

IP OPTICAL NETWORK CONTROL PLANE
PERFORMANCE ANALYSIS AND DEPLOYMENT

By

QIANG SONG

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of
the requirements for the degree of Doctor of Philosophy.

THE CITY UNIVERSITY OF NEW YORK

2007

UMI Number: 3283623

Copyright 2007 by
Song, Qiang

All rights reserved.

UMI[®]

UMI Microform 3283623

Copyright 2008 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

©2007

QIANG SONG

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Prof. Ibrahim Habib

Date

Chair of Examining Committee

Dean Mumtaz Kassir

Date

Executive Officer

Prof. Leonid Roytman (CUNY, CCNY EE Dept.)

Prof. Truong-Thao Nguyen (CUNY, CCNY EE Dept.)

Prof. Mohamed Zahran (CUNY, CCNY EE Dept.)

Prof. Nirwan Ansari (NJIT, ECE Dept.)

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

2007

Abstract

IP OPTICAL NETWORK CONTROL PLANE PERFORMANCE ANALYSIS AND DEPLOYMENT By

QIANG SONG

Adviser: Professor Ibrahim Habib

The thesis is focused on the design, performance analysis, and deployment of the control plane for future IP optical networks. The control plane is needed to enable cost-effective bandwidth-on-demand (BoD) services via connection-oriented bandwidth guaranteed pipes. Firstly, we analyzed the performance of the Generalized Multi-Protocol Label Switching (GMPLS) control plane and assess its value added proposition vis-à-vis the centralized network management (CNM) plane. Comparative simulation analysis was conducted to understand the limitations of using the control/management planes in enabling future connection-oriented services. We identified operating regions where the benefits of using the control plane to provision BoD services is maximized in terms of minimizing the network resources required to meet the demands of the services. Secondly, we proposed and analyzed a heuristic algorithm to schedule two types of BoD services: book-ahead and immediate request ones subject to quality of service constraints such as minimizing call blocking rate while maximizing network utilizations.

Finally, we designed and deployed a GMPLS software engine for the Circuit-switched High-speed End-to-End Transport Architecture (CHEETAH) test-bed, called the CHEETAH Virtual Label Switching Router (CVLSR). CVLSR provides a cost-efficient dynamic sharing mechanism to enable non-GMPLS switches (e.g., Ethernet switches, routers and other cross-connects) to participate in the dynamic setup and release of end-to-end bandwidth guaranteed connections. In addition, CVLSR is equipped with a novel function called dynamic network clustering which could enable a broad range of applications such as remote visualization, grid computing and e-learning among others.

Acknowledgments

Firstly, I would like to convey my deepest gratitude to my mentor Professor Ibrahim Habib for his guidance, support and academic advice. I sincerely appreciate his detailed mentoring and closed supervision.

Next, I would like to thank the Examining Committee for their help, efforts, and supports with the thesis. I would also like to thank Sprint and Nortel for their funding for the analytical part of the thesis, and NSF for its funding for design and deployment of the control plane for an experimental IP optical network. Special thanks go to colleagues from UVA, ORNL, MCNC, and Dragon team, for their helps in creating the CHEETAH test-bed.

Finally, my deepest appreciation goes to my family, who give me their love and supports all the time.

To my family

Table of Contents

Abstract.....	iv
Acknowledgments.....	vi
Table of Contents.....	viii
List of Tables.....	xii
List of Figures.....	xiii
Abbreviations.....	xvii
Chapter 1 Introduction.....	1
1.1 Motivations.....	1
1.2 Objectives and Key Findings.....	3
1.3 Thesis Structure.....	8
Chapter 2 Control Plane Architecture.....	10
2.1 New Services Demand Advances in Network Control and Management.....	10
2.2 Overview of the Traditional Centralized Network Management....	12
2.3 IP Optical Control Plane.....	15
2.4 Research Methodology.....	17

Chapter 3 The Value added Proposition of GMPLS Control Plane in IP	
Optical Networks	22
3.1 Distributed Dynamic Routing	22
3.1.1 OSPF Convergence Time	22
3.1.2 OSPF Overhead Traffic	23
3.1.3 The Effect of Traffic Load	26
3.1.4 The Effect of Call Inter-Arrival Time.....	27
3.1.5 The Impact of Bandwidth Granularity	30
3.1.6 Cost Advantages of Distributed Dynamic Routing over Static Routing	32
3.1.7 Distributed Versus Centralized Approaches.....	36
3.2 Connection Setup	40
3.2.1 The Impact of Propagation Delay	41
3.2.2 The Impact of RSVP Message Processing and OXC Cross-connect Delay	42
3.3 Protection and Restoration.....	44
3.3.1 Protection and Restoration Schemes.....	44
3.3.2 Protection and Restoration Analysis.....	48
3.3.2.1 Dynamic vs. Static Routing	48
3.3.2.2 M:N Shared Protection Performance Evaluation.....	50
3.3.2.3 Dynamic Restoration (none-protection)	53
3.3.2.4 Pre-emptable Traffic	55
3.3.2.5 Business Driver Analysis.....	57
3.3.3 The impact of Node Degrees on Protection Performance	60
3.3.4 Suggested NSFNET Modification for Protection to Be Efficient	62
3.3.5 Protection and Restoration In Distributed Vs. Centralized Approaches.....	65
Chapter 4 Book-Ahead Reservation for Connection Oriented Services	68

4.1	Book-ahead Reservation Mechanism	69
4.2	Heuristic Algorithm	70
4.3	Comparison between the FASP with SPF and SWP	73
4.4	Comparison between the ILP approach and the heuristic algorithm 75	
4.5	Performance Analysis	77
4.5.1	The optimum total number of slots and slot length	77
4.5.2	The impact of the call granularity	81
4.5.3	Cost advantage of BA over IR	82
4.5.4	Designated start time book-ahead calls	83
4.5.5	Gold and economy classes of BA	84
4.5.6	Variable bandwidth reservation	87
4.5.7	The impact of IR calls on BA calls	89
Chapter 5 Design of Control Plane in CHEETAH Network.....		95
5.1	CHEETAH Network.....	95
5.2	CVLSR DESIGN AND IMPLEMENTATION	98
5.2.1	Parameter configuration.....	99
5.2.2	Routing Module	100
5.2.3	Signaling Module.....	101
5.2.3.1	RSVP Process	102
5.2.3.2	Connection Admission Control.....	102
5.2.3.3	Switch Fabric Control.....	103
5.2.3.4	Network Cluster Management	103
5.3	CVLSR Deployment.....	108

Chapter 6 Conclusions and Future Directions	110
6.1 Summary of Thesis	110
6.2 Conclusion	112
6.3 Future Works	116
References.....	118
Publications.....	120

List of Tables

Table 2.1 Confidence Interval of Call Blocking Ratio	19
Table 3.1 Protection and Restoration Schemes Comparison	45

List of Figures

Fig. 2.1 A Simplified View of the Centralized Network Management Plane.	13
Fig. 2.2 Simplified overview of the control plane approach.....	15
Fig. 2.3 Node architecture.....	17
Fig. 2.4 NSFNET Topology.....	20
Fig. 2.5 16 nodes, 32 links (node degree=4).....	21
Fig. 2.6 16 nodes, 40 links (node degree=5).....	21
Fig. 3.1 Max. OSPF link throughput vs. different topologies.....	25
Fig. 3.2 Distributed dynamic routing vs. static routing at different traffic loads per node.	25
Fig. 3.3 Differential advantage of dynamic vs. static routing at different traffic loads per node.....	27
Fig. 3.4 Distributed dynamic vs static routing with different call inter-arrival times.	28
Fig. 3.5 Differential advantage of dynamic vs static routing with different call inter- arrival times.	28
Fig. 3.6 Call blocking ratio vs. traffic load and call granularity.....	31
Fig. 3.7 Network utilization vs. call granularity.	32
Fig. 3.8 Number of wavelengths per link vs. traffic load	34
Fig. 3.9 Wavelengths saved in distributed dynamic routing vs. static routing	34
Fig. 3.10 Capacity improvement in distributed dynamic routing vs. static routing.....	35
Fig. 3.11 Efficiency gain G vs. centralized database update interval	37

Fig. 3.12 Efficiency gain G approach vs. call inter-arrival time.....	38
Fig. 3.13 The impact of RSVP signaling propagation delay	41
Fig. 3.14 Impact of connection setup delay with other delay parameters considered	42
Fig. 3.15 Performance comparison for protection schemes in dynamic vs. static routing	49
Fig. 3.16 Performance improvement using dynamic routing compared with static routing.	50
Fig. 3.17 The impact of number of wavelengths per link on the performance of protection schemes.....	50
Fig. 3.18 Number of wavelengths needed to keep call blocking rate less than 10% in each scheme.....	51
Fig. 3.19 M:N shared protection performance evaluation	52
Fig. 3.20 Dynamic restoration performance vs. traffic load	54
Fig. 3.21 The impact of call inter-arrival time on dynamic restoration performance.....	54
Fig. 3.22 Low priority traffic blocking ratio vs. traffic load.....	55
Fig. 3.23 Low priority traffic preempted ratio vs. traffic load.....	56
Fig. 3.24 Low priority traffic admitted vs. traffic load.....	57
Fig. 3.25 Percentage of capacity increased in tomorrow vs. today's PL services	58
Fig. 3.26 Percentage of revenue increased in tomorrow vs. today's PL services.....	59
Fig. 3.27 Impact of number of links per node on the protection schemes	60
Fig. 3.28 Total number of wavelengths needed vs. number of links per node	61
Fig. 3.29 NSFNET after modification (red bold links are added links)	62
Fig. 3.30 Performance comparison before and after NSFNET Modification.....	63
Fig. 3.31 Performance improvement vs. number of wavelengths per link.....	64

Fig. 3.32 The impact of central database-update-delay on M:N protection in centralized dynamic routing approach.....	66
Fig. 3.33 The impact of central database-update-delay on dynamic restoration in centralized dynamic routing approach	67
Fig. 4.1 Advance reservation mechanism.....	69
Fig. 4.2 Pseudo C code of FASP	71
Fig. 4.3 Topology used in the ILP approach.....	72
Fig. 4.4 Call blocking vs. number of wavelengths per link.	73
Fig. 4.5 Network utilization vs. number of wavelengths per link.....	74
Fig. 4.6 Percentage of performance improvement of FASP vs. SPF.....	74
Fig. 4.7 Call blocking probability of ILP approach vs. heuristic method.....	76
Fig. 4.8 Network utilization of ILP approach vs. heuristic method.....	76
Fig. 4.9 Blocking probability vs. number of slots used by the scheduler	78
Fig. 4.10 Network utilization vs. number of slots used by the scheduler	79
Fig. 4.11 Blocking probability vs. slot length and the total number of slots	79
Fig. 4.12 Optimum operating region in terms of slot length and total number of slots....	80
Fig. 4.13 Blocking probability vs. call granularity.	81
Fig. 4.14 Percentage of cost savings of BA over IR.	83
Fig. 4.15 Restricted start time scheduling.....	84
Fig. 4.16 Call blocking probability vs. different proportional of service classes	85
Fig. 4.17 Network utilization vs. different proportional of service classes	86
Fig. 4.18 Revenue vs. different proportional of service classes	87
Fig. 4.19 Call blocking vs. max. number of wavelengths can be reserved.....	88

Fig. 4.20 Network utilization vs. max. number of wavelengths can be reserved	89
Fig. 4.21 The blocking probability of BA calls	91
Fig. 4.22 The blocking probability of IR calls.....	91
Fig. 4.23 The network utilization.....	92
Fig. 5.1 CHEETAH network Test-bed	97
Fig. 5.2 The CVLSR software architecture	98
Fig. 5.3 RSVP messages functional flow diagram	101
Fig. 5.4 Dynamic cluster illustration.....	105
Fig. 5.5 CVLSR deployment in CHEETAH network.....	108

Abbreviations

BA	Book-ahead
BoD	Bandwidth-on-demand
CHEETAH	Circuit-switched High-speed End-to-End Transport ArcHitecture
CVLSR	CHEETAH Virtual Label Switch Router
CNM	Centralized Network Management
DRAGON	Dynamic Resource Allocation via GMPLS Optical Networks
DWDM	Dense Wavelength Division Multiplexing
EPL	Enhanced Private Line
FASP	First Available Shortest path
GMPLS	Generalized Multi-Protocol Label Switching
ILR	Integer Linear Programming
IR	Immediate Reservation
LMP	Link Management Protocol
LSDB	link state database
LSP	Label Switching Path
NSFNET	National Science Foundation Network
SPF	Shortest Path First
OSPF	Open Shortest Path First

OSS	Operating Support System
OVPN	Optical Virtual Private Networks
OXC	Optical Cross-connect
SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical Network
RSVP	Reservation Protocol
TE	Traffic Engineering

Chapter 1

Introduction

In this chapter, firstly, the motivations of this research and development work are provided. Next, the objectives of the thesis and key findings are presented, and followed by a description of the thesis structure.

1.1 Motivations

Current optical transport networks, dominated by synchronous optical network/synchronous digital hierarchy (SONET/SDH), have traditionally used centralized network management for connection control. The drawback of using centralized approach is that it requires a complex Operating Support System (OSS), and as the network getting bigger, scalability becomes an issue. In addition, the process of integrating multiple vendors' equipments into a single OSS can be costly and time consuming. It normally takes weeks or months to finish an end-to-end provision, and it is impossible to provide services based on real-time traffic patterns.

As the development of Dense Wavelength Division Multiplexing (DWDM) and Optical Cross Connect (OXC), the concept of optical control plane has been raised to address these issues. The optical control plane is designed to provide faster provisioning, effective traffic engineering, and efficient protection and restoration.

In addition, new services, such as Bandwidth-on-Demand (BoD), Enhanced Private Line (EPL) and Optical Virtual Private Networks (OVPN), can be provided.

It has been a coconscious in recent years to use Generalized Multi-protocol Label Switching (GMPLS) as the control plane protocols [1][2][3][4]. GMPLS is a suite of protocols that includes OSPF-TE for routing [5], RSVP-TE [6] for signaling and LMP [7] for auto-discovery. The current GMPLS protocol is not mature, several organizations, such as IETF, ITU and OIF, are working together to standardize it.

However, adopting a control plane is a huge undertaking for any carrier because of the incurred expenditures. In addition to capital expenditure (CAPEX) of new optical switches, there are other costs incurred in changing the carrier's operations model. Carriers today have separate networking departments: one is for IP routing infrastructure and services, the other is for optical transport and services. Introducing a GMPLS control plane into the transport network requires that these two organizations adopt a converged and unified operation model. The new operation model requires justification via a solid business case such as increasing revenues from new services as well as decreasing operations costs.

Despite the recent developments [8][9][10][11] in the standards and research communities, the efficiency and value added proposition of the IP optical control plane have not been fully analyzed, and further more, our experience with the deployment of GMPLS-based control plane in either production or experimental networks are rather limited.

The above issues not only drew attentions from network operators and network infrastructure vendors, but also from federal government agencies. Significant amount of funding is recently provided for the research and development of the GMPLS network. Taking our case for example, the analytical work of this thesis is funded by Sprint and Nortel to examine the trade-off of the newly proposed distributed IP optical control plane vs. the traditional centralized approach. The GMPLS test-bed development in this thesis, CHEETAH network [12], is funded by the National Science Foundation (NSF) and the Department of Energy (DoE). Other recently developed GMPLS experimental high performance optical networks include DRAGON[13] (Dynamic Resource Allocation via GMPLS Optical Networks), and HOPI (Hybrid Optical Packet Infrastructure) [14]. The main intention of these networks is to foster the development of experimental control plane solutions for connection-oriented service provisioning through the optical-based networks.

1.2 Objectives and Key Findings

The objectives of this thesis are to investigate the efficiency of traffic engineering and the value-added proposition of using GMPLS control plane for the connection-oriented service provisioning in the optical transport network. Our focus is not to propose new protocols or algorithms, but rather to provide guidance to network operators on how to maximize the benefits of deploying a GMPLS control plane in terms of minimizing the call blocking and maximizing the resource utilization.

The questions that will be addressed in the thesis include:

- How well is OSPF-TE fit for the control plane of optical transport network?
- What is the impact of the RSVP-TE signaling delay?
- What is the limitation of the GMPLS control plane in terms of services it can provide?
- Is the distributed control plane approach always better than the centralized management approach in terms of the types of services?
- What is the benefit of using GMPLS control plane in the protection and restoration?
- How well is the scheduling reservation system compared with non-scheduling system, and what is the impact of the immediate reservations on the book-ahead ones?
- How is the complexity of adopting the control plane in the real networks?

Our key findings through NSFNET simulations are:

- Dynamic routing utilizing OSPF-TE is not always efficient compared with static routing for the purpose of realizing better links efficiencies, or decreasing call blocking probabilities. There exists an operating region in which it has better performance over static routing, i.e. limiting traffic loads to less than 0.8 Erlang/per wavelength. In addition, the OSPF convergence delay would have an impact on the type of services the control plane could provide. If the per node average call inter-arrival time is shorter than 1 minute, the efficiency (in terms of call blocking

ratio) of distributed dynamic routing starts to deteriorate due to the OSPF convergence delay.

- The number of wavelengths per link has impact on the performance of dynamic routing. The smaller the number of wavelengths, the more evident that dynamic routing performs better than static routing. As the number of wavelengths per link becomes more than 30, the performance difference between dynamic routing and static routing tends to be a constant. Compared with static routing, dynamic routing can save at least 15% wavelengths per link, or increase the traffic load by at least 20%.
- In BoD services, the level of call granularity plays an important role to achieve high network utilization, i.e. increases the operator's revenues. There is an optimal region for the call granularity to achieve maximum utilization and keep the network scalable. In NSFNET, the optimal call granularity is 1/60 of the total link capacity.
- Distributed network has its advantages in boosting capacity over centralized network mainly in the bandwidth-on-demand (BoD) services, i.e. shorter inter-arrival times and holding times. As the inter-arrival times and holding times become longer, distributed network gradually loses its advantages. For enhanced-private-line (EPL) services, both distributed and centralized approaches have the same performance. Even in the BoD services, only if the OSS database updating delay is

longer than certain value, does distributed approach perform better than centralized approach.

- The connection setup delay is the limiting factor in determining the types of BoD services a carrier offers. The call blocking ratio will increase quickly if the average call inter-arrival time is less than tens of seconds.
- Assuming only single link failure happens, M:N shared protection can achieve 100% recovery at the expense of 10% increase in the number of wavelengths compared to non-protection schemes; while saves up to 48% wavelengths compared to the 1:1 dedicated protection.
- The network works more efficient in terms of total number of wavelengths needed if it has a higher average node degree. Through NSFNET simulations, we show that by increasing the average node degree of NSFNET from 2.50 to 4.06, 36.0% of the total number of wavelengths can be saved in the shared protection scheme and it can be translated to a 56% increase in the return on investment ratio.
- In tomorrow's EPL services which several classes of services are offered, the pricing structure for the newly added services should be justified by the capacity saving of the services so that the total revenue enhancement is positive. In the assumed price structure which M:N private line service is priced 20% lower than 1:1 private service, at least 30% of the traffic load need to be maintained as M:N services (the rest are 1:1 services) to obtain the same revenues as that provided by the

100% 1:1 private line services. The revenue enhancement can be up to 44% and capacity enhancement can be up to 75% if 10% 1:1 and 90% M:N private line service scenario is adopted.

- In order to achieve lowest blocking probability in the advance reservation scheme, the look-ahead time in the advance reservation should be 7 times of the average call holding time. To minimize the number of slot in the reservation system, the slot length should be 3 units time.
- The heuristic path calculation algorithm developed in this thesis can reduce blocking by 36% compared with the traditional SPF algorithm.
- If the initial reservation failed, allowing the recourses to be reserved at two times of the requested bandwidth will maximize the network performance.
- To reduce the impact of the immediate reservation requests to the book-ahead ones, the holding times of the immediate requests should be known to the scheduler.
- To maximize the network utilization, distributed scheduling should be considered and the holding time of the immediate request should be short.
- The GMPLS engine (CVLSR) we developed provide a cost-efficient manner to control the non-GMPLS equipment, for example, the Giga-bit Ethernet switch, to enable more users to participate the dynamic end-to-end bandwidth guaranteed connection setup. By reusing the RSVP

signaling, a novel dynamic clustering function is added to the CVLSR to enable two groups of end-hosts to establish dedicated connections between the local and remote sites. The clustering function provides support for new services, such as e-learning, grid computing and videoconference etc.

1.3 Thesis Structure

The rest of the thesis is organized as follows. Chapter 2 provides the overviews of both the traditional centralized OSS network management approach and the newly developed distributed GMPLS control plane approach. The main components and features of the GMPLS control plane are illustrated. Next, the research methodology used to analyze the performance of the GMPLS control plane is presented.

In Chapter 3, the benefits and optimum operating regions of the control plane are identified for the connection-oriented service provisioning in terms of routing, signaling, protection and restoration. The value added propositions of the distributed control plane vis-à-vis the centralized network management plane are analyzed and quantified.

Chapter 4 proposes and analyzes a heuristic algorithm to schedule two types of BoD services: book-ahead and immediate request ones. Operating regions are identified where the benefits of using the control plane to provision BoD services is maximized in terms of minimizing the network resources required to meet the demands of the services. The impacts of immediate reservations on the book-ahead reservations are analyzed.

Chapter 5 provides the design and deployment of a GMPLS engine (CVLSR) for the CHEETAH test-bed control plane. The significance of the CVLSR is discussed. The concept of the network cluster which could provide video-conferencing, e-learning and grid-computing is also presented in this chapter.

The thesis ends with Chapter 6, which presents a summary of the thesis, our major contributions and the future work.

Chapter 2

Control Plane Architecture

In this chapter, firstly the new services and their requirements related to the advanced network control and management are presented. Secondly, service-provisioning using the traditional centralized OSS network management is overviewed and its limitations are outlined. Thirdly, to address these limitations, the distributed GMPLS control plane approach is recommended and the major building blocks of the control plane are illustrated. Finally, the research methodology to analyze the control plane performance and its value-added proposition is introduced.

2.1 New Services Demand Advances in Network Control and Management

Recently there is a continues growth in the demand for end-user initiated, bandwidth guaranteed and dedicated connection services, such as e-Science applications, grid computing, IP TV, high-definition video conferences and video games, etc.. The above applications can be classified into two main classes: a) those requiring high capacity pipes (e.g., multi Gbps) for file transfers of large datasets, and b) those requiring streaming of low bandwidth pipes (e.g., few Mbps) for visualization and steering purposes. Both classes may be further sub-divided

according to short-duration or long-duration applications. In addition, applications may require immediate or scheduled services. In all cases, these emerging applications have specific challenging networking requirements that can be summarized as follows.

1) Most important is cost-effectiveness. Demands for large or even small capacity pipes will not be substantial unless they are economically attractive to users. Offering users quality guarantees at affordable prices, implies that network providers maximize the productivity of the network assets without sacrificing quality. This requires new modes of operations, in which end-to-end communications paths are not statically provisioned. Large capacity pipes have been traditionally statically provisioned which is a very costly solution. Statistical sharing of these paths among many users will significantly cut the cost down. Obviously, there will be a chance of blocking. However, the network resources (interfaces, ports, wavelengths, and time-slots) can be appropriately sized to maintain an acceptable rate of blocking. This circuit switching sharing mode is similar to what we have today in telephone networks.

2) Users now require high degree of flexibility for their communications needs. The applications require variable amounts of bandwidths pipes and for varying durations. Except for requests scheduled for a future time, there is a need for “application-driven” provisioning of communications paths. This requires equipping the end-user devices with capabilities for signaling such requests to the next hop without consulting a centralized network management system.

3) Inter-domain and inter-networking dynamic provisioning is increasingly important. Supercomputers, instrumentation tools, and other storage facilities form grids that cross networks demarcations. Furthermore, these networks are heterogeneous in the sense that networks' elements have different switching and multiplexing capabilities such as wavelength capable, time-slot capable, packet capable, or layer 2 switching capable. Hence, provisioning of heterogeneous paths across multiple networks is an important requirement.

4) There is a need to schedule short-term and long-term, as well as low capacity and large capacity connections. This requirement will ensure that the networking resources are efficiently utilized in response to the applications' demands. Scheduling is needed not just to allocate and reserve resources, but also to trigger dynamic reconfiguration of the network resources (e.g., time-slots at varying granularities) to meet these dynamic demands.

5) Finally, dynamic provisioning of large bandwidth pipes also require distributed fault management mechanisms, for reliability purposes, capable of localizing faults and providing fast restorations, even across multiple networks' domains.

Clearly, the above requirements draw a fundamentally new picture of networking that is quite different from what is happening in the telecommunications industry today.

2.2 Overview of the Traditional Centralized Network Management

Today, users interact with networks in a client-server model. Users provide the carrier with requirements for "static networking demands". The carrier, then, uses a

centralized network management (CNM) system to provision static paths to fulfill the customer request. This is done via an IP-VPN platform, ATM permanent virtual circuits (PVC), or point-to-point leased lines.

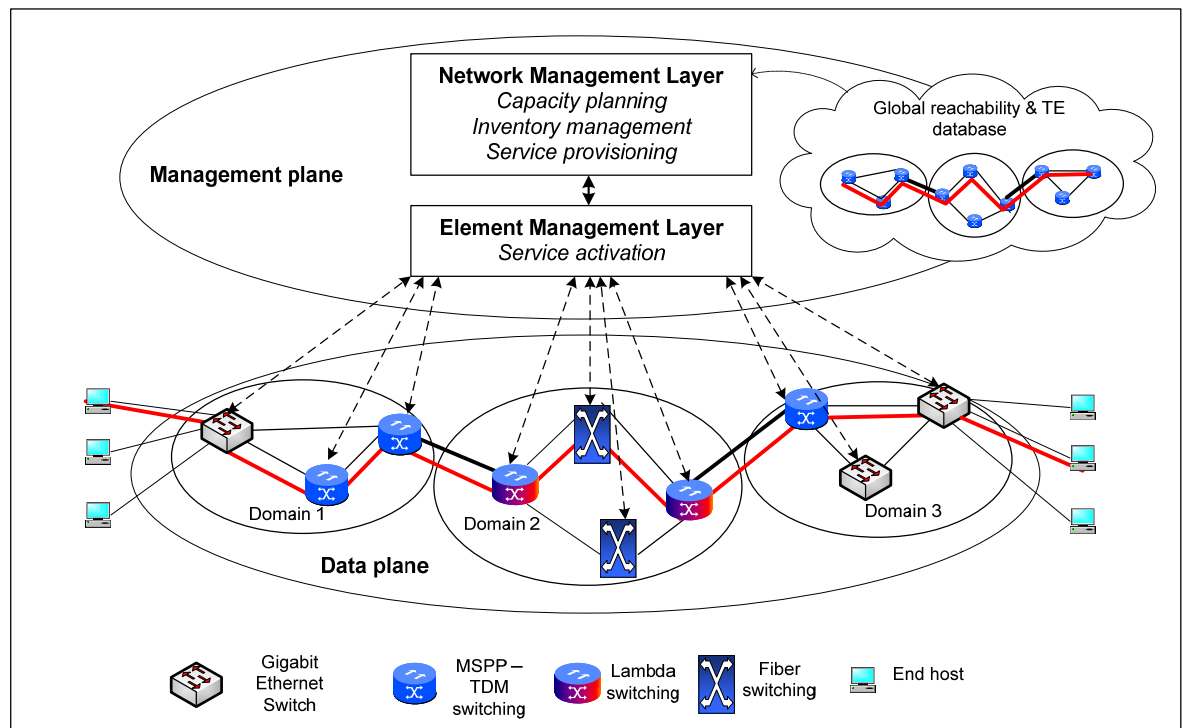


Fig. 2.1 A Simplified View of the Centralized Network Management Plane.

Fig. 2.1 shows an outline of the management plane often deployed in the CNM system. For simplicity, we have omitted many functions such as the service management layer, and other network management processes, for details see [15]. What is important to outline here, is that the element management layer is the layer responsible for activating (i.e., cross-connecting) the actual network element (switch, multiplexer, circuit within a switch) based upon requests from the upper network management layer. It relays to the upper layer information about each network element and its status of operation such as its configuration, connectivity,

alarms, and other physical layer measurements for performance and capacity management. The element management layer also gathers network topology information and sends it to the network management layer where it is used for functions like network planning and provisioning of services. It is also used to update important databases such as the network inventory database where the network assets are maintained and cross-correlated with services they support. In many cases, this process is not fully automated, thus leading to errors in determining which network resources are available for provisioning new services and which ones are already committed.

The service activation and provisioning process is the one responsible for actual configuration of the network resources (i.e., selection of routes, ports in switches, generating activation message). Upon receiving a connection request from end-host, the network management module calculates the end-to-end path, generates the activation signal, and then sends it to each network element through the element management layer.

The provisioning process is quite complex and often requires manual intervention. Clearly, the centralized approach is not scalable for “Bandwidth-on-Demand” (BoD) services where requests for service are frequent and require near real time provisioning. These services require immediate scheduling similar to how telephone calls.

2.3 IP Optical Control Plane

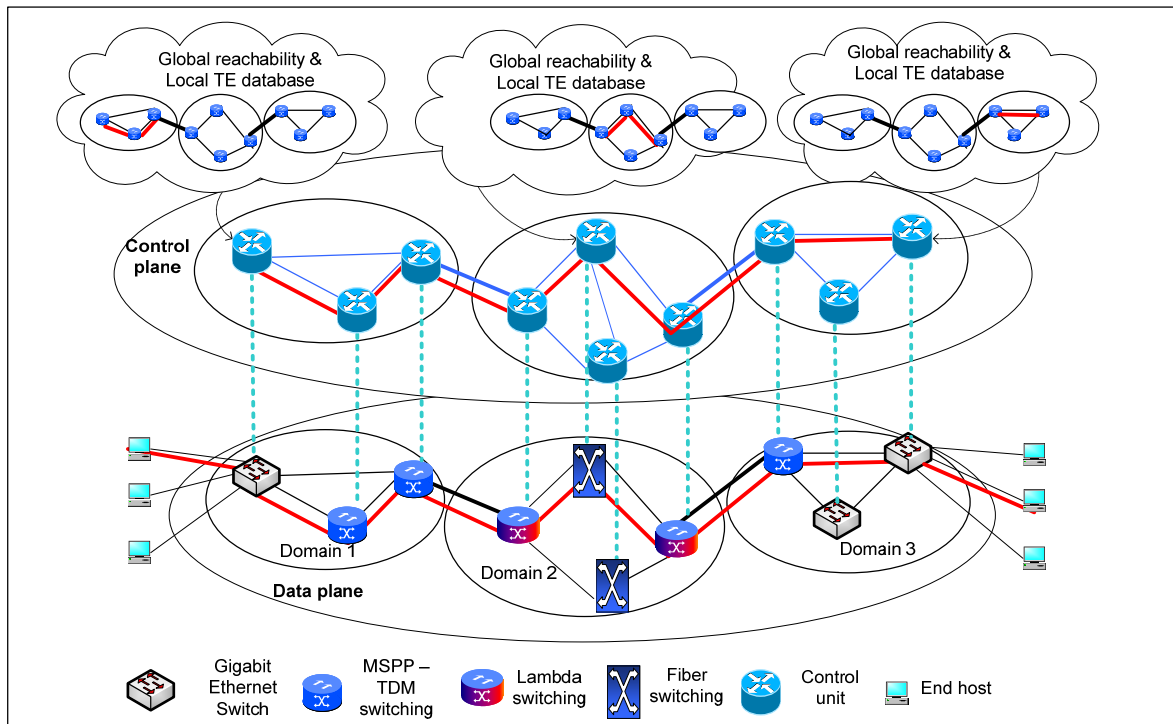


Fig. 2.2 Simplified overview of the control plane approach

The GMPLS control plane extends the Multi-Protocol Label Switching (MPLS) architecture [16] to encompass time-division (e.g., SONET/SDH, PDH, G.709), wavelength (lambdas), and spatial switching (e.g., incoming port or fiber to outgoing port or fiber). The GMPLS control plane (shown in Fig. 2.2) is based upon three main pinnacles: routing, signaling and auto-discovery [8]. Each network element is controlled by GMPLS control unit that has OSPF-TE for dynamic routing, RSVP-TE for signaling, and LMP module for auto-discovery of links, and ports, and control plane/data plane correlation. The new capabilities enabled by the GMPLS control plane, such as distributed dynamic routing, connection control, shared protection, rapid provisioning, auto-discovery, and distributed fault

management enabled, etc., overcome the limitations of the traditional centralized management approach, and make real-time provisioning not only feasible but also cost-effective.

