado fluxo não está sendo obedecido e, a partir daí, desviar esses fluxos que não estão sendo garantidos para caminhos alternativos que estejam disponíveis na rede. A Figura 13 ilustra uma visão geral da arquitetura DQMA.

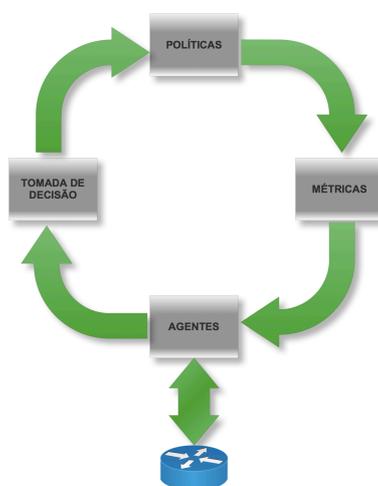


Figura 13 Visão geral da arquitetura do DQMA [57].

Na arquitetura proposta, observam-se os aspectos relacionados às métricas, às políticas, às tomadas de decisão e ao agente. As métricas são as medidas de QoS coletadas pelo agente no nó em que está inserido. As políticas são informações do contexto de QoS que deve ser obedecido para as classes de serviços da rede OBS. O contexto de cada classe contém limiares com os máximos ou mínimos valores de métricas que devem ser garantidos.

A tomada de decisão é baseada na utilização dos recursos da rede e na qualidade de serviço experimentada pelos fluxos pertencentes a diferentes classes de serviço em um determinado momento. Essa tomada de decisões trata das ações que são realizadas mediante possíveis violações de políticas (e.g., quebra de contexto), sendo que essas ações podem ser: o envio de um alarme para a borda da rede ou o roteamento de fluxos através da engenharia de tráfego do GMPLS. O problema dessa alternativa é que as rotas alternativas são fixas e limitadas, ou seja, existe uma quantidade fixa de rotas para uma classe de serviço. Isto pode ser ruim em cenários de alta intensidade de tráfego, já que essas rotas podem não ser suficientes para um grande número de requisições.

O objetivo do trabalho em [59] foi a proposta de uma arquitetura para o estabelecimento de rotas que atenda aos requisitos de qualidade de uma tarefa em um ambiente de grade, onde a arquitetura é baseada no paradigma de comutação de rajadas ópticas e no plano de controle GMPLS com a utilização de um componente responsável por armazenar as informações sobre os recursos do seu AS. Esse componente, denominado Servidor GOBS (*Grid Optical Burst Switching*), recebe consultas sobre a disponibilidade de recursos para processar uma determinada tarefa e informa uma possível rota que atende as exigências de uma requisição. Com a informação obtida a partir dessa resposta, a reserva efetiva dos recursos é feita através

da sinalização GMPLS. A Figura 14 ilustra onde o servidor GOBS está localizado dentro da arquitetura proposta.

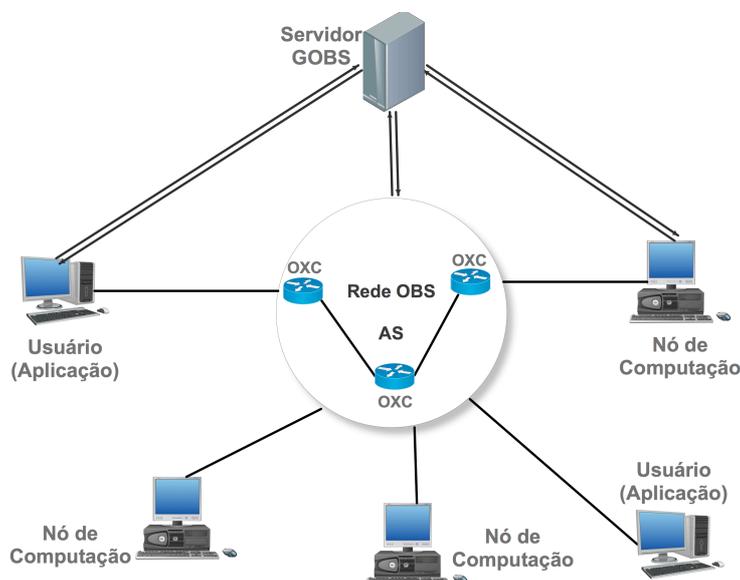


Figura 14: Servidor GOBS[59].

É interessante citar o comportamento do Servidor GOBS quanto a seleção de rotas, pois, como o Servidor GOBS atua apenas em um AS, era comum que ocorresse escassez de recursos e neste caso o mesmo retornava ao usuário uma aproximação quando não conseguia calcular a rota que atendia a todos os requisitos da tarefa ao invés de simplesmente descartar a requisição. Para isso foi definido um parâmetro chamado RDT (*Route Discovery Threshold*) que definia o limite para o qual uma métrica pode ser alterada pra que novos caminhos sejam encontrados nas próximas requisições, todavia, o RDT possuía um limiar baixo, pois não seria possível atrasar em demasia o cálculo de seleção da rota, visto que a rajada já havia sido gerada, então para solucionar este problema o autor definiu um destino padrão para quando o limiar do RDT esgotasse.

Este artigo serviu de base para o desenvolvimento desta dissertação, porém, o problema dessa alternativa é que o roteamento interdomínio é uma questão que não

foi abordada, uma vez que grades normalmente são compostas de recursos que podem estar localizados em domínios administrativos distintos, limitando assim as suas opções de busca de recursos, o que aumenta o número de requisições não garantidas assim como o acréscimo de requisições descartadas pelos nós de computação.

3.4 Roteamento *Anycast*

Um conjunto de algoritmos para a determinação do destino e para resolução de contenção e para deflexão de rajadas é proposto em [31]. A contenção em redes OBS ocorre quando mais de uma requisição competem por um mesmo comprimento de onda da fibra no núcleo da rede. Já a deflexão é uma solução para o problema da contenção que se caracteriza pela seleção imediata de outro comprimento de onda para uma rajada que concorre por um comprimento de onda já ocupado, até que se encontre um recurso disponível. Os algoritmos propostos são classificados em algoritmos de determinação de destino e algoritmos de deflexão de rajadas.

Os algoritmos de determinação de destino são:

- Atribuição de destino relaxada (*soft destination assignment*): Aqui o destino é determinado de forma aleatória e as rajadas serão encaminhadas pela menor rota que leva a esse destino. No entanto, a requisição é processado no primeiro nó ao longo da rota que possui recursos disponíveis.
- Atribuição de destino estrita (*hard destination assignment*): Nesse caso o destino também é selecionado aleatoriamente, porém caso o mesmo não seja capaz de processar a requisição outro destino é escolhido. Essa alternativa nada mais é de que a comunicação ponto-a-ponto tradicional (*unicast*).

Os algoritmos de deflexão de rajadas são:

- Deflexão randômica (*random port deflection*): No caso de uma contenção ou ausência de recursos para processar uma requisição a rajada é encaminhada para uma porta aleatória, caso não seja possível encontrar uma porta disponível, a rajada é descartada.
- Deflexão ponderada (*weighted port deflection*): Aqui a seleção da porta de saída é baseada na verificação de que porta é a menos congestionada de acordo com a probabilidade de bloqueio medida.
- Deflexão baseada na disponibilidade de recursos de grade (*weighted Grid-resource availability deflection*): Nesse caso a escolha da porta de saída é feita com base uma função que avalia a disponibilidade de recursos capazes de processar a requisição.

Dessa forma, é possível realizar várias combinações entre os algoritmos de determinação de destino e de deflexão de rajadas. O problema dessa abordagem é não considerar as métricas de rede na determinação do destino e, além disso, não há a preocupação em oferecer garantias de desempenho para as aplicações.

3.5 Considerações Finais

Esta seção apresentou algumas das propostas da literatura relacionadas aos temas dessa dissertação. A criação de um plano de controle para as redes ópticas em grades vem sendo objeto de discussão na comunidade nos últimos anos. Diversas iniciativas para a construção de planos de controle para as rede ópticas vêm sendo implementadas por grupos de pesquisa ao redor do mundo, sendo que a computação em grade vem sendo considerada fundamental para guiar novas decisões no projeto das redes avançadas de próxima geração.

A ideia dessas propostas de plano de controle é tornar a gerência e a operação da rede mais dinâmicas e transparentes para o usuário/aplicação, reduzindo assim a intervenção humana na reserva e na manutenção dos recursos da rede, o que é importante para ambientes onde recursos são inseridos e removidos com grande frequência, como uma grade.

A comutação de rajadas ópticas apresenta uma série de características que a tornam vantajosa para ser utilizada em grades e diversos esforços de padronização vem sendo conduzidos para a definição de uma arquitetura de rede Grid-OBS. No entanto, ainda não há uma massificação no seu uso em iniciativas práticas devido principalmente aos custos associados às tecnologias envolvidas.

É possível observar que na maioria das propostas apresentadas não há uma preocupação explícita com o problema da qualidade de serviço e, conseqüentemente, com o oferecimento de garantias de desempenho para as aplicações de grade. Os trabalhos que tratam de qualidade de serviço em sua grande parte não levam em consideração a existência de recursos disponíveis na rede, que poderiam ser utilizados em caso de cenários de alta intensidade de tráfego, onde as

técnicas propostas podem não se mostrar capazes de garantir o cumprimento de restrições absolutas de desempenho, sendo que grades se caracterizam por serem constituídas por uma grande quantidade de recursos distribuídos.

O uso de engenharia de tráfego é uma alternativa viável para grades, uma vez que esse modelo muitas vezes envolve redes de grande porte, com uma grande quantidade de nós e enlaces, muitas vezes pertencentes a redes nacionais e continentais já estabelecidas. A engenharia de tráfego permite que se utilize recursos disponíveis para que a rede tenha uma melhor utilização como um todo. Uma aplicação da engenharia de tráfego é na implementação de qualidade de serviço, já que nem sempre os esquemas tradicionais podem garantir o desempenho desejado.

Capítulo 4

Arquitetura para o Oferecimento de Conexões Determinísticas Multidomínio em Redes GOBS Baseadas no GMPLS

Este capítulo apresenta a arquitetura proposta para o provisionamento automático de conexões com garantias de desempenho em redes baseadas na comutação de rajadas ópticas que utilizam um plano de controle GMPLS em um ambiente de grade computacional. A Seção 4.1 discute aspectos relacionados ao oferecimento de qualidade de serviço em redes OBS baseadas no GMPLS. A Seção 4.2 trata de conexões determinísticas e seleção de recursos em redes que compõem ambientes de grade. Na Seção 4.3 é descrito o Servidor GOBS, componente responsável pela divulgação dos seus recursos intradomínio para o GOBS Raiz. O componente principal da arquitetura proposta, o Servidor GOBS Raiz, é descrito na Seção 4.4. Na Seção 4.5 o algoritmo para a seleção de rotas determinísticas é apresentado. A Seção 4.6 descreve a arquitetura de monitoração utilizada para atualizar as estatísticas de utilização de recursos. A Seção 4.7 descreve o funcionamento da proposta no que se refere ao processo de estabelecimento de conexões determinísticas. As modificações sugeridas na sinalização GMPLS para aproveitar as características da comutação OBS são descritas na Seção 4.8. Por fim, a Seção 4.9 conclui o capítulo.

4.1 – Qualidade de Serviço em Redes OBS

Baseadas no GMPLS

Oferecer qualidade de serviço em redes ópticas apresenta algumas dificuldades decorrentes das limitações existentes nas tecnologias de transmissão ópticas, principalmente no que se refere à ausência de estruturas de armazenamento temporárias que sejam viáveis em termos de custo e complexidade de implantação. A maioria dos esquemas de diferenciação de serviços propostos para as redes eletrônicas é baseada no esquema de armazenamento e encaminhamento (*store and forward*) de pacotes e no gerenciamento de buffers eletrônicos como as filas dos roteadores IP. Já a comutação óptica trabalha com a abordagem de encaminhamento direto (*cut-through*) que não necessita de armazenamento temporário nos equipamentos já que o uso de estruturas de armazenamento eletrônico implica na conversão do sinal óptico para eletrônico em cada dispositivo de comutação o que resulta em reduções nas taxas de transmissão, e, por essa razão, mecanismos de QoS adaptados para redes puramente ópticas se fazem necessários.

Considerando que a arquitetura GMPLS é uma excelente opção para integrar o protocolo IP com a tecnologia WDM e o paradigma de comutação de rajadas ópticas apresenta vantagens dentro do cenário tecnológico atual, a integração das duas tecnologias que resultou na arquitetura denominada LOBS (*Labeled Optical Burst Switching*) [60] proporciona uma diminuição na complexidade de gerenciamento nas redes ópticas e abre possibilidades para um controle melhorado dos recursos que compõem a rede.

Dentro dessas novas possibilidades, a diferenciação de serviços é um aspecto que pode ser agregado às facilidades oferecidas pelo plano de controle GMPLS, principalmente quando se fala em engenharia de tráfego. Não se pode esperar que os mecanismos tradicionais para oferecer QoS em redes OBS sejam suficientes para atender os mais variados cenários de tráfego que podem surgir na rede, além disso, os mecanismos atuais nem sempre são capazes de garantir um nível de serviço absoluto onde as métricas de QoS deverão obedecer a valores estritos.

A proposta desta dissertação resulta do fato de que o plano de controle GMPLS oferece funcionalidades interessantes que podem auxiliar os mecanismos de diferenciação de serviços já existentes para redes OBS, aumentando a possibilidade de se atingir níveis de QoS mais precisos, o que pode ser de importância fundamental para as aplicações de grade.