Through distributed OSPF-TE dynamic routing, each GMPLS control unit within every switch builds a global topology view (reach-ability and TE information) of the network within its specific routing area, path computation can be done via source routing CSPF (constraint shortest path first) and a strict explicit route can be computed to the routing-area boundary. Inter-domain routing does not include TE information (for scalability reasons), hence, only reach-ability information is available. In this case, loose routing, or hop-by-hop routing is used to find the next hop between different domains. Here, we define a domain in the same context as in [17], where a domain is a collection of network elements with the same address space or path computation responsibility.

The path or next hop information then can be communicated to the signaling module. Through RSVP messages, the capacity of each link along the path is verified to see if the requested bandwidth and switching capability can be satisfied. If yes, the control unit will send activating messages to the switched to reserve the required bandwidth. After all the reservations are done on the links along the path, the connection is established, and the data can be sent through the path.

In the DWDM system, the number of parallel data links in a physical link can be huge. A physical cable link may consist of tens of fibers and a fiber may consist of hundreds of wavelengths. For TDM system, like SONET, a data transmission link then can be one of the time slots in a single wavelength. In order to manage

these parallel data links, Link management Protocol is created to maintain control channel connectivity, verify the physical connectivity of the data links, correlate the link property information, suppress downstream alarms, and localize link failures for protection/restoration purposes in multiple kinds of networks [7].

2.4 Research Methodology

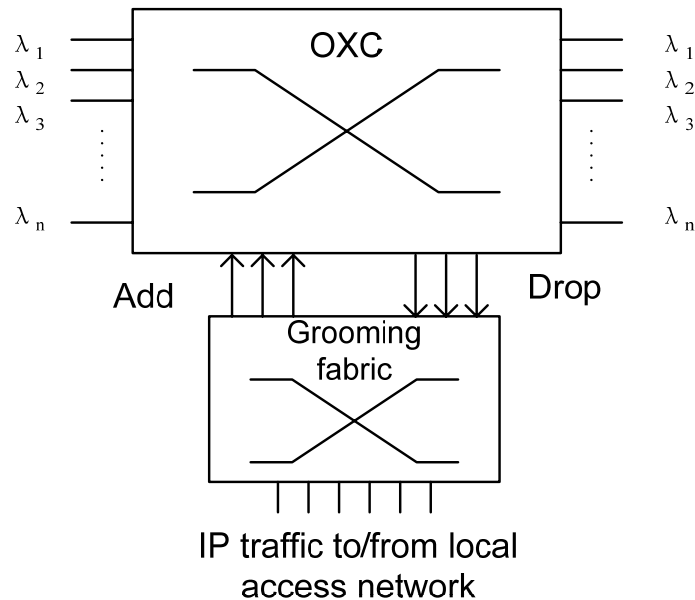


Fig. 2.3 Node architecture

Fig. 2.3 illustrated an OXC architecture, which is composed of optical cross-connect (OXC) and traffic grooming fabric. The grooming fabric grooms the traffic from the local access network and adds them to the OXC for switching. The network simulation tool OPNET is used to build an OXC node simulator. For simplicity, the OXC node simulator is assumed to have wavelength conversion capabilities and only the groomed traffic in the backbone network is simulated. For the reason that most of the applications, such as large file transfers and movie downloads are not bi-directional, only unidirectional label switching paths (LSPs)

are carried out in our simulations. Three functions are simulated in the OXC node: GMPLS engine, call generator, and data collector.

The GMPLS engine has two sub-modules: OSPF-TE for distributed routing and RSVP-TE for signaling. The OSPF-TE sub-module supports OSPF Version 2 (RFC 2328 [18]), The OSPF Opaque LSA Option specification (RFC 2370 [19]), and TE Extensions to OSPF Version 2 (RFC 3630 [20]). The RSVP-TE sub-module supports RSVP – Version 1 Functional Specification (RFC 2205 [21]), RSVP – Version 1 Message Processing Rules (RFC 2209[22]), and the Use of RSVP with IETF Integrated Services (RFC 2210[23]). The LMP is not explicitly simulated, but its functionalities are assumed implicitly in the simulations.

The call generator simulates call requests from a node to other randomly selected nodes in the network. The call request parameters, such as call holding time, inter-arrival time and requested bandwidth, are adjustable to simulate BoD and EPL services' requests. The calls' holding and inter-arrival times are both set to follow exponential distributions. In the rest of the thesis, if not specified parameters such as traffic load, call holding and inter-arrival times, always represent traffic from each node.

The data collector reports two network performance indicators, which are frequently used in this thesis for the connection-oriented network performance analysis. One is the call blocking probability P_b ; the other is the network utilization U . The definitions of P_b and U are:

$$P_b = N_{blocked} / N_{total} \quad (1)$$

$N_{blocked}$: total number of blocked calls in the network

N_{total} : total number of calls generated in the network

$$U = \frac{\sum_{i=1}^{N_{total}} \sum_{j=1}^K H_i L_j}{KWT} \quad (2)$$

H_i : the i th call holding time

L_j : the j th link occupying indicator $L_j = \begin{cases} 1 & \text{if } j\text{th link is used by the } i\text{th call;} \\ 0 & \text{otherwise;} \end{cases}$

K : total number of links

W : number of wavelengths per link, each connection requests one wavelength

T : total simulation time

Sufficient number of calls is generated to ensure that P_b and U reach their stable values before measurements are collected. Twenty simulations with different seeds have been used to get an accurate value. Table 2.1 lists the confidence intervals of the call blocking ratios.

Table 2.1 Confidence Interval of Call Blocking Ratio

Call block ratio (mean)	0.001	0.01	0.1
Standard deviation	0.0002	0.0004	0.002
95% confidence interval	0.001 +/- 0.0001	0.01 +/- 0.0002	0.1 +/- 0.001

In the thesis, we use the term distributed dynamic routing to mean that each node keeps a global view of the network topology within its routing domain. Traffic engineering (TE) database is updated as the link bandwidth changes. Path computation via Constraint Shortest Path First (CSPF) is done in each node to calculate the end-to-end path upon receiving a call request. There is no centralized network element for path computation or TE algorithm. To evaluate the efficiency of distributed dynamic routing, static (hop-by-hop) routing is introduced for

comparison. By static routing, we mean that no bandwidth information is updated; hence, each node maintains a static routing table to provide fixed next-hop information.

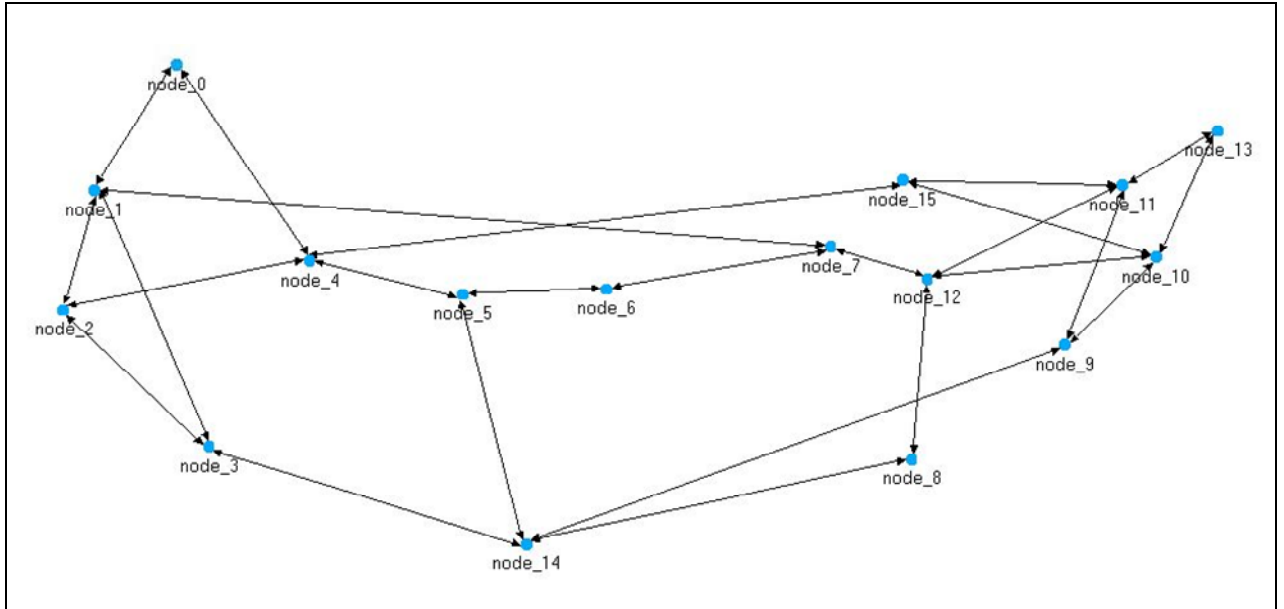


Fig. 2.4 NSFNET Topology

Three network topologies have been studied in our simulations. In topology 1, we used the NSFNET (Fig. 2.4) to obtain results based upon a real network topology. NSFNET has 16 nodes and 25 links with an average node degree of 3. Topology 2 (Fig. 2.5) and 3 (Fig. 2.6) are randomly generated topologies used to study a network with various node degrees. For comparison purposes, both of them have the same number of nodes as that of NSFNET. Topology 2 has 32 links with a node degree of 4, and Topology 3 has 40 links with a node degree of 5.

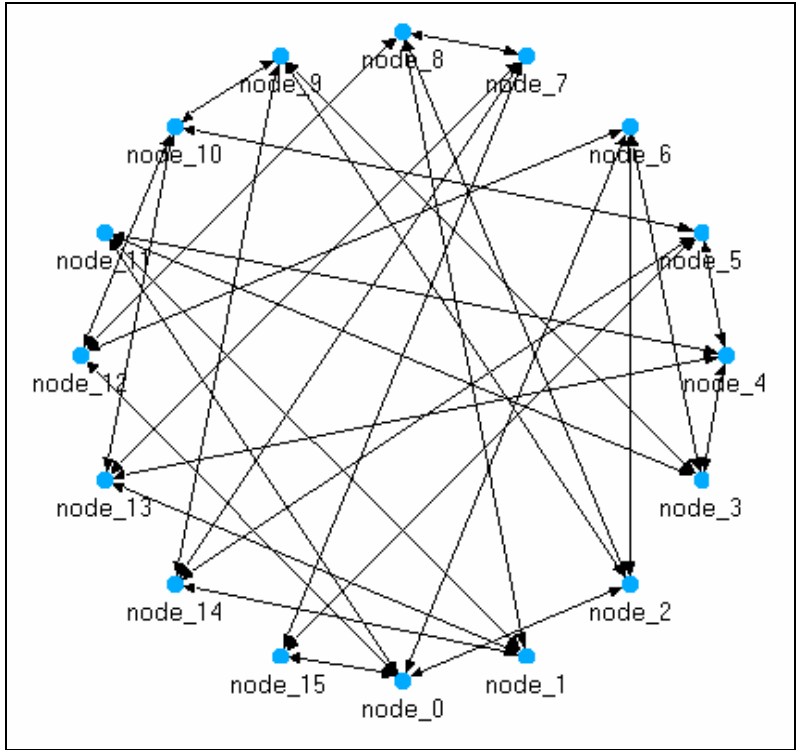


Fig. 2.5 16 nodes, 32 links (node degree=4)

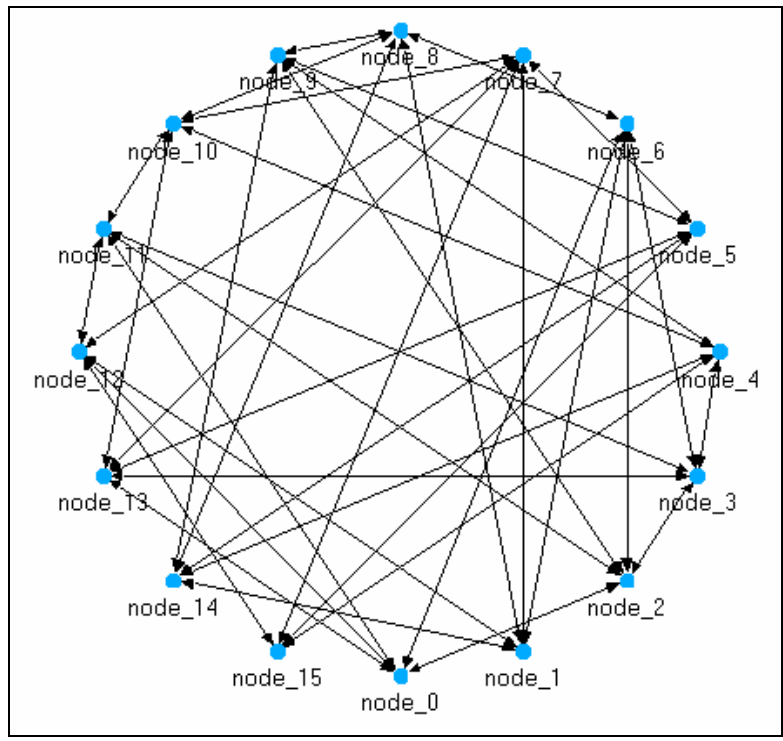


Fig. 2.6 16 nodes, 40 links (node degree=5)

Chapter 3

The Value added Proposition of GMPLS Control Plane in IP Optical Networks

In this chapter,* we quantify the benefit of the GMPLS distributed control plane and compare it with the centralized network management plane. Control plane functionalities such as dynamic routing, user-initiated end-to-end connection setup, and protection and restoration in IP optical networks are analyzed, and the optimal operating regions to maximize their values are identified.

3.1 Distributed Dynamic Routing

In the following subsections, we focus our analysis first on the efficiency of distributed dynamic routing in the GMPLS control plane in terms of convergence time, overhead traffic, traffic load, call inter-arrival time and call bandwidth granularity; and then discuss the cost advantage of distributed dynamic routing in the optical transport network. Finally, we compare distributed dynamic routing with centralized dynamic routing.

3.1.1 OSPF Convergence Time

The OSPF convergence time is measured from the time a topology change occurs until the link state databases of the affected nodes have been updated and

* The work presented in this chapter was funded by Sprint and Nortel.

synchronized with this new information. This, of course, includes the time needed to advertise topology updates.

The measured OSPF convergence time in NSFNET is around 10 to 15 seconds. Such a long convergence delay is certainly not favorable for dynamic routing to be effective since path computation would be done based on out-dated information. Theoretical limits for link-state routing protocols should be in the order of link propagation time scales, (i.e., tens of milliseconds). Some methods have been introduced to reduce the convergence time [24], such as giving higher priority to LSA propagation over SPF computation, and queuing hello packets in front of data packets among others, however, the reliability of these modifications has not been tested. In addition to these methods, other means of reducing the convergence time, in optical transport networks, is of course to limit the routing area domain to few nodes (in the order tens of nodes). The point to observe here is that in spite of the relatively long convergence times that we have in our simulations, we shall see in our results, that there are advantages to deploying dynamic routing for bandwidth on demand services. For private line services, however, we shall see that the advantages are not as significant as in bandwidth on demand.

3.1.2 OSPF Overhead Traffic

The following simulations are used to quantify the control channel OSPF-TE overhead traffic in networks with different topologies in order to understand the impact of scalability. The number of nodes in the network has been varied from 21 to 200, and the number of links connected to each node has been varied from 2 to

20. The nodes are randomly connected to form a mesh network, which is similar to network show in Fig. 2.6. Fig. 3.1 shows the maximum OSPF link throughput. We can see that the OSPF traffic increases almost exponentially as the number of links increases in the network. When we consider both TDM and DWDM scenarios, one link in our simulation may contain thousands of parallel component links (e.g., wavelengths, or time-slots). If these component links are not bundled together, OSPF will have scalability problems. One major function of LMP is to bundle these parallel component links into one or several TE-Links. The results clearly demonstrate the importance of such functionality. Fig. 3.1 demonstrates that the number of nodes in a single OSPF routing area should be limited to tens of nodes and that the number of links (could be TE LMP bundled links) should be within similar bounds. These findings will be further validated through simulations in following sections. We would like to emphasize that the figures tell us that although we can tolerate the OSPF traffic overhead, we need to limit the number of nodes at least in order to minimize the OSPF convergence time which is the main performance bottleneck parameter.

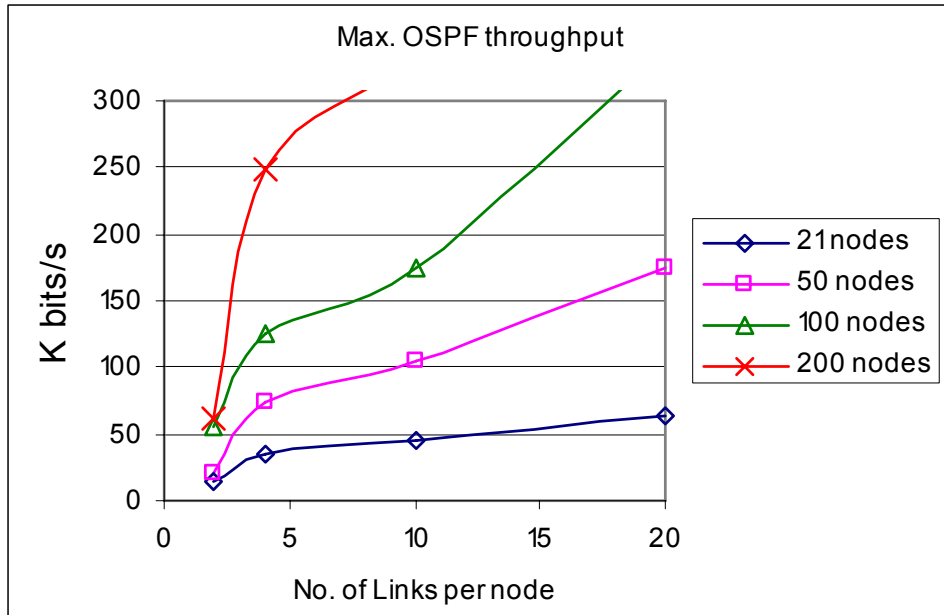


Fig. 3.1 Max. OSPF link throughput vs. different topologies.

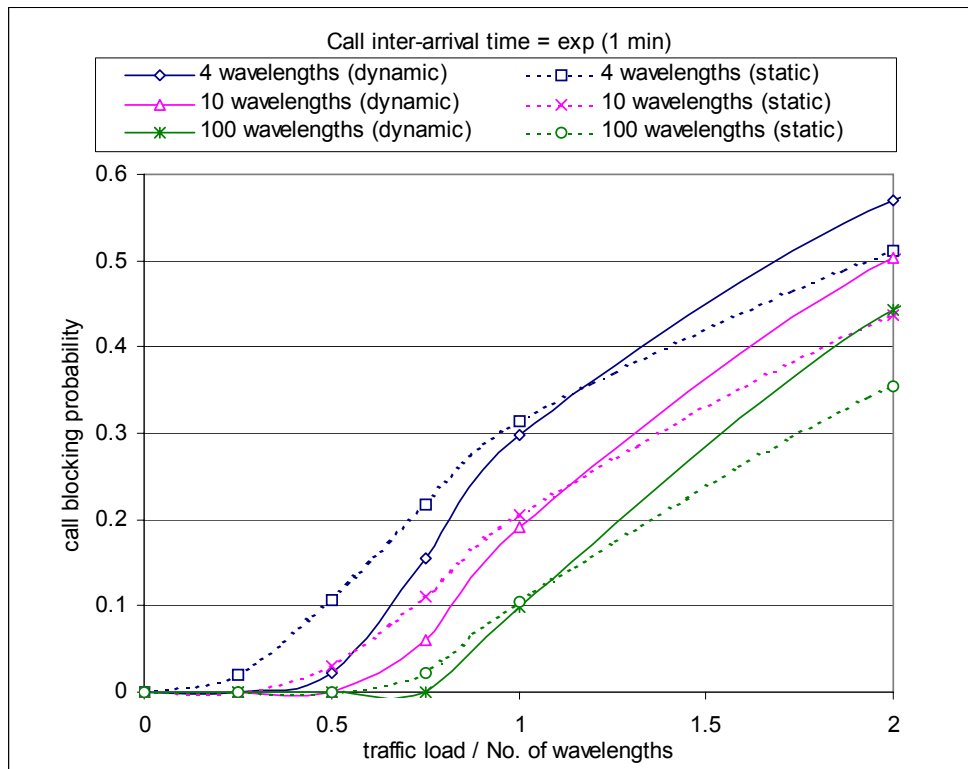


Fig. 3.2 Distributed dynamic routing vs. static routing at different traffic loads per node.

3.1.3 The Effect of Traffic Load

Two parameters are used to depict the traffic characteristics at each node: call inter-arrival time and call holding time where:

$$L = H/I \quad (3)$$

L: Traffic load (Erlang at each node)

H: Average call holding time

I: Average call inter-arrival time

Fig. 3.2 shows how the call blocking probability varies with traffic loads under both distributed dynamic and static routing. The traffic inter-arrival time is exponential distributed with mean values of 1 minute. We can see that distributed dynamic routing loses its efficiency when traffic loads increase above a certain value. For 4 pairs of wavelengths per link, this value is around 4 Erlang; whereas for 10 pairs of wavelengths per link, this value is around 10 Erlang; the curve of 100 pairs of wavelengths shows the same phenomenon. It is about 1 Erlang of traffic per wavelength, i.e., the wavelength is always occupied by the traffic. Obviously, the point here is that distributed dynamic routing will not provide any advantages when the network is over-loaded; hence, we need to operate the links at values of less than 0.8 Erlang per wavelength.

Fig. 3.3 is another view of Fig. 3.2. In order to easily compare the performance of dynamic versus static routing. From these figures, we can see that the best performance of distributed dynamic routing occurs when traffic loads are less than 0.8 Erlang/per wavelength.

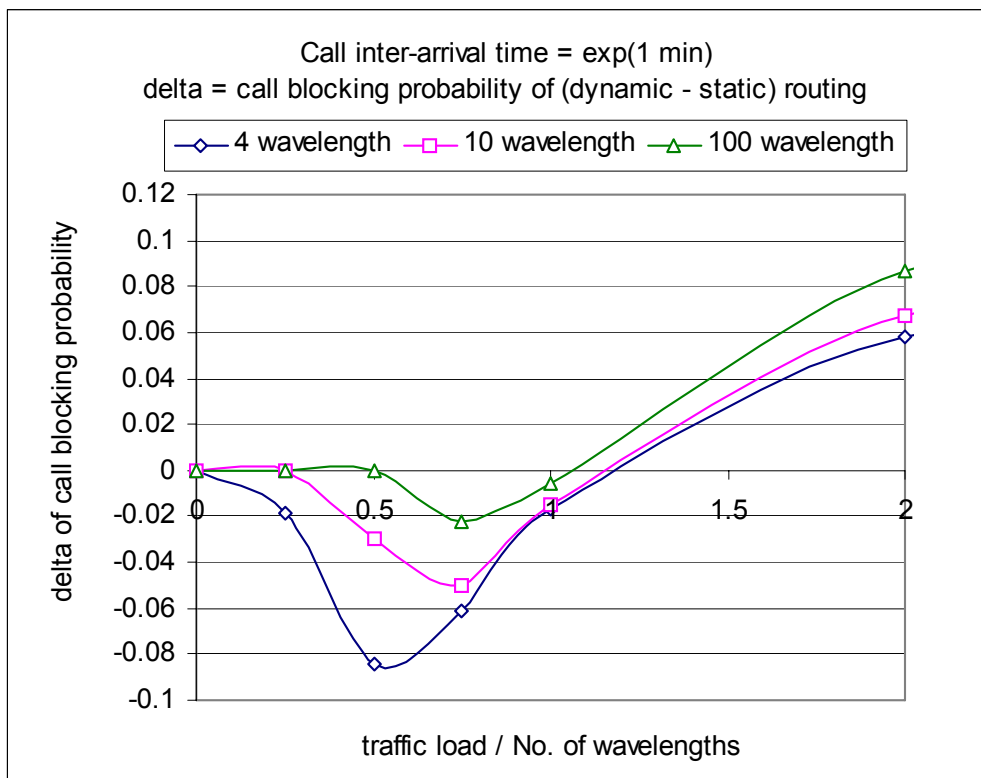


Fig. 3.3 Differential advantage of dynamic vs. static routing at different traffic loads per node.

3.1.4 The Effect of Call Inter-Arrival Time

Fig. 3.4 shows the call blocking probability with the same traffic load (2 Erlang for 4 pairs of wavelengths), but with varying inter-arrival times. We can see that when the inter-arrival time is reduced down into the milliseconds range, distributed dynamic routing does not have any advantages. This is readily understood since when the inter-arrival time is in the milliseconds range, the OSPF link state databases are, almost, always outdated due to the delay of both the OSPF convergence time, as well as the delay due to RSVP signaling. Therefore, for such types of services (BoD) with extremely short inter-arrival times, we need to use extremely fast signaling (preferably hardware implementation) as well as other

measures to decrease the OSPF convergence time (decreasing the amount of information to be flooded, small number of nodes per routing area, etc.) or keep the link state database updated as quickly as possible.

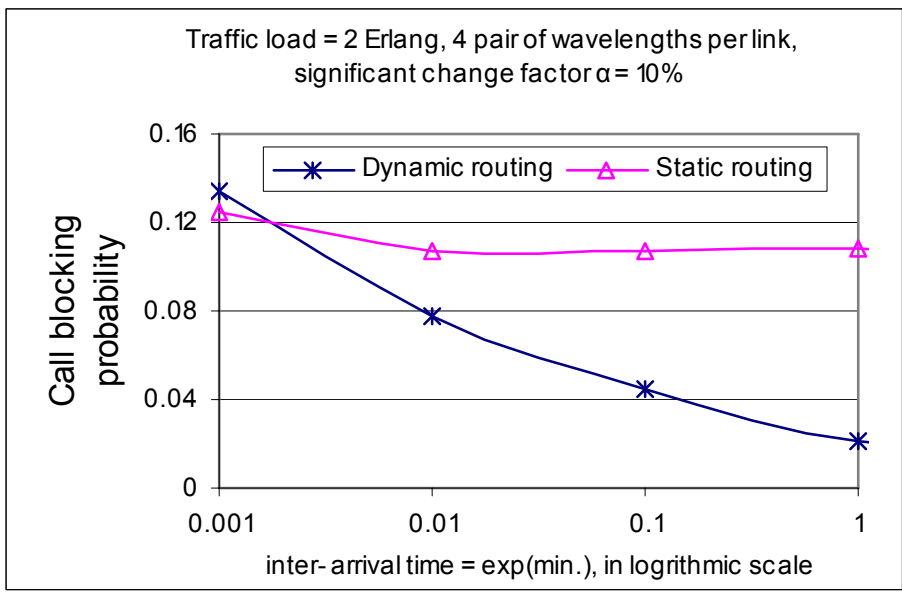


Fig. 3.4 Distributed dynamic vs static routing with different call inter-arrival times.

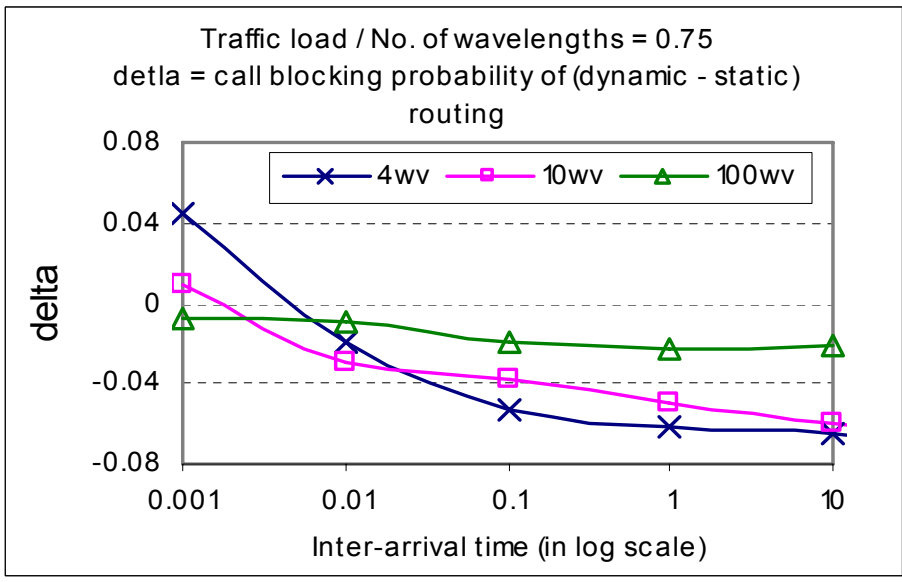


Fig. 3.5 Differential advantage of dynamic vs static routing with different call inter-arrival times.

Fig. 3.5 shows the same trends of Fig. 3.4 but for Traffic Load/wavelengths=0.75, which is a value favorable for distributed dynamic routing as previously explained. Here, the Y-axis shows the parameter delta, which is the difference between computing the call blocking using distributed dynamic routing and static routing. Positive values of delta imply that the call blocking probability is higher for distributed dynamic routing and vice versa. Clearly, there is a range for inter-arrival times where distributed dynamic routing provides maximum advantages over distributed static routing. This range depends upon the traffic load to wavelength ratio (less than 0.8). From the figure, the inter-arrival time has to be above the milliseconds range for wavelengths that are above 10 per link. Interestingly, as the inter-arrival time increases above the one-minute range, there is no additional gain achieved by using distributed dynamic routing. The gain curve stays flat, indicating that although there is a gain for using distributed dynamic routing, this gain saturates for large values of inter-arrival times. A typical characteristic of private lines in comparison to BoD services is that the ratio of their inter-arrival times is extremely large. Hence, this curve tells us that for private line services, the gain of using dynamic over static tends to saturate.

It is interesting to note that in Fig. 3.5, the 100 wavelength curve (around 10 minutes inter-arrival time) can be construed as an indication of private line services, whereas the 4 wavelength (around 0.1 up to 1 minute) is an indication of the BoD service. This is because for the 100 wavelength, and for a traffic load/wavelength ratio of 0.75, the load is 75 Erlang, which means that the holding time is 750 minutes for an inter-arrival time of 10 minutes. Clearly, those values

are indications of long-term pipes. If we examine the 4 wave-length curve at an inter-arrival time of one minute, then we obtain a holding time of 3 minutes. Clearly those values are an indicative of BoD services where both the holding times and inter-arrival times are much less than PLs. The curve reinforces our conclusions that the gain achieved by deploying dynamic bandwidth for BoD services is significantly more than that for PL services.

3.1.5 The Impact of Bandwidth Granularity

Fig. 3.6 depicts the call blocking ratio B verses the traffic load A and the call granularity N_d according to Erlang B formula [25] in a single link point-to-point network.

$$B(N_d, A) = \frac{A^{N_d} / N_d!}{\sum_{k=0}^{N_d} A^k / k!} \quad (4)$$

The requested bandwidth from each call equals to $1/N_d$ of the whole link capacity. Clearly, traffic with finer call granularities leads to lower blocking ratio than that with coarser granularities. Additionally, higher utilization can be achieved for finer granularity calls at a certain call blocking rate.

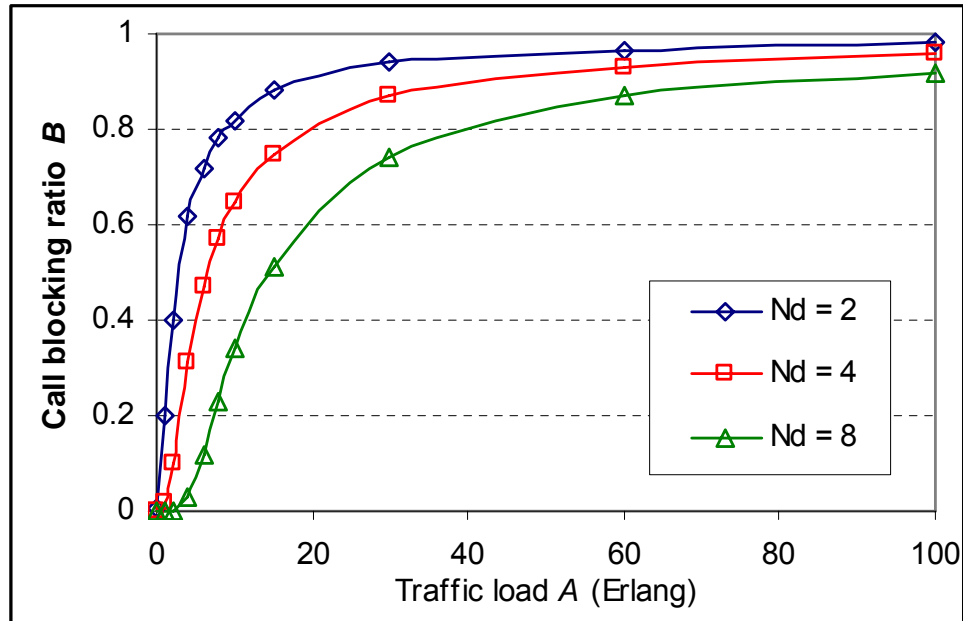


Fig. 3.6 Call blocking ratio vs. traffic load and call granularity.

In our simulations, to study the impact of call granularity, a link is divided into different number of equal capacity channels. Each call requests the bandwidth of one channel. Thus, a channel could represent a wavelength in the lambda switching capable network or a time slot in the TDM switching capable network. Fig. 3.7 shows that traffic with finer granularity calls can achieve higher network utilization than traffic with coarser granularity calls. The network utilization tends to saturate as the number of channels in a link reaches 60. Further increase in the number of channels will not result in a significant gain in the network utilization. This implies that the optimal granularity for each call request should be in the OC-3 level for a network with OC-192 links or the number of wavelengths should be around 60 in a WDM network in which the call granularity is in per wavelength level.

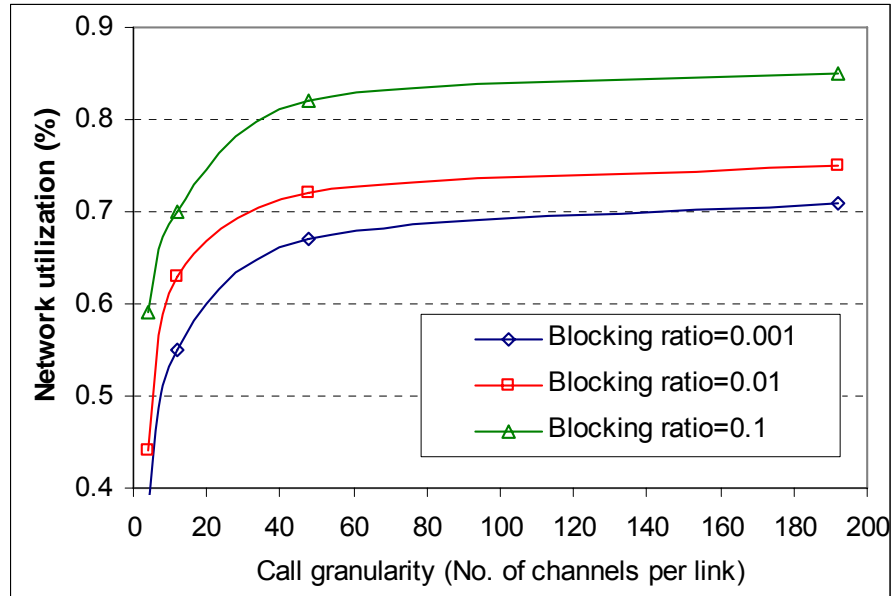


Fig. 3.7 Network utilization vs. call granularity.

3.1.6 Cost Advantages of Distributed Dynamic Routing over Static Routing

In these simulations, the traffic load per node is varied from 2 to 200 Erlangs, and the call blocking ratio is maintained at 0.01 to obtain the number of wavelengths needed to support the traffic load. The call inter-arrival time at each node is set to 10 minutes so that the OSPF convergence delay does not impact the call blocking ratio. The call holding time is varied from 20 to 2000 minutes to generate 2 to 200 Erlangs traffic load. The above parameters are chosen to reflect a special type of BoD service called Provisioned Bandwidth (PB) where connections last for few minutes to few hours (e.g., a broadcast event or large data transfers encountered in eScience applications). This is in contrast to BoD services where calls' inter-arrival times are in the seconds or fractions time range and holding times may be as small as fractions of seconds (e.g., fast file transfers) similar to

telephony calls of today, see [26] for a detailed analysis of the impact of OSPF convergence delay on the call blocking ratio.

Two parameters are introduced to quantify the advantages of distributed dynamic routing over static routing. One is the cost saving factor S which is the percentage of wavelengths saved to accommodate a certain amount of traffic load.

$$S (\%) = (N_s - N_d) / N_s * 100\% \quad (5)$$

Where N_s is the number of wavelengths needed in static routing, N_d is the number of wavelengths needed in dynamic routing.

The other parameter is the capacity improvement factor C which is the percentage of increase in traffic load accommodated by certain number of wavelengths.

$$C (\%) = (L_d - L_s) / L_s * 100\% \quad (6)$$

Where L_s is the traffic load accommodated in static routing, L_d is the traffic load accommodated in dynamic routing. Note that in the above definitions, one may replace the number of wavelengths by the number of time-slots (e.g., such as in SONET TDM links).

Fig. 3.8 shows that the number of wavelengths increases linearly as the traffic load increases for both dynamic and static routing. At a fixed amount of traffic load, the network with higher node degree (topology 3) requires fewer wavelengths than that with lower node degree (topology 1 and 2). The advantages of dynamic routing compared to static, in terms of number of wavelength per link, become more evident with the increase in traffic load. However, it is clear that increasing

the node degree is a much more effective measure in terms of the cost-effectiveness of the network assets required to accommodate a certain traffic load.

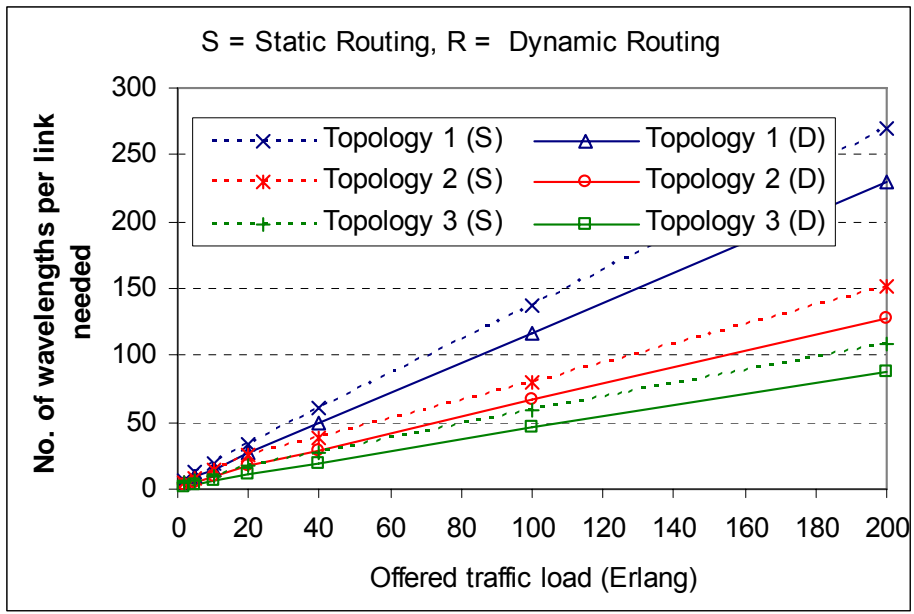


Fig. 3.8 Number of wavelengths per link vs. traffic load

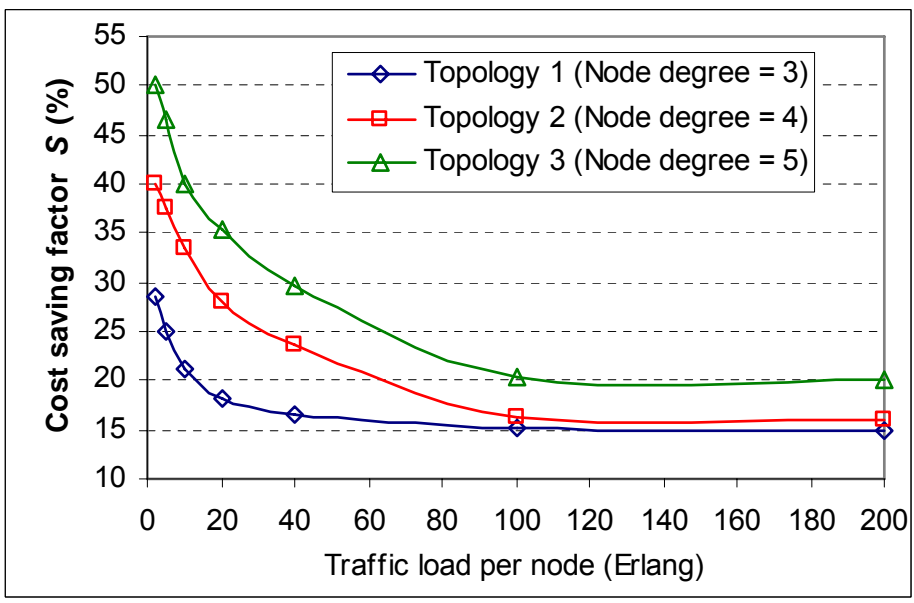


Fig. 3.9 Wavelengths saved in distributed dynamic routing vs. static routing

Fig. 3.9 depicts the cost saving factor S of distributed dynamic routing over static routing. This percentage indicates the cost savings to carriers. The network

with higher node degrees exhibits higher savings than those with lower node degrees. For dynamic routing, a network with higher node degree means there are more alternative paths that could be chosen for routing the traffic. In the same network, as the traffic load increases, the cost saving factor S drops. This is because the number of wavelengths also increases as the traffic load increases, and static routing tends to have more choices in the wavelength selection at each hop, therefore its blocking ratio is closer to that of dynamic routing. As the traffic load increases above a certain value (about 30, 90 and 110 Erlang for networks with node degrees being 3, 4 and 5 respectively), the percentage of savings tends to saturate (around 15%, 16% and 20%).

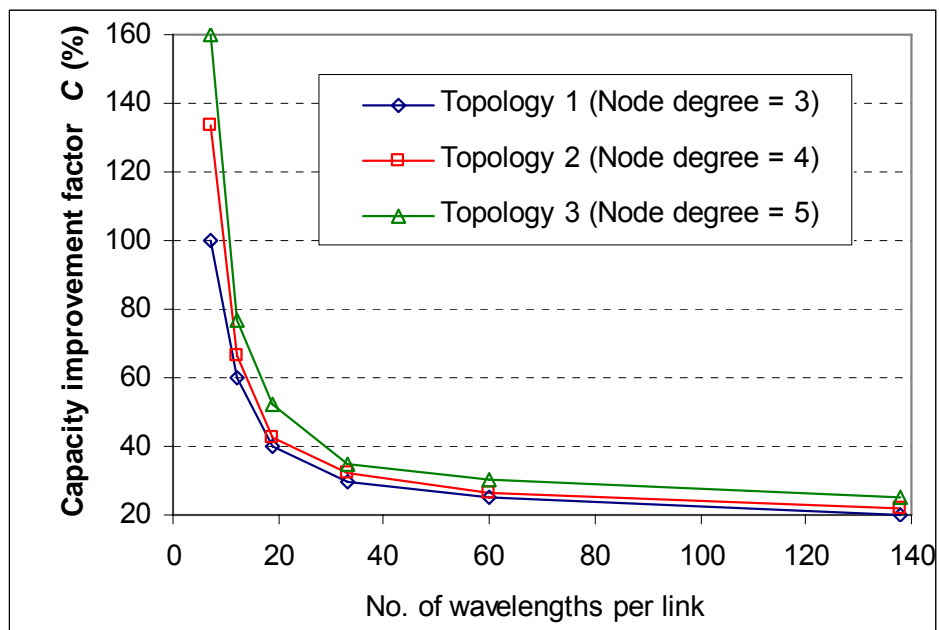


Fig. 3.10 Capacity improvement in distributed dynamic routing vs. static routing

Fig. 3.10 shows the percentage of capacity improvement in traffic load for distributed dynamic routing approach compared to the static one. This is an indication of the revenue increase to carriers obtained by deploying distributed

dynamic routing. The percentage of capacity improvement decreases sharply as the number of wavelengths increases. This happens for the same reason that we have mentioned previously for Fig. 3.9, which is that static routing has more choices in the wavelength selection at each hop as the number of wavelengths increases. Therefore, the traffic load accommodated in static routing is closer to that of dynamic routing. For example, in NSFNET, at 7 wavelengths per link, the capacity improvement is 100%; whereas, at 19 wavelengths per link it is only 40%. As the number of wavelength increases above 30, the improvement in traffic load tends to be constant around 20%.

The above results imply that traffic load could increase at least 20% (i.e., 20% or more increase in revenue) if the distributed dynamic routing is used. Similarly, the number of wavelengths could be reduced at least 15% (i.e., 15% or more cost reduction). Both cost savings and capacity improvement could be maximized by deploying dynamic routing in networks with less than 30 wavelengths per link.

3.1.7 Distributed Versus Centralized Approaches

A simplified centralized routing agent is built to compare dynamic routing implemented in a centralized agent with the same implemented in a distributed approach. By centralized dynamic routing we mean that the path computation is performed in a centralized network element such as centralized network management (CNM). Additionally, in the CNM approach, a centralized routing database keeps track of the global topology of the network across multiple routing domains. The motivation for this analysis is to show the scalability limitations of CNM approaches in supporting services such as BoD, as well as to further identify

the value added proposition and limitations of distributed dynamic routing under different traffic load scenarios and for different types of services.

The centralized agent has a centralized routing database that stores the global topology and TE information and is updated periodically by the Element Management System (EMS) in each node. Upon receiving a call request from the ingress node, the agent performs CSPF path computation and sends the message containing the computed path to the ingress node. After receiving the message, the ingress node starts to use signaling protocol (e.g. RSVP-TE) to set up the path.

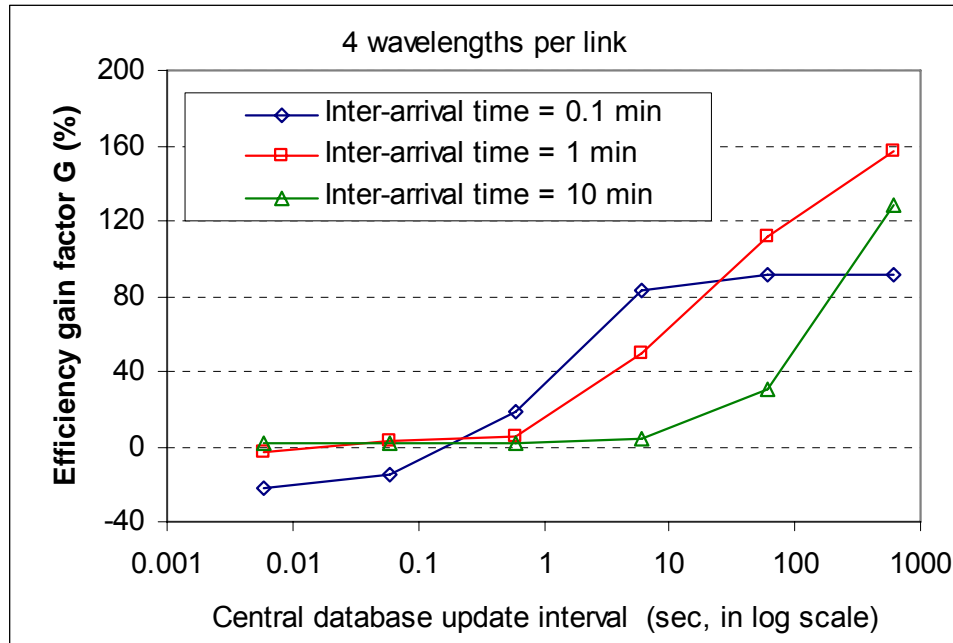


Fig. 3.11 Efficiency gain G vs. centralized database update interval

Fig. 3.11 quantifies the efficiency gain G of distributed dynamic routing over centralized dynamic routing versus the centralized database update interval. The efficiency gain is defined as

$$G (\%) = (L_d - L_c) / L_c * 100\% \quad (7)$$

Where L_c and L_d are the traffic load at 0.01 call blocking ratio in the centralized and

distributed approaches, respectively. The centralized database update interval is an indication of the time-delay needed to update the centralized routing table. This parameter depends upon many factors such as the size of the network (e.g., number of nodes, links, wavelengths per link, etc.), calls' inter-arrival times, as well as the type and efficiency of the algorithms used in updating the centralized routing database. Obviously, this is a dependant complex parameter which requires extensive analytical as well as empirical modeling to quantify, accurately. Our objective here is not to quantify this parameter, but rather to use it as a differentiator parameter that distinguishes centralized approaches from distributed ones. Thus, we use it as an independent parameter in these simulations.

The distributed approach performs better in terms of the amount of traffic load that it can support, as long as the centralized approach has prolonged database update intervals. The shorter the central database update interval, the closer is the performance of centralized approach to distributed approach.

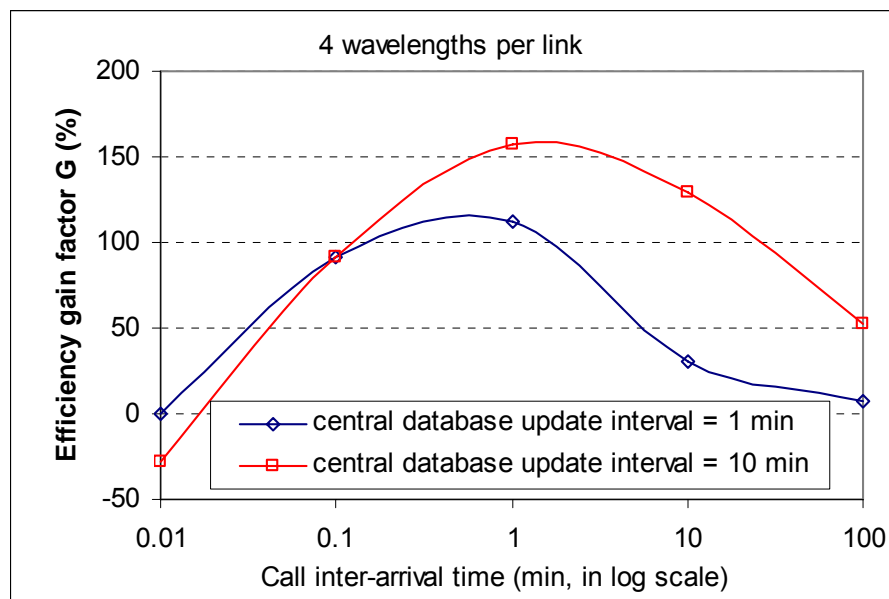


Fig. 3.12 Efficiency gain G approach vs. call inter-arrival time

As shown in Fig. 3.12, the efficiency gain factor varies with different call inter-arrival times and exhibits a bell shape curve. The distributed approach only exhibits significant gain over centralized approach for certain types of traffic. Out of this range, distributed approach shows no notable advantages. Consider the curve of 1-minute central database updating interval; this range is from 0.01 to 100-minute call inter-arrival times, which falls into the BoD services range. For calls with longer call inter-arrival times, which emulate the PL services, distributed approach exhibits little or no efficiency improvement. Note for calls with very short inter-arrival times, e.g., less than 0.01 minute, distributed dynamic routing approach performs no better than centralized approach in terms of traffic load it accommodated. At such short inter-arrival times, the CSPF path computations are based on outdated databases due to the OSPF convergence delay. At the same traffic load, the call blocking ratio of dynamic routing is even higher than that of static routing [26]. This phenomenon is further confirmed in the centralized approach where the network with 10-minute update interval (more towards static routing) can accommodate more traffic than that with 1-minute update interval (more towards dynamic routing), suggesting that dynamic routing (both distributed or centralized) adds no value for such types of traffic. Therefore, static routing is recommended for both approaches. Due to the scalability issue of centralized approach- as the network becomes larger- distributed static routing, i.e. hop-by-hop routing, is more efficient than CSPF-based approaches.

3.2 Connection Setup

We assume that a request for connection setup starts at time T_1 at node 3 (see g. 2.4), and the destination is to node 13. RSVP starts signaling setup at T_1 . In the worst case scenario, if the RSVP signaling can not find enough resources at node 3 at the time the RESV message arrives to node 3, the call will be blocked and a TEAR message will be issued to release all the resources that have been already reserved along previous OXCs (i.e., the link from node 14 to node 9; node 9 to node 10; and node 10 to 13). Assuming a request for a second call arrives at time T_2 at node 14 and the destination is node 9, this call will use one of the resources mentioned above (i.e. the link from node 10 to node 9). If the call inter-arrival time is very short, i.e., the time interval T_2-T_1 is close to, or shorter than, the first call's setup delay then it is highly likely that the second call would be blocked since that resource might not have been released by the first call. This will lead to a higher call blocking rates.

The connection setup up delay T_s is the time required to setup an optical path. It mainly consists of RSVP signaling propagation delay T_{prop} , path calculation delay T_d , RSVP message processing delay T_p , and OXC cross-connect delay T_x .

$$T_s = T_{prop} + T_d + T_p + T_x \quad (8)$$

In this section we discuss the impact of the connection setup delay on the call blocking ratio for traffic with short inter-arrival times (from hundreds of milliseconds to tens of seconds). As we demonstrated previously, for such type of traffic, the OSPF convergence delay also attributes to the higher call blocking ratio

if dynamic routing is applied. In order to identify the impact only introduced by the connection setup delay, static routing is used in the simulations.

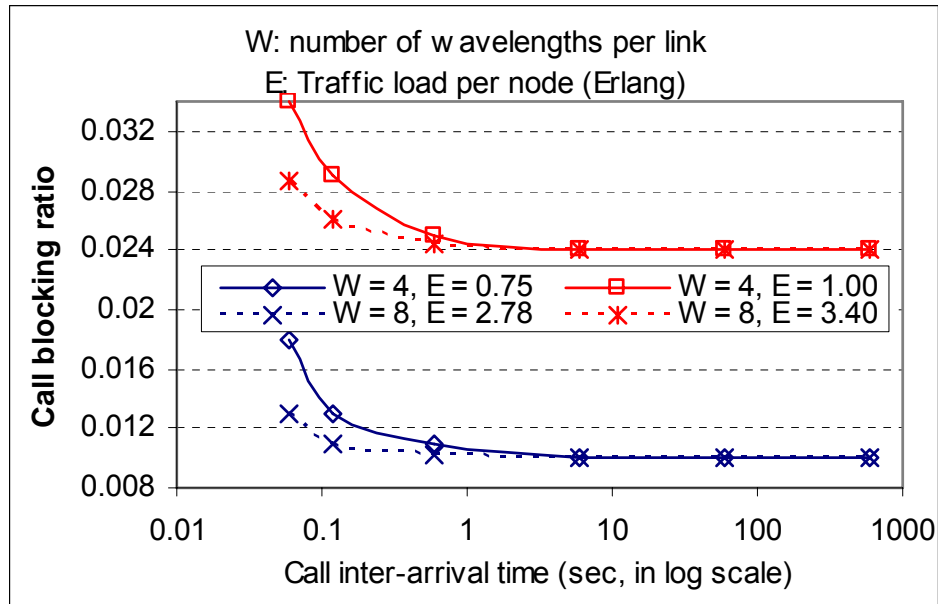


Fig. 3.13 The impact of RSVP signaling propagation delay

3.2.1 The Impact of Propagation Delay

In the simulations of Fig. 3.13, only the propagation delay is considered. The study provides the smallest blocking ratio a carrier could achieve at a fixed traffic load with different call inter-arrival times. To analyze the impact of connection setup delay on traffic with short inter-arrival times, the call inter-arrival time is varied from 0.06 to 600 seconds while keeping the traffic load constant. For the two 4-wavelength curves with different traffic loads, the call blocking ratios start to increase as the call inter-arrival time becomes shorter than 3 seconds. Similar observation is obtained with the two 8-wavelength curves. In 8-wavelength curves, however, the call blocking ratios start to increase as the call inter-arrival time becomes less than 0.6 second. The above result implies that the connection setup

delay limits the calls with very short inter-arrival time the carrier could provide, and in order to reduce the limitation, more wavelengths should be added.

3.2.2 The Impact of RSVP Message Processing and OXC Cross-connect Delay

In Fig. 3.14, the RSVP message processing delay and OXC cross-connect delay are considered. Since the static routing is used, the path computation delay is actually the routing table lookup delay and is ignored in our simulations. The OXC cross-connect time is introduced in a serial-transaction manner (the OXC will forward the RSVP signaling message to the next OXC only after it finishes its cross connection).

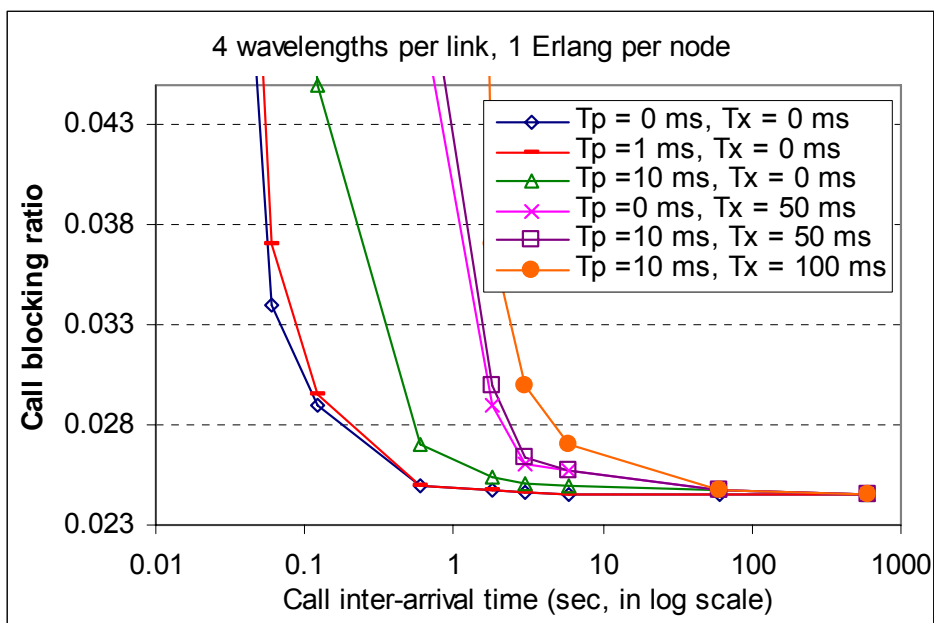


Fig. 3.14 Impact of connection setup delay with other delay parameters considered

In order to observe the impact introduced by only the RSVP message processing delay, the OXC cross-connect time is set to 0, the per node RSVP message process delay T_p is set to 1 and 10 ms respectively. Results show that if the per node RSVP processing delay is at 1 ms level, the impact of RSVP message processing delay is quite limited. As T_p increases to 10s, its impact on the call blocking ratio starts to appear as the call inter-arrival time is less than 6 seconds. To investigate the impact of OXC cross-connect delay, first, we set the OXC cross-connect delay to 50 ms, and we vary the RSVP message processing delay from 0 to 10 ms. These two curves in Fig. 3.14 are almost identical, showing that the impact of the connection setup delay is now dominated by the OXC cross-connect delay. Next, we increase the OXC cross-connect delay from 50 ms to 100 ms and keep the RSVP message processing delay to 10 ms. Consider the curve of 100 ms OXC cross-connect delay for example, the impact of setup delay starts to be visible at 100-second call inter-arrival time.

In summary, the connection setup delay limits the number of calls with very short inter-arrival time that the carrier could provide. The network performance in terms of call blocking ratio will significantly deteriorate as the call inter-arrival time is shorter than tens of seconds. The major contributors to the drop in performance are signaling propagation delay, RSVP message processing delay and OXC cross-connect delay. Among them, the current level of OXC cross-connect delay is the dominating factor.

3.3 Protection and Restoration

In this section, we benchmark the protection and restoration capabilities of the GMPLS control plane in an optical mesh network. Three protection and restoration schemes are described, and followed by the performance analysis to address the cost savings by utilizing control plane to perform the protection and restoration services.

3.3.1 Protection and Restoration Schemes

Current synchronous optical networks (SONET) use bi-directional ring topologies for protection and restoration functionalities. Although this recovery method is fast and reliable, it comes at the expense of unutilized network resources. The development of GMPLS technology, and new optical cross-connect switches (OXC) have made it possible to use mesh topology for the optical transport network. GMPLS introduces a distributed control plane concept to solve the scalability issue caused by the centralized OSS management in the traditional SONET ring network. The GMPLS control plane is a protocol suite that has Open Short Path First with Traffic Engineering (OSPF-TE) for distributed dynamic routing, Resources Reservation Protocol with Traffic Engineering (RSVP-TE) for signaling, and Link Management Protocol (LMP) for auto-discovery. References [27][28][29] provide the framework and architecture for protection and restoration in mesh networks using GMPLS control plane. Various protection and restoration algorithms have been studied in [30][31][32][33]. The promise of the optical control plane is that it could make the recovery time comparable with SONET while utilizing the network resources more efficiently and also reducing the

operators operational cost. Reference [34] outlines the benefit of using control plane to do the distributed fault management. Recent studies [35][36] try to model and quantify the overall operational savings brought by the GMPLS control plane. However, due to the protocol complexity and lack of network experiments, the overall effectiveness and restrictions of applying GMPLS control plane in optical mesh network for protection and restoration have not been fully explored from a network operator point of view.

The protection and restoration for mesh network can be path based, link based, or segment based [37][38][39][40]. In this thesis, the protection and restoration schemes are path-based and categorized by the resources reservation methods for the back-up paths. In the protection schemes, the wavelengths are pre-reserved for the backup paths; whereas in restoration schemes, the wavelengths are not pre-reserved for the backup paths. Three types of protection and restoration schemes are defined in our simulations. Their major differences in the methods of backup path setup are listed in Table 3.1.

Table 3.1 Protection and Restoration Schemes Comparison

	Backup path calculation	Wavelength reservation	OXC cross-connect	EPL service type
1:1 dedicated protection	Yes	Yes	Yes	Class 1
M:N shared protection	Yes	Yes	No	Class 2
Dynamic restoration (No protection)	No	No	No	Class 3

1:1 dedicated protection: Both of the working path and backup path are calculated upon receiving the connection request. When setting up the backup path, resources are reserved and the OXC is pre-cross-connected along the backup path. When a failure occurs, upon receiving the failure notification, the source node

switches the traffic to the already established backup path. The 1:1 protection is class 1 type of service. It offers 100% recovery and shortest recovery time among the three types of protection and restoration schemes.

M:N shared protection: The working path and backup path are calculated upon receiving connection request. When setting up the backup path, resources are reserved but the OXC is not cross-connected along the backup path. Newly admitted backup paths may share the pre-reserved wavelengths of existing back-up path if their working paths are shard risk link group (SRLG) disjointed. When a failure occurs, upon receiving the failure notification, the source node signals the backup path setup and cross-connects OXCs along the pre-calculated backup path. In this thesis, the SRLG concept is introduced in the backup path computation to guarantee 100% recovery upon a single failure. Although this shared protection can offer 100% recovery, since the backup path needs to be setup after the failure, the recovery time is much longer than the 1:1 dedicated protection scheme. We categorize the shared protection into class 2 type of service.

Dynamic restoration (None-protection): There are no pre-calculated backup paths and pre-reserved resources. Upon a failure, after receiving the failure notification, the source node calculates the restoration path and signals the path setup. In our simulations, to make the recovery rate (total number of restored light paths divided by the total number of failed light paths) more trustful, we sequentially fail every link after the network reaches its stable state and then recover it. The dynamic restoration belongs to class 3 type of service. It does not

guarantee 100% recovery, and has almost the same restoration time as M:N protection if the backup path calculation time is ignored.

The pre-emptable traffic is defined as the least priority type of traffic. It can be pre-empted when the link it is using is needed by traffic with any higher level of service classes, such as 1:1protected, shared protected, and dynamic restored traffic.

OPNET is used to build a model for an OXC node, in which OSPF dynamic routing ([41][42][43][44] list more advanced routing methods) with wavelength conversion and link recovery were simulated.

In the simulation, we assume only single link failure occurs (studies on multi-link failure can be found in [45][46][47]). In order to guarantee 100% recovery in shared protection, first, the working path and backup path should be SRLG disjointed; second, if two working paths are SRLG disjointed, their back paths can share the same wavelengths. In our simulations, all the wavelengths within a fiber link belong to the same SRLG. We use the link number to denote the SRLG number. Each link keeps a record of the links' numbers it is protecting, and floods this information to other nodes of the network. When using Dijkstra-algorithm to do backup path computation, first, the links along the working path are pruned to guarantee that the working path and backup path SRLG are disjointed; second, the links that are protecting the links along the working path are also pruned to guarantee that the newly calculated backup paths are SRLG disjointed with other backup paths already existed in the network whose working paths share the same links. If there is already a wavelength reserved in the link along the existing backup

paths, the wavelength can be shared by this newly calculated backup path. In this way we created the first backup wavelengths group. If the backup path could not be found in the first backup wavelengths group, the second backup wavelengths group can be created. If the backup path could not be found in the n th backup wavelengths group (here n equals to the total number of wavelengths per link), the call request is blocked.

3.3.2 Protection and Restoration Analysis

In the NSFNET simulations, we setup the following simulation parameters to keep the results more realistic:

CSPF path computation time $T_d = 0.4$ ms (measured with 1.7 G Pentium 4 processor using Dijkstra's Algorithm).