O roteamento explícito oferecido pelo GMPLS possibilita a definição de rotas alternativas às determinadas por técnicas baseadas na escolha do melhor caminho, que são comumente usadas na Internet, já que muitas vezes essas rotas não são suficientes para atender uma alta demanda de tráfego sem degradar o desempenho das classes de serviço como um todo. Isso geralmente ocorre em cenários de intensidade de tráfego elevada onde nem sempre a diferenciação na admissão das rajadas será suficiente para garantir níveis de serviço adequados, principalmente quando se deseja QoS absoluta, isto é, com garantias de desempenho.

Desta forma, para se atingir um nível de desempenho absoluto em cenários de alta intensidade de tráfego, propõe-se a utilização do roteamento explícito para desviar fluxos de rajadas que pertençam a classes de alta prioridade com requisitos estritos de qualidade. Dessa forma, espera-se que o aumento da quantidade de recursos disponíveis para uma classe resulte em um melhor serviço para as suas

rajadas. Dependendo do nível de serviço desejado, é possível determinar quantos e quais fluxos de rajadas deverão adotar uma nova rota e, assim, obter o melhor desempenho para as classes prioritárias.

4.2 – Conexões Determinísticas e Seleção de Recursos

Um aspecto importante no contexto da computação em grade é a necessidade de se proporcionar às aplicações níveis de serviço que estejam de acordo com as exigências das mesmas. Portanto, oferecer Qualidade de Serviço (QoS) às aplicações e aos usuários da grade é fundamental para que os serviços sejam cada vez mais atrativos para os diferentes perfis de usuários, contribuindo assim para a disseminação do paradigma de computação em grade e, conseqüentemente, para a redução dos custos de implementação e manutenção desses serviços. Por essa razão, as redes ópticas de próxima geração são a melhor alternativa para esse cenário em virtude de sua capacidade de transmissão de grandes volumes de dados em altas velocidades.

A qualidade de serviço em grades computacionais possui vários níveis. Um nível importante está relacionado à alocação de recursos de computação para executar uma determinada tarefa. É importante garantir que as tarefas de maior prioridade tenham uma disponibilidade maior de recursos de computação reservados para si [61]. Outro nível diz respeito à capacidade que a infraestrutura de comunicação que suporta a grade deve ter para que as aplicações funcionem adequadamente, principalmente aquelas que possuem requisitos temporais estritos e exigem tempo de resposta rápido.

Por essa razão, as redes ópticas de próxima geração são as candidatas ideais para esse cenário em virtude de sua capacidade de transmissão de grandes volumes de dados em altas velocidades. Mesmo assim, é desejável o cumprimento de certas garantias de desempenho que são exigidas pelas aplicações de grade. Logo, para isso, é necessário que se possa prover conexões ópticas capazes de oferecer um determinado nível de serviço para os fluxos de tráfego. Por exemplo, uma determinada tarefa pode requerer um caminho cuja probabilidade de perda não seja maior que um certo limiar, definido com base nas características da aplicação.

Uma grade computacional é geralmente constituída por uma grande quantidade de recursos, inclusive recursos de rede. Do ponto de vista do usuário, nem sempre importa quem vai executar a tarefa e nem por onde os dados serão encaminhados, o que importa é que a tarefa seja executada de acordo com as restrições associadas a ela. Dessa forma, qualquer recurso em princípio pode ser alocado a uma tarefa.

Por essa razão, uma questão importante nas grades computacionais é a definição de que recursos devem ser reservados para viabilizar a execução de uma determinada tarefa. Assim, a descoberta e seleção de recursos é um tópico que desperta bastante interesse na área [62]. Com relação às conexões determinísticas é possível notar que a descoberta e a seleção de recursos podem ser estendidas aos componentes de rede, e devem levar em consideração as características da aplicação. Por exemplo, através da monitoração dos níveis de utilização dos enlaces é possível decidir por uma rota menos congestionada para atender uma determinada requisição de uma aplicação e, assim, garantir que a tarefa não seja prejudicada por atrasos ou perdas excessivas.

A engenharia de tráfego proporcionada pela arquitetura MPLS/GMPLS é uma opção atrativa para auxiliar na alocação dessas rotas determinísticas uma vez que conta com protocolos de roteamento e sinalização que permitem a definição de LSPs (*Label Switched Paths*) de forma explícita ou baseada em certas restrições [63].

4.3- Servidor GOBS

O Servidor GOBS inicialmente foi proposto em [59] para viabilizar a escolha de uma rota que atenda os requisitos de uma aplicação de grade em seu AS. Para tal este servidor fica responsável por armazenar todas as informações referentes aos recursos que compõem a grade local. Essas informações são estruturadas em seu conteúdo e este será responsável por determinar qual a melhor rota que atende os requisitos de uma tarefa. Em seguida, os protocolos de roteamento e sinalização do GMPLS irão efetuar a reserva efetiva dos recursos para a tarefa. A estrutura sugerida para o servidor GOBS é apresentada a seguir e é exemplificada na Tabela 3:

- *Tipo*: Especifica se é um nó de grade ou de rede.
- *Processamento*: Representa a capacidade de processamento que um determinado nó possui medida em GFLOPS (10^9 operações de ponto flutuante por segundo), aplicável aos nós de grade.
- *Armazenamento*: Representa a capacidade de armazenamento que um determinado nó de grade possui em memória principal, medida em bytes. Aplicável aos nós de grade.
- *Bloqueio*: Probabilidade de bloqueio medida de forma “*on-line*” em função da quantidade de requisições bloqueadas em relação ao total de requisições. Aplicável aos nós de rede.

Tabela 3: Exemplo de entrada no servidor GOBS.

Nó	Tipo	Processamento	Armazenamento	Bloqueio
0	Grade	3GFLOPS	512 GB	-
1	Rede	-	-	0.002
2	Rede	-	-	0.005
3	Grade	1 GFLOP	1024GB	-
4	Rede	0	-	0.001

Vale ressaltar que deverá existir um servidor GOBS responsável por um AS particular da rede. Um AS neste contexto deve ser entendido como um subconjunto de uma topologia maior. Em redes com múltiplos AS deverá existir um Servidor GOBS responsável por cada AS.

4.4 – Servidor GOBS Raiz

Quando uma requisição de grade é solicitada e esse recurso não estiver disponível em seu AS, o servidor GOBS enviava através de um circuito intradomínio para um destino padrão estático onde os dados eram descartados. Este procedimento causa sérios problemas para tais requisições de grade, tal qual atraso, o aumento da probabilidade de bloqueio, acréscimo de requisições não garantidas, timeout do processo, entre outros. Para resolver estes problemas o servidor GOBS Raiz foi inserido nesta arquitetura.

Este componente é responsável a fazer o monitoramento de todos os servidores GOBS comuns que a ele forem associados. A comunicação entre estes dois agentes é pré-estabelecida sendo que toda vez que um novo AS for incluído na topologia, o seu servidor GOBS precisará ser registrado no servidor GOBS Raiz com as informações

específicas relacionadas a sua localização na rede, tal qual os seus nós de borda para que o servidor GOBS Raiz saiba quem são seus AS vizinhos, quais rotas ele poderá encaminhar como resposta ao seu solicitante e principalmente a capacidade total de processamento e armazenamento livres, assim como o nível de bloqueio atualizado. A estrutura sugerida para entrada de informações no servidor GOBS Raiz é apresentada a é exemplificada na Tabela 4:

Tabela 4: Exemplo de entrada no servidor GOBS Raiz.

AS	Nó de borda	Tipo de recurso	Processamento	Armazenamento	Bloqueio
1	3, 5.	Grade	15 GFLOPS	512 TB	-
		Rede	-	-	0.002
2	13, 16	Grade	10 GFLOPS	128 TB	-
		Rede	-	-	0.03

4.5 – Seleção de Rotas

A seleção de rotas deve levar em consideração parâmetros da rede e da grade. Neste caso, utiliza-se a probabilidade de bloqueio experimentada por fluxos de uma determinada classe de serviço em um enlace, o nível de utilização de um enlace, bem como a disponibilidade de processamento e armazenamento de um nó da grade. Essa seleção de recursos de rede tem por objetivo auxiliar nas decisões de engenharia de tráfego e minimizar o bloqueio de conexões em um ambiente dinâmico e com grande disponibilidade de recursos como a grade.

Como não há a determinação de qual destino a priori será responsável por processar a tarefa, pode se considerar que a abordagem de reserva de recursos adotada segue o modelo *anycast* [31]. O destino será fixado a partir do momento em

que a rota for calculada como resposta a consulta feita ao servidor GOBS Raiz. Então, a sinalização GMPLS, através do roteamento explícito, fará a reserva dos recursos.

Um algoritmo de busca para viabilizar a seleção da rota para uma determinada tarefa é proposto e apresentado. O objetivo desse algoritmo é retornar uma resposta ao nó de OBS de onde partiu a requisição. Essa resposta deverá conter a informação de uma rota explícita que atende os requisitos da tarefa que servirá como entrada para que o nó de borda execute a sinalização OBS/GMPLS e efetue a reserva dos recursos. O funcionamento do algoritmo de seleção de rotas é ilustrado na Figura 15.

Seleção de rotas	
1:	Determinar origem, marcar origem como verificado.
2:	A partir da origem, verificar quais próximos recursos alcançáveis. Incluir a origem no vetor de rota explícita. Analisar o primeiro recurso.
3:	Verificar o tipo do nó.
4:	Se for nó de grade e não está marcado como verificado:
5:	Verificar se existe capacidade de processamento/armazenamento disponível no período desejado, marcar o nó como verificado. Em caso afirmativo, incluir esse nó no vetor de rota explícita. Devolver o vetor de rota explícita.
6:	Se for nó de rede e não está marcado como verificado:
7:	Verificar se os níveis de serviço estão dentro das especificações definidas e marcar o nó como verificado. Em caso afirmativo, incluir o nó no vetor de rota explícita.
8:	Continuar com a busca em largura (repetir a partir do passo 3) até que se chegue a um nó de borda que executa a tarefa (passo 5).
9:	Se nenhum caminho com as especificações definidas for encontrado, solicitar recurso ao GOBS Raiz.
10:	Verificar se existe capacidade de processamento/armazenamento disponível no período desejado, em algum dos AS correspondentes aos GOBS comuns associados. Em caso afirmativo, incluir esse nó no vetor de rota explícita. Devolver o vetor de rota explícita.
11:	Verificar se os níveis de serviço estão dentro das especificações definidas e marcar o caminho como verificado. Em caso afirmativo, incluir o nó no vetor de rota explícita.
12:	Repassar ao GOBS comum do AS com o recurso disponível a reserva dos níveis de serviço para a requisição solicitante.
13:	Repassar ao GOBS comum o vetor de rota a seguir para a origem da requisição

Figura 15: Algoritmo de seleção de rotas.

Onde *degree* representa o fator de mudança da métrica que afetará as próximas buscas por rotas e *metric* é o valor da métrica em questão. O fluxograma da Figura 16 apresenta um resumo do funcionamento do algoritmo proposto.

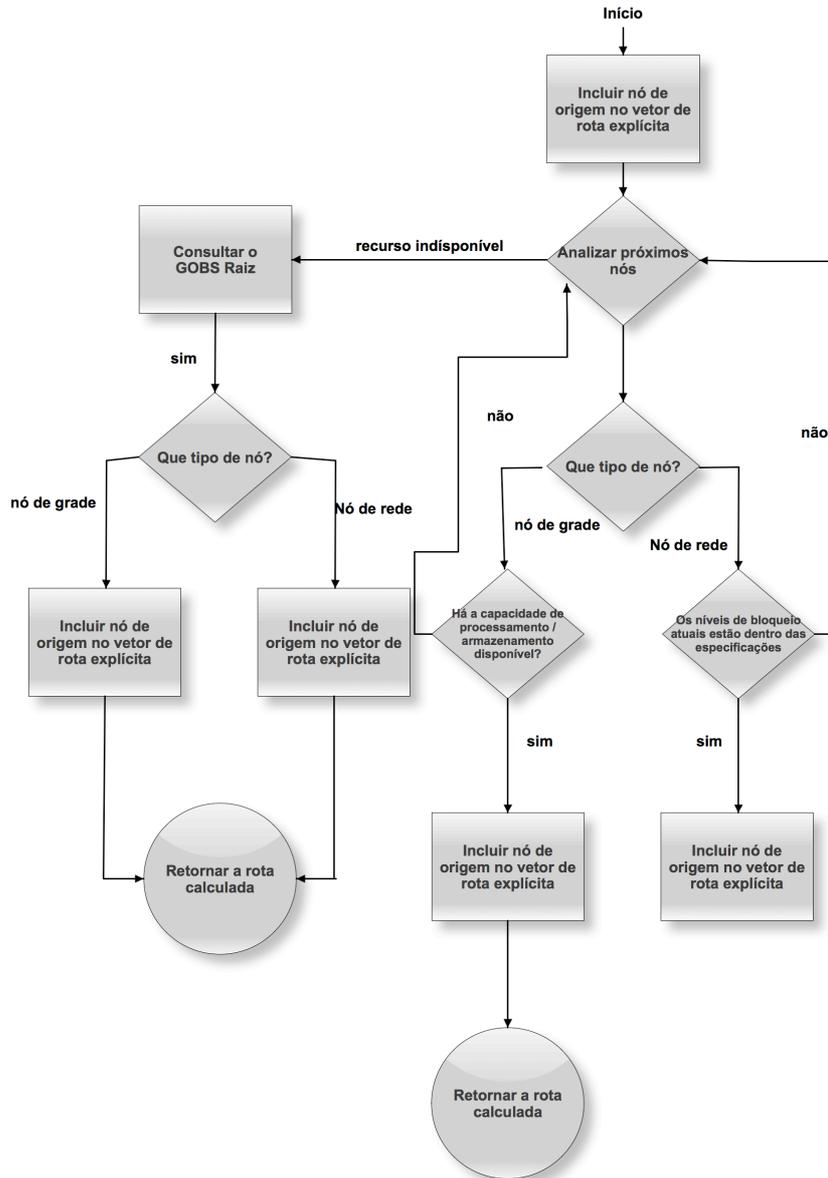


Figura 16: Fluxograma de seleção de rotas.

4.6 – Monitoramento de Recursos

Para possibilitar a monitoração dos níveis de qualidade de serviço experimentados por classes de rajadas ópticas em enlaces específicos da rede utiliza-se um agente denominado DQMA (*Dynamic QoS Management Agent*) proposto em [58]. De uma forma geral, o DQMA coleta estatísticas em tempo real dos fluxos de cada classe de rajada e as compara com uma tabela que contém os níveis de serviço definidos previamente (contexto). Existem dois tipos de agente: o de núcleo, responsável por coletar as estatísticas e enviar alarmes para a borda na ocorrência de quebras de contexto e o de borda que toma uma decisão de engenharia de tráfego no recebimento de uma mensagem de alarme.

O DQMA será o componente responsável por atualizar as estatísticas dos recursos no Servidor GOBS. Para isso, o DQMA deverá ser estendido para levar em consideração também os recursos de computação. Neste trabalho, a funcionalidade de alarme do agente DQMA não é considerada, uma vez que uma decisão de engenharia de tráfego é feita para cada tarefa com base nas informações presentes no Servidor GOBS.

4.7 – Funcionamento da Proposta

Em linhas gerais, a proposta funciona da seguinte forma: no momento da submissão de uma tarefa (*job*) a mesma é encapsulada em uma rajada óptica e as informações referentes aos parâmetros de QoS, bem como a demanda de processamento e armazenamento da tarefa, são codificadas no pacote de controle (BCP). Antes de a rajada ser enviada, uma consulta é feita ao Servidor GOBS Raiz com o objetivo de se obter uma rota que satisfaça os requisitos da tarefa. A rota obtida é então repassada para o GOBS comum solicitante através do protocolo de sinalização. O protocolo RSVP-TE permite o roteamento explícito através do uso de Objetos de Roteamento Explícito

(*ERO – Explicit Route Objects*) [63] que são codificados dentro das mensagens PATH do protocolo. Após o tempo de ajuste, a rajada é enviada para o destino escolhido, com base na rota calculada pelo Servidor GOBS Raiz.

A Figura 17 ilustra a ideia geral da arquitetura de seleção e estabelecimento de conexões determinísticas proposta para esta dissertação.

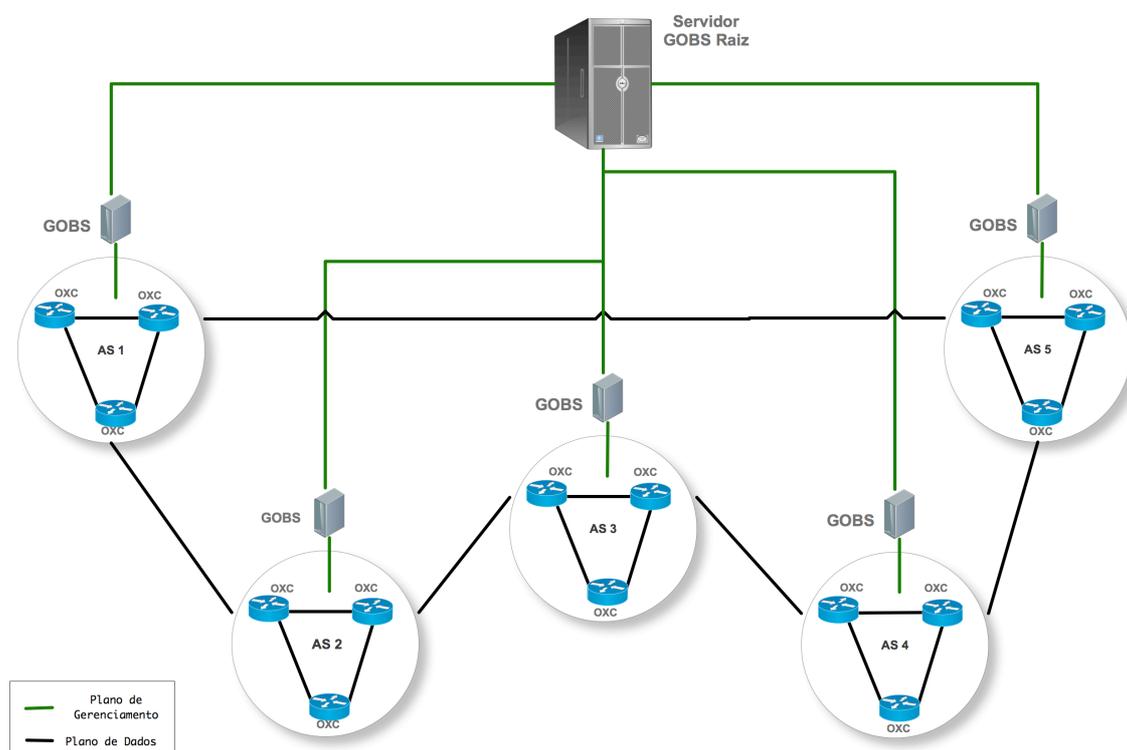


Figura 17: Arquitetura para o provisionamento de conexões determinísticas.

Os nós de rede OBS são responsáveis por solicitar as rotas para o Servidor GOBS de acordo com os requisitos definidos na tarefa, os nós de núcleo encaminham as rajadas com base no caminho selecionado. O Agente DQMA coleta estatísticas nos nós OBS e as repassa para o Servidor GOBS.

A Figura 18 ilustra a integração entre os diferentes planos (gerenciamento, controle, dados) e serviços que compõem a arquitetura proposta para o provisionamento de conexões determinísticas em redes Grid OBS.

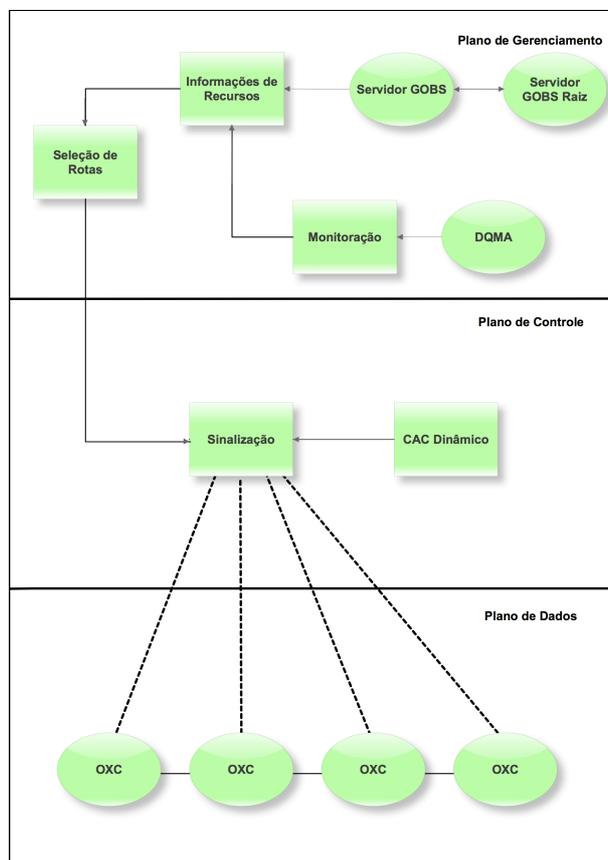


Figura 18: Integração entre os planos na arquitetura proposta.

O plano de gerenciamento é composto dos serviços de Seleção de Rotas, Informações de Recursos e Monitoração. A Seleção de Rotas interdomínio utiliza os dados oferecidos pelo serviço de Informações de Recursos, que é representado pelo Servidor GOBS Raiz. As informações sobre os recursos são atualizadas pelo serviço de Monitoração, que é realizado pelo agente DQMA.

No plano de controle destaca-se o serviço de Sinalização que é responsável por reservar os recursos ao longo da rota determinística fornecida pelo serviço Seleção de Rotas. Na sinalização das rotas se faz necessário o uso do serviço CAC Dinâmico que é

o mecanismo de controle de admissão que fará a diferenciação das classes de serviço em função do número de comprimentos de onda disponíveis para cada uma. O plano de dados consiste dos equipamentos de comutação propriamente ditos (OXC's).

4.8 – Modificações na Sinalização GMPLS

Uma questão importante diz respeito à necessidade de se realizar modificações nos mecanismos de sinalização do GMPLS para tornar o processo de reserva de recursos mais adequado ao OBS reduzindo, assim, a latência de sinalização. Esta preocupação surge do fato de que o controle da distribuição de rótulos normalmente empregado pelos protocolos de sinalização no contexto MPLS é o controle ordenado[63]. No controle ordenado, a mensagem de sinalização (*e.g.*, *RSVP Path Messages*) passa por todos os nós ao longo da rota até chegar ao destino. Após isso é que os rótulos são distribuídos (configurados) no sentido inverso garantindo a criação de um LSP consistente.

É interessante destacar nesta proposta a alternativa implementada onde o GMPLS oferece para minimizar os altos tempos de reserva é a utilização de rótulos sugeridos (*suggested labels*). Com um rótulo sugerido é possível agilizar o processo de reserva uma vez que o nó pode configurar um rótulo (*e.g.*, comprimento de onda, no contexto óptico), no momento em que recebe a sinalização. Dessa forma, não seria necessário esperar todos os nós serem configurados para enviar uma rajada. Isso resulta em tempos de sinalização menores comparados aos mecanismos de sinalização tradicionais e aproveita melhor a característica de ausência de confirmação inerente ao OBS.

Por essa razão, a arquitetura proposta nesta dissertação implementa um modelo de sinalização que procura seguir a ideia do rótulo sugerido. Esta sinalização funciona

da seguinte forma: quando o nó recebe a mensagem de sinalização ele tenta imediatamente alocar os recursos necessários (*lambdas*) para a rajada. Se existir um comprimento de onda disponível para a classe da rajada, este é reservado sem nenhum tipo de confirmação. Caso contrário, a sinalização é repassada para os demais nós, mas não é processada, isto é, não há novas tentativas de alocação de recursos.

Depois que a rajada passa por um nó onde há um comprimento de onda reservado para a mesma, os recursos são liberados para requisições futuras. Se nenhum comprimento de onda for encontrado para a rajada, a mesma é descartada. Além disso, se faz necessário que a sinalização (BCP) inclua o tempo de ajuste, o tamanho da rajada, a classe de serviço e as respectivas demandas de processamento, armazenamento e de qualidade de serviço da tarefa.

4.9 – Considerações Finais

Este capítulo apresentou a arquitetura proposta para o provisionamento automático de conexões determinísticas multidomínio em redes Grid OBS baseadas no GMPLS. O elemento principal dessa arquitetura é o componente denominado Servidor GOBS Raiz que é responsável por armazenar as informações sobre recursos de rede e de computação disponíveis em todos os AS com GOBS comuns associados a ele e por calcular caminhos que sejam capazes de oferecer um nível de desempenho absoluto para as rajadas em um ambiente de grade.

Um algoritmo de seleção de rotas que objetiva selecionar caminhos que satisfaçam os requisitos da aplicação foi proposto. O roteamento explícito é parte fundamental da proposta, pois é através dessa funcionalidade que os recursos ao longo da rota serão alocados de fato. O RSVP-TE fornece essa característica através do uso de Objetos de Roteamento Explícito.

Além disso se fez necessário a utilização e adaptação da arquitetura de monitoração proposta em [60] com o objetivo de se evitar inconsistências nas informações armazenadas no Servidor GOBS.

Capítulo 5

Simulação e Análise dos Resultados

Este capítulo tem o objetivo de avaliar o funcionamento da arquitetura proposta para o provisionamento automático de conexões determinísticas multidomínio em redes Grid-OBS. A metodologia utilizada para a análise é a simulação computacional. A Seção 5.1 apresenta o ambiente de simulação empregado nesta avaliação. A Seção 5.2 descreve o cenário proposto para as simulações. Os resultados obtidos são discutidos na Seção 5.3. A Seção 5.4 conclui o capítulo.

5.1 – Adaptações Necessárias no Simulador

Para viabilizar pesquisas na área de redes existem alternativas como modelagem analítica, implementação de protótipo e simulação. Em virtude dos custos associados à implantação de um protótipo real para a computação em grade é interessante, em um primeiro momento, a realização de estudos através de ambientes computacionais que permitam a simulação de diferentes cenários bem como o teste de soluções propostas dentro do contexto apresentado.

O uso de ferramentas de simulação é fundamental para avaliar novas propostas antes da implantação de um ambiente real de produção para que se tenha um melhor planejamento na definição dos requisitos que um protótipo real deve possuir bem como corrigir possíveis erros e limitações do modelo.

A ferramenta de simulação utilizada para viabilizar a análise da proposta é o ns-2 [65]. Essa opção se deu pelo fato de que o ns-2 é uma ferramenta bastante difundida na comunidade, permite a adição de novos módulos e é livre. O ns-2 trabalha com a linguagem de scripts OTcl para a construção das simulações sendo que o seu núcleo é desenvolvido em C++. O ns-2 permite a simulação de diferentes protocolos e tecnologias de rede como as várias versões do TCP, diversos protocolos de roteamento, redes locais, redes sem fio, etc. Existe também uma comunidade ativa de usuários que contribui no desenvolvimento de novos módulos para suprir as demandas que surgem. Por exemplo, no contexto de redes ópticas tem-se o OWNS [66] desenvolvido para proporcionar a simulação de redes WDM.