RSVP per node processing time $T_p = 1$ ms (measured with 1.7 G Pentium 4 processor).

OXC cross-connect time $T_x = 100$ ms (provided by one of the OXC manufactures).

In the following simulations, dynamic routing is used except when static routing is explicitly stated. For SRLG, only the simple conventional graph-presentable type is simulated. The more complex SRLGs, such as fibers coming from different links but sharing the same conduit terminating at a switch, are not considered.

3.3.2.1 *Dynamic vs. Static Routing*

The performances of dynamic routing in 1:1 and M:N protections have similar characteristics as those in none-protection schemes, i.e., there is a traffic load range

from 0.2 to 0.75 Erlangs per wavelength for dynamic routing to be effective, as shown in Fig. 3.15. Outside this traffic load range, dynamic routing has no advantage over static routing in the protection schemes. Take M:N protection case for example, this effective load range is from 0.25 to 0.75 Erlangs per wavelength, and the best performance for dynamic routing happens at 0.5 Erlangs per wavelength. Fig. 3.15 shows the performance comparison of dynamic routing with static routing. Within the effective traffic load range, call blocking ratio of dynamic routing is up to 10% lower than that of static routing in 1:1 and M:N shared protection schemes.

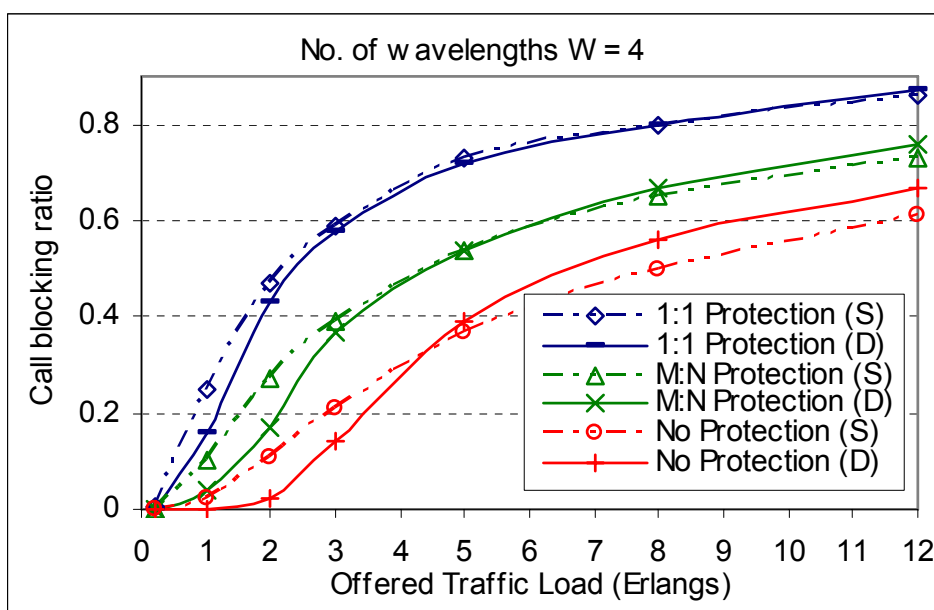


Fig. 3.15 Performance comparison for protection schemes in dynamic vs. static routing

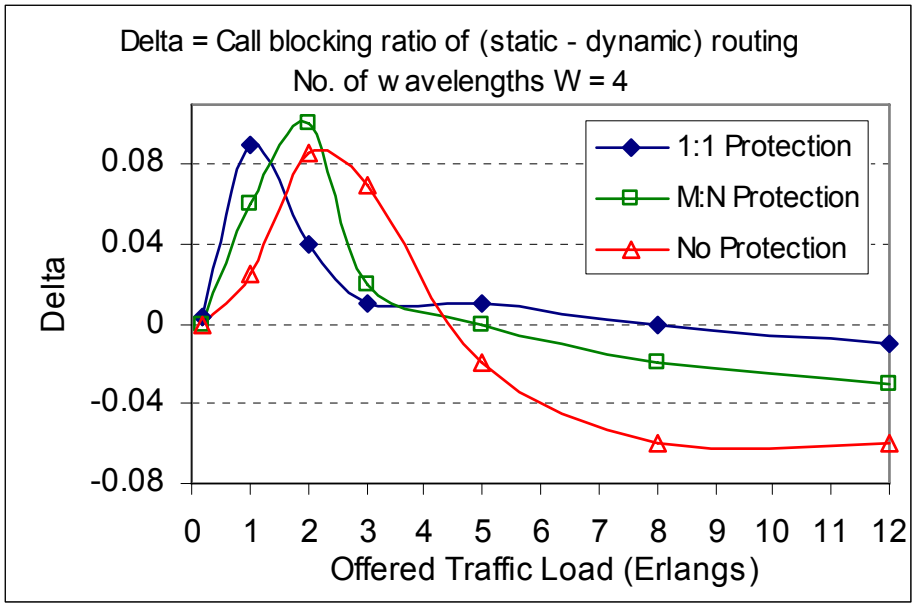


Fig. 3.16 Performance improvement using dynamic routing compared with static routing.

3.3.2.2 M:N Shared Protection Performance Evaluation

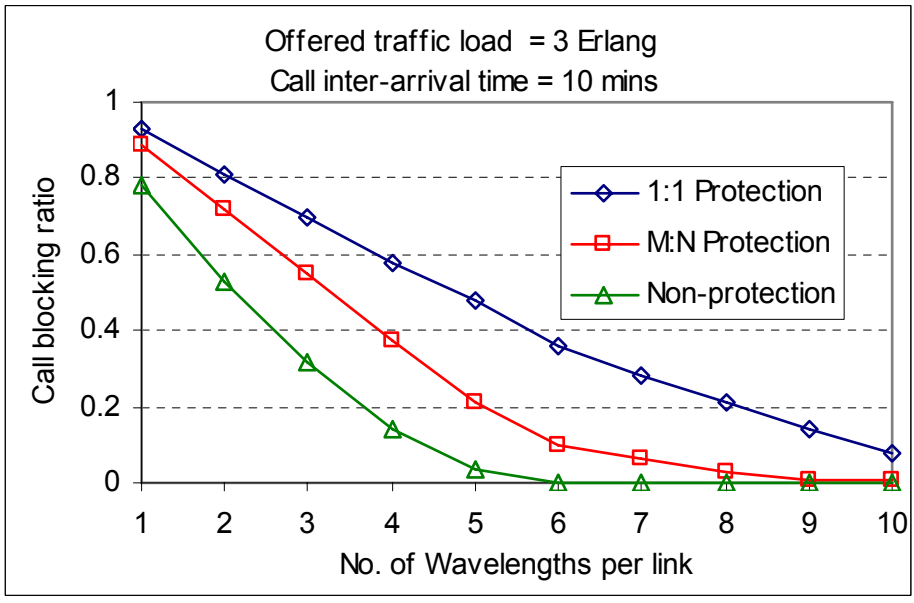


Fig. 3.17 The impact of number of wavelengths per link on the performance of protection schemes

Fig. 3.17 provides the relationship between the call blocking ratio and the number of wavelengths per link at the fixed traffic load. Obviously, M:N shared protection has lower call blocking ratio than 1:1 dedicated protection. To compare the performance of M:N shared protection with 1:1 dedicated protection and no protection schemes in terms of number of wavelengths needed, we increase the traffic load to observe the number of wavelengths needed in each scheme while keeping the call blocking ratio less than 0.1.

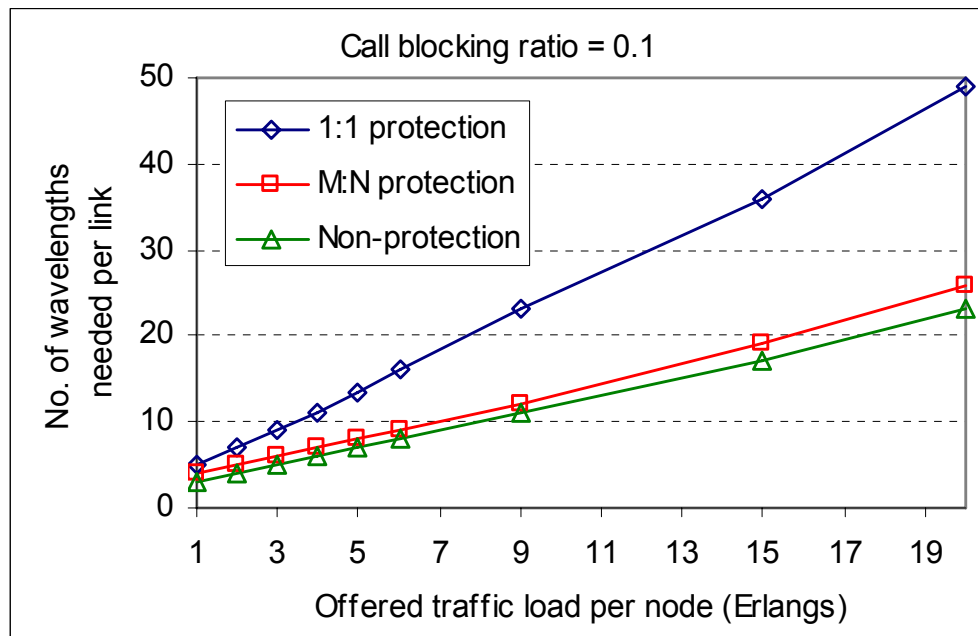


Fig. 3.18 Number of wavelengths needed to keep call blocking rate less than 10% in each scheme.

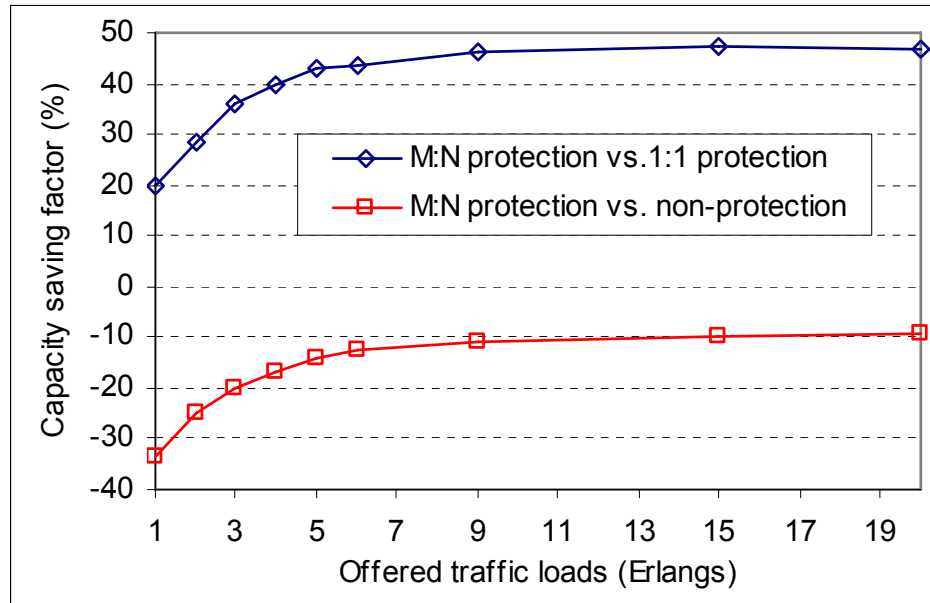


Fig. 3.19 M:N shared protection performance evaluation

As indicated in Fig. 3.18, the number of wavelengths needed in each scheme increases almost linearly with the increase of traffic loads. Fig. 3.19 depicts the capacity saving factor x for M:N protection compared with 1:1 protection and none-protection schemes. X is the percentage of number of wavelengths saved in M:N protection compared with 1:1 and none-protection. For example, at 15 Erlangs offered traffic load per node, 19, 36, 17 wavelengths per link are needed in M:N, 1:1 and none-protection networks respectively. The shared protection saves 48% of wavelengths compared with 1:1 protection; whereas the shared protection needs extra 10% of wavelengths compared with none-protection. Both of the capacities saving factors remain fixed as the traffic load further increases. So for shared protection, it is better for the network to operate at traffic loads at least of 15 Erlangs per node to take advantage of the maximum savings. Increasing the Erlang

load will of course imply that the number of wavelengths also must be increased, thus the cost will be higher.

3.3.2.3 *Dynamic Restoration (none-protection)*

Fig. 3.20 shows that if the offered traffic load is lower than 0.25 Erlangs per wavelength, 100% recovery can be obtained statistically without any protection scheme by using dynamic restoration, however it is not economic feasible for the operator to operate at such a small traffic load. At the same offered traffic load per wavelength, the network with more wavelengths per link will have a higher recovery rate than that of fewer wavelengths per link. Obviously, the higher the call blocking ratio the lower the recovery rate is. The deterioration of the recovery rate is more quickly than the increase of the call blocking ratio. At the 4-wavelength curve, as the traffic load increases from 0.25 to 0.75 Erlang, the call blocking ratio increases by only 1%, whereas the recovery rate decreases by 36%. It indicates that it may be more important to determine traffic load operation range based upon the recovery rate rather than the call blocking ratio.

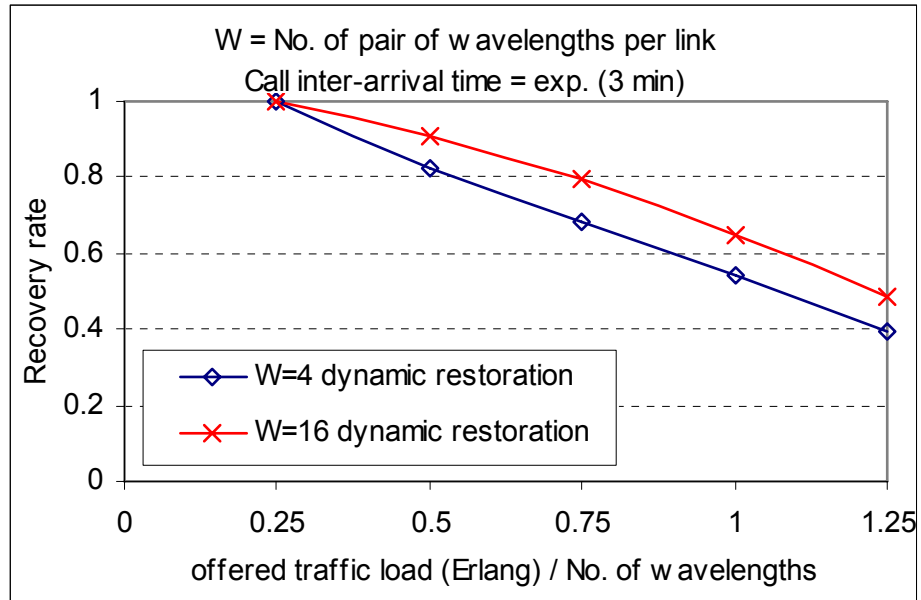


Fig. 3.20 Dynamic restoration performance vs. traffic load

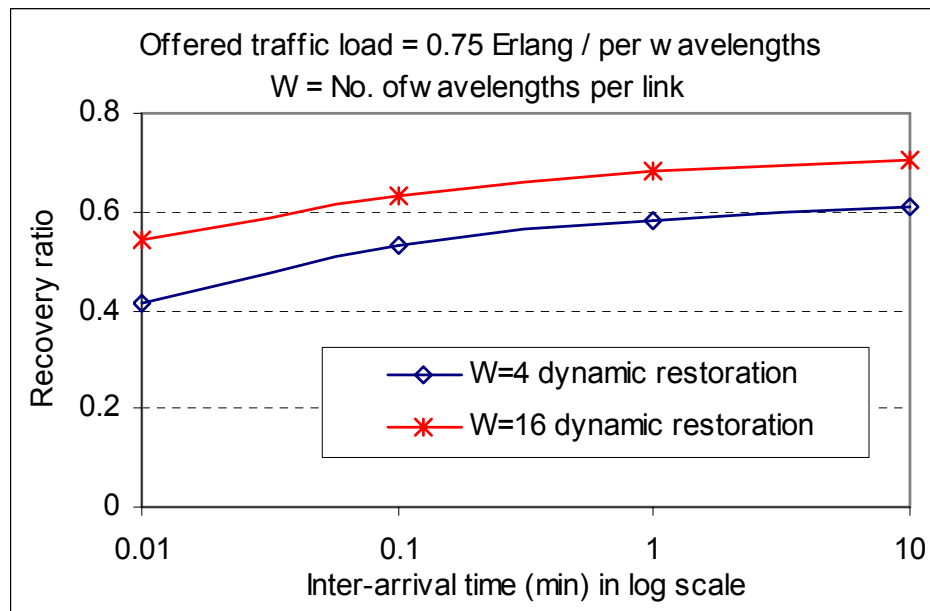


Fig. 3.21 The impact of call inter-arrival time on dynamic restoration performance

Fig. 3.21 shows the impact of call inter-arrival time on the performance of dynamic restoration. As the call inter-arrival time becomes smaller than 0.5 min,

similar to the performance deterioration of dynamic routing, the recovery rate starts to deteriorate.

3.3.2.4 Pre-emptable Traffic

To analyze the pre-emptable traffic in different protection and restoration schemes, we create two types of traffic in the network. One is high priority traffic; the other is low priority traffic. The low priority traffic is not protected, and is pre-emptable when the high priority traffic needs to use the wavelengths being occupied by the low priority traffic. We load the network with 0.82, 1.5, and 2.7 Erlangs high priority traffic for 1:1, M:N and no protection respectively, such that the call blocking ratio is 0.1 for each traffic load. At the same time, we load 0.5 ~ 9.5 Erlangs low priority traffic in the network.

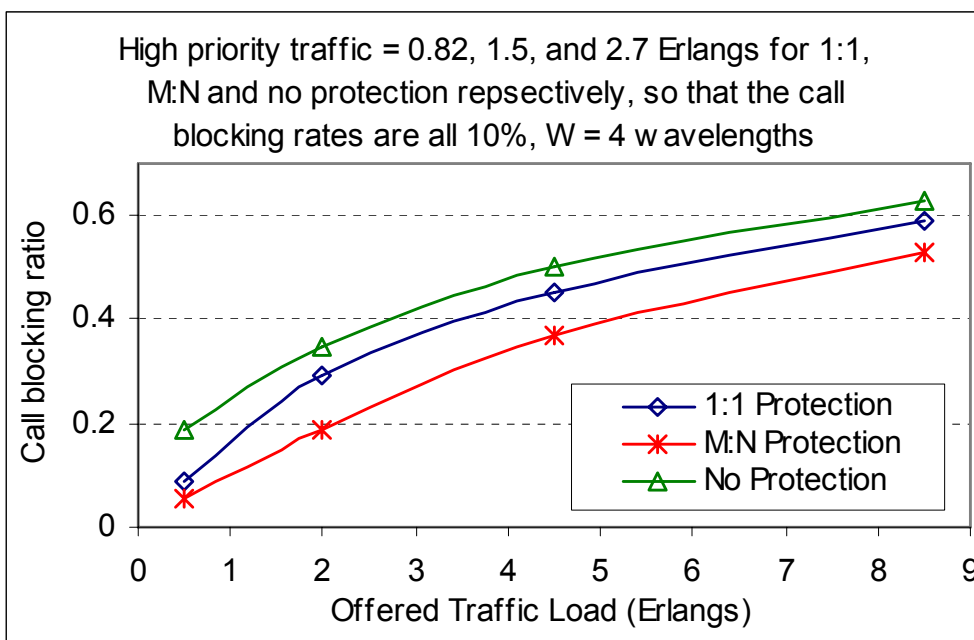


Fig. 3.22 Low priority traffic blocking ratio vs. traffic load

Fig. 3.22 shows an 8% lower call blocking ratio in M:N protection compared to the 1:1 protection. Whereas in the Fig. 3.23, there is 8% higher call pre-empted rate

(total number of pre-empted calls divided by total number of admitted calls) in the M:N protection compared to the 1:1 protection. Fig. 3.24 shows the actual traffic loads admitted in the networks. The M:N protection and 1:1 protection network have almost the same admitted low priority traffic load. In all the above three figures, the non-protection curves are added for reference. For low priority traffic, non-protection has the highest call blocking ratio and call pre-empted rate, and the lowest admitted traffic loads.

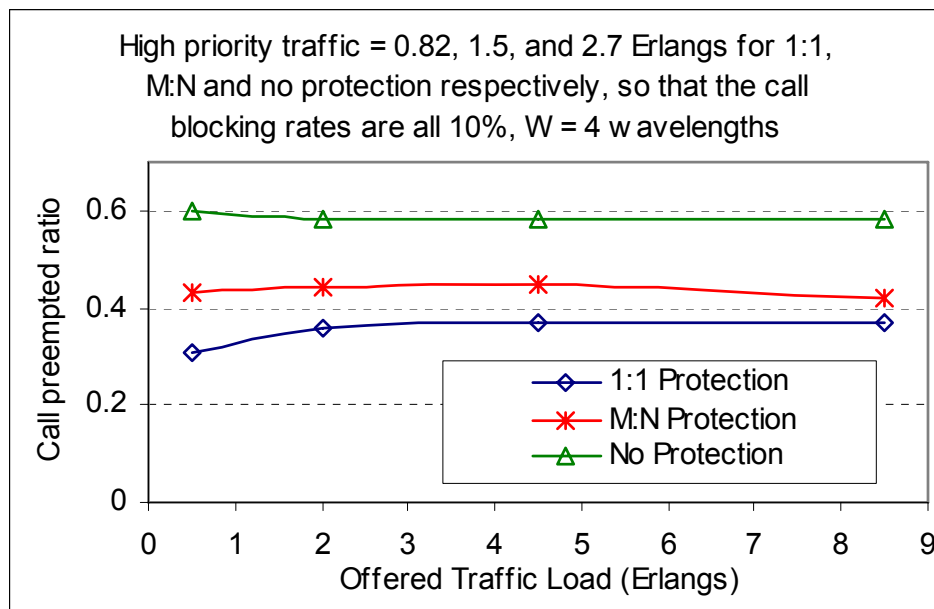


Fig. 3.23 Low priority traffic preempted ratio vs. traffic load

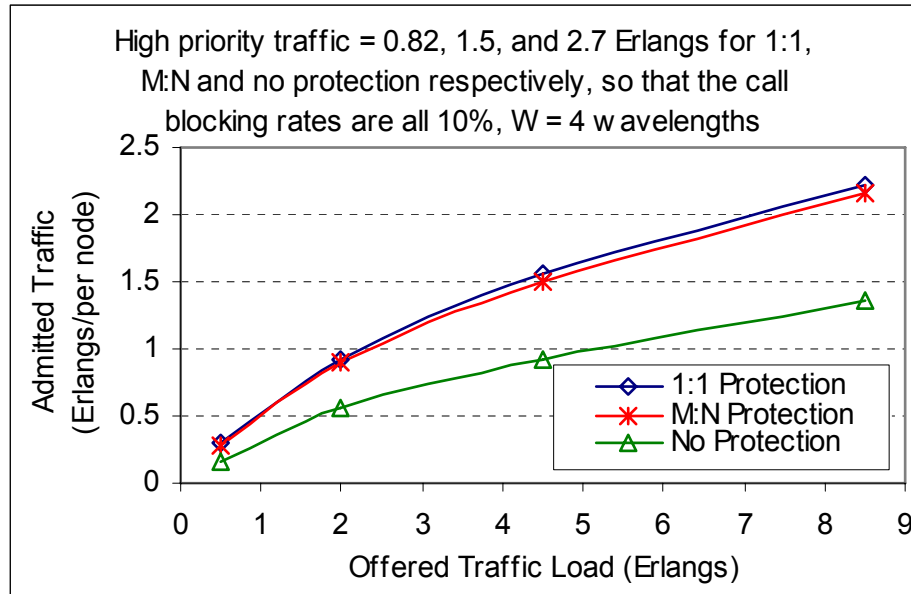


Fig. 3.24 Low priority traffic admitted vs. traffic load

3.3.2.5 Business Driver Analysis

The saving factor x mentioned earlier in sub-section 3.3.2.2 is an important parameter for pricing different classes of services to boost revenues. As shown in Fig. 3.25 and Fig. 3.26, by implementing M:N protection or no protection services, the network capacity can be greatly improved; however, the increase in the capacity does not always result in the increase of the revenue (please see the blue and pink curves), if the prices of the new services are not justified by the capacity savings of the services. I.e. the percentage of the price drop of the new service should not exceed the percentage of the capacity saving of the service.

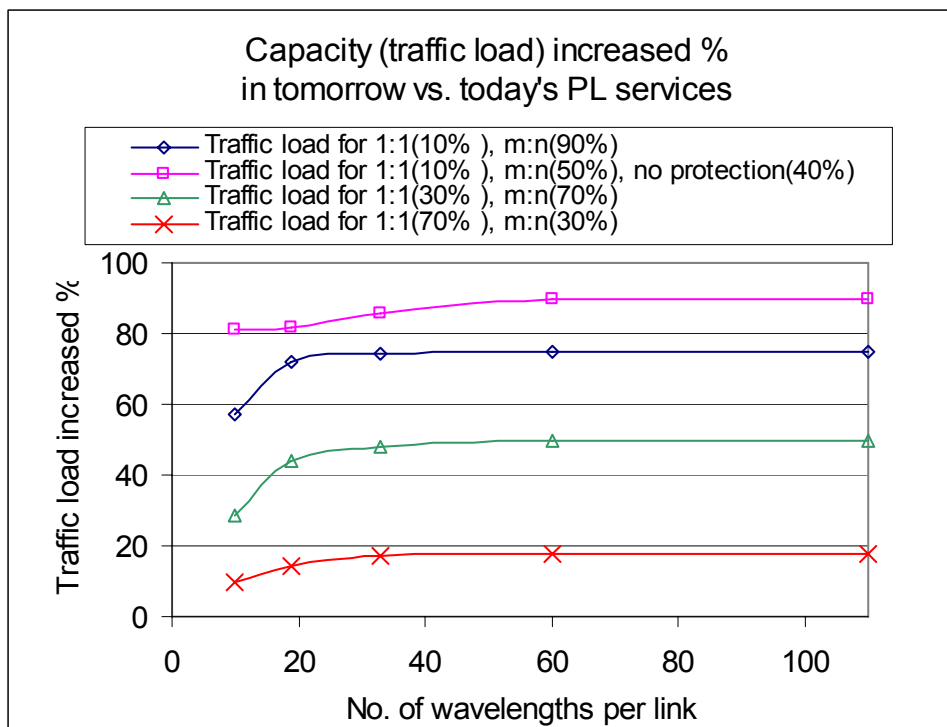


Fig. 3.25 Percentage of capacity increased in tomorrow vs. today's PL services

In our simulations, we define today's PL services as 100% 1:1 protection services, and define tomorrow's services as mixed classes of services which employ 1:1, M:N and no-protection schemes at the same time. In Fig. 3.26, An aggressive price structure is assumed as the following:

M:N protection service is priced 20% lower than 1:1 protection service.

No protection service is priced 40% lower than 1:1 protection service, i.e. 25% lower than M:N protection service.

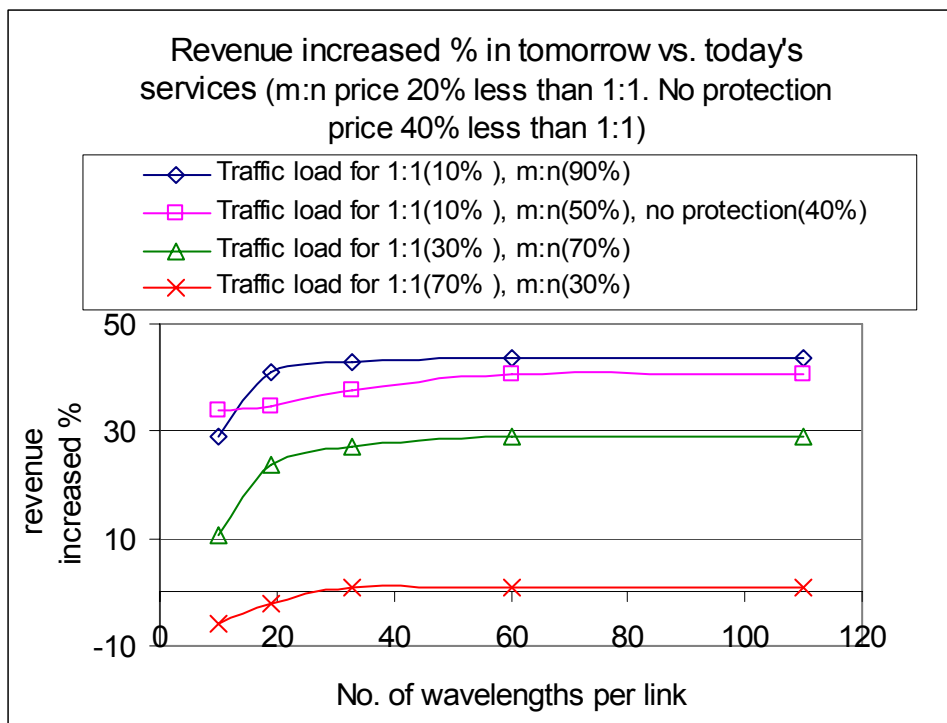


Fig. 3.26 Percentage of revenue increased in tomorrow vs. today's PL services

In Fig. 3.25, the pink curve represents three classes of PL services, 10% 1:1, 50% M:N, and 40% no protection services. The reason why we chose a small portion for 1:1 services is that, M:N protection can also provide 100% recovery assume single failure occurs as the 1:1 protection does, the drawback is that it has longer restoration time than 1:1 protection. The pink curve provides the greatest capacity improvement (90%); however, in Fig. 3.26, it does not always offer the greatest revenue enhancement (40% compared with 44% of the dark blue curve in the high end of traffic load), which is lower than the dark blue curve (10% 1:1, 90% M:N). This is because, 13% capacity saving for no-protection service compared with M:N protection service can not justify the 25% price difference between the two kinds of services. This tells us that if the price difference between no-protection and M:N protection services, for example below 13%, is not an

attractive market price, then it is meaningless to offer no-protection services; except in the situation that the network has few wavelengths per link, for example less than 10, i.e. the capacity saving is more than 25% for no-protection service over M:N protection service (see Fig. 3.19), is meaningful that the no-protection service added (Fig. 3.26). The green and light blue curves in Fig. 3.26 also tell us, beside the price factor, the proportion of the new services should be greater than certain value (in this case, the value for M:N protection services is 30%) so that the revenue enhancement can be positive.

3.3.3 The impact of Node Degrees on Protection Performance

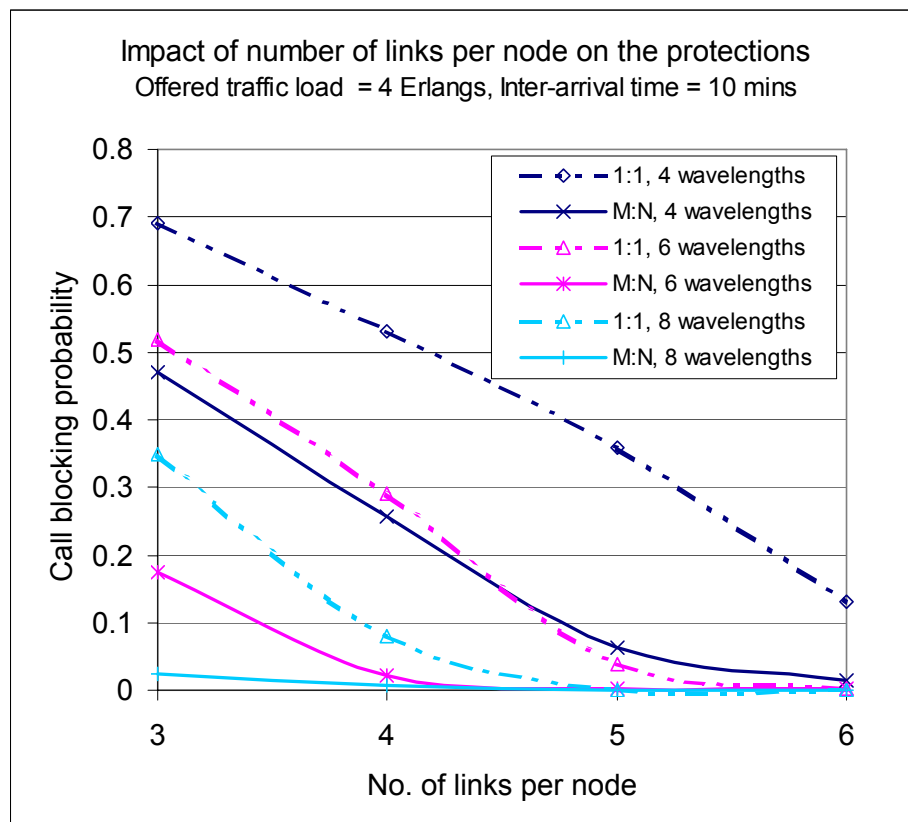


Fig. 3.27 Impact of number of links per node on the protection schemes

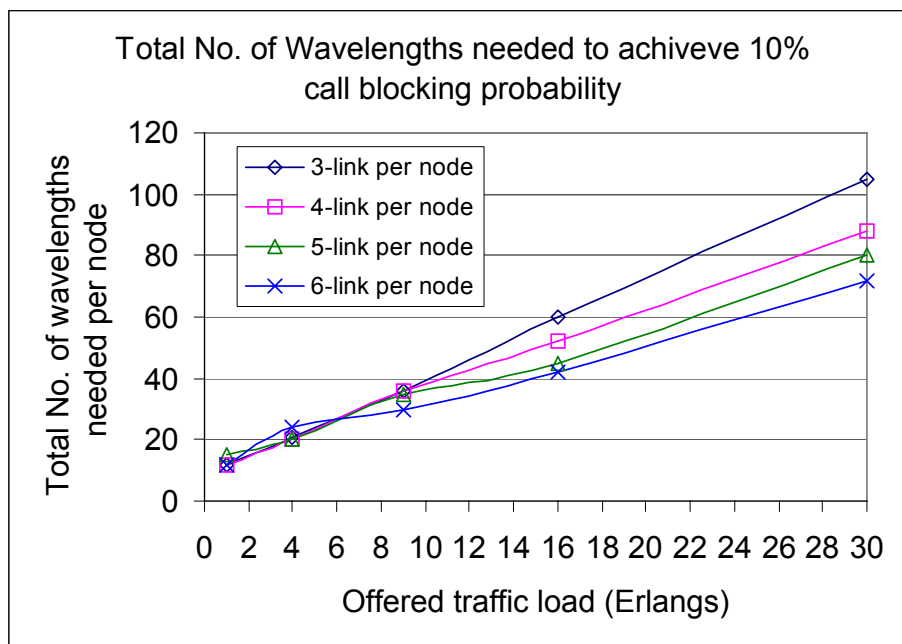


Fig. 3.28 Total number of wavelengths needed vs. number of links per node

Fig. 3.27 shows call blocking probability of the 1:1 and the M:N protection schemes in the networks with variable number of links per node. In the simulations, we fix the number of nodes (16 nodes, to compare with the 16-node NSFNET) and number of links connected to each node (from 3 to 6), and then let the computer randomly generate the topologies. As shown in Fig. 3.27, in order to decrease the call blocking probability, one way is to increase the number of wavelengths per link; the other way is to increase the node degree. In

Fig. 3.28, we keep the call blocking probability less than 10% to obtain the total number of wavelengths needed in the network. After the traffic load is more than 9 Erlangs, for the same call blocking rate, the network that has higher number of links per node needs lower total number of wavelengths. In the next sub-section, we demonstrate that increasing the node degree could be more cost-effective

(evaluated by total number of wavelengths needed) than simply increasing the number of wavelengths per link.

3.3.4 Suggested NSFNET Modification for Protection to Be Efficient

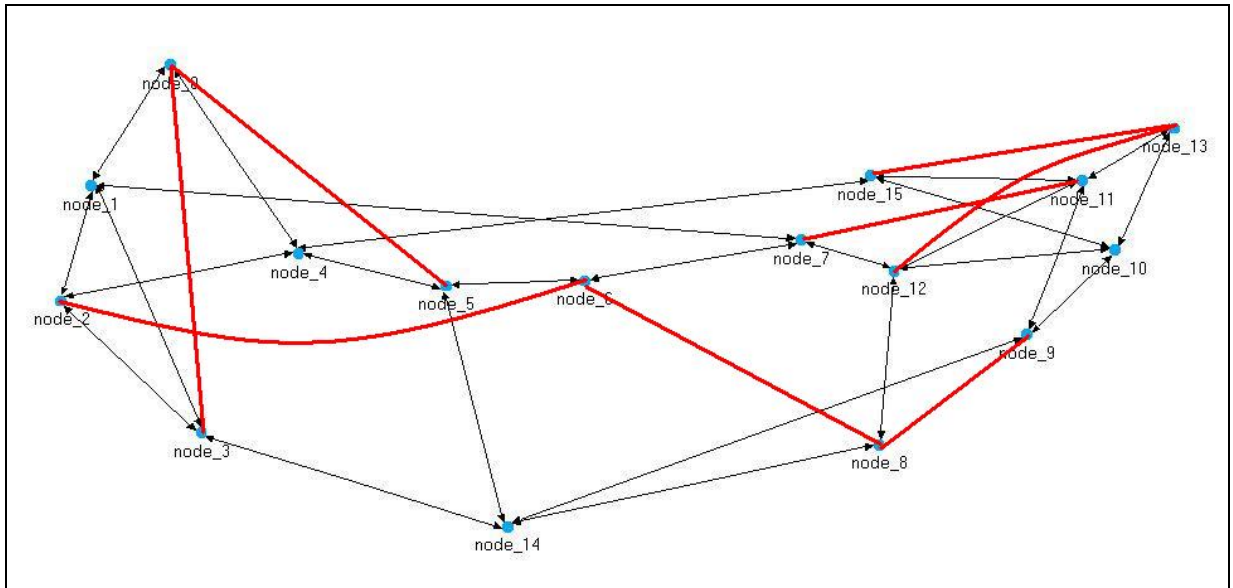


Fig. 3.29 NSFNET after modification (red bold links are added links)

From the analysis of the previous sub-section, we conclude that increasing the node degree of the network can improve the performance of the protection schemes. In this sub-section, we try to modify the NSFNET to observe the performance improvement. In Fig. 3.29, by adding 8 links (the red bold links) to the existing NSFNET topology, the average node degree is increased from 2.50 to 4.06. Fig. 3.30 depicts the call blocking probabilities of 1:1 and M:N protection schemes at different number of wavelengths per link before and after the topology modification (the none-protection curves are added for reference). Fig. 3.31 gives the performance improvement at different number of wavelengths. Take the curves of M:N protection for example, this modification can lead up to 26% decrease in the call blocking rate, and the maximum decrease occurs at 4 wavelength per link.

In Fig. 3.30, the blocking rate for M:N protection in the modified network is 10% at 4 wavelengths per link. If we don't do the topology modification, we can increase the number of wavelengths per link and still achieve this 10% call blocking rate. Fig. 3.30 tells us that the number of wavelengths per link needed is 6. The following simple calculation shows that by increasing the average node degree of NSFNET from 2.50 to 4.06, 36% of the total number of wavelengths can be saved in the shared protection scheme and it can be translated to a 56% increase in the return on investment ratio.

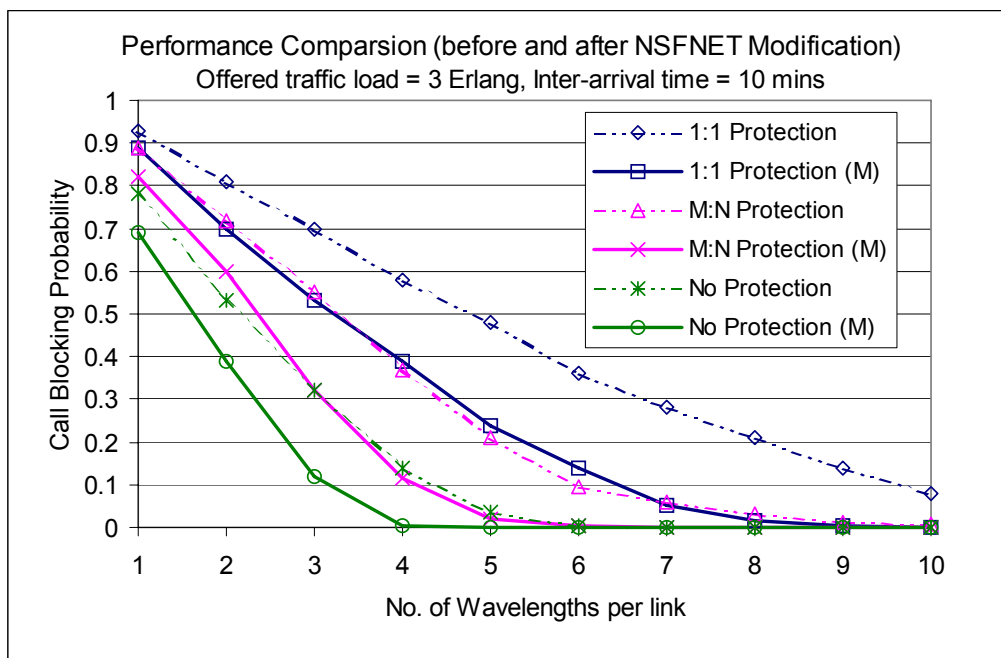


Fig. 3.30 Performance comparison before and after NSFNET Modification

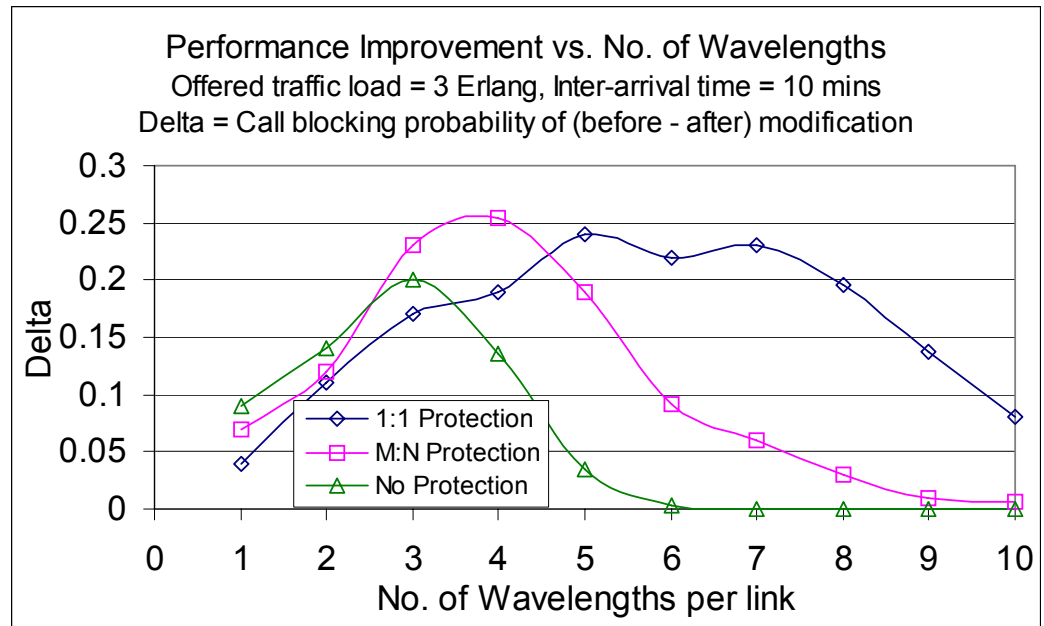


Fig. 3.31 Performance improvement vs. number of wavelengths per link

Scenario 1: adding 8 links (bi-directional) to the existing 4 wavelengths per link network.

Scenario 2: adding 2 wavelengths per link to the existing 4 wavelengths per link network.

N1: Total extra number of wavelengths needed in scenario 1. $N1 = 4 \times 8 \times 2 = 64$

N2: Total extra number of wavelengths needed in scenario 2. $N2 = 2 \times 25 \times 2 = 100$

Δc : Percentage of number of wavelengths saved in scenario 1 compared with scenario 2.

$$\Delta c = (N2 - N1)/N2 = 36\% \quad (9)$$

C: cost per wavelength.

R: Extra revenue brought by the network expansion. Since both scenarios have the same 10% call blocking rates after the expansion, they both have the same R.

Δ ROI: The percentage of increase in return on investment of scenario 1 compared with scenario 2.

$$\Delta\text{ROI} = (R/N1*C - R/N2*C)/(R/N2*C) = N2/N1 - 1 = 56\% \quad (10)$$

3.3.5 Protection and Restoration In Distributed Vs. Centralized Approaches

A simplified centralized management agent, like OSS, is built to compare the distributed approach with the centralized approach in terms of protection and restoration. The centralized agent has a centralized routing database that stores the global topology and TE information and is updated by the Element Management System (EMS) in each node. Upon receiving a call request from the ingress node, the agent performs working and backup path computation and sends the message containing the computed path to the ingress node. After receiving the message, the ingress node starts to use signaling protocol (e.g. RSVP-TE) to set up the path. The centralized database update interval is an indication of the time-delay needed to update the centralized routing table. This parameter depends upon many factors such as the size of the network (e.g., number of nodes, links, wavelengths per link, etc.), calls' inter-arrival times, as well as the type and efficiency of the algorithms used in updating the centralized routing database. Obviously, this is a dependant complex parameter which requires extensive analytical as well as empirical modeling to quantify, accurately. Our objective here is not to quantify this parameter, but rather to use it as a differentiator parameter that distinguishes centralized approaches from distributed ones. Thus, we use it as an independent parameter in these simulations. We assume the central database-update-delay increases linearly with the increase of call inter-arrival rate:

$$\text{Database-update-delay } D = k * \text{Inter-arrival rate} = k / \text{Inter-arrival time} \quad (11)$$

Where k is a factor that can be varied to observe the impacts of different delay values.

The Impact of Database-update-delay on Protection and Restoration

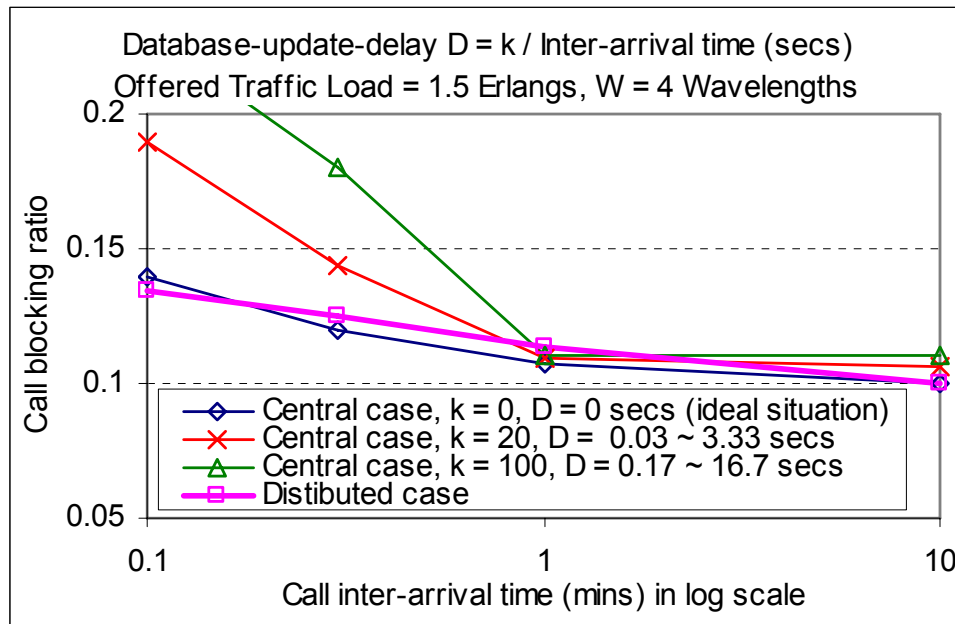


Fig. 3.32 The impact of central database-update-delay on M:N protection in centralized dynamic routing approach

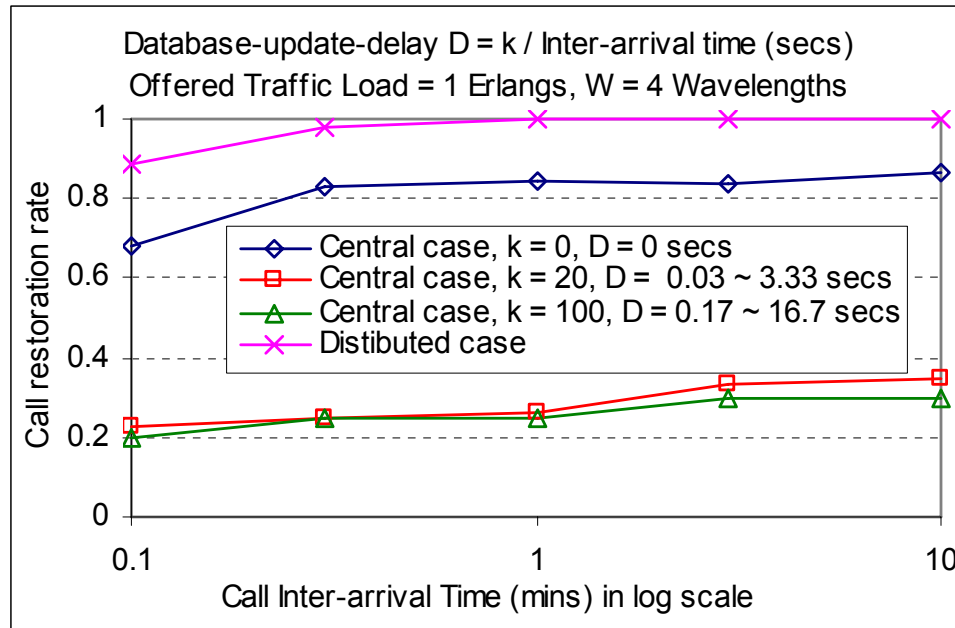


Fig. 3.33 The impact of central database-update-delay on dynamic restoration in centralized dynamic routing approach

Fig. 3.32 shows the impacts of central database-update-delay on the call blocking probability of M:N protection. $K = 0$ curve is added for reference. Even though the database-update-delay is 0 sec ($k = 0$ curve), the performance (call blocking probability) of centralized network is worse than that of distributed network if the call inter-arrival time is shorter than 0.3 min; Fig. 3.33 shows the impacts of database-update-delay on the call blocking probability of dynamic restoration. At $K = 0$, restoration rate of centralized dynamic routing approach is 15% worse than that of distributed dynamic routing approach. Clearly, centralized dynamic routing approach is not suitable for protection and restoration purpose for the calls with shorter inter-arrival times.

Chapter 4

Book-Ahead Reservation for Connection Oriented Services

Our discussions in the previous chapters about the connection provisioning are all for the immediate requests. How to provide effective reservations for these immediate requests has been the focus of the research community so far. Little attention has been paid on the book-ahead reservations. However, as indicated in section 3.1.5, the level of call granularity has a big impact on the network performance in terms of call blocking and network utilization. The finer the call granularity, the better the network performance. In order to reduce the degradation of network performance caused by the coarser granularity calls, these calls are better to be scheduled.

In this chapter, we first describe the reservation mechanism used in this thesis. Then we introduced a heuristic algorithm for path computation and compare it with both the Integer Linear Programming (ILP) optimization approach [48] and the shortest path first (SPF) method. Finally, the characteristics of the scheduling system and the impact of the immediate reservations (IR) on the book-ahead (BA) reservations are analyzed.

4.1 Book-ahead Reservation Mechanism

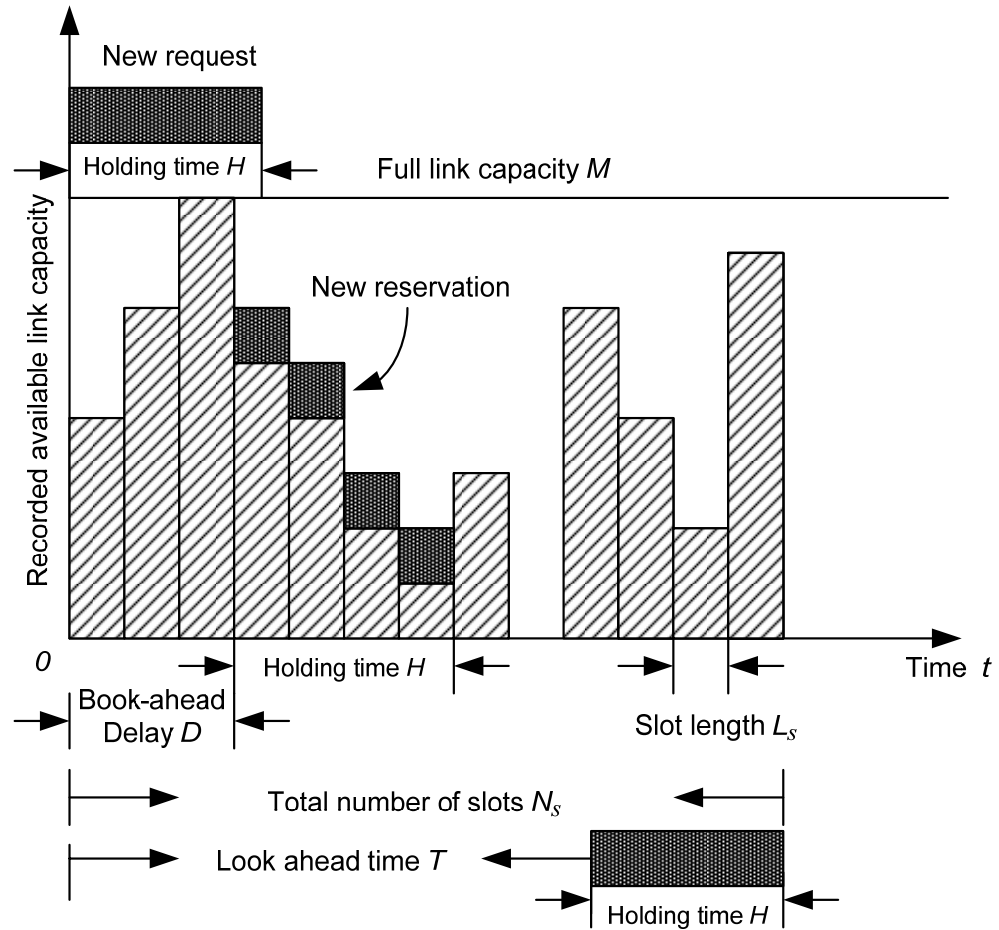


Fig. 4.1 Advance reservation mechanism

In the thesis, the time-slot based advance reservation method [49][50] [51] is used. The future time is divided into time slots (Fig. 4.1). The unreserved link capacity of each slot is stored in the link state database (LSDB). Therefore, the future link capacity is represented by a series of discrete values which could be used for future path calculations. The LSDB can be either centralized or distributed. How to disseminate this information is out of the scope of this chapter. We assume that our path calculations for book-ahead reservations are based on accurate LSDB.

Please refer to Chapter 3 for the impact of the database updating delay on the network performance in both centralized approach and distributed approach.

The following notations are use in this chapter:

The network is denoted as a directed graph $G = (V, E)$, where V is the set of nodes and E is the set of edges (links).

N_s : Total number of slots used in the scheduling system to store link capacity information.

L_s : The slot length in unit time.

M : Total link capacity

T : Look ahead time, i.e. the call can only be scheduled to start within T , and

$$T = N_s * L_s - H \quad (12)$$

D : Book ahead delay.

4.2 Heuristic Algorithm

The path computation for the book-ahead reservation could be done either on-line or off-line depending on the applications. Due to its complexity and time-consuming, the ILP optimization approach, which is a special case of multi-commodity flow problem, could be used for off-line path optimization for a network with only several nodes, and it quickly becomes unmanageable for network with 20 nodes [52]. Therefore, a heuristic on-line algorithm, called first available shortest path (FASP) algorithm, is developed to obtain a sub-optimal solution on a call-by-call basis. FASP shows better performance than the standard shorted path first algorithm (SPF) and shortest-widest path algorithm (SWP) [53] which chooses a feasible shortest path among those links with maximum residual

capacity. The pseudo C code of FASP is shown in Fig. 4.2 and described as the following steps:

```

1:   t_start = 0;
2:   Path_Vector (E1, E2, ..., Ep) = SPF_Path_Compute (Source, Destination, G(V, E))
3:   if (Path_Vector == NULL)
4:       return NULL; //the path cannot be found, the call is blocked
5:   for (Ei = E1; Ei <= Ep; Ei++){
6:       Ei_status = 0; //Set status of link Ei "not satisfied"
7:       for (ti = t_start; ti < T; ti++){
8:           for (tj = ti; tj < (ti+H); tj++){
9:               //In link Ei, no continuous time slots for holding time H available
10:              //starting from ti, try starting from next t(i+1) slot
11:              if (available bandwidth of Ei < requested bandwidth){
12:                  t_start = tj;
13:                  break;
14:              }
15:          }
16:          //Find continuous time slots for holding time H in link Ei, check next
17:          //link starting from slot tj
18:          t_start = tj;
19:          Ei_status = 1; //Set status of link Ei "satisfied"
20:          break;
21:      }
22:      //Link Ei does not satisfy the request, remove it from E, re-compute the path
23:      if (Ei_status == 0){
24:          remove Ei from E;
25:          goto 2;
26:      }
27:  }
28:  //Every link of the path passes the bandwidth test, the call is admitted.
29:  return Path_Vector (E1, E2, ..., Ep);

```

Fig. 4.2 Pseudo C code of FASP

Step 1: Dijkstra's shortest path (DSP) algorithm is applied to obtain the shortest path which contain link E_1, E_2, \dots, E_p .

Step 2: Starting from the first link in the calculated path and first slot, determining if there is a series of continuous slots (total length is equal to the call holding time) and whose available capacities can satisfy the request.

Step 3: If at certain time slot t_i , the link does not satisfy the request, go to step 2 and check again, but this time starting from $t(i+1)$ slot. The search is looped until the continuous slots if find, record the starting slot t_j , then go to Step2 starting the search for the next link from the t_j slot.

Step 4: If during the searching at certain link, the end of the book-ahead slot (total-slot length – call holding length) is reached without finding the consecutive

slots, this link is pruned from the network graph and go back to step 1 to calculate the next shortest path.

Step 5: If the last link in the path is reached and the common continuous slots are found in all the links along the path, the search is done and the reservation is made. Otherwise, the call is blocked.

Using the Pentium 4 processor with 3.6 GHz CPU, the average path calculation delay obtained by FASP for NSFNET (with 50 wavelengths per link and 1 Erlang per wavelength) is 0.61 milliseconds with the variance and standard deviation being 0.0037 and 0.061 respectively. The average number of SPF loop count in the FASP is 1.1 with the variance and standard deviation being 0.19 and 0.44 respectively.

Two topologies are used in our performance analysis. For ILP approach, a 5-node network shown in Fig. 4.3. is used. For FASP heuristic algorithm, NSFNET topology (Fig. 2.4) is used.

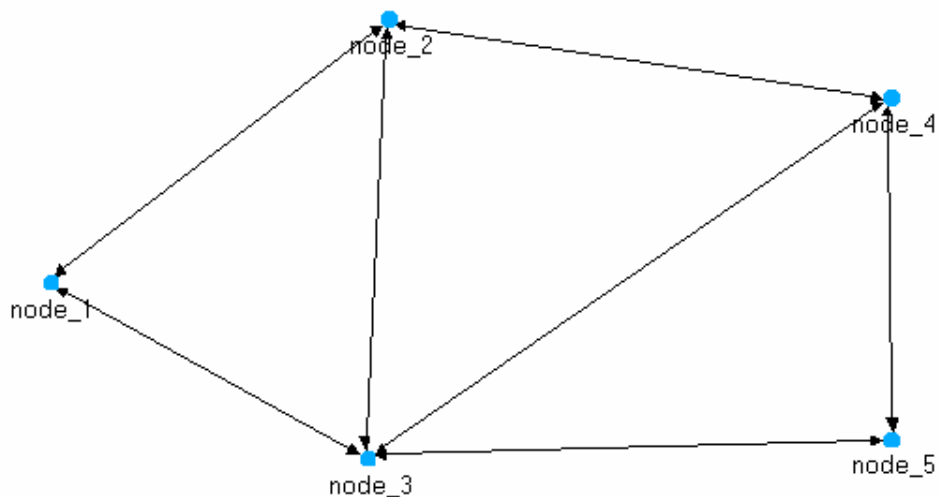


Fig. 4.3 Topology used in the ILP approach

The requested bandwidth of each call is $1/M$. $1/M$ could be a single wavelength, 10G bps, or 1 Gbps, depending on the capacity of the line card installed in the OXC. In this chapter, $1/M$ denotes a single wavelength, i.e., the total number of the wavelengths per link is M .

4.3 Comparison between the FASP with SPF and SWP

The comparison of FASP to SPF and SWP in terms of call blocking and network utilization is shown in Fig. 4.4 and Fig. 4.5 respectively. The parameters used in the simulations are as follows:

Offered traffic load is 1 Erlang per wavelength. $H= 50$ unit, $N_s = 200$, $L_s= 1$ unit.

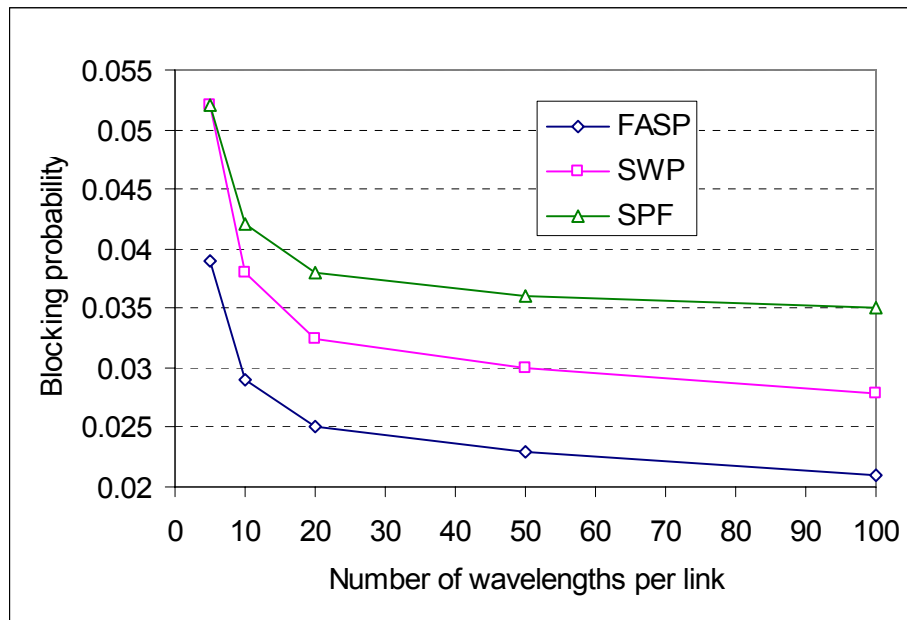


Fig. 4.4 Call blocking vs. number of wavelengths per link.