No entanto nem sempre essas extensões são compatíveis entre si, por questões que variam desde a versão do simulador na qual as mesmas foram desenvolvidas, as particularidades existentes no projeto de cada uma delas e a impossibilidade de integrá-las facilmente em um cenário homogêneo. Por essa razão, e pela inexistência e dificuldade de acesso a outros módulos que pudessem ser utilizados na análise, foram desenvolvidas extensões para caracterizar a proposta, sendo que as principais são:

- Um agente para representar a funcionalidade de um nó de borda OBS que executa as funções de montagem de rajadas, mapeamento de tarefas de grade (*jobs*) para rajadas e sinalização (tempo de ajuste).
- Modificações no nó MPLS com a adição de estruturas para representar os comprimentos de onda da fibra e a implementação do controle de admissão dinâmico de rajadas.
- Um agente DQMA para coletar as estatísticas de bloqueio em um determinado nó da rede.

- Um componente para representar o Servidor GOBS com a função de busca necessária para calcular a rota intradomínio.
- Um componente para representar o Servidor GOBS Raiz com a função de armazenar os dados referentes aos recursos disponíveis de um AS distinto com o seu servidor GOBS comum associado a ele capaz de realizar busca necessária para calcular a rota interdomínio.
- Um agente para representar um nó de computação de grade. Esse nó deve conter informações sobre a capacidade de processamento e armazenamento total e disponível.
- Um gerador de tráfego responsável por gerar tarefas de grade, com as suas respectivas demandas de processamento, armazenamento e o limiar de bloqueio para a classe de serviço a qual pertence.

O nó desenvolvido por [57] para auxiliar na análise da arquitetura proposta é denominado *LOBS Node (Labeled OBS Node)* e é ilustrado na Figura 19.

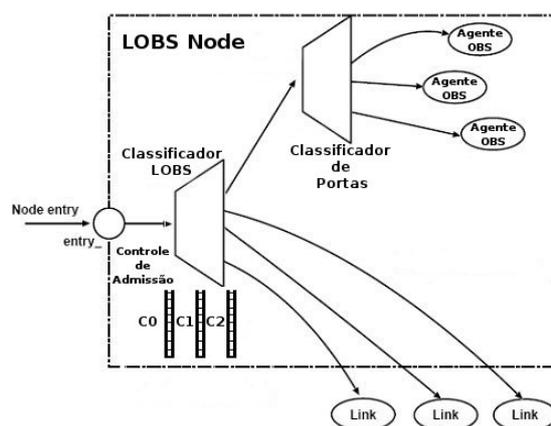


Figura 19: Estrutura do LOBS Node[57].

O classificador LOBS é o ponto de distribuição responsável por encaminhar a informação para o próximo componente de simulação. Quando o pacote (BCP ou

rajada) chega ao nó, o classificador LOBS o analisa para verificar se o mesmo tem o nó como destino final. Caso o pacote esteja endereçado ao próprio nó, ele é repassado para o classificador de portas que vai encaminhá-lo ao agente OBS correspondente. Caso contrário o pacote é encaminhado para os outros nós da rota através dos enlaces ópticos.

Quando o pacote é um BCP, o *LOBS Node* executa o mecanismo de controle de admissão de rajadas para verificar se existe um comprimento de onda disponível para a classe de serviço informada no BCP. Caso exista um comprimento de onda disponível, este é reservado imediatamente, caso contrário o BCP é descartado. Se o pacote for uma rajada óptica o classificador verifica se existe um comprimento de onda referente ao BCP. Após essa verificação e encaminhamento da rajada, o comprimento de onda é liberado para novas requisições. Caso não haja comprimento de onda reservado para a rajada, a mesma é descartada.

No agente OBS o processo de mapeamento de tarefas de grade é feito de uma forma um-para-um, ou seja, cada tarefa é convertida em uma rajada correspondente. Essa opção se deu pelo fato de que o esse mapeamento um-para-um proporciona uma simplificação da operação da rede, eliminando problemas como a ordenação dos pacotes [46].

O agente OBS envia BCPs integrados às mensagens de sinalização usadas para reservar os caminhos. Vale ressaltar quando uma rota selecionada já foi utilizada anteriormente, a mensagem de sinalização cumprirá somente a função de BCP, reservando comprimentos de onda, quando possível, para a rajada.

O agente GRIDJOB é responsável por gerar tarefas individuais. A aplicação JOBGGENERATOR utiliza os serviços do agente GRIDJOB para gerar um conjunto de tarefas. O agente GRIDJOB, no recebimento de um BCP, atualiza os valores de processamento e armazenamento disponíveis naquele nó. Quando a tarefa em si é

processada, esses contadores são restaurados para as próximas requisições. O funcionamento do agente GRIDJOB é apresentado na Figura 20.

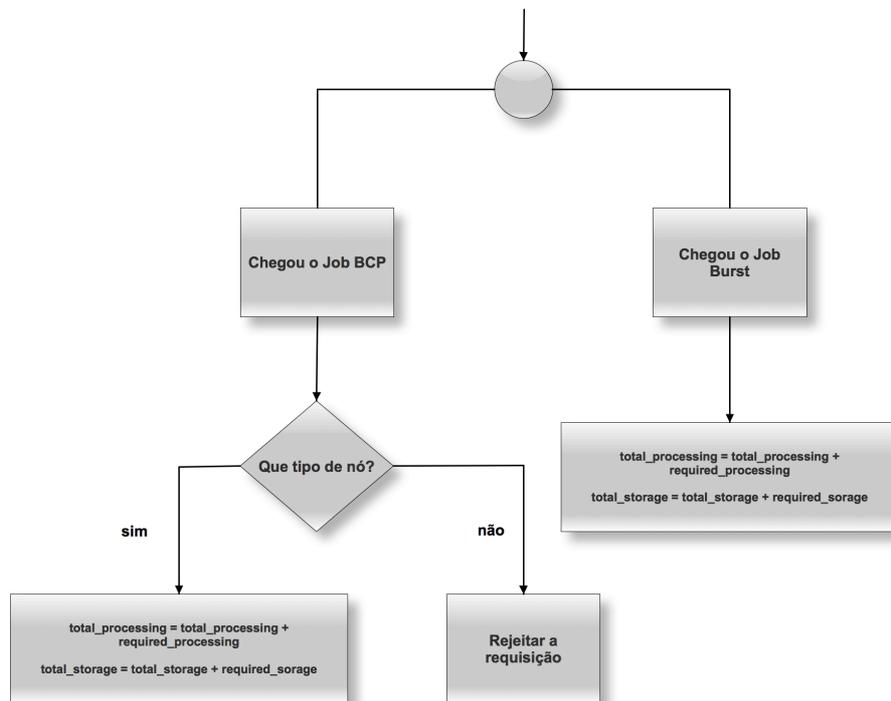


Figura 20: Agente GRIDJOB.

A interação entre os principais componentes desenvolvidos para o simulador com o intuito de caracterizar a proposta está ilustrada na Figura 21.

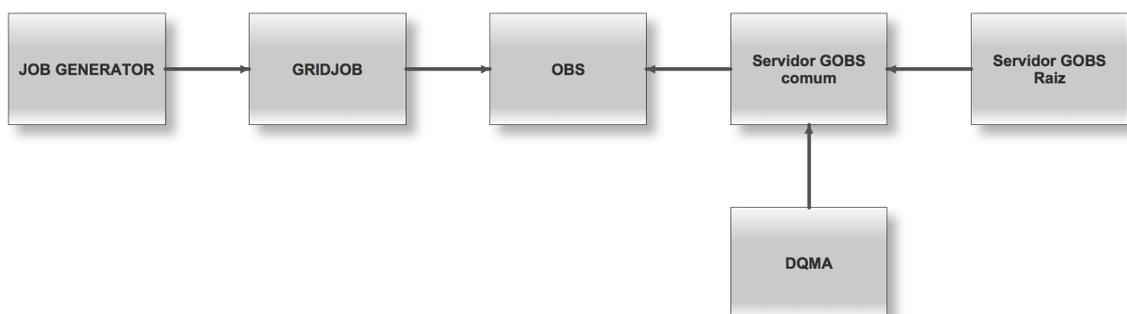


Figura 21: Interação entre os componentes desenvolvidos para o NS-2.

A aplicação JOBGGENERATOR, através do Agente GRIDJOB envia tarefas que por sua vez são processadas e transformadas em rajadas no *LOBS Node* através do Agente OBS. O Agente OBS solicita rotas para o servidor GOBS comum que por sua

vez recebe as informações atualizadas sobre os recursos intradomínio do Agente DQMA e quando estes encontrarem-se insuficientes para a requisição o servidor GOBS Raiz repassará as informações atualizadas sobre os recursos interdomínio.

5.2 – Cenário da Simulação

Para avaliar esta proposta, foi utilizado um cenário onde a topologia consiste de uma rede hipotética com 5 AS com a variação de 3 a 2 caminhos disjuntos como podem ser visto na figura 22.

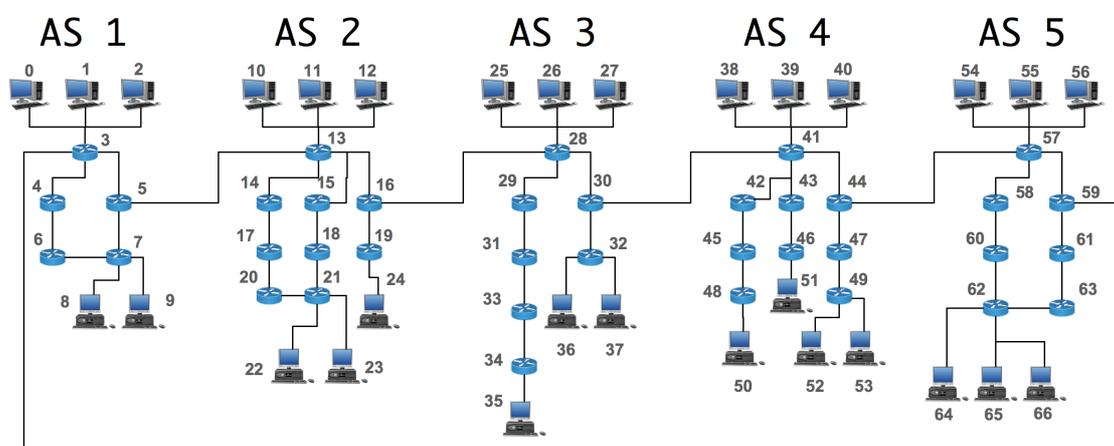


Figura 22: Topologia utilizada nas simulações.

Nesta topologia foram implementados dois modelos de simulação. O primeiro é o cenário denominado como interdomínio estático, onde o servidor GOBS Raiz não está. Quando a requisição é solicitada e recurso a ela vinculado não estiver disponível em seu AS, o servidor GOBS simplesmente enviará através de um circuito interdomínio para um destino padrão estático no AS vizinho onde os dados serão descartados. A figura 23 ilustra os caminhos interdomínio estáticos deste modelo em destaque na cor verde.

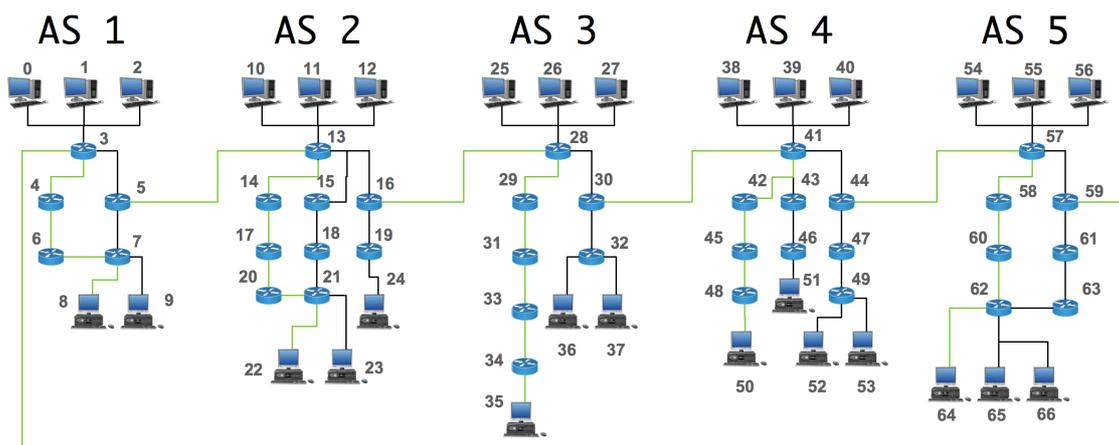


Figura 23: Modelo de simulação 1 com os seus circuitos interdomínio estático em destaque (verde).

O segundo modelo de simulação é o tema desta dissertação que apresenta o servidor GOBS Raiz que foi descrito no tópico 4.4, onde o mesmo atuará como um gerenciador de requisições de recursos interdomínio e disponibilizará os mesmos através de consultas dinâmicas.

Os nós foram distribuídos conforma pode ser visto na tabela 5, onde existem os geradores das tarefas (*jobs*), os nós de computação que processam as tarefas. Os demais (3 a 19) são nós de rede (comutadores ópticos). Os enlaces possuem capacidade de 10 *gigabits* por segundo e retardo de propagação de 1 milissegundo [67] [68]. Esta topologia foi intencionalmente escolhida por proporcionar caminhos disjuntos, que por sua vez permitem a visualização de maneira mais clara do impacto na mudança nas rotas selecionadas.

Tabela 5: Disposição dos nós na topologia.

Geradores de tráfego	Nós de grade	Nós de rede
0, 1, 2, 10, 11, 12, 25, 26, 27, 38, 39, 40, 54, 55, 56.	8, 9, 22, 23, 24, 35, 36, 37, 51, 52, 53, 64, 65, 66.	3, 4, 5, 6, 7, 13, 14, 15, 16, 17, 18, 19, 20, 21, 28, 29, 30, 31, 32, 33, 34, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 57, 58, 59, 60, 61, 62, 63.