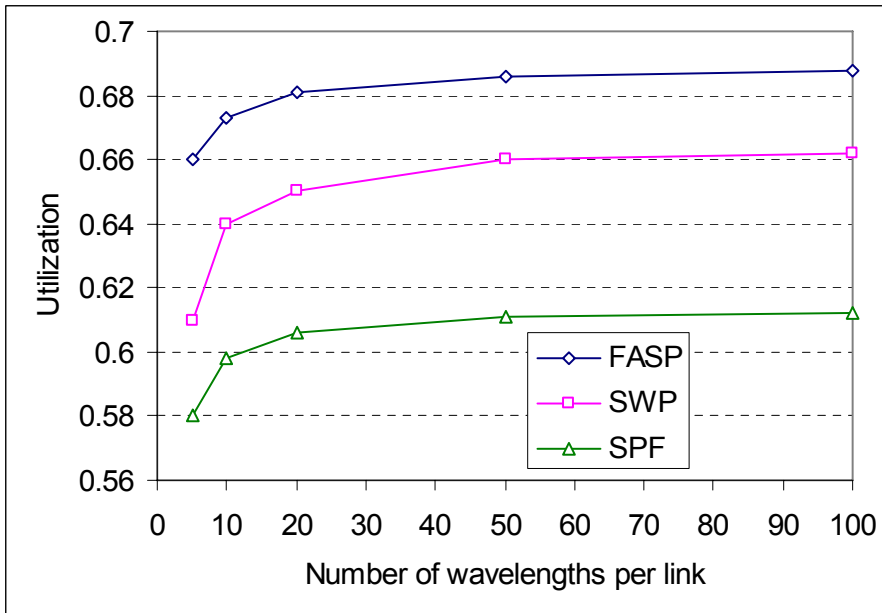


Fig. 4.5 Network utilization vs. number of wavelengths per link

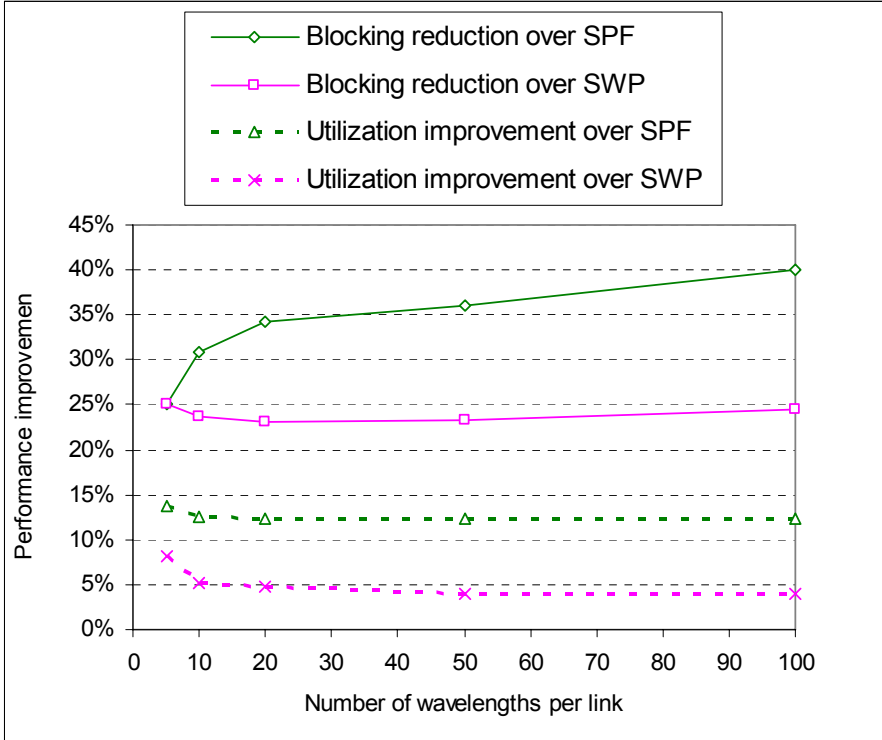


Fig. 4.6 Percentage of performance improvement of FASP vs. SPF

Fig. 4.6 shows the percentage of performance improvement of FASF compared with that of SPF, her we define:

$$\text{Blocking reduction over SPF (\%)} = \text{Blocking of (SPF - FASP)} / \text{Blocking of SPF} \quad (13)$$

$$\text{Blocking reduction over SWP (\%)} = \text{Blocking of (SWP - FASP)} / \text{Blocking of SWP} \quad (14)$$

$$\text{Utilization improvement over SPF (\%)} = \text{Utilization of (FASP - SPF)} / \text{Utilization of SPF} \quad (15)$$

$$\text{Utilization improvement over SWP (\%)} = \text{Utilization of (FASP - SWP)} / \text{Utilization of SWP} \quad (16)$$

We can see that the more the number of wavelengths, the lower the call blocking of the FASP compared with SPF and SWP. For example, if number of wavelengths is 50, FASP performs 36% and 23% better than SPF and SWP respectively. This number increases 40% and 24% respectively if the number of wavelengths increases to 100. In terms of utilization, however, as the number of wavelengths increases, the utilization improvement tends to be a stable value, which is around 12% and 4% respectively.

4.4 Comparison between the ILP approach and the heuristic algorithm

The ILP optimization problem is solved using GNU linear programming kit. The optimization is based on 1 Erlang per wavelength per node network traffic, with the requested bandwidth being $1/n$ of the total link capacity, n is changed from 1 to 20.

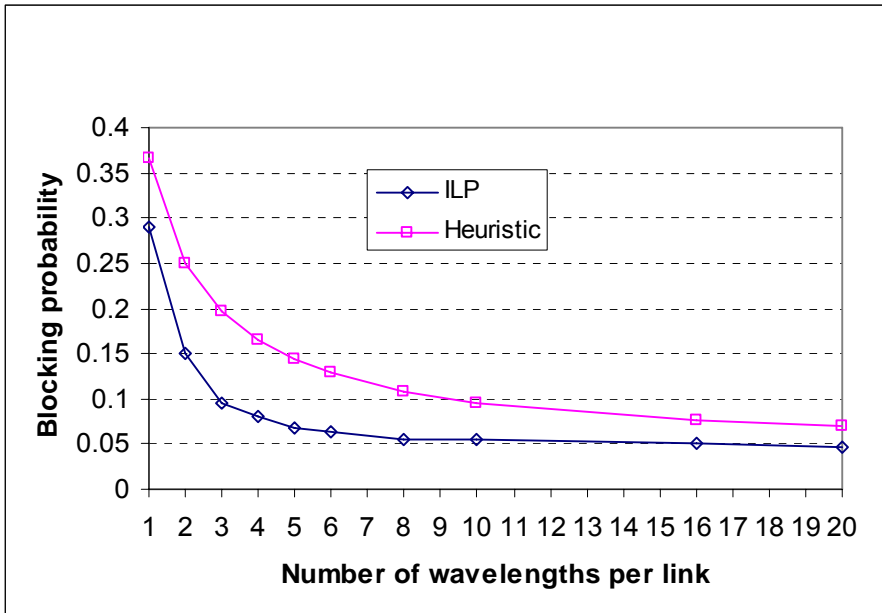


Fig. 4.7 Call blocking probability of ILP approach vs. heuristic method

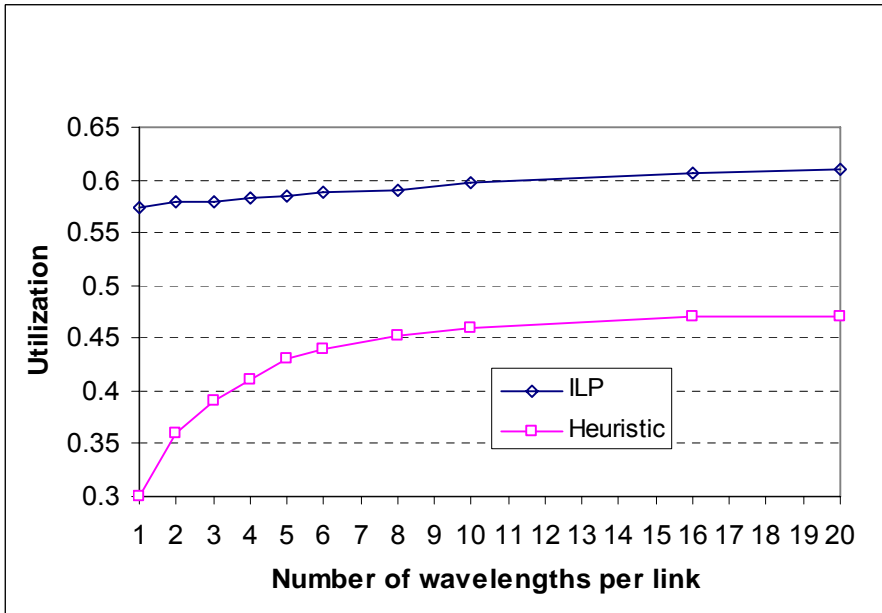


Fig. 4.8 Network utilization of ILP approach vs. heuristic method

Fig. 4.7 shows the comparison of the call blocking between the ILP approach and the heuristic approach. As the number of wavelengths per link increases, the call blocking of FASP tends to be more closed to that of ILP. For example, at 20

wavelengths per link, the difference of blocking between heuristic and IPL is only 2%. The network utilization comparison (Fig. 4.8) shows that as the number of wavelength per link increases, the difference between the two approaches tends to be constant. From 8 wavelengths per link and above, the ILP approach is about 14% higher than the heuristic approach.

4.5 Performance Analysis

In this sub-section, FASP is used to study the characteristics of the scheduling system. The network topology used in the simulations is NSFNET (Fig. 2.4).

4.5.1 The optimum total number of slots and slot length

The first issue in designing the advance reservation system is to determine how much of the total number of slots should be used and how long each slot should be. These parameters determine the size of the storage used in the computer and the amount of overhead traffic that will be generated to transmit the stored information.

Fig. 4.9 shows the blocking probability vs. number of slots used by the scheduler. For comparison, the offered traffic load per load is normalized to 1 Erlang per wavelength, and FASP algorithm is applied. Network scenarios of 50 and 20 wavelengths per link with different call holding times are simulated. First, by using FASP, we can achieve a lower blocking probability around 2% if we have sufficient number of slots N_s (slot length is $L_s = 1$ unit) in the scheduler, i.e., the look-ahead time is long enough. Second, the blocking probability tends to reach to its minimum if the look-ahead time T is longer than certain value, beyond that, no

noticeable reduction in the blocking probability will be observed. Take the 50 wavelengths/50 unit call holding time curve for example, this value is about 7 times of the average call holding time (350-unit look ahead time). Fig. 4.10 shows that the network utilizations of the above scenarios exhibit the same trends, and the highest utilization that can be reached is around 71%.

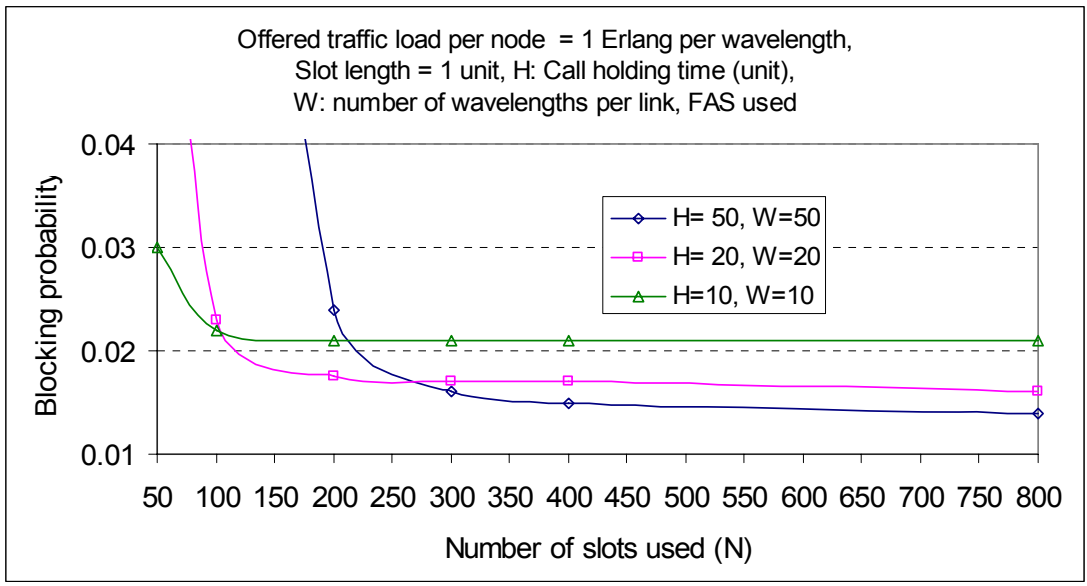


Fig. 4.9 Blocking probability vs. number of slots used by the scheduler

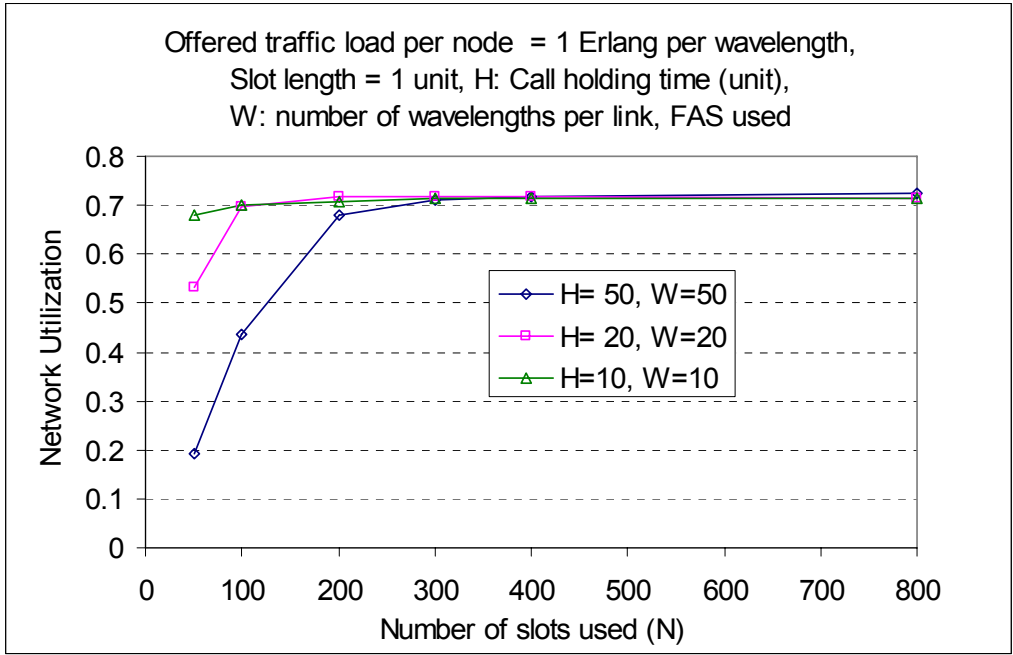


Fig. 4.10 Network utilization vs. number of slots used by the scheduler

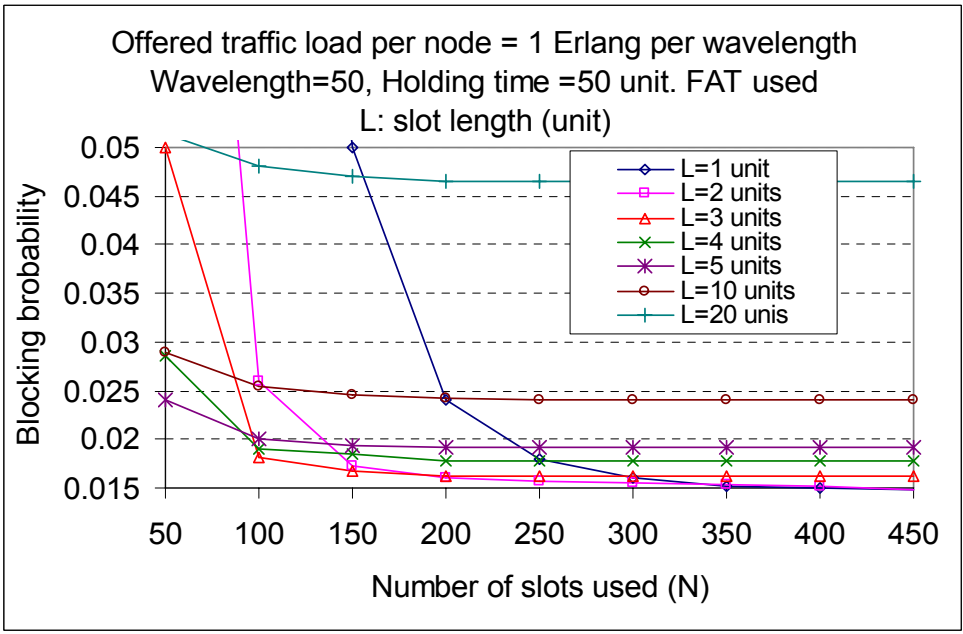


Fig. 4.11 Blocking probability vs. slot length and the total number of slots

In order to study the impact of the slot length L_s on the blocking probability, the L_s is varied from 1 unit to 20 unit. The results are shown in Fig. 4.11. The lowest

call probability is obtained when slot length is 1 time unit, however it requires more number of slots. As the slot length increased, the blocking probability is also increased, but the required total number of slots to reach the stable state is reduced. However, the product of total number of slots N_s and slots length L_s to reach the stable state remains the same, Take the $L_s=1$ curve for example, when N_s is 350, the blocking probability reaches the lowest point, $N_s * L_s$ equals to 350. for the $L_s=2$ curve, N_s is about 175, $N_s * L_s$ is 350. As the slot length is further increased above 5 time units, the blocking probability will increases significantly, which is out of the realistic operating region. Fig. 4.12 is plotted to better understand the optimum operating region in terms of slot length and the total number of slots. If 2% blocking probability is set as an acceptable value, the smallest N_s is 100, where L_s is 3.

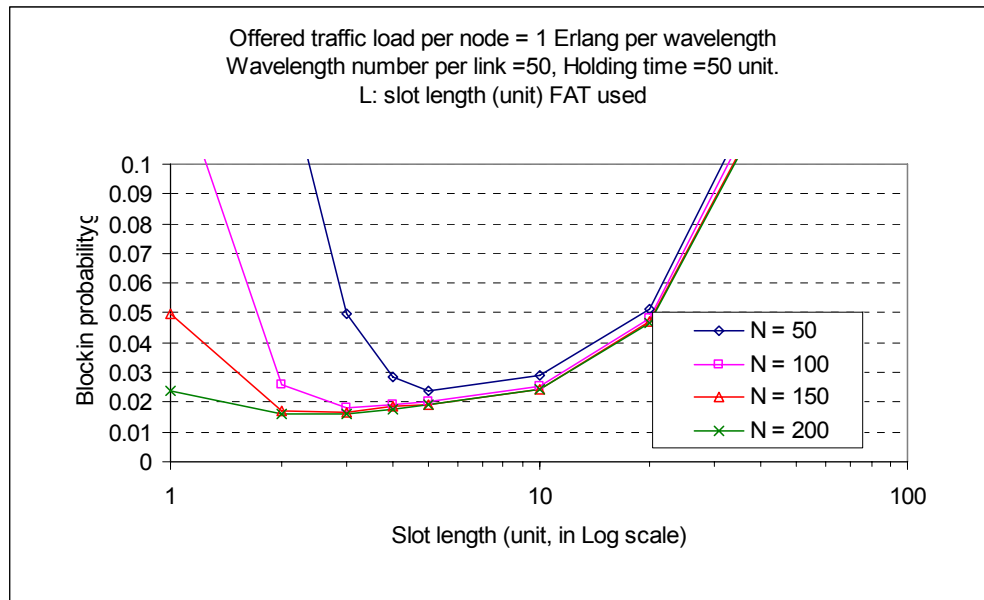


Fig. 4.12 Optimum operating region in terms of slot length and total number of slots

4.5.2 The impact of the call granularity

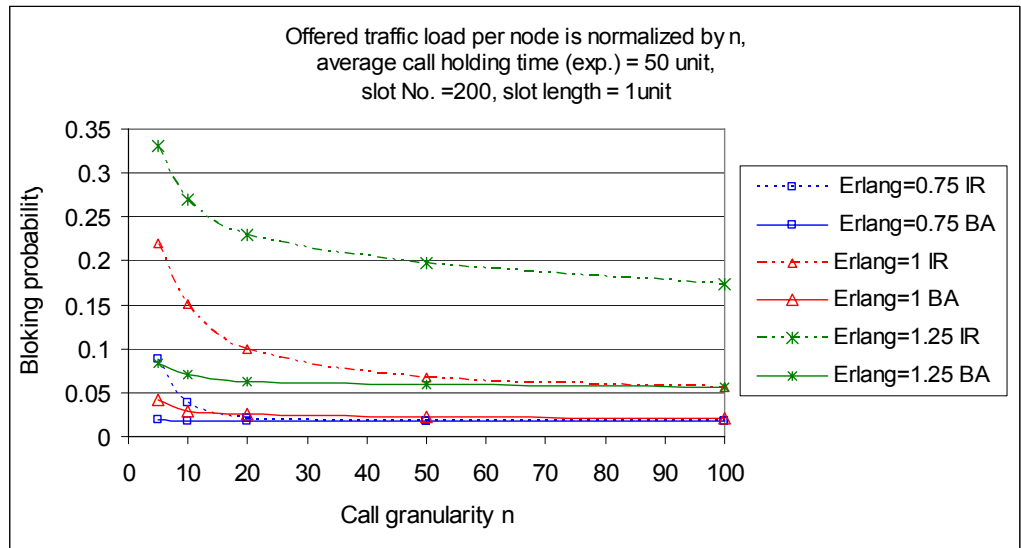


Fig. 4.13 Blocking probability vs. call granularity.

Fig. 4.13 depicts the call blocking probability vs. the call granularity for both the book-ahead (BA) calls and immediate (IR) calls at different traffic loads that is normalized by n . CSPF is used in the path computation for IR calls. At coarser granularity level, say n is less than 20, the call blocking probability decreases radically as n gets bigger. However, as n further increases to the finer granularity level, the reduction of the call blocking probability becomes small. As n gets bigger than certain value, take the 1 Erlang curve for example, this value is 60, the reduction of the call blocking probability becomes unnoticeable. Compared with IR calls, the improvement in terms of call blocking for BA calls is more dramatic in the coarser granularity level than those in the finer granularity level. It is also noticed that BA has the significant advantage over IR in the coarser call granularity region and higher offered traffic load. In the finer call granularity region, if the

network is under loaded, for example, at 0.75 Erlang per wavelengths per node, IR performs almost as good as BA in terms of call blocking probability.

4.5.3 Cost advantage of BA over IR

In DWDM system, the higher the number of wavelengths per link, the higher the cost of the wavelength multiplexer/demultiplexer and the transponders. Therefore, reducing the number of wavelengths used in the network could translate into significant cost savings. Fig. 4.14 depicts the percentage of wavelengths saved in BA scenario compared with IR. In this simulation, the number of wavelengths is obtained for both scenarios under certain offered traffic load while keeping the blocking probability below 2%. In BA the book-ahead time is set to be seven times of the average call holding time. From the figure, we can see that the percentage of savings is higher in the lighter traffic load that requires less number of wavelengths than that in the heavier traffic load that requires more number of wavelengths. As the traffic load reaches over 60 Erlang, the percentage of savings tends to be constant, which is about 17% in this case.

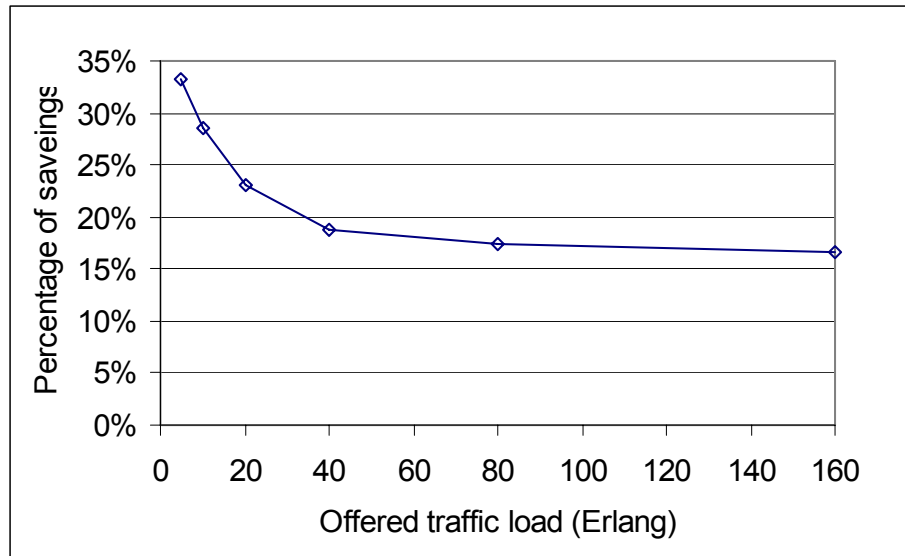


Fig. 4.14 Percentage of cost savings of BA over IR.

4.5.4 Designated start time book-ahead calls

In the previous sections, we use the first available time slot in our simulations for reserve delay-insensitive calls. However, there are applications that require specific starting time or a starting time within an acceptable period. This scenario is depicted in Fig. 4.15. The preferred call starting slot is uniform distributed between 0 and maximum starting slot number. If the call could not be reserved at the preferred starting slot, the user may specify how many slots after the preferred slot it may accept. The call blocking probability with different values of accepted slots are shown in Fig. 4.15. Obviously, the more the number of accepted slots, the lower the call blocking probability. However, there is a region in terms of maximum starting slot where lower blocking could be achieved. In this case, the maximum starting slot number should be less than two-third of the total slot number used in the reservation system.

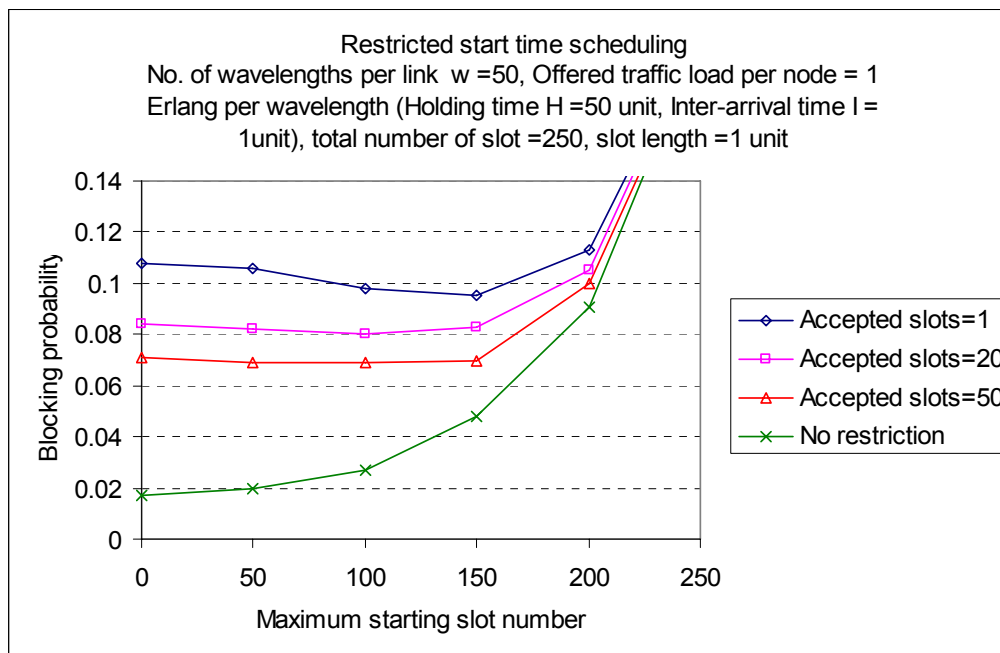


Fig. 4.15 Restricted start time scheduling

4.5.5 Gold and economy classes of BA

Introducing different levels of services is a way to boost the revenue of the service provider. In this sub-section, we study two classes of BA services. One is called gold that is illustrated in the previous section. For gold BAs, the customers specify the start time of BA calls. The other is called economy that the start time of the BA calls are not specified by the customers, but is arranged by the scheduler which could be placed anywhere in the book-head time. In the simulations, the start times of the gold services are randomly selected in the book-ahead time, and maximum 10 time-slot shifts from the start time is accepted. Fig. 4.16 shows the call blocking probabilities of gold, economy and the total calls as the gold service increases its proportion of the total calls. As the proportion of gold calls increases, the blocking of economy calls decreases and goes to almost zero if the gold calls

goes to 35% of the totals. This is due to the high blockings of the gold services decreasing the network utilization (shown in Fig. 4.5), which leaves more resources for economy services. Although the total call blocking increases and the network utilization decreases after introducing the gold services, the revenue could be significantly enhanced by charging the customer more for the gold services.

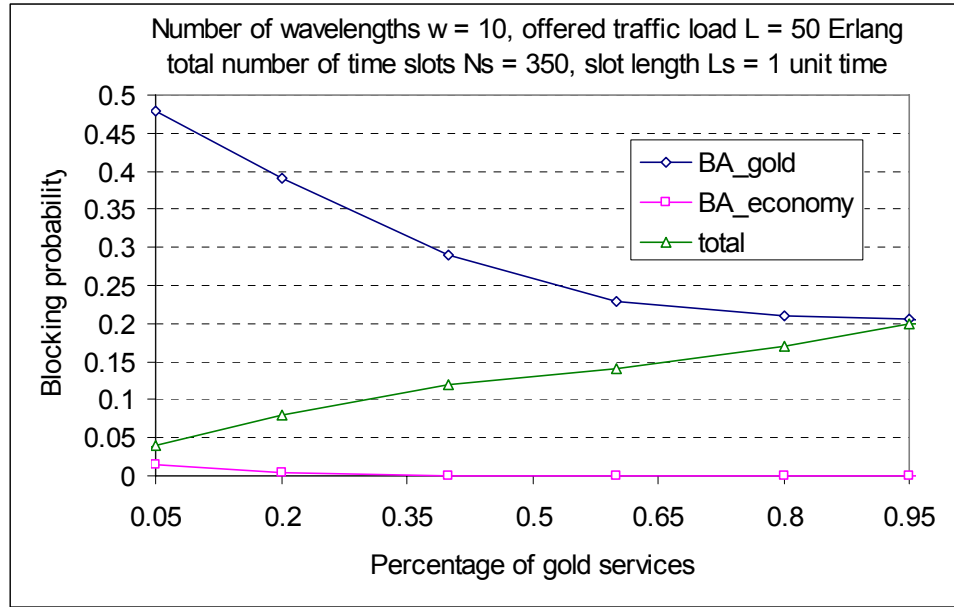


Fig. 4.16 Call blocking probability vs. different proportional of service classes

Revenue of only economy services:

$$R_{economy_only} = P_{economy} * L * (1 - PB_{economy_only}) \quad (17)$$

Where $P_{economy}$ is the price of the economy service, L is the offered traffic load, and $PB_{economy_only}$ is the blocking probability.

Revenue of both economy and gold services:

$$R_{both} = P_{economy} * L * (1 - G_p) * (1 - PB_{economy}) + P_{gold} * L * G_p * (1 - PB_{gold}) \quad (18)$$

Where $P_{economy}$ and P_{gold} are the prices for economy and gold services respectively. $PB_{economy}$ and PB_{gold} are the blocking probabilities for economy and gold services respectively. G_p is the percentage of the gold services.

The percentage of the revenue enhancement after introducing gold services is

$$R_E (\%) = (R_{both} - R_{economy_only}) / R_{economy_only} \quad (19)$$

R_E with different percentage of gold services is plotted in Fig. 4.18.



Fig. 4.17 Network utilization vs. different proportional of service classes

The higher the price for the gold services, the higher the total revenue. If the price of gold services is 1.3 times of the economy services, the break-even happens at gold calls is 45% of the total calls. The break-even point falls to 20% if the price of the gold services is 1.5 times of the economy services. If the price of the gold services is 2 times of the economy services, the revenue could be enhanced by 20% if the gold services is kept to 40% of total services.

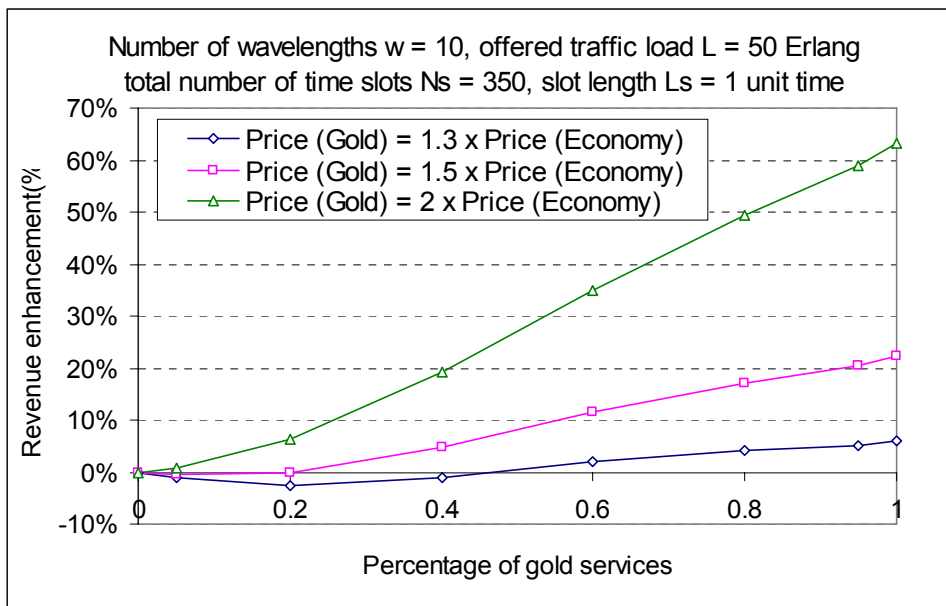


Fig. 4.18 Revenue vs. different proportional of service classes

4.5.6 Variable bandwidth reservation

From Fig. 4.1, we see that the bandwidth utilization gaps exists due to the advance reservation. If these gaps can be filled by certain applications, the network utilization will be enhanced. Delay insensitive data transfers can be such kind of applications, in which the size of the transferred data is fixed, as long as the data is transferred before a deadline, the time it takes to transfer the data is not a big concern. In such a case, during reservation, the call holding time (the data transfer time) and the reserved bandwidth can be manipulated so that the call can be fit in the gaps. For example, a call's original holding time is X for one wavelength reservation, if the number of reserved wavelengths is changed to 2 for this call, the call holding time then changed to $X/2$. Fig. 4.19 and Fig. 4.20 depict the call blocking and network utilization respectively in accordance to the maximum number of wavelength can be reserved. The variable wavelength reservation obeys

the following algorithm: first, single wavelength reservation is tried, if it failed, then reduce the call holding time by half and 2-wavelength reservation is tried. The call is blocked until the maximum-wavelength reservation fails.