Os *jobs* possuem tamanhos médios de 1,5 *megabytes* [68] distribuídos exponencialmente. A demanda de processamento de uma tarefa é uma fração do total disponível em um nó de simulação também distribuída exponencialmente com uma média de 60% do total, escolhida para proporcionar bloqueios pela ausência de recursos de computação. O total de capacidade de processamento e armazenamento disponível em cada nó de computação é de 15 GFLOPS e 1 *gigabyte* respectivamente. As capacidades de processamento e armazenamento escolhidas são semelhante às encontradas em computadores pessoais atuais [69]. O tempo de ajuste das rajadas é de 3 milissegundos. O tráfego foi modelado de acordo com uma distribuição de Poisson exponencialmente distribuído.

São definidas três classes de serviço para esta análise: classe 0, classe 1 e classe 2, sendo que a classe 0 é a de mais alta prioridade, a classe 1 possui prioridade intermediária e classe 2 é a de melhor esforço.

Foi utilizado um mecanismo de controle de admissão semelhante ao proposto em [54] onde existe um número máximo de comprimentos de onda reservado para cada classe de serviço. A Tabela 6 ilustra a distribuição de comprimentos de onda para cada classe de serviço. Cada fibra possui um total de 8 comprimentos de onda, valor normalmente utilizados em sistemas CWDM (*Coarse WDM*) [70].

Tabela 6: Número de comprimentos de onda por classe.

Classe	Número de comprimentos de onda
Classe 0	5
Classe 1	3
Classe 2	1

A probabilidade de bloqueio máxima definida para cada classe de serviço é mostrada na Tabela 7. Esses valores foram escolhidos a partir de testes preliminares que demonstraram que os mesmos são representativos para a análise da proposta e também são próximos de valores encontrados na literatura [71].

Tabela 7: Probabilidade de bloqueio máxima permitida por classe.

Classe	Probabilidade de bloqueio
Classe 0	0.01
Classe 1	0.05
Classe 2	0.1

5.3 – Resultados Obtidos

Os resultados referentes à probabilidade de bloqueio, atraso fim-a-fim, processamento e armazenamento médio utilizados, e número de requisições rejeitadas nos nós de grade serão apresentados nesta subseção. O objetivo é avaliar de maneira comparativa, com base nos dois modelos de simulação descritos na subseção 5.1, o impacto da aplicação do Servidor GOBS Raiz e do mecanismo de seleção de rotas proposto na obtenção de garantias de qualidade de serviço em redes Grid OBS e na utilização dos recursos de computação disponíveis. Todas as simulações foram conduzidas com 100 repetições para cada variação na carga de *erlangs* totalizando 1000 repetições e foi utilizado um intervalo de confiança de 95% em relação à média das amostras coletadas.

A primeira análise é referente à probabilidade de bloqueio experimentada pelos nós que recebem a maior parte das requisições. Escolheu-se analisar a princípio os nós 3, 13, 28, 41 e 57, pois estes são os nós mais vezes selecionados, conforme mostra a Figura 24. No modelo de simulação 1 esses nós compõem a rota padrão, isto é, a rota escolhida quando a rajada é enviada ao destino padrão de cada AS, já no modelo de

simulação 2, que vem ser a proposta desta dissertação os mesmos nós atuam como a rota para todas as requisições interdomínio que até então são descartados no modelo 1.

A carga na rede varia de 30 a 300 *erlangs*. A carga em *erlangs* é definida pelo produto do número médio de requisições que chegam à rede pelo tempo médio que as requisições permanecem na rede [66].



Figura 24: Número de requisições por nó.

Vale destacar que outros nós apresentaram um número elevado de requisições como por exemplo os nós 5, 16, 30, 44 e 59, entretanto decidiu-se não apresentar os resultados dos mesmos porque dentro das simulações foi analisado o grau de importância de cada um desses 10 nós com maior número de requisições. Como foi descrito no item 5.2, todas as requisições passam obrigatoriamente pelos nós 3, 13, 28, 41 e 57, pois eles são diretamente o próximo salto dos nós responsáveis por gerar as tarefas, já os nós 5, 16, 30, 44 e 59 tem apenas, quando comparado com os 5 primeiros nós, uma figuração de saída ou entrada do AS solicitante ao AS provedor, desconsiderando as requisições intradomínio.

As Figuras 25, 26, 27, 28 e 29 mostram a probabilidade de bloqueio comparando com os 3 modelos de qualidade de serviço nos nós 3, 13, 28, 41 e 57 respectivamente.

Nó 3

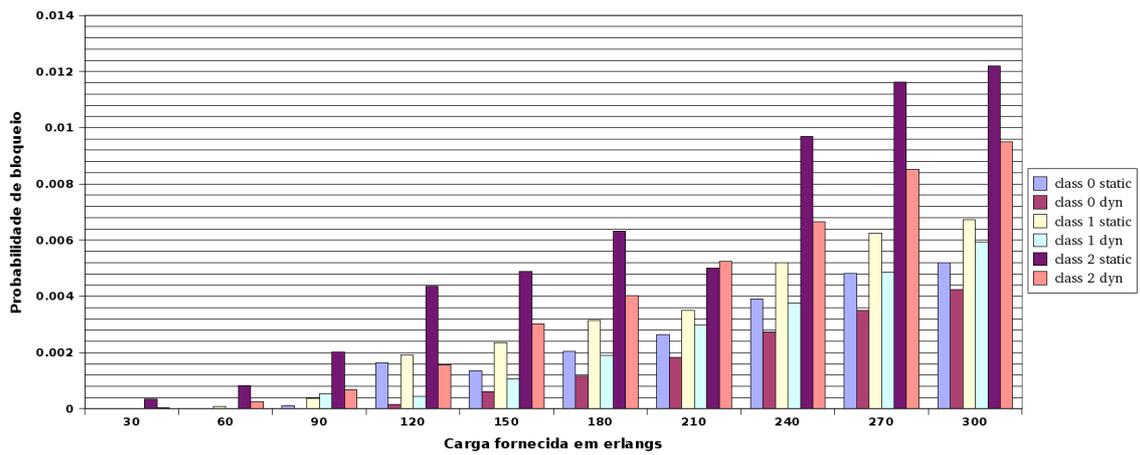


Figura 25: Probabilidade de bloqueio x Classes de serviço (Nó 3).

Nó 13

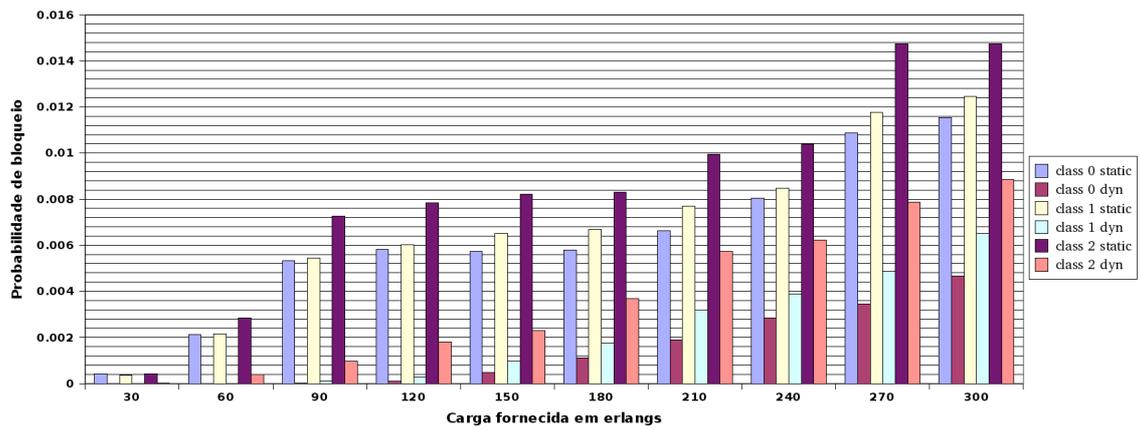


Figura 26: Probabilidade de bloqueio x Classes de serviço (Nó 13).

Nó 28

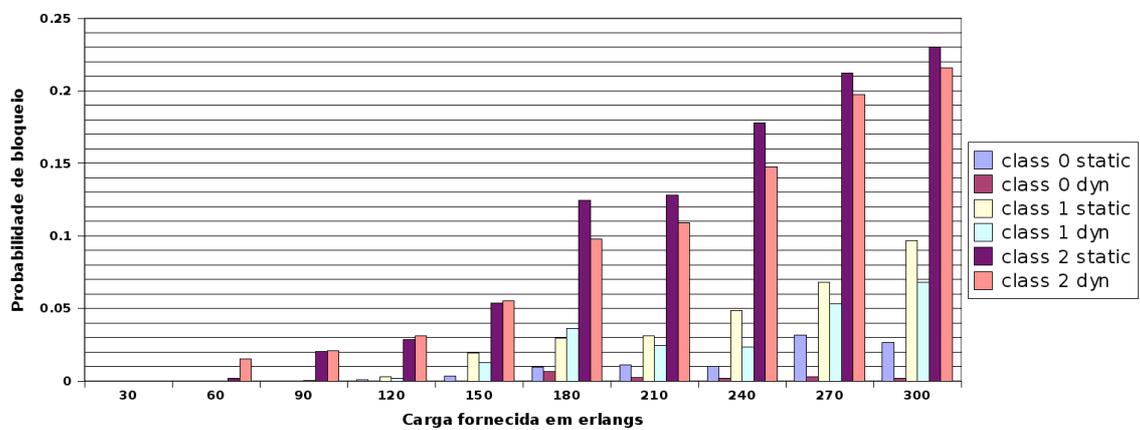


Figura 27: Probabilidade de bloqueio x Classes de Serviço (Nó 28).

Nó 41

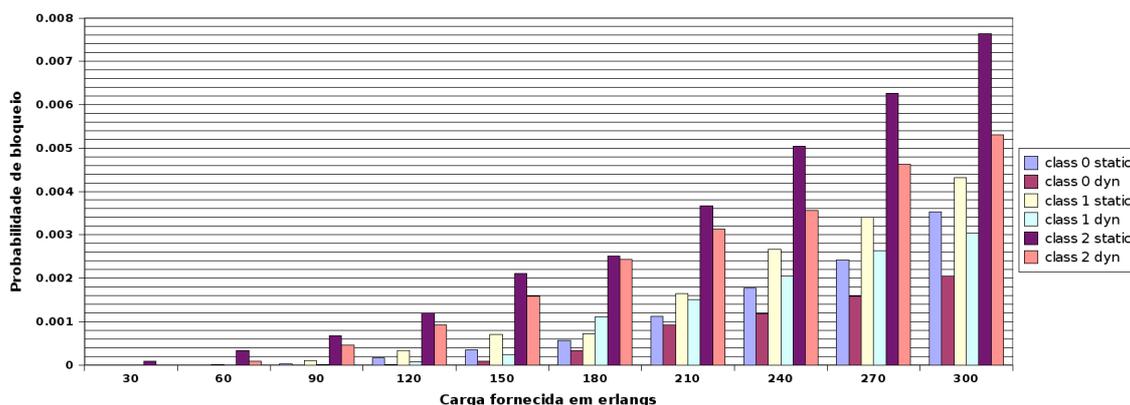


Figura 28: Probabilidade de bloqueio x Classe de serviços (Nó 41).

Nó 57

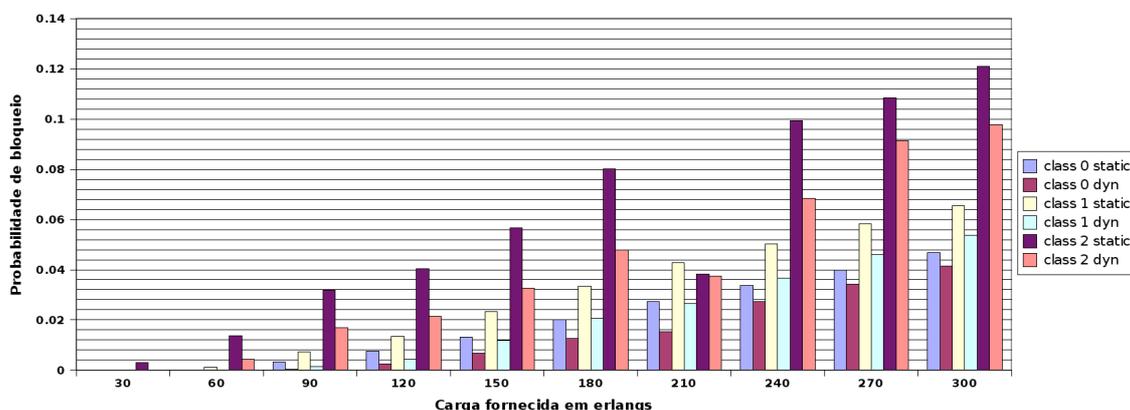


Figura 29: Probabilidade de bloqueio x Classe de serviços (Nó 57).

Segundo os dados acima apresentados é notório a melhora quanto a redução da probabilidade de bloqueio no modelo de simulação 2 (interdomínio dinâmico) quando comparado ao modelo de simulação 1 (interdomínio estático) como pode ser visto nos nós acima analisados, lembrando que estes são comuns a todas as possíveis rotas intradomínio, uma vez que são o próximo salto dos nós que geram tráfego (vide tabela 5), assim como rotas interdomínio de entrada no AS provedor do recurso.

Apenas nestes 5 nós foram totalizados 300 iterações sendo 60 para cada nó, a qual foram distribuídas para cada classe de serviço 20 iterações. Com base nestes resultados foi possível demonstrar o aproveitamento do modelo de simulação com

interdomínio dinâmico de aproximadamente 98% onde o mesmo foi superior em 294 iterações.

É importante também frisar que em alguns casos o modelo de simulação com interdomínio dinâmico teve requisições bloqueadas a mais quando comparado com o modelo de simulação com intradomínio estático, entretanto, nestes casos os níveis de probabilidade de bloqueio foram superiores em apenas 0,01%, como pode ser visualizado na figura 27, mais precisamente na carga de 180 *erlangs* com as classes de serviço 1 .

Para estes nós o desempenho de todas as classes foi garantido durante toda simulação.

Outra análise realizada é referente à utilização dos recursos de computação através da verificação das médias de utilização dos recursos de processamento e armazenamento em todos os nós de grade na topologia apresentada. As Figuras 30 e 31 ilustram o uso dos recursos de processamento e armazenamento para os nós de computação da grade pelos 2 modelos de simulação.

Pela figura nota-se que a arquitetura proposta nesta dissertação proporciona uma redução considerável na carga de processamento, uma vez que as tarefas são melhor distribuídas entre os nós da grade caracterizando o caráter *anycast* da proposta.

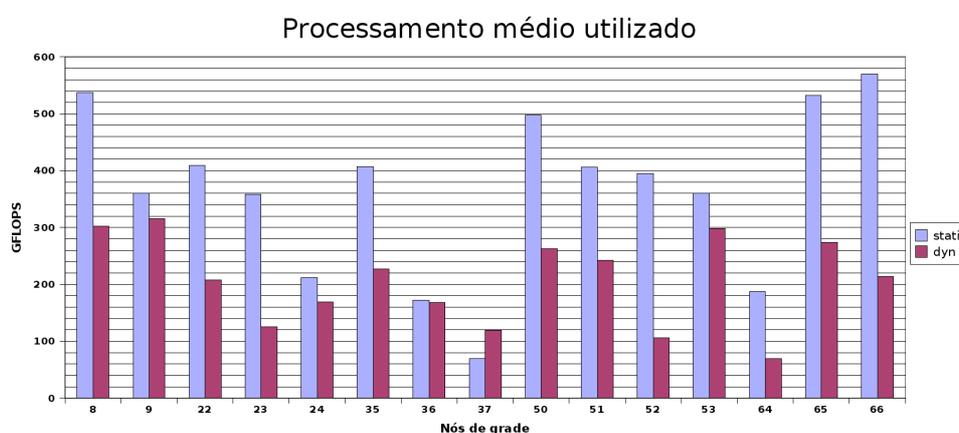


Figura 30: Processamento médio utilizado.

Para o espaço de armazenamento observa-se que no modelo de interdomínio dinâmico a carga é melhor distribuída entre os nós evitando assim a sobrecarga e consequentemente a indisponibilidade de um recurso de armazenamento em um nó de computação em conjunto como ocorre no modelo de simulação interdomínio estático, devido à alta carga de tráfego oferecida à rede e ao espaço de armazenamento limitado disponíveis nos nós. É possível observar um incremento mais significativo na utilização do espaço de armazenamento nos nós 8, 24, 36, 51 e 64 devido o fato de que mais rajadas são encaminhadas para esses nós em decorrência da seleção de rotas, todavia em nenhum dos casos ocorreu sobrecarga de espaço de armazenamento porque o GOBS Raiz realiza um balanceamento de recursos disponíveis nos AS através das informações disponibilizadas pelo GOBS comum a ele associados .

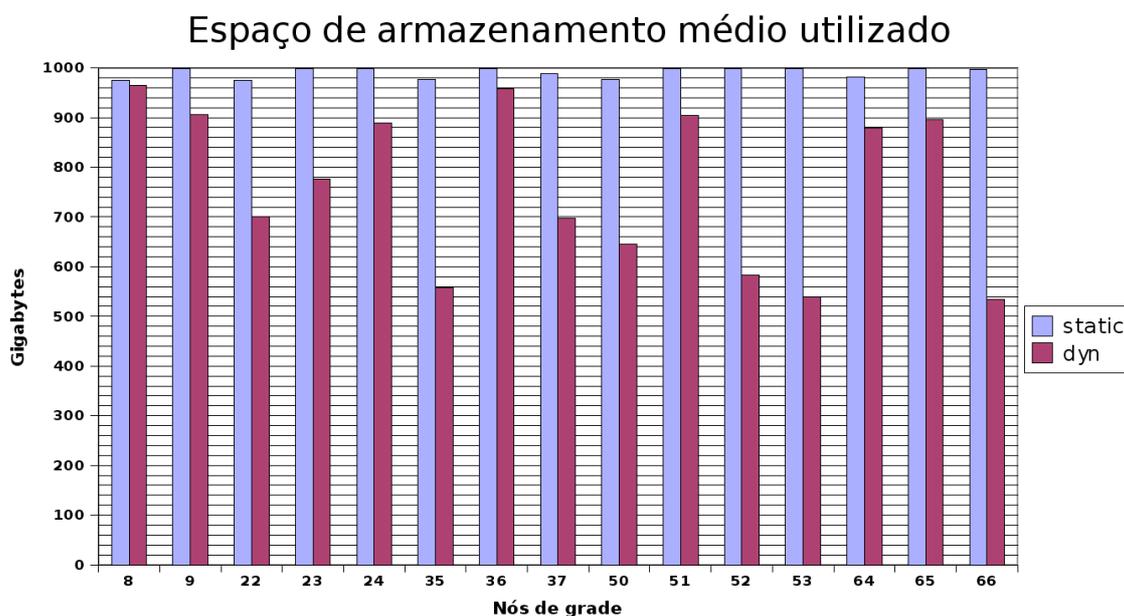


Figura 31: Espaço de armazenamento médio utilizado.

É possível verificar na figura 32 o valor do atraso do modelo interdomínio dinâmico mais elevado em alguns nós em relação ao modelo interdomínio estático. Isto ocorre devido no modelo de interdomínio dinâmico o requisitante possui um tempo de

espera maior pois o usuário faz a sua requisição ao GOBS comum e quando tal recurso encontrar-se indisponível o mesmo fará a consulta ao GOBS Raiz, diferente do modelo interdomínio estático que já reconhece toda a sua topologia e o seu usuário acaba esperando menos tempo na obtenção da resposta para aonde seguir, pois em alguns dos nós abaixo existe o comportamento do destino padrão para quando a requisição desejada não puder ser atendida.

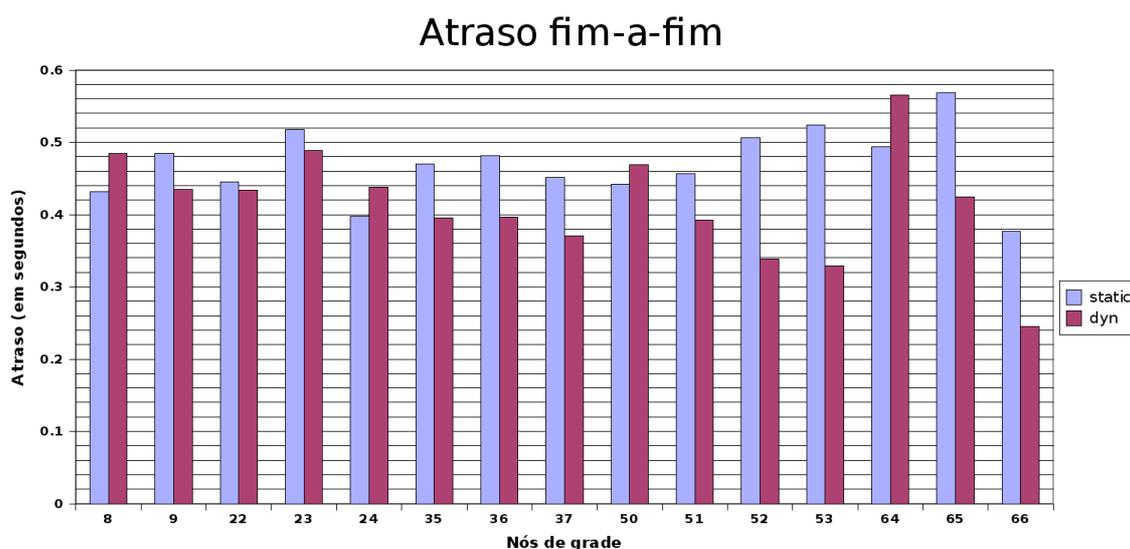


Figura 32: Atraso fim-a-fim.

Em seguida, a Figura 33 apresenta o número de requisições descartadas exclusivamente devido à ausência de recursos de computação (processamento e armazenamento) nos nós da grade.

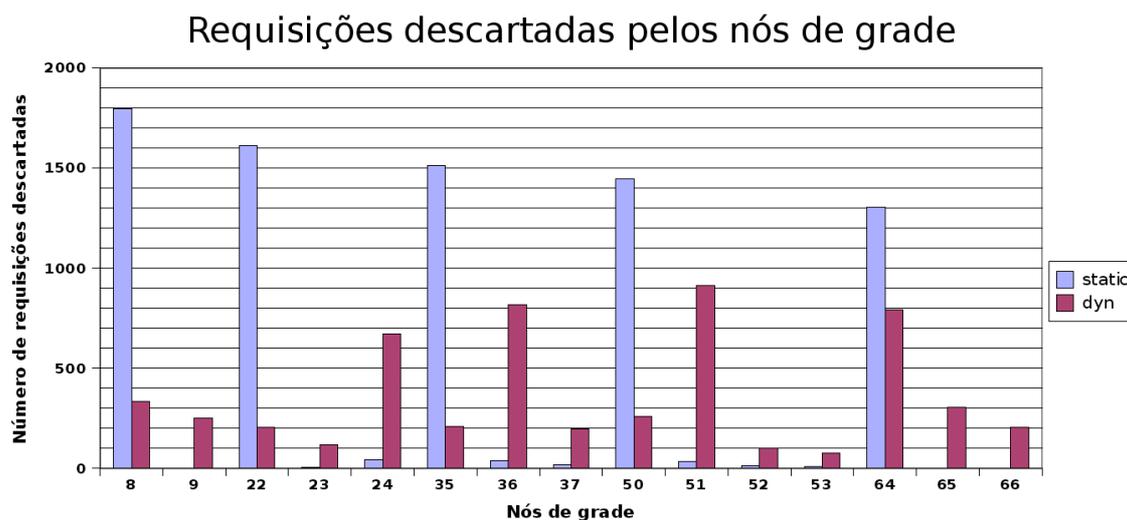


Figura 33: Número de requisições descartadas nos nós de computação.

Vale lembrar que tais requisições são descartadas exclusivamente pela ausência de recursos de computação (*e.g.*, processamento/armazenamento) suficientes para atender a uma determinada tarefa.

Na sua maioria, 10 dos 15 nós de grade, a proposta desta dissertação obteve valores superiores de requisições descartadas em comparação com o modelo de interdomínio estático. Isto porque no modelo interdomínio dinâmico ocorre um maior e mais frequente balanceamento da carga através da resposta do GOBS Raiz na tentativa de uniformizar a disponibilidade de recursos nos nós de computação da grade, fazendo com que haja tais descartes. Porém, é importante destacar que o número de requisições descartadas no modelo interdomínio estático nos nós 8, 22, 35, 50 e 64 é superior ao modelo proposto porque estes nós são o destino padrão dos AS5, AS1, AS2, AS3 e AS4 respectivamente.

A partir de agora serão apresentados resultados comparando a média geral da probabilidade de bloqueio em todos os nós da topologia apresentada entre os o modelo de simulação interdomínio estático e o modelo proposto nesta dissertação.

As Figuras 32, 33 e 34 apresentam resultados referentes à probabilidade de bloqueio experimentada por cada classe de serviço utilizando os dois esquemas de gerenciamento de qualidade de serviço propostos.

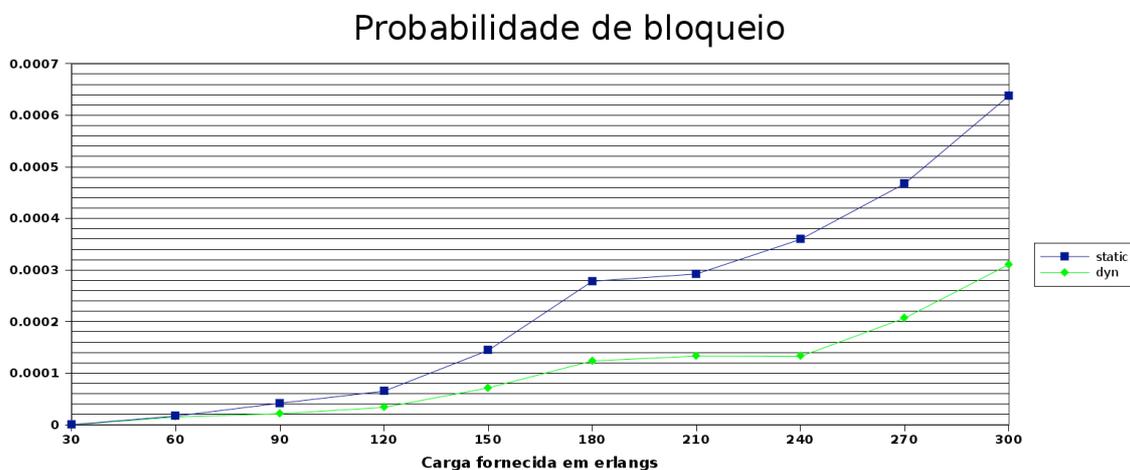


Figura 34: Probabilidade de bloqueio (static x dyn - Classe 0).