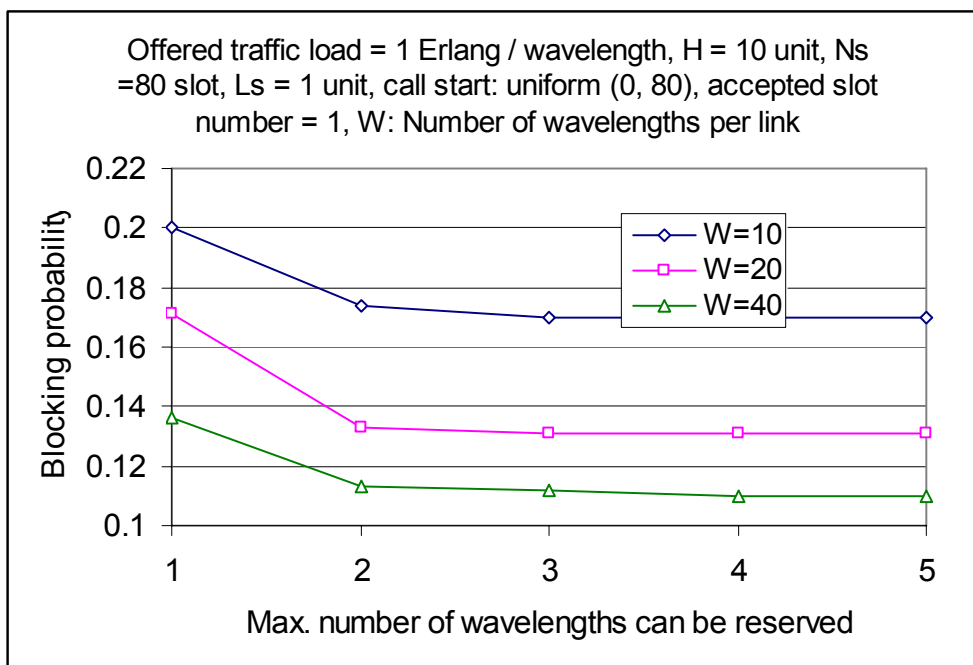


Fig. 4.19 Call blocking vs. max. number of wavelengths can be reserved

As indicated in these figures, the optimum value of maximum number of wavelengths that should be reserved is 2, at which the call blocking and network utilization reach their almost-optimum value. Take the 40-wavelength curve for example, if the maximum number of reservable wavelengths changed from 1 to 2, the call blocking decreases from 13.6% to 11.3%, which accounts for 16.9% reduction, while the network utilization increases from 50.8% to 63.2%, which accounts for 24.4% improvement. Further increase the value of maximum number of reservable wavelengths, the call blocking and network utilization do not show noticeable improvement.

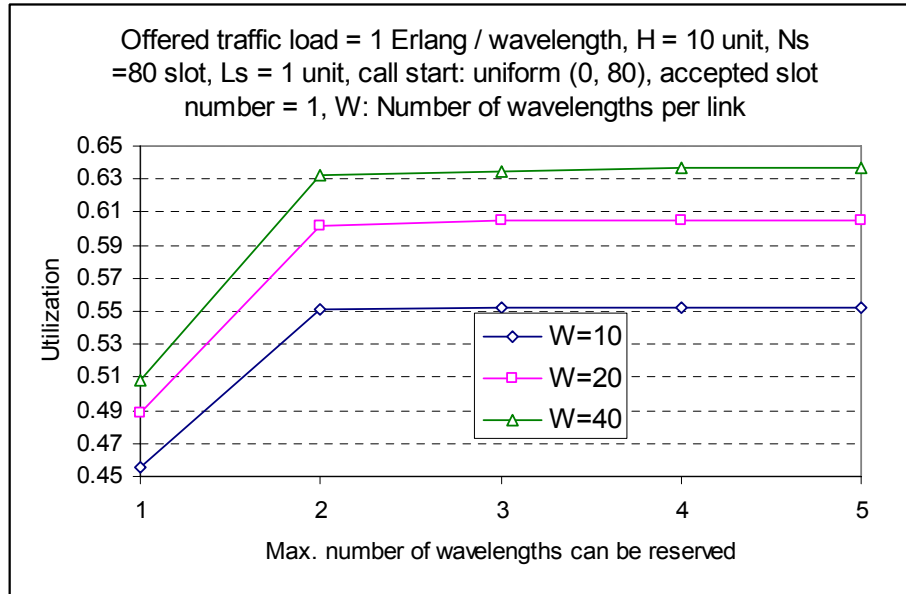


Fig. 4.20 Network utilization vs. max. number of wavelengths can be reserved

4.5.7 The impact of IR calls on BA calls

In the IP optical network, both IR and BA calls may exist at the same time. Since the IR calls tend to be those with short inter-arrival times and short holding times, the number of these calls is much more than that of BA calls. It could cause scalability problem if the IR calls' information is updated in the centralized scheduler. Therefore, for centralized scheduling approach, we assume the scheduler does not have any information about the IR calls. The IR calls at each node are set up using the existing GMPLS control plane where the routing is done by OSPF-TE and signaling is done by RSVP-TE. Please note, the node itself does not have the BA scheduling information, therefore, the IR calls may take the network resources that have been reserved by the BA calls that have not been initiated yet. In such a case, the BA calls will be blocked.

For comparison, a distributed scheduling approach is introduced where each node contains a PCE. The PCE has a global traffic engineering database which contains not only the current but also the future link utilization information reserved by the PCE. This database could be updated by the OSPF-TE. The IR call holding time is assumed to be known in advance. Upon the arrival of an IR call, the scheduler checks the network resource for future period of this call's holding time. If the resources have not been allocated by the BA call, the IR call is admitted and its link utilization is recorded by the PCE.

In the simulation, the following network and traffic parameters are adopted. Total number of wavelengths per link is 10. Total number of slots used by the scheduler is 80 slots and the slot length is 1 unit time. The offered traffic load of BA calls is normalized to 1 Erlang per wavelength with the average call holding time being 10 unit time, and the average call inter-arrival time being 1 unit time. The offered traffic load of IR calls is also normalized to 1 Erlang per wavelength per node, however, to study the impact of different types of IR calls on the BA calls, the average call holding time is varied from 1 to 100 unit time, accordingly the average call inter-arrival time is varied from 0.1 to 1 unit time. To understand the impact of the IR calls to the BA calls, the scenario that only BA calls exists is simulated, whose blocking probability is 2.5%.

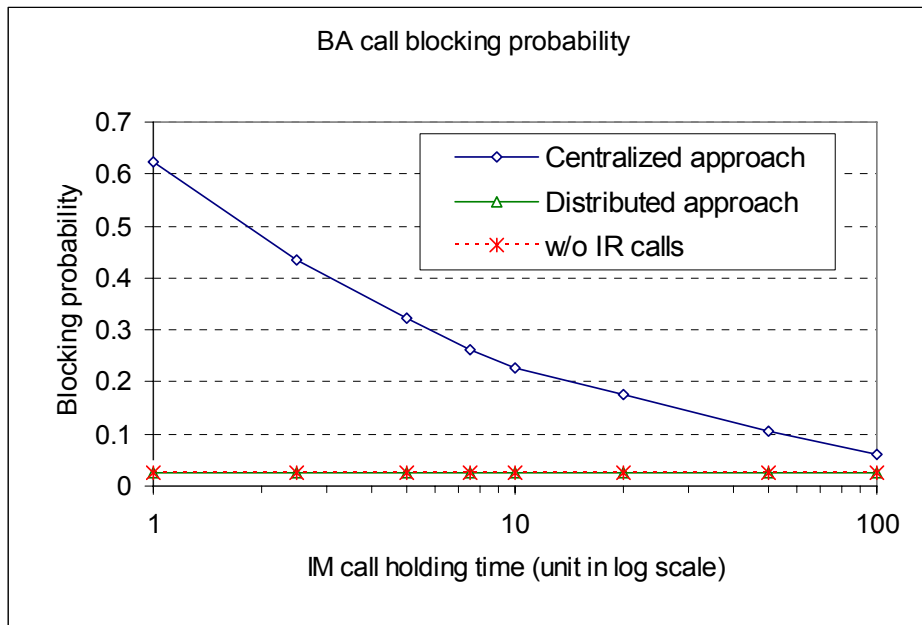


Fig. 4.21 The blocking probability of BA calls

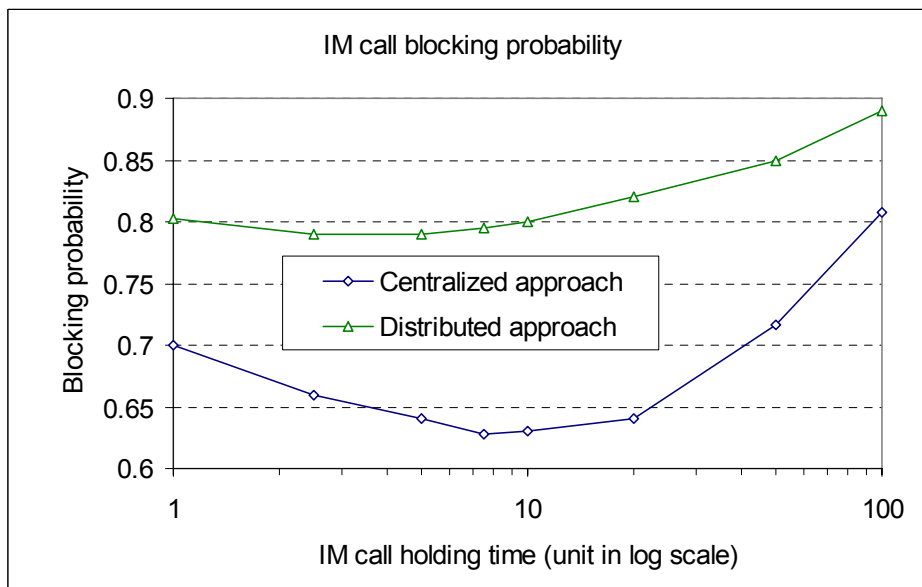


Fig. 4.22 The blocking probability of IR calls

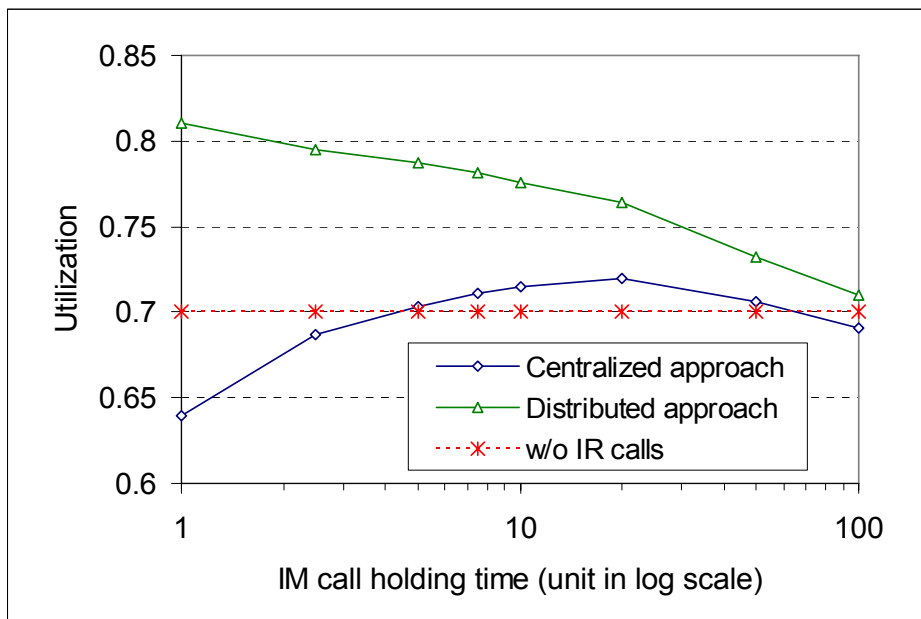


Fig. 4.23 The network utilization

Fig. 4.21 shows the blocking probability of BA calls of the three scenarios. As expected, the centralized approach has the highest blocking, since the stored future link availability can be altered by the IR calls. There are two reasons the scheduled BA calls are blocked. On one hand, the future link capacity can be changed just before the path computation is made due to the IR calls, which made the path computation based on the un-updated link capacity information. On the other hand, just before the BA calls are initiated, the IR calls may occupy the resources that reserved for the BA calls, resulting in the BA calls being blocked. The higher call blocking is mainly caused by the former reason. The more the IR calls, and the shorter the IR call holding time, the less inaccurate of the centralized future link utilization database and the higher the call blocking. In the distributed approach, since each node has both the IR and BA calls' information, the IR calls can also be viewed as the scheduled calls, with the book-ahead time equals to zero. The

blocking for BA calls stays almost the same as that of the scenario no IR calls exist. Only if the IR call holding time is longer than that of the BA calls, the BA call blocking started to increase by 1%, which is unnoticeable in the figure. The increase in blocking of the BA calls is due to the reduced network resources that introduced by the “scheduling” of IR calls.

Fig. 4.22 and Fig. 4.23 depict the call blocking probability for IR calls and the total network utilization respectively. Note that in these scenarios, the BA calls’ traffic load stays the same, therefore, the variation of the network utilization is purely introduced by the IR calls. For the reason that the IR calls can take the resources whenever they are available, IR calls in centralized approach has the lower call blocking probability than that of distributed approach. However, the IR calls also blocks many BA calls and the utilization gain obtained by admitting IR calls could not compensate the loss introduced by blocking BA calls. Therefore, Centralized approach has the lower network utilization.

In distributed approach, since only the gaps in the link utilization could be allocated for IR calls, BA calls will not be blocked. The blocking for IR calls is high, but due to the gaps are filled with IR calls, the network utilization is high, and the shorter the holding time of the IR call, the better the utilization. It is also noticed that the IR blocking curves of the two scenarios in Fig. 4.22 exhibit more or less the U shapes. At the region where the IR call holding time is slightly smaller than that of BA call, the IR call shows lower call blocking. This is because that the average link utilization gap lengths left by the BA calls’ reservation are shorter than

the average holding time of the BA calls, and the IR calls with the slightly smaller holding time have more chances to fill in these gaps, resulting in lower blocking.

In centralized approach, at the region where the IR calls' holding time is greatly smaller than that of BA calls, which means the IR call inter-arrival time reduces and the number of calls increases, but the number of gaps stays the same, this leads to only a small portion of IR calls fill in the gaps, resulting in higher blocking. Due to the admitted IR calls' holding time is short, their contribution to the network utilization does not compensate the loss by blocking the BA calls, and the network utilization is reduced. At the region where the IR calls' holding time is greatly larger than that of BA calls, the chances for IR calls to fill in the gaps are also reduced, resulting in higher call clocking as well. For the same reason mentioned above, if the IR call holding time is further increased above certain value, in our case, the value is around 60 unit time (6 times of the BA call holding time), the utilization gain obtained by IR calls does not compensate the loss of blocking the BA calls, the network utilization becomes less than the scenario without the IR calls.

Chapter 5

Design of Control Plane in CHEETAH Network

In this chapter*, we describe the design and implementation of a GMPLS software engine that is used to enable non-GMPLS switches (e.g., Ethernet switches, routers and other cross-connects) to participate in the dynamic setup and release of end-to-end bandwidth guaranteed connections. We have deployed this engine in the experimental CHEETAH optical network, and hence we call it the Cheetah Virtual Label Switching Router (CVLSR). It allows users to share the network resources dynamically and makes it possible to extend the connections across the local area network to the end-users. The CVLSR engine is also equipped with a function called dynamic network clustering which is needed to enable a broad range of applications such as remote visualization, grid computing and e-learning.

5.1 CHEETAH Network

Recent advances in optical networking technologies have made it feasible to provision bandwidth on-demand (BoD) services directly to the end-users. Specifically, the GMPLS based distributed control plane is a possible enabling technology that is used to set up and release dynamically provisioned deterministic

* The work presented in this chapter was funded by NSF ANI-0312376, NSF ANI-0335190 grants, and carried out together with my colleague Ms. Zhaoming Li.

bandwidth end-to-end connections across layers' 3, 2 or 1 networking devices. Recently, several experimental optical networks have been deployed to study and demonstrate the efficacy of GMPLS in enabling BoD services including e-Science, large data transfers (e.g., multi-Gigabytes and above) among others. The Circuit-switched High-speed End-to-End Transport ArchItecture (CHEETAH) is an example of such experimental efforts.

The CHEETAH project is a comprehensive effort to develop the infrastructure and networking enabling technologies to support a broad class of eScience projects and specifically the Terascale Supernova Initiative (TSI) [54]. Its data plane consists of 10 Gbps Ethernet and OC-192 SONET connections interconnecting several "enterprise networks" sites like the supercomputer at ORNL with the Centaur laboratory at North Carolina State University (NCSU) where eScience applications run, see Fig. 5.1. Other connections include several 1 Gbps Ethernet links to enterprise networks users as outlined in the figure. Three Sycamore SN16000 SONET cross-connect switches (SONET-XC) form the "backbone" of the network, whereas users, in each enterprise network, gain access to the network via gigabit Ethernet switches such as the Foundry switch FEX424.

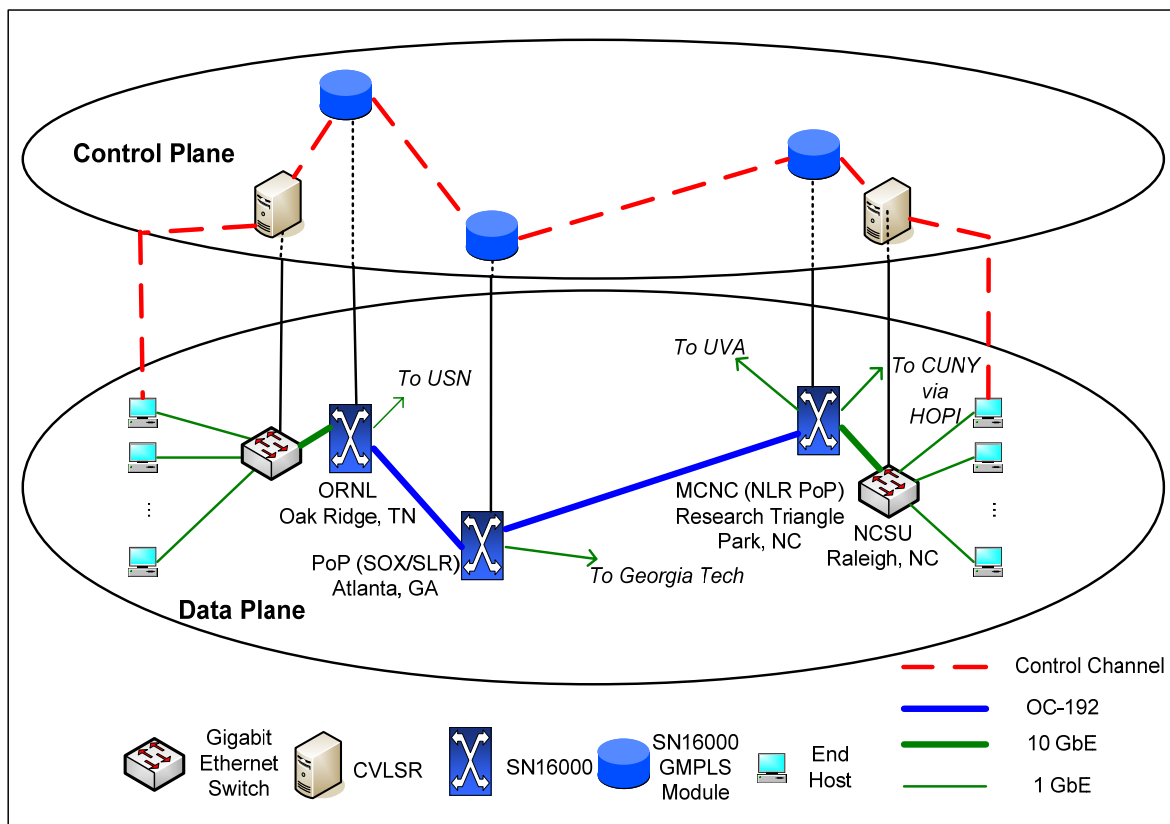


Fig. 5.1 CHEETAH network Test-bed

CHEETAH uses a distributed GMPLS control plane to enable application-initiated, on-demand dynamic provisioning of end-to-end circuits at various granularities (e.g., 10 Mbps to 10 Gbps). An important premise is the dynamic sharing of the backbone large bandwidth pipes by many users, thus cutting down the expensive cost of networking at these capacities (e.g., 10Gbps). Each end-host has two interface cards: a primary 1 Gbps for data plane connectivity, and a secondary one for TCP/IP connectivity as well as for transport of control plane signaling messages. Each end-host is equipped with an RSVP signaling engine in order to generate PATH messages requests to the GMPLS engine on the next hop.

In order to extend the end-to-end connection to the end user, it is necessary for the Ethernet switch to be equipped with a GMPLS engine, which we call the

Cheetah Virtual Label Switching Router (CVLSR). It is developed based upon the Dragon VLSR engine, which is based upon an open source OSPF [55] and RSVP engines [56]. The concept of CVLSR can be extended to include any router, or cross connect switch.

5.2 CVLSR DESIGN AND IMPLEMENTATION

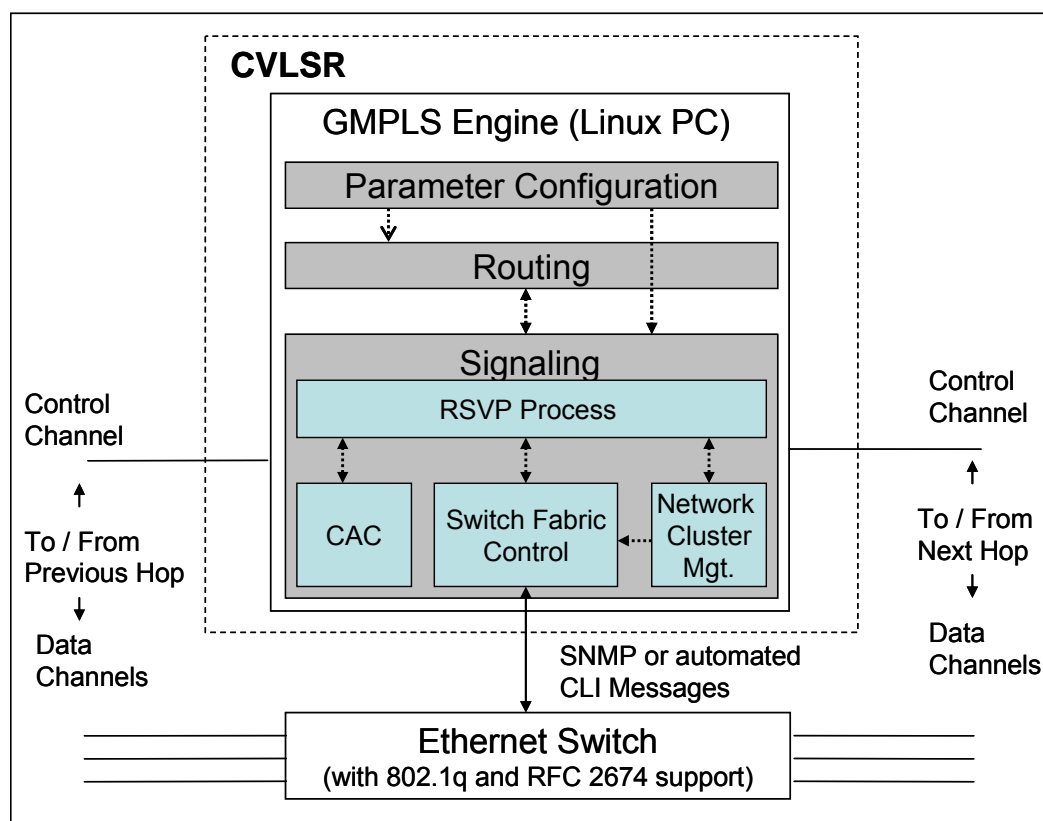


Fig. 5.2 The CVLSR software architecture

We recall that the main capabilities of the GMPLS control plane are routing, signaling and auto-discovery [57]. In our implementation, the CVLSR contains three basic modules (shown in Fig. 5.2): parameter configuration, routing and signaling. Routing is based upon Open Shortest Path First - Traffic Engineering (OSPF-TE) whereas signaling uses Resource ReSerVation Protocol - Traffic

Engineering (RSVP-TE). The CVLSR does not have LMP capabilities; hence, physical and TE links' configuration parameters are obtained through manual configuration. The CVLSR also includes a function that we call dynamic network cluster function that is needed to enable a broad range of applications, such as remote visualization, grid computing and e-Learning. The network cluster management is implemented as a sub-module in the signaling module. In the following sub-sections, we describe the details of each of the modules.

5.2.1 Parameter configuration

The Parameter configuration module is a user interface for the administrator to configure network parameters for CVLSR, such as the local and remote control channel ID (IP address or interface index in case of unnumbered interface), local and remote data-interface ID, and TE information including the interface switching capabilities and maximum reservable bandwidth, etc. These configuration parameters are stored in configuration files and read into corresponding data structures by the routing and signaling modules. The following relationship mappings are obtained from these data structures.

- 1) Control/data interfaces mapping. The administrator enters the data interfaces IDs. The same is done for the control channels. Each control channel is then correlated to a number of parallel data channels.

- 2) Data interface ID/port number mapping. Similar procedure to the above happens here, except that it is now an association of a single port number to a single data interface ID.

3) Local/remote interface mapping. This mapping provides the association between the local and remote interfaces for the control and data channels.

4) Cluster member ID/port number mapping. This is the association between the cluster member ID and the physical port number of the Ethernet switch.

Since the end-hosts in CHEETAH are not equipped with OSPF, in order to avoid overloading the processor, additional TE information (control channel and data channel interface ID) of the end-hosts' IP addresses are also defined in the parameter configuration module.

5.2.2 Routing Module

The CVLSR uses OSPF to build its routing table and traffic Engineering (TE) link state database. The implementation supports the following IETF requests for comments: RFC 2328 [18] , RFC 2370 [19], RFC 3630 [28], and RFC 4203 [5].

The TE link state database is available for intra-area routing, whereas only the reachability is available for inter-area routing. In the latter case, hop-by-hop routing is used to find the next hop according to the routing table. In the former, an end-to-end path is computed via CSPF using the TE-link state database. Note that in source routing intra-area OSPF-TE is used hence all TE-links are configured in the same OSPF routing area.

The module processes requests from signaling module and returns the corresponding results. For example, it returns the local data interface ID according to the remote data interface ID and vice versa. Similarly, it returns the control interface ID according to the data interface ID and vice versa. It also returns the loopback address of the current CVLSR for RSVP to construct the HOP object. It

also updates the bandwidth information of the related TE LSA in the link state database based on the RSVP bandwidth reserve/release notifications. The module also advertises the end-hosts reachability information to other GMPLS engines. Finally, it supports control channel separation, which is defined in RFC 3945 [8]. Hence, a single control channel could be associated with multiple parallel TE-links.

5.2.3 Signaling Module

The signaling module is composed of the following four sub-modules: RSVP process, switch fabric control, connection admission control (CAC), and network cluster management.

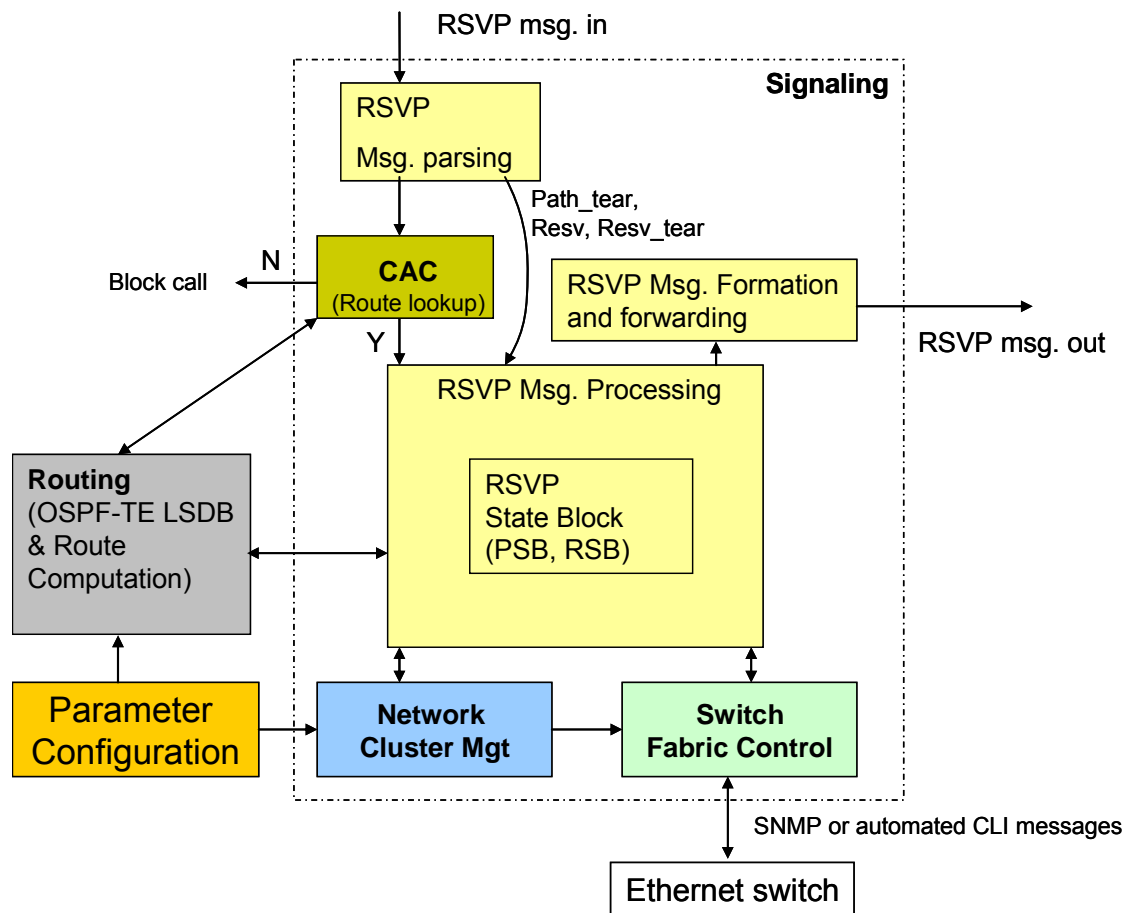


Fig. 5.3 RSVP messages functional flow diagram

Fig. 5.3 illustrates the RSVP messages functional flow diagram among the above-mentioned sub-modules. An object-relationship model is used for data exchange among these sub-modules. The data exchanges with the routing module, such as routing computation, interface mapping, and TE-Link state database updating is realized through a socket.

5.2.3.1 RSVP Process

The functionalities of RSVP process sub-module include (a) Sending, receiving, parsing, processing and constructing RSVP messages; (b) building RSVP state blocks; (c) sending queries to the routing module to obtain the next hop IP address, for hop-by-hop routing, or explicit route object (ERO), for source routing; (d) informing the routing module to update its database when bandwidth is reserved/released; and (e) communicating with routing module to obtain the manually configured parameters that are read by the routing module from the OSPF configuration file.

Our implementation supports the following set of IETF RFCs: 2205 [21], 2209 [22], 2210 [23] and 3473 [6]. In an Ethernet switch, sub 1 Gbps connections are setup/released using tagged VLANs. Hence, RSVP must support Layer 2 switching capability (L2SC) which is used to identify sub 1Gbps connections, whereas fiber switching capability (FSC) is used for full 1Gbps port mapping.

5.2.3.2 Connection Admission Control

Upon receiving a PATH request, CAC decides whether to accept or reject the request based upon the availability of the bandwidth and the switching capability of the TE-links.

In the hop-by-hop routing, CAC sends a query to the OSPF-TE link state database asking for the TE information of the links between the Ethernet switch and the neighboring SONET-XC switch. If the unreserved bandwidth satisfies the bandwidth requirement, and the link switching capability matches the “switching type” in the Lable_Rsrequest object in the PATH message, the PATH request is granted and the connection is admitted. Otherwise, the request is rejected. In source routing, CAC sends a request to the OSPF routing module to obtain the ERO through CSPF path calculation with the bandwidth and switching capability constraints. If an ERO is successfully returned, the request is granted. Otherwise, the request is rejected.

5.2.3.3 Switch Fabric Control

To setup a VLAN, the switch fabric control sub-module sends SNMP messages to the MIB of the Ethernet switch to place the input and output ports of the switch into a tagged or untagged VLAN. Following the VLAN setup, the switch fabric control sub-module sends control messages to the switch to put the VLAN in a higher priority queue in order to guarantee the bandwidth of the VLAN as well as rate-limit the input port/VLAN so that the VLAN traffic does not exceed the its specified bandwidth (i.e., policing).