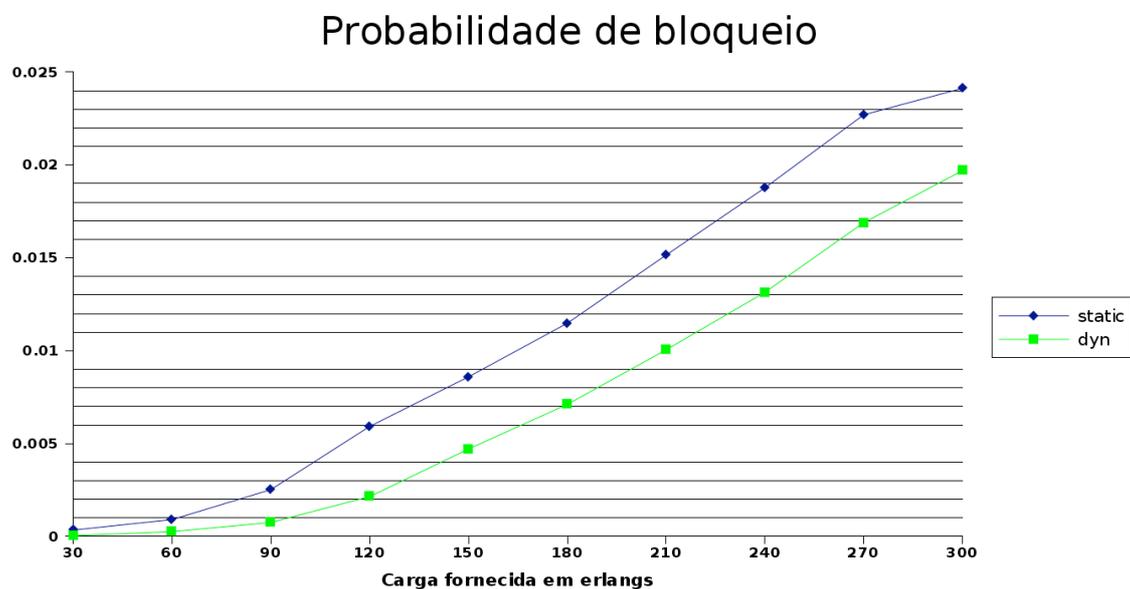


Figura 35: Probabilidade de bloqueio (static x dyn - Classe 1).

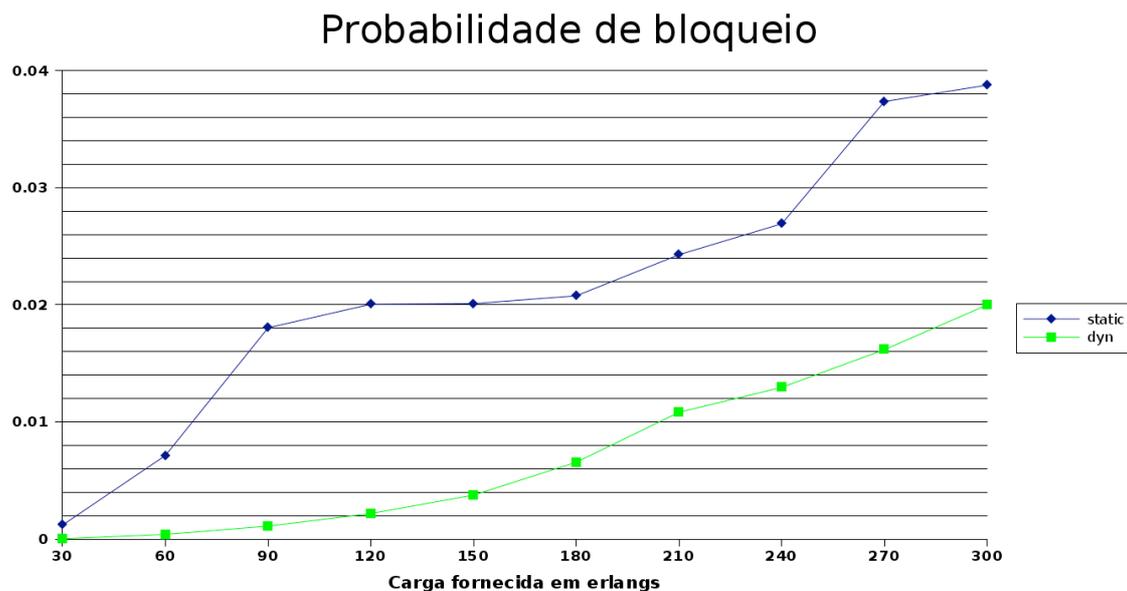


Figura 36: Probabilidade de bloqueio (static x dyn - Classe 2).

É possível observar que, de maneira geral, o modelo de simulação interdomínio dinâmico apresenta um melhor desempenho comparado ao modelo de simulação interdomínio estático. Isso se deve ao fato de que o modelo de simulação interdomínio dinâmico se adapta mais rapidamente à dinâmica da rede uma vez que as rajadas são tratadas de forma independente.

5.4 Considerações Finais

Esta subseção conclui o capítulo 5 que apresentou a metodologia utilizada para a análise da arquitetura proposta nesta dissertação assim como o ambiente de simulação usado que valida a mesma. As extensões desenvolvidas para o simulador necessárias para caracterizar a proposta, tal qual os cenários utilizados na avaliação e os principais resultados obtidos foram apresentados e discutidos.

Com base nos resultados apresentados é possível provar que a proposta é capaz de proporcionar níveis absolutos de qualidade de serviço conforme os parâmetros adotados [71] para as rajadas dentro do contexto apresentado.

A importância do Servidor GOBS Raiz foi comprovada durante este capítulo através dos resultados das simulações apresentadas que expôs o melhoramento da probabilidade de bloqueio e na melhor utilização dos recursos de computação.

Capítulo 6

CONCLUSÕES E TRABALHOS FUTUROS

Este capítulo apresenta as conclusões gerais do trabalho e apresenta algumas possibilidades de trabalhos futuros. A Seção 6.1 lista as principais contribuições da dissertação, a Seção 6.2 apresenta as conclusões gerais do trabalho. Finalmente, a Seção 6.3 apresenta os trabalhos futuros.

6.3 Contribuições

As principais contribuições deste trabalho de acordo com os objetivos definidos na introdução estão listadas a seguir:

- A criação de um modelo para o provisionamento multidomínio dinâmico de conexões determinísticas em redes GOBS, combinando funcionalidades de controle centralizadas e distribuídas.
- A definição de uma metodologia para o cálculo de rotas que leva em consideração os parâmetros de rede e de computação definidos pelo usuário/aplicação, e que retorna uma aproximação quando não encontra uma rota exata para a requisição.
- Modificações na sinalização GMPLS para integrá-la com a sinalização OBS e agilizar o processo de reserva.
- Desenvolvimento de extensões para o simulador ns-2 que caracterizam a proposta.

6.2 – Conclusões

Ao término desta dissertação, conclui-se que cada vez mais faz-se necessário uma alta quantidade de recursos de computação, pois, a cada dia que passa, mais e mais aplicações que exigem grande poder de processamento, memória e armazenamento. Entretanto os custos de tais recursos são inviáveis para a grande maioria dos que o necessitam, por isso, o crescimento de pesquisas para o aprimoramento da computação em grade é um fato e a cada dia vem ganhando mais e mais força, em um contexto onde a colaboração entre pesquisadores de diversas áreas do conhecimento é cada vez mais intensa.

Atualmente o método de comutação de rajadas ópticas permanece como a melhor opção para viabilizar a computação em grade, porém a comutação de rajadas ópticas precisa ser ainda melhor adaptada para esse modelo, visto que toda sinalização OBS não possui resposta. Nesse contexto, surgem vários desafios para as redes que devem ser resolvidos com o objetivo de adequá-las ao modelo de computação em grade, porém o paradigma ainda se mostra atrativo no contexto de grades, por apresentar características como a granularidade variável das informações e a separação entre dados de controle e de aplicação.

Quando se trata de aplicações em grade é importante frisar qual o tipo de dados que podem ser gerados, por isso garantias de desempenho nas conexões oferecidas são uma questão fundamental, o que é exigido por certos tipos de aplicações científicas e por aplicações baseadas em um grande volume de conteúdo multimídia.

Redes ópticas possuem como uma das suas principais vantagens a possibilidade de tráfego de grande quantidade de dados, porém, pelo fato de em redes ópticas tratar-se de luz, não será possível contar com buffers similares aos existentes nas redes eletrônicas. Qualidade de serviço é um problema importante em redes puramente

ópticas, e mesmo assim esta dissertação apresentou uma arquitetura para o provisionamento dinâmico multidomínio de conexões determinísticas, onde o desempenho de todas as classes de qualidade de serviço foram garantidos durante toda simulação baseadas no paradigma de comutação de rajadas ópticas em um ambiente de grade.

A engenharia tráfego permite que se possa oferecer qualidade de serviço em redes ópticas. A característica de encaminhar dados para rotas ociosas, reduzindo, assim, perdas excessivas em enlaces específicos da rede presente em protocolos de sinalização do GMPLS (*e.g.*, RSVP-TE) foi incorporada e adaptada à proposta uma vez que se utiliza o roteamento explícito para reservar os comprimentos de onda nos nós da rota selecionada pelo Servidor GOBS Raiz.

Os resultados mostraram-se satisfatório, logo que provaram que a proposta é capaz de melhorar e garantir o desempenho todas as classes de serviço, e proporciona um uso mais eficiente dos recursos de computação, através do uso da engenharia de tráfego.

A proposta desta dissertação combina o contexto centralizado e distribuído procurando atender as diferentes necessidades exigidas de um plano de controle para grades. O Servidor GOBS Raiz é o componente centralizado que possui uma visão geral acerca de todos os recursos dos AS a ele associados através do GOBS comum em cada um deles, o que proporciona um controle de rotas de maneira precisa e dinâmica. As funcionalidades de sinalização e monitoração são implementadas de forma distribuída, proporcionando um rápido estabelecimento de caminhos determinísticos e uma rápida atualização das informações no Servidor GOBS Raiz através do servidor GOBS comum em cada AS, contribuindo assim para que as informações armazenadas se mantenham consistentes.

6.3 – Trabalhos Futuros

Uma das principais preocupações para uma atualização desta arquitetura é a adição de novas métricas para a seleção de rotas, como atraso e variação do atraso, permitindo que as rotas selecionadas possam atender de forma mais precisa os requisitos impostos pelas aplicações de grade. Com base nesses dados novas formas de seleção de caminhos serão implementadas e com isso tornando ainda mais precisos os resultados diminuindo assim o número de requisições grid rejeitadas, probabilidade de bloqueio, entre outras.

Criar métodos de comunicação entre servidores GOBS Raiz sobre trocas de informações de recursos de grade dos AS associados a eles. Diminuindo assim a centralização da proposta com o aumento do seu dinamismo.

Propor uma solução integrada ao GOBS Raiz de predição de falhas em redes GOBS utilizando inteligência artificial, tendo como base métricas da camada física para antecipar problemas que podem ocasionar perdas de tarefas submetidas para a grid.

Tendo em vista as limitações de desempenho existentes no ns-2, é importante buscar adaptações a solução apresentada através de métodos de simulação matemática, que podem apresentar dados mais precisos e relevantes, assim como reformular a proposta para outras ferramentas de simulação tão conceituadas quanto o NS, como o OPNET, OMNET, ou até mesmo para a sua versão mais atualizada, o NS-3. Com isso, buscando sempre buscar métodos de adquirir resultados mais próximos a realidade.

Por fim, implementar a solução proposta nesta dissertação em um ambiente real (*testbed*) para a validação desta arquitetura em um ambiente de produção, todavia, seria provavelmente necessário realizar adaptações no cenário desenvolvido tendo vista o alto custo de uma rede óptica com comutadores OBS. Neste caso, a implementação seria centrada nas funcionalidades de controle e gerenciamento.

Referências Bibliográficas

- [1] FOSTER, I. **Grid: A New Infrastructure for 21st Century Science**. Physics Today, v. 42, n. 2, fev. 2002.
- [2] JUKAN, A.; KARMOUS-EDWARDS, G. **Optical Control Plane for the Grid Community**. IEEE Communications Surveys & Tutorials, v. 9, n. 3, p. 30-44, 2007.
- [3] TRAVOSTINO, F.; MAMBRETTI, J.; KARMOUS-EDWARDS, G. **Grid Networks: Enabling Grids with Advanced Communication Technology**. 1st ed. John Wiley and Sons, 2006.
- [4] GURUSAMY, M.; MURHTY, C. S, R. **WDM Optical Networks: Concepts, Design, and Algorithms**. Prentice Hall, 2001.
- [5] GOFF, D. **Fiber Optic Reference Guide**. 3rd ed. Focal Press, 2002.
- [6] VOKKARANE, V. M.; JUE, J. P. **Optical Burst Switched Networks**. 1st ed. Springer, 2005.
- [7] QIAO, C. e YOO, M. Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet. Journal of High Speed Networks (JHSN), v. 8, n. 1, p. 69-84, ago. 1999.
- [8] TANWIR. S. **Network Resource Scheduling and Management of Optical Grids**. 69 f. Master's Thesis, Department of Computer Science, North Carolina State University, Raleigh, North Carolina, United States. 2007.
- [9] NEJABATI, R. **Grid Optical Burst Switched Networks (GOBS): Informational Track**. [S.l.]: Open Grid Forum, Grid High-Performance Networking Research Group, abr. 2008.