5.2.3.4 Network Cluster Management

A dynamic network cluster is defined as a group of end-hosts in a single location sharing a dedicated bandwidth connection with one or more end-hosts in another single location (see Fig. 5.4). An end-host that is authorized to use the clustering function, can initiate (or join) a new (or existing) cluster connection

without any prior manual configurations. A network cluster connection is thus a point-to-point one whereby each point is an Ethernet switch or similar switching device. A single end-host can be a participant in multiple clusters. For example, end-host A2 can be a member of both Clusters 1 and 2 simultaneously. This is achieved using globally tagged VLANs. Once a cluster connection is established, new members could be added and existing members could be removed. The clustering function is needed in order to enable applications requiring several end-hosts to share the same connection. This scenario arises in grid-computing, e-learning, and video-on-demand applications among others. Provisioning connections (at the application level) for such applications could be problematic unless the network can provide a functionality such as the network cluster one. The setup and release of the cluster connection would then relieve the application from the burden of performing a similar function at the application layer.

To implement the dynamic network cluster management we have defined three new RSVP objects: CLUSTER, VLAN_ID_Set and GLOBAL_VLAN_ID with the unknown class. The node that does not recognize these objects ignores these objects but forwards them, unexamined and unmodified, in all messages resulting from this message. The CLUSTER object stores the hosts' IP addresses of both the local and remote clusters. The end-host who initiates the cluster setup determines the initial members of the cluster and hence constructs the CLUSTER object in the PATH message and sends it to the CVLSR. This CLUSTER object is saved in the corresponding LSP path state and forwarded in any refresh message resulting from that state. After receiving the RESV message, the path state is

examined. If the content of the CLUSTER is empty, normal procedure is performed, i.e. only the input port and output port are placed in the same VLAN. Otherwise, the content of the CLUSTER object is checked against cluster member ID/port number mapping to find out the port numbers connected to the cluster members. These ports are placed in the same VLAN. The network cluster management module then calculates the bandwidth for each cluster member and informs the switch fabric control module to perform rate-limiting for these ports accordingly.

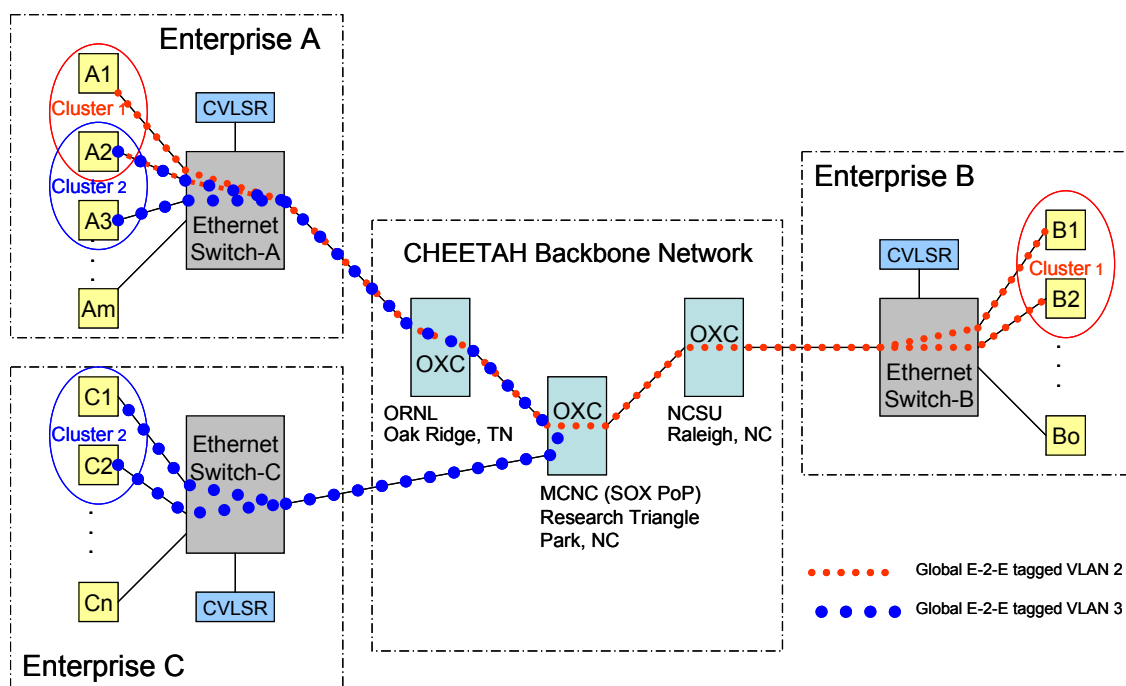


Fig. 5.4 Dynamic cluster illustration

The VLAN_ID_Set and GLOBAL_VLAN_ID objects are used to generate the VLAN ID for global end-to-end VLAN across the CHEETAH network. The VLAN_ID_Set object in the PATH message is used to store the working VLAN IDs which cannot be assigned for new VLANs. It limits the VLAN ID choices of a

downstream node to a set of acceptable VLAN IDs. Each node generates its own outgoing VLAN_ID_Set, based on the combination of the incoming VLAN_ID_Set and its own working VLAN_ID_Set. The receiver of a VLAN_ID_Set must restrict its choice of a VLAN ID to a one that is not in the VLAN_ID_Set. The VLAN_ID object is placed into the RESV message to instruct the upstream nodes to use the same VLAN ID specified in the VLAN_ID object for the VLAN setup.

An Operational Example:

1) A user decides who should be in the cluster in both the local enterprise and the remote enterprise. The end-host gets these hosts; IP addresses through OCS/DNS queries. These IP addresses are stored in the CLUSTER object of the Path message and sent to the local CVLSR. One of the IP addresses of the remote cluster member is used as the destination IP address in the Path message.

2) The local CVLSR receives the Path message and save the CLUSTER object in the Path state. Assume at the time the Path message is received, the CVLSR is using two VLANs (VLAN 2 and VLAN 3) for other CHEETAH circuits. VLAN ID 2 and 3 are stored into the VLAN_ID_Set object and is sent by the Path message to the SN16k.

3) The SN16k finds out that the Path message it received has an unknown object (CLUSTER object) with class number x11bbbbbb. It forwards the Path message without modifying the Cluster object (according to unknown object process rules defined in RFC 2205 [21]) to the downstream SN16k. The downstream SN16k will

perform the same procedure as the upstream SN16k did and forward the Path message to the remote CVLSR.

4) The remote CVLSR receives the Path message and save the CLUSTER object in the Path state. Assume at the time the Path message is received, the remote CVLSR is using three VLANs (VLAN 4, 5 and 6) for other CHEETAH circuits. VLAN ID 4, 5 and 6 are put into the VLAN_ID_Set object and is sent by the Path message to the receiver CHEETAH host.

5) The receiver CHEETAH host gets the Path message, selects a VLAN ID 7 that is not included in the VLAN_ID_Set which has VLAN ID from 2 to 6, puts this VLAN ID 7 in the "VLAN_ID" object in the Resv message, and sends it to the upstream node.

6) The remote CVLSR receives the Resv message, examines the Path State, if the CLUSTER object stored in the Path state is empty, the RSVPD performs normal reservation. Otherwise, RSVPD check the IP addresses in the Cluster object against the hosts' IP addresses in the cheetah.conf, if there is a match, the port of the Ethernet switch that connected to that host is located. The switch fabric control module in the CVLSR sends SNMP messages to place all the ports connected to the cluster members to the same VLAN, places the VLAN is a higher priority queue to guarantee the requested bandwidth, and rate-limit the port/VLAN.

7) The local CVLSR does the similar procedures as the remote CVLSR and sends the Resv message the sender host.

8) The sender host receives the Resv message. The cluster setup is finished.

5.3 CVLSR Deployment

Fig. 5.5 shows the CVLSR deployment in the CHEETAH network. Currently the GMPLS Overlay model is used, however, the Peer model is also supported. The former model is used whenever it is desired to hide the details of the CHEETAH topology from the enterprise switch. The latter is used whenever the switch is within the same administrative domain of the CHEETAH network. Between the Sycamore nodes in the CHEETAH network, Sycamore uses a proprietary GMPLS based routing and signaling protocol, “BroadLeaf”, to setup connections between those nodes. Hence, the backbone network’s topology information and links bandwidth information are not advertised, thus hop-by-hop routing is used by the CVLSR to find the next hop. Out-of-band control channels are setup by IPIP tunnels across the Internet. IPsec is used to secure the control messages, i.e. the OSPF and RSVP messages.

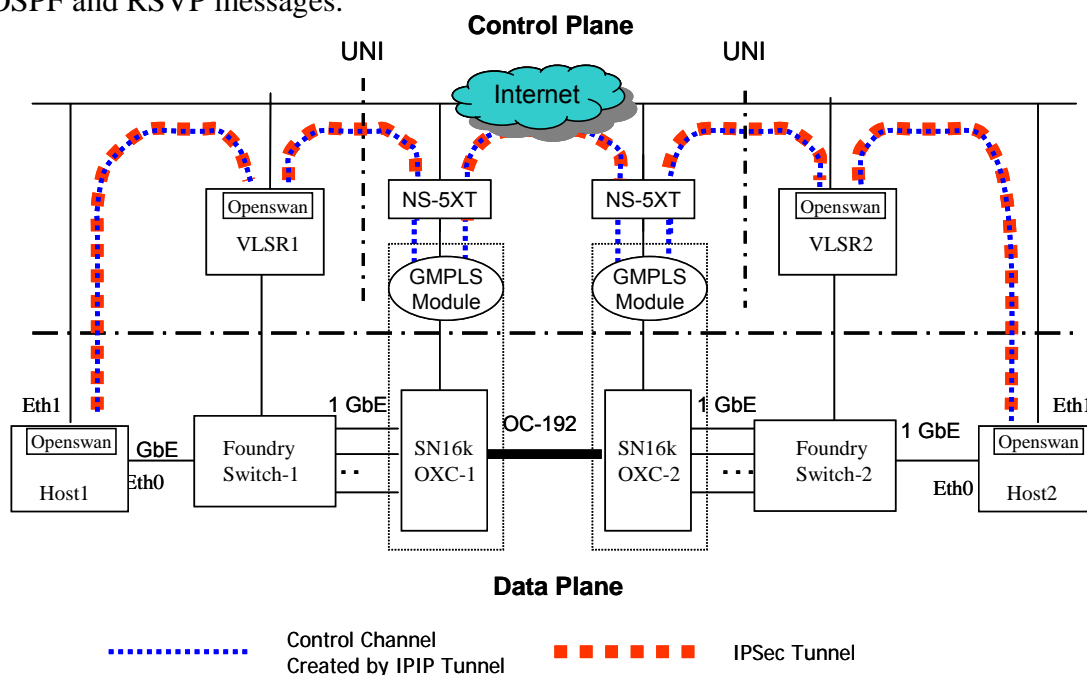


Fig. 5.5 CVLSR deployment in CHEETAH network

Deploying the CVLSR in the CHEETAH control plane consists of the following steps:

1) Setup control channels. The control channels are used for OSPF-TE and RSVP-TE control messages, such as OSPF Hello messages, OSPF link state advertisements, and RSVP PATH and RESV messages, etc. The control channels are created by configuring IPIP tunnels between the control Interfaces. These control interfaces could be configured to be numbered or unnumbered interfaces. The unnumbered interface configuration has the benefits of reducing management overhead and saving the IP addresses resources. In the CHEETAH network, the SN16000 uses unnumbered interfaces for both control and data channel interfaces. Due to the Linux kernel limitation, the control interface of the CVLSR has to be numbered interface.

2) Configure the data channel Interface (TE-link configuration). Unnumbered interface configuration is used for CVLSR. Each pair of data interfaces between the CVLSR and the host or SN16000 represents one TE-link. The TE-LSAs containing the TE-links' available bandwidth and switching capability are flooded by the link state advertisement.

3) Configure other parameters, such as control/data interfaces mapping, data interface ID/port number mapping, Local/remote interface mapping and Cluster member ID/port number mapping.

Chapter 6

Conclusions and Future Directions

In this chapter, we begin with the summary of the thesis, and then present our main contributions in the conclusion part. Finally, we list several areas that worth being investigated in the future.

6.1 Summary of Thesis

The thesis consists of two parts: the analytical part, and design and development part. The analytical part (Chapter 3, 4) focuses on the performance analysis of the GMPLS control plane. The limitation, the optimum operation region and the value added proposition of the control plane are identified. The design and development part (Chapter 5) emphasizes on establishing a GMPLS network control plane which support end-host initiated bandwidth guaranteed connection. The design of the GMPLS engine (CVLSR) and the control plane setup are highlighted.

The thesis begins with Chapter 1, in which the motivation of the thesis is presented. The author augured that the traditional centralized OSS network management approach could not satisfy new functional requirements to providing new services, e.g. bandwidth-on-demand, grid computing, and high-definition video conferencing, etc., which are introduced by the fast growing of the Internet. These new functional requirements, e.g. fast provisioning, auto-discovery and rapid

restoration, could be realized by the GMPLS control plane. However, due to the limited research on the performance of the GMPLS control plane and the cost issues, the operators are hesitated to embrace it into their network operations.

To better understand the GMPLS control plane, in Chapter 2, the centralized OSS approach is reviewed. The building blocks of GMPLS protocol suits and their functionalities are highlighted. Network simulation is used to qualify the cost savings and revenue enhancement brought by the GMPLS control plane. The simulation methodology is introduced in this chapter.

The simulation results of the GMPLS control plane are presented in Chapter 3. The routing, signaling, and protection and restoration are analyzed respectively. In the routing part, the OSPF-TE overhead traffic, the impact of the traffic load and the call inter-arrival time on the call blocking probability are quantified. The advantages of dynamic routing over static routing, the impact of the granularity on the dynamic routing and the comparison of distributed routing with the centralized routing are presented. In the signaling part, the impact of the connection setup delay on the network performance is illustrated. Several delay contributors, such as propagation delay, signal processing delay and OXC cross-connection delay are analyzed respectively. In the protection and restoration part, different protection and restoration schemes are described. The performance characteristics and benefits of using control plane in mesh network for protection and restoration services are obtained through simulations. The distributed protection and restoration approach is compared with the centralized approach.

Chapter 5 provides an analysis of the book-ahead reservation in the connection-oriented networks. Book-ahead reservation is not yet a feature in the standard GMPLS control plane. However, the network resources can be greatly enhanced by the book-ahead reservations. A heuristic path computation approach is applied to the time-slot based book-ahead advance reservation system for the analysis. Several book-head scenarios, such as different call granularity reservation, designated start-time reservations and variable reservations are discussed. The impact of the immediate reservations on the book-ahead reservation is also analyzed.

Chapter 6 presented our experience in setting up a GMPLS network test-bed. The focus of this chapter is the CVLSR design and deployment. The building modules of the CVLSR and their functionalities are described. The architecture of the test-bed control plane is presented.

6.2 Conclusion

It is important for carriers to quantify the value-added proposition of deploying GMPLS based control plane. Deploying a complex control plane could result into an expensive proposition with little value-added. The capabilities of the control plane will depend on the set of services offerings that a carrier wishes to enable. Therefore, the advantages of the control plane functionalities, such as distributed dynamic routing, connection control, shared protection, rapid provisioning, auto-discovery, and distributed fault management, should be clearly quantified. Two main goals for maximizing these benefits are identified: 1) decreasing operations costs and 2) increasing revenues via enhancing the efficiency of existing network assets as well as via enabling new services. Our research and development on the

GMPLS control plane reveal that (numerical results are based on NSFNET topology):

Distributed dynamic routing utilizing OSPF-TE is not always efficient for the purpose of realizing better links efficiencies, or decreasing call blocking probabilities. There exists an operating region in which it has better performance over static routing, i.e. the traffic loads should be less than 0.8 Erlang/per wavelength. The number of wavelengths per link also has an impact on the performance of distributed dynamic routing. The smaller the number of wavelengths, the better the distributed dynamic routing performs over static routing. As the number of wavelengths per link increases above 30, the performance difference between distributed dynamic routing and static routing tends to be minimum and constant. Compared with static routing, distributed dynamic routing can save 15% wavelengths per link, or increase the traffic load by 20%.

The OSPF convergence delay limits the type of services the carriers could provide. As the average call inter-arrival time is shorter than 1 minute, the impact of the OSPF convergence delay on the performance (in terms of call blocking ratio) of the distributed dynamic routing cannot be ignored.

In BoD services, the level of call granularity plays an important role to achieve high network unitization, i.e. increases the operator's revenues. There is an optimal region for the call granularity to achieve maximum utilization and keep the network scalable. This optimal call granularity is 1/60 of the link bandwidth.

Distributed approach has its advantages in boosting capacity over centralized approach mainly in the bandwidth-on-demand (BoD) services, i.e. shorter inter-arrival times and holding times. As the inter-arrival times and holding times become longer, distributed approach gradually loses its advantages. For EPL services, both distributed and centralized approaches exhibit almost the same performance. If a carrier is interested only in EPLs, then doing routing via a dynamic distributed approach is not a useful proposition. It is recommended to utilize a hybrid approach to routing that combines a centralized routing (done off-line in the OSS) for capacity planning, and long-term traffic engineering. Detailed local routing decisions inside a vendor sub-network can be implemented using dynamic routing subject to the routing constraints provided by the OSS.

The RSVP-TE message processing delay and OXC cross-connection delay should be kept to a minimum to reduce the connection setup delay, which has an impact on the network call blocking ratio. If the call inter-arrival time is within tens of seconds, the call blocking ratio will dramatically increase.

Assuming only single failure occurs, shared protection can achieve 100% recovery at the expense of 10% extra wavelengths over the non-protection scheme, and save up to 48% of wavelengths over the 1:1 dedicated protection. Both 1:1 and shared protections have the same capacity to accommodate the pre-emptable traffic. The drawback of M:N protection is that its restoration time is much longer than that of 1:1 protection. For the dynamic restoration algorithm, it is not economically feasible to have a 100% recovery. At a blocking rate of 10%, we can operate at an efficiency factor of 0.75 Erlang/wavelength for a 4-wavelength-per-link network.

At this point, the recovery rate will be approximately 65%. Increasing Node degrees is more efficient than increasing the number of wavelengths.

In order to achieve lowest blocking probability for the book-ahead reservations, the look-ahead time in the advance reservation should be 7 times of the average call holding time. The heuristic path calculation algorithm developed in this thesis can reduce blocking by 36% compared with the traditional SPF algorithm. If the initial reservation failed, allowing the recourses to be reserved at two times of the requested bandwidth will maximize the network performance. To maximize the network utilization, distributed scheduling should be considered and the holding time of the immediate request should be short.

To provide the end-to-end connection-oriented services in a cost-effective manner, dynamic sharing of bandwidth is critical. Our implantation of the GMPLS engine, CVLSR, provides non-GMPLS devices, such as Ethernet switches, with capabilities to dynamically provision of guaranteed bandwidth pipes. The CVLSR offers a scalable and cost-efficient solution for adding users to the network while adding users directly to PoP nodes will quickly exhaust the capacity of the ports. It provides the capability to connect CHEETAH to HOPI network and other networks reachable via HOPI. In addition, the CVLSR offers a means for simplifying the implementation of signaling heterogeneous paths by relieving the end-hosts from extra routing and signaling processing burden. The TE-database and RSVP state block maintained by CVLSR provides a suitable environment to add peering functions across multiple networks, such as path computation, accounting, security, and administration. Finally by utilizing dynamic setup and control of VLANs

within the switch, dynamic clusters applications at sub-1Gbps can be realized to further increase the flexibility of the connection control and support more services, such as e-learning, grid-computing and video-conferencing and video-gaming.

6.3 Future Works

The research and development of the IP optical network started only in the recent years. There are still a lot of issues that are worth of being investigated. Based on the work presented in this thesis, we illustrated several areas for future work.

- Inter-area and inter-domain traffic engineering. So far, our routing analysis is based on a single OSPF routing area. The TE information can only be flooded within the single routing area. Although the IETF has published several RFCs regarding to the general requirements on inter-area traffic engineering [58], inter-domain traffic engineering [59][60] and inter-autonomous System traffic engineering [61], How to disseminate TE information and what kind of TE information should be disseminated in the CHEETAH network remains a question.
- Signaling for heterogeneous path setup. The current signaling implementation in CVLSR only supports the path setup within a single domain. As indicated in Chapter 5, CHEETAH network is across several network domains. Certain links within the CHEETAH network has to be static configured. This certainly limited the network flexibility and reachability. There are several current IRTF drafts regarding to the requirements and procedures in setting up path across multi-region and

multi-layer networks [62][63]. Implementing these requirements and procedures in the CVLSR for heterogeneous path setup will significantly enhancement the functionalities of CVLSR.

References

- [1] Greg M. Bernstein, Jennifer Yates and Debanjan Saha, "IP-Centric Control and Management of Optical Transport Networks", IEEE Communications Magazine, Oct. 2000.
- [2] Chadi Assi, etc., "Optical networking and real-time provisioning- an integrated vision for the next-generation Internet", IEEE Network, Jul./Aug. 2001.
- [3] Guangzhi LI, Jennifer Yates, etc., "Control Plane Design for Reliable Optical Networks", IEEE Communications Magazine, Feb. 2002.
- [4] Ayan Banerjee, John Drake, and Jonathan P. Lang, etc., "Generalized Multiprotocol Label Switching: An Overview of Routing and Management Enhancements", IEEE Communications Magazine, Jan. 2001.
- [5] K. Kompella, Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, Oct. 2005.
- [6] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, Jan. 2003.
- [7] J. Lang, "Link Management Protocol (LMP)", RFC 4202, Oct. 2005.
- [8] E. Mannie, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, Oct. 2004.
- [9] OIF-UNI-01.0, "User Network Interface (UNI) 1.0 Signaling Specification", OIF, Oct. 2001.
- [10] ITU-T G.7714, Generalized automatic discovery techniques, Nov. 2001.
- [11] ITU-T Recommendation G.8080N.1304, "Architecture for the Automatically Switched Optical Network (ASON)". Jul. 2001.
- [12] <http://cheetah.cs.virginia.edu/about.html>.
- [13] <http://dragon.east.isi.edu/twiki/bin/view/Main/WebHome>

-
- [14] <http://networks.internet2.edu/hopi>.
- [15] Telecom Operations Map, Approved Version 2.1, TeleManagement Forum, March 2000.
- [16] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, Jan. 2001.
- [17] A. Ayyangar, J. Vasseur, "Inter domain GMPLS Traffic Engineering - RSVP-TE extensions," IETF Internet draft, work in progress, Feb. 2005.
- [18] J. Moy, "OSPF Version 2", RFC 2328, Apr. 1998.
- [19] R. Coltun, "The OSPF Opaque LSA Option", RFC 2370, Jul. 1998.
- [20] D. Katz, K. Kompella, D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, Sep. 2003.
- [21] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, Sep. 1997.
- [22] R. Braden, L. Zhang, "Resource ReSerVation Protocol (RSVP) --Version 1 Message Processing Rules", RFC 2209, Sep. 1997.
- [23] J. Wroclawski, "The Use of RSVP with IETF Integrated Services", RFC 2210, Sep. 1997.
- [24] C. Alaettinoglu, etc., Towards Milli-Second IGP Convergence, draft-alaettinoglu-ISIS-convergence-00, Nov. 2000.
- [25] R. B. Cooper, Introduction to Queuing Theory, 2nd edition. New York, North Holland, 1981.
- [26] Qiang Song, Ibrahim Habib, Wesam Alanqar, "Simulation Studies on Performance Evaluation of Distributed Dynamic Routing in GMPLS Optical Networks", IEEE GLOBECOM '05, St. Louis, MO, Nov, 2005
- [27] Ding Zhemin; Hamdi, M.; "On the design of a protection/provisioning framework in IP/WDM optical networks", ICC 2005.
- [28] P. Ho and H. T. Mouftah, "A Framework for Service-Guaranteed Shared Protection in WDM Mesh Networks," IEEE Communications Magazine, Feb. 2002.

-
- [29] Sunil Gowda; Sivalingam, K.M.; “Protection mechanisms for optical WDM networks based on wavelength converter multiplexing and backup path relocation techniques”, INFOCOM 2003.
- [30] Hailemariam, A.G.; Ellinas, G.; Stern, T.; “Localized failure restoration in mesh optical networks”, OFC 2004.
- [31] Shengli Yuan; Jue, J.P.; “Dynamic lightpath protection in WDM mesh networks under risk-disjoint constraint”, GLOBECOM 2004.
- [32] Tornatore, M.; Canhui Ou; Jing Zhang; Pattavina, A.; Mukherjee, B.; “Efficient shared-path protection exploiting the knowledge of connection-holding time”, Optical Network Design and Modeling, Feb. 2005.
- [33] Jun Zheng; Baoxian Zhang; Mouftah, H.T.; “Dynamic path restoration based on multi-initiation for GMPLS-based WDM networks”, ICC 2004.
- [34] Puype, B.; Vasseur, J.-P.; De Maesschalck, S.; Colle, D.; Lievens, I.; Pickavet, M.; Demeester, P.; “Benefits of GMPLS for multilayer recovery”, IEEE Communications Magazine, Jul. 2005.
- [35] Verbrugge, S.; Pasqualini, S.; Jager, M.; Iselt, A.; Kirstadter, A.; Chahine, R.; Colle, D.; Pickavet, M.; Demeester, P.; “Modeling operational expenditures for telecom operators”, ONDM 2005.
- [36] Pasqualini, S.; Kirstadter, A.; Iselt, A.; Chahine, R.; Verbrugge, S.; Colle, D.; Pickavet, M.; Demeester, P.; “Influence of GMPLS on network providers' operational expenditures: a quantitative study”, IEEE Communications Magazine, July 2005.
- [37] Jian Wang; Sahasrabudhe, L.; Mukherjee, B.; “Path vs. subpath vs. link restoration for fault management in IP-over-WDM networks: performance comparisons using GMPLS control signaling”, IEEE Communications Magazine, Nov. 2002.
- [38] Dahai Xu; Yizhi Xiong; Chunming Qiao; “Novel algorithms for shared segment protection” IEEE Journal on Selected Areas in Communications, Oct. 2003.
- [39] Saradhi, C.V.; Ng Chee Kong; Gurusamy, M.; “Fast and resource efficient segment-based failure recovery in WDM optical mesh networks”, IEEE MILCOM 2004.
- [40] Jing Fang; Sivakumar, M.; A.K.; Sivalingam, K.M.; “On Partial Protection in Groomed Optical WDM Mesh Networks”, Dependable Systems and Networks, 2005.
- [41] Casetti, C.; Mardente, G.; Mellia, M.; Manufo, M.; Lo Cigno, R.; “On-line routing optimization for MPLS-based IP networks”, HPSR 2003.

-
- [42] Xiaowen Chu, Jiangchuan Liu, "DLCR: a new adaptive routing scheme in WDM mesh networks", ICC 2005.
- [43] Cinkler, T.; Mesko, D.; Viola, G.; Tapolcai, J.; "Routing with partially disjoint shared path (PDSP) protection", NGIN 2005.
- [44] Necker, M.; Gauger, C.; Bodamer, S.; "A new efficient integrated routing scheme for SDH/SONET-WDM multilayer networks", OFC 2003.
- [45] Fumagalli, A.; Tacca, M.; Wu, K.; Vasseur, J.-P.; "Local recovery solutions from multi-link failures in MPLS-TE networks with probable failure patterns", GLOBECOM 2004.
- [46] Schupke, D.; Prinz, R.; "Performance of path protection and rerouting for WDM networks subject to dual failures", OFC 2003.
- [47] Hongsik Choi; Subramaniam, S.; "Hyeong-Ah Choi; Loopback recovery from double-link failures in optical mesh networks", IEEE/ACM Transactions on Networking, Dec. 2004 Dec. 2004.
- [48] Zhaoming Li, "Design and Deployment of A GMPLS Control Plane in IP Optical Networks", doctoral thesis, City University of New York, 2007.
- [49] Virtamo, J.T., "A model of reservation systems", IEEE Transactions on Communications Volume 40, Issue 1, Jan. 1992 Page(s):109 – 118.
- [50] Lars-Olof Burchard, "On the performance of computer networks with advance reservation mechanisms", ICON2003. The 11th IEEE International Conference on Networks, 2003.
- [51] O. Schel'en and S. Pink, "Resource Sharing in Advance Reservation Agents," Journal of High Speed Networks, Special issue on Multimedia Networking, vol. 7, no. 3-4, 1998.
- [52] A. Assad, "Multicommodity network flows: A survey," Networks, vol. 8, pp. 37–92, 1978.
- [53] S. Suri, M. Waldvogel, and P. R. Warkhede, "Profile-based routing: A new framework for MPLS traffic engineering", Proc. of the 2 International Workshop on Quality of future Internet Services (QofIS'01), pp. 138-157, Portugal, September 2001.
- [54] <http://www.phy.ornl.gov/tsi>
- [55] <http://www.zebra.org/>

-
- [56] <http://www.kom.tu-darmstadt.de/en/downloads/software>
- [57] Ibrahim W. Habib, Qiang Song, Zhaoming Li, and Nageswara S. V. Rao, "Deployment of GMPLS Control Plane for Grid Applications in Experimental High Performance Networks," *IEEE Communications Magazine*. Mar. 2006.
- [58] J.-L. Le Roux, J.-P. Vasseur, J. Boyle, "Requirements for Inter-Area MPLS Traffic Engineering", RFC 4105, Jun. 2005.
- [59] A. Farrel, J.-P. Vasseur and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", RFC 4726, Nov. 2006.
- [60] A. Farrel, A. Ayyangar, JP. Vasseur, "Inter domain Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering - RSVP-TE extensions", IETF draft, draft-ietf-ccamp-inter-domain-rsvp-te-06.txt, work in progress, Apr. 2007.
- [61] R. Zhang, J.-P. Vasseur, "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", RFC 4216, Nov. 2005.
- [62] Kohei Shiomoto, Dimitri Papadimitriou, Jean-Louis Le Roux, Martin Vigoureux, and Deborah Brungard "Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN), IETF draft, draft-ietf-ccamp-gmpls-mln-reqs-03.txt, work in progress, Apr, 2007.
- [63] K. Shiomoto, R. Rabbat, A. Ayyangar, A. Farrel, and Z. Ali "Procedures for Dynamically Signaled Hierarchical Label Switched Paths" IETF draft, draft-ietf-ccamp-lsp-hierarchy-bis-02.txt, work in progress, Apr. 2007 .

Publications

1. Qiang Song, Zhaoming Li, I. Habib, "Advance Reservation in IP Optical Transport Networks for Connection Oriented Services", in preparation.
2. Qiang Song, Zhaoming Li, I. Habib, "CHEETAH Virtual Label Switch Router Design and Deployment in GMPLS Optical Networks", **IEEE Global Telecommunications Conference, GLOBECOM '06**, Page(s):1 – 6.
3. X. Zhu, X. Zheng, M. Veeraraghavan, Z. Li, Q. Song, I. Habib, and N. S. V. Rao, "Implementation of a GMPLS-based Network with End Host Initiated Signaling",

IEEE International Conference on Communications, ICC06, Jun, 2006. Volume 6, June 2006 Page(s):2710 – 2716.

4. Ibrahim W. Habib, Qiang Song, Zhaoming Li, and Nageswara S. V. Rao, “Deployment of GMPLS Control Plane for Grid Applications in Experimental High Performance Networks”, **IEEE Communications Magazine**. Mar. 2006. Volume 44, Issue 3, March 2006 Page(s):65 – 73.
5. Qiang Song, Z. Li, I. Habib, and W. Alanqar, "Value-added proposition of the GMPLS control plane in IP optical networks", **JOURNAL OF OPTICAL NETWORKING**, Volume 4, page(s): 838-855 (2005).
6. Qiang Song; Habib, I.; Alanqar, W., “On the performance evaluation of distributed dynamic routing in GMPLS optical networks”, **IEEE Global Telecommunications Conference, GLOBECOM '05**. Volume 4, 28 Nov.-2 Dec. 2005 Page(s):2067 – 2071.
7. Qiang Song, Ibrahim Habib, Wesam Alanqar, "GMPLS Control Plane Auto-discovery Evaluation and Its Interfacing with OSS in IP Optical Network", poster presentation, **Ninth IFIP/IEEE International Symposium on Integrated Network Management, IFIP/IEEE IM '05**, Nice, France, May, 2005.
8. Qiang Song; Habib, I.; Alanqar, W., “Performance evaluation of connection setup in GMPLS IP optical network”, **Optical Fiber Communication Conference, OFC/NFOEC '05**, Volume 3, 6-11 March 2005 Page(s):3 pp. Vol. 3.
9. Qiang Song; Zhaoming Li; Habib, I., “Using link bundling for efficient utilization of dynamically provisioned optical circuits”, **Conference on Optical Network Design and Modeling, ONDM'05**. Feb. 7-9, 2005 Page(s):205 – 211.