- [10] MANNIE E. **Generalized Multi-Protocol Label Switching (GMPLS) Architecture**. RFC 3945. [S.l.]: Internet Engineering Task Force, Network Working Group, out. 2004.
- [11] HABIB, I.W, et al. **Deployment of the GMPLS Control Plane for Grid Applications in Experimental High-Performance Networks**. IEEE Communications Magazine, v. 44, n. 3, p. 65- 73, mar. 2006.
- [12] FOSTER, I. **What is the Grid? A Three Point Checklist**. Disponível em: <<http://www.fp.mcs.anl.gov/~foster/Articles/Grid/WhatIsTheGrid.pdf>>, jul. 2009. Acesso em: nov. 2010.
- [13] FOSTER, I.; KESSELMAN, C.; TUECKE, S. **The anatomy of the Grid: Enabling scalable virtual organizations**. International Journal on Supercomputer Applications, v. 15, n. 3, p. 200-222, 2001.
- [14] BATISTA, D. **Engenharia de Tráfego Multi-Camada para Grades**. 2006. 118 f. Dissertação de Mestrado, Instituto de Computação, Universidade Estadual de Campinas, Campinas, São Paulo, Brasil.
- [15] OGF. **Open Grid Forum**. Disponível em: <<http://www.ogf.org>>. Acesso em: dez. 2010.
- [16] FOSTER, I. et al. **The Open Grid Services Architecture, Version 1.5**. [S.l.]: Open Grid Forum, Open Grid Services Architecture Working Group, jul. 2006.
- [17] LHC. **Large Hadron Collider**. Disponível em: <<http://lhc.web.cern.ch/lhc/>>. Acesso em: dez. 2010.
- [18] EVLBI. **The European VLBI Network**. Disponível em: <<http://www.evlbi.org/>>. Acesso em: dez. 2010.

- [19] OPTIPUTER. **Optical networking, Internet Protocol, computer storage, processing and visualization Technologies.** Disponível em: <<http://www.optiputer.net/>>. Acesso em: jan. 2011.
- [20] CINEGRID. **CineGrid.** Disponível em: <<http://www.cinegrid.org/>>. Acesso em: jan. 2011.
- [21] MUKHERJEE, B. **Optical WDM Networks.** 1st ed. Springer, 2006.
- [22] ABELÉM, A. **Difusão Seletiva em Inter-Redes IP Baseadas em Redes Ópticas.** 2003. Tese de doutorado. Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, Brasil.
- [23] NSI-WG. **Network Service Interface Working Group.** Open Grid Fórum, Disponível em: <<http://forge.gridforum.org/sf/projects/nsi-wg>>. Acesso em: set. 2010.
- [24] ENLIGHTENED COMPUTING. **Enlightened Computing Project.** Disponível em: <<http://www.enlightenedcomputing.org/>>. Acesso em: set. 2010.
- [25] DRAGON Project. **Dynamic Resource Allocation via GMPLS Optical Networks.** Disponível em: <<http://dragon.maxgigapop.net>>. Acesso em: dez. 2010.
- [26] **PHOSPHORUS PROJECT.** Disponível em: <<http://www.ist-phosphorus.eu/activities.php>>. Acesso em: jul. 2008.
- [27] GENI.NET. **Global Environment for Network Innovations.** Disponível em: <<http://www.geni.net>>. Acesso em: jan. 2011.
- [28] FIRE INITIATIVE. **Future Internet Research and Experimentation.** Disponível em: <<http://cordis.europa.eu/fp7/ict/fire/>>. Acesso em: fev. 2011.
- [29] AKARI Project. **AKARI Architecture Design Project for New Generation Network.** Disponível em: <<http://akari-project.nict.go.jp/>>. Acesso em: nov. 2010.

- [30] YOO, M.; QIAO, C.; DIXIT, S. **Optical Burst Switching for Service Differentiation in the Next-Generation Optical Internet**. IEEE Communications Magazine, v. 39, n. 2, p. 98-104, fev. 2001.
- [31] DE LEENHEER, M. et al. **Anycast Algorithms Supporting Optical Burst Switched Grid Networks**. In: INTERNATIONAL CONFERENCE ON NETWORKING AND SERVICES (ICNS). Proceedings... Silicon Valley, United States, 2006, p. 63.
- [32] WEI, J. Y.; MCFARLAND, R. I. **Just-in-time Signaling for WDM Optical Burst Switching Networks**. Journal of Lightwave Technology, v. 18, n. 12, p. 2019–2037, dez. 2000.
- [33] NEJABATI, R et al., **Programmable Optical Burst Switched Network: A Novel Infrastructure for Grid**. In: 5TH IEEE/ACM INTERNATIONAL SYMPOSIUM ON CLUSTER COMPUTING AND THE GRID (CCGRID). Proceedings... Cardiff, United Kingdom, 2005, p. 993-999.
- [34] XIONG, Y.; VANDENHOUTE, M. e CONKOYA, H. **Control Architecture in Optical Burst Switched WDM Networks**. IEEE Journal of Selected Areas in Communications, v. 18, n. 10, p. 1838-1851, out. 2000.
- [35] GLIF. **Global Lambda Integrated Facility**. Disponível em: <<http://www.glif.is/>>. Acesso em: ago. 2010.
- [36] ROSEN, E.; VISWANATHAN, A. e CALLON, R. **Multiprotocol Label Switching Architecture**. RFC 3031. Janeiro, 2001.
- [37] BERNSTEIN, G.; RAJAGOPALAN, B.; SAHA, D. **Optical network control: architecture, protocols, and Standards**. 1st ed. Addison-Wesley, 2003.

- [38] KATZ, D.; KOMPELLA, K.; YEUNG, D. **Traffic Engineering (TE) Extensions to OSPF Version 3**. RFC 5329. [S.l.]: Internet Engineering Task Force, Network Working Group, 2008.
- [39] AWDUCHE, D. et al. **Fast Reroute Extensions to RSVP-TE for LSP Tunnels** RFC 4090. [S.l.]: Internet Engineering Task Force, Network Working Group, 2005.
- [40] JAMOUSSE, B., et al. **Constraint-Based LSP Setup using LDP**. RFC 3212. [S.l.]: Internet Engineering Task Force, Network Working Group, 2002.
- [41] LANG, J. **Link Management Protocol (LMP)**. RFC 4204. [S.l.]: Internet Engineering Task Force, Network Working Group, 2005.
- [42] BERGER, L. (editor). **Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description**. RFC 3471. [S.l.]: Internet Engineering Task Force, Network Working Group, 2003.
- [43] MORROW, M.; SAYEED, A. **MPLS and Next-Generation Networks: Foundations for NGN and Enterprise Virtualization**. 1st ed. Cisco Press, 2006.
- [44] LANG, J.; RAJAGOPALAN, P. **Generalized Multi-Protocol Label Switching (GMPLS) Recovery Functional Specification**. Internet Draft, draft-ietf-ccamp-gmpls-recovery-functional-03.txt. [S.l.]: Internet Engineering Task Force, Network Working Group, 2009.
- [45] VIANA, J. **Tolerância à Falhas em Redes Ópticas de Nova Geração**. 2006. Dissertação de Mestrado, Departamento de Engenharia Elétrica e de Computação, Universidade Federal do Pará. Belém, Pará, Brasil.
- [46] DE LEENHEER, M. et al. **An OBS-based Grid Architecture**. In: WORKSHOP ON HIGH-PERFORMANCE GLOBAL GRID NETWORKS (GLOBECOM 2004). Proceedings... Dallas, United States, 2004, p. 390-394.

- [47] OKI E., TAKEDA T., LE ROUX J.-L., FARREL A., **Framework for PCE-based inter-layer MPLS and GMPLS traffic engineering**. IETF Internet Draft. <http://tools.ietf.org/id/draft-ietf-pce-inter-layer-frwk-10.txt>, March 2009.
- [48] CHAMANIA M., CHEN X., ADMELA J., RAMBACH F., HOFFMANN M. . **An Adaptive inter-domain PCE framework to improve resource utilization and reduce inter-domain signaling**. Optical Switching and Networking, Elsevier, vol. 6, issue 2, pp. 259-267, jun. 2009.
- [49] YANNUZZY M., MASHIP-BRUIIN X., SANCHEZ-LOPEZ S., and MARINTORDERA E. . **OBGP+: An improved path-vector protocol for multi-domain optical networks**. Optical Switching and Networking, Elsevier, vol. 6, issue 2, pp. 111-119, Abr. 2009.
- [50] YANNUZZY M., MASHIP-BRUIIN X., FABREGO G., SANCHEZ-LOPEZ S., SPRINTSON A., ORDA A. . **Toward a new route control model for multi-domain optical networks**. IEEE Communications Magazine, vol. 46, no. 6, pp. 104-111, Jun 2008.
- [51] BESHIR A., YANNUZZI M., KUIPERS F. **Inter-Domain Routing in Optical Networks with Wavelength Converters**. IEEE International Conference on Communications, ICC2010. Maio, 2010.
- [52] LEHMAN, T.; SOBIESKI, J.; JABBARI, B. **DRAGON: a Framework for Service Provisioning in Heterogeneous Grid Networks**. IEEE Communications Magazine, v. 44, n. 3, p. 84-90, mar. 2007.
- [53] VOKKARANE, et al. **Generalized Burst Assembly and Scheduling Techniques for QoS Support to Optical Burst-Switched Networks**. In: IEEE GLOBAL TELECOMMUNICATIONS CONFERENCE (GLOBECOM). Proceedings... Taipei, Taiwan, China, p. 2747–2751, 2002.

- [54] ZHANG, Q. et al. **Absolute QoS Differentiation in Optical Burst-Switched Networks**. IEEE Journal on Selected Areas in Communications, v. 22, n. 9, p. 1781-1795, nov. 2004.
- [55] LONG, K.; TUCKER, R.; WANG, C. **A New Framework and Burst Assembly for IP DiffServ over Optical Burst Switching Networks**; In: IEEE GLOBAL TELECOMMUNICATIONS CONFERENCE (GLOBECOM), Proceedings... San Francisco, United States, p. 3159-3164, 2003.
- [56] LAZZEZ, A. et al. **A Dynamic QoS-based Scheme for Admission Control in OBS Networks**. In: IEEE INTERNATIONAL CONFERENCE ON COMMUNICATIONS (ICC). Proceedings... Glasgow, Scotland, p. 449-454, 2007.
- [57] ESTEVES, R. et al. **Qualidade de Serviço Absoluta em Redes OBS Baseadas no GMPLS**. In: 25º SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS. Anais... Belém, Pará, Brasil, p. 577-590, 2007.
- [58] FARIAS, F.; ESTEVES, R.; ABELÉM, A. **Gerenciamento Dinâmico de Qualidade de Serviço em Redes OBS Baseadas no GMPLS**. In: 26º SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS. Anais... Rio de Janeiro, Brasil, p. 523-536, 2008.
- [59] ESTEVES, R.; FARIAS, F.; NUNES, A.M.; ABELÉM, A. **Provisionamento Automático de Conexões Determinísticas para Redes GOBS**. In: 27º SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS. Anais... Recife, Brasil, p. 481-494, 2009.
- [60] QIAO, C. **Labeled optical burst switching for IP-over-WDM integration**. IEEE Communications Magazine, v. 38, n. 12, p. 104-114, set. 2000.
- [61] LI, M.; BAKER, M. **The Grid: Core Technologies**. 1st ed. John Wiley and Sons, 2005.

- [62] LIU, C. et al. **Design and Evaluation of a Resource Selection Framework for Grid Applications.** In: 11th IEEE INTERNATIONAL SYMPOSIUM ON HIGH PERFORMANCE DISTRIBUTED COMPUTING (HPDC). Proceedings... Edinburgh, Scotland, p. 63-72, 2002.
- [63] AWDUCHE, D. et al. Requirements for Traffic Engineering over MPLS. RFC 2702. [S.l.]: Internet Engineering Task Force, Network Working Group, 1999.
- [64] MAGALHÃES, M.; CARDOZO, E. **Introdução à Comutação IP por Rótulos Através de MPLS.** Livro texto dos Minicursos. In: 19º SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, Florianópolis, Santa Catarina, Brasil, cap. 3, 2001.
- [65] NS-2. **Network Simulator.** Disponível em: <<http://www.isi.edu/nsnam/ns/>>. Acesso em: jan. 2010.
- [66] WEN, B. et al. **Optical Wavelength Division Multiplexing (WDM) Network Simulator (OWns): Architecture and Performance Studies.** SPIE Optical Networks Magazine Special Issue on Simulation, CAD, and Measurement of Optical Networks, mar. 2001.
- [67] GAUGER, C. **Viability and Performance of Optical Burst Switching.** In: 9th EUROPEAN CONFERENCE ON NETWORKS AND OPTICAL COMMUNICATIONS (NOC). Proceedings... Eindhoven, The Netherlands, 2004,
- [68] CHEN, Y.; TANG, W.; VERMA, P. K. **Latency in Grid over Optical Burst Switching with Heterogeneous Traffic.** In: HIGH PERFORMANCE COMPUTING AND COMMUNICATIONS (HPCC 2007). Proceedings... Houston, United States, p. 334-345.
- [69] INTEL CORPORATION. **Intel microprocessor export compliance metrics.** Disponível em: <<http://www.intel.com/support/processors/sb/CS-023143.htm>>. Acesso em: dez. 2010.

- [70] PACKETLIGHT NETWORKS. **DWDM over CDWM**. Disponível em: <http://www.packetlight.com/technology/dwdm_over_cwdm/>. Acesso em: mar. 2010.
- [71] FODOR, G.; RÁCZ, S.; TELEK, M. **On providing blocking probability and throughput guarantees in a multi-service environment**. International Journal of Communication Systems, v. 14, n. 4, p. 257-285, jan. 2002.
- [72] WANG, X.; MORIKAWA, H.; AOYAMA, T. **A deflection routing-compatible wavelength assignment algorithm for burst switched photonic networks**. In: IEICE GENERAL CONFERENCE. Proceedings... Japan, p. 439, 2001.
- [73] INFORMATION SCIENCES INSTITUTE (ISI). **NARB and RCE Architecture Version 2.0**. Disponível em < <http://dragon.east.isi.edu/twiki/pub/DRAGON/NARB/narb-rce-architecture-v2.0.pdf> >. Acesso em: jun. 2010.
- [74] FOSTER, I.; KESSELMAN, C. (editors). **The Grid 2: Blueprint for a New Computing Infrastructure**. 2nd ed. Morgan Kaufmann, 2003.
- [75] PAN, Y. et al. **Parallel and Distributed Processing and Applications**. Springer, 2005.