QoS Guarantees in IP over WDM, Restricted, λ, and Integrated Burst Switching *

Chi-Hsiang Yeh
Dept. of Electrical and Computer Engineering
Queen's University
Kingston, ON K7L 3N6, Canada
yeh@ee.queensu.ca

Emmanouel A. Varvarigos
Dept. of Electrical Engineering and Informatics
University of Patras
Patras, Greece
manos@ceid.upatras.gr

Abstract

During the past few years, Internet Protocol (IP) with MultiProtocol Label Switching (MPLS) over wavelength division multiplexing (WDM) has emerged as the paradigm of choice for the future Internet core. WDM burst switching (WBS), also referred to as optical burst switching (OBS), is a new switching technique that combines some advantages of circuit switching and packet switching and may be suitable for future WDM networks due to its flexible granularity. Previous protocols for QoS provisioning in WBS networks, however, are based on prioritization and cannot provide QoS guarantees. In this paper, we apply the *scalable resource reservation protocol (SRRP)*, a signaling protocol developed for the future Internet, for session-by-session or connection-by-connection resource reservation in WDM-based burst-switched networks, including WBS, λ burst switching, restricted burst switching, and integrated burst switching. We propose the *scheduled reservation mechanism* and *geographical area protection (GAP) technique* for providing efficient and reliable QoS guarantees in these WDM-based burst-switched networks, with overheads that are considerably smaller than those of previous reservation and protection techniques. In particular, we propose the *3O strategy*: *overscheduling*, *overreservation*, and *overbooking* to increase the probability for successful and efficient scheduling, to guarantee QoS for connections with bursty traffic, and to increase the utilization of expensive resources.

1 Introduction

A common protocol stack for the Internet core is IP over asynchronous transfer mode (ATM) over synchronous optical networks (SONET) over WDM. Although IP/ATM/SONET/WDM has achieved certain success and wide-spread adaptation due to the fast layer-2 forwarding, QoS-guarantee, and traffic management capability of ATM, reliability of SONET,... etc, many mechanisms implemented for ATM are redundent to the IP-centric protocol stack (since ATM was originally designed to replace, rather than working under IP). Also, SONET switches are relatively ineffieicnt and expensive. As a result, a lot of recent efforts have been devoted to the removal the ATM and SONET layers from the protocol stack in order to increase efficiency and reduce costs. In particular, MPLS is recently among the most noticeable developments in the networking field. Many desirable features of ATM have been incorporated into MPLS, including fast layer-2 forwarding, traffic engineering, and connection-oriented service that may enable

*This research is supported in part by the Computer Techology Institute, Patras, Greece.

QoS guarantees in IP networks. A major difference between MPLS and ATM is that MPLS is designed to work with IP, instead of replacing IP. Also, MPLS is more scalable due to its capability to aggregate sessions with the same destination node. As a result, IP/MPLS over WDM has emerged as the paradigm of choice for the future optical Internet. However, several important issues are still not resolved in current solutions to IP over WDM (IPoW). In particular, nondisrupted QoS guarantees in IPoW in the presence of faults, which can be achieved in SONET based on its 1+1 protection, are still open. Also, the current version of RSVP, the standard for signaling in the Internet, is not scalable due to its processing overheads and session-by-session reservations (without aggregation), while the achievable throughput for QoS traffic can still be improved and the required latency can still be reduced. Moreover, the supports for differentiated service are still under development and interworking between wireless access networks and optical backbone network is still not well understood.

In this paper, we apply the *scalable resource reservation protocol (SRRP)* to IPoW networks to address the aforementioned unresolved issues. In particular, we present the *slotted reservation mechanism* and the *penalty-based adaptable reservation mechanism* for saclable and adaptive reservations based on SRRP, which are important to broadband optical networks and mobile wireless networks, respectively. We also present the *differentiated admission control mechanism* and the *probabilistic admission control mechanism* for admission control with differentiated service quality according to the QoS requirements, while achieving high utilization. Moreover, we present the *conservative transmission discipline*, the *aggressive transmission discipline*, and the *optimistic transmission discipline* for transmissions with small latency and/or small buffer requirements. We show through simulations that SRRP can achieve considerably smaller call blocking rates as compared to previous reservation protocols.

Most previous work concerning IPoW assumes either circuit switching or photonic packet switching. In circuit switching, a connection is fully established before any data are transmitted. Currently, circuit switching is more suitable to WDM than packet switching, since the latter will lead to higher overheads due to its finer granularity. Also, when combined with wavelength routing, optical/electronic/optical (O/E/O) conversion is not required in circuit switched networks, removing the electronic bottleneck. However, the times required to set up and tear down a circuit are large so that a circuit is usually kept for a long time; also, once a circuit is set up between a pair of nodes, the capacity cannot be used by other connections even though the circuit is idle, leading to low utilization and efficiency. To solve these problems and to combine the desirable features of circuit switching and packet switching, several researchers proposed to use WDM burst switching (WBS), also called optical burst switching (OBS), as a potential alternative for switching in WDM networks. In WBS, a burst is typically composed of a number of packets, so the granularity of WBS is between circuit switching and packet switching, leading to reasonably small overhead as in circuit switching and moderate to high utilization as in packet switching. Previous protocols for QoS provisioning in WBS networks [**?**], however, used prioritization on a burst-by-burst basis and cannot provide QoS guarantees. Moreover, WBS networks require full wavelength-conversion capability at every network node, which may be prohibitively expensive and may not be necessary, especially for underutilized nodes.

In this paper, we introduce WDM-based switching techniques, including λ *burst switching (λBS)*, *restricted-wavelength-conversion burst switching (RBS)* (*restricted burst switching* for short), and *integrated burst switching (IBS)*. λBS is a burst switching technique where no wavelength-conversion capability is required; RBS is a burst switching technique where only partial wavelength-conversion capability is required; IBS allows heterogeneous nodes and links belonging to WBS, RBS, or λBS. These WDM-based burst switching techniques may require implementation costs that are considerably lower than those required by WBS networks. We apply SRRP to these WDM-based burst switching techniques, on top of its applications to wireless mobile networks, Ethernet, and conventional optical networks [**?**]. In particular, we propose the *scheduled reservation mechanism* for SRRP to reserve the required resources for WDM-based burst switched networks, the *priority-based aggressive reservation interactive scheme (PARIS)* to better utilize the network resources and to provision QoS probabilistically in a differentiated manner, and the *circumscribed protection technique* for noninterrupted QoS guarantees in the presence of faults. Due to the unique, desirable, characteristics of SRRP for solving the aforementioned challenging issues and its applicability to various networking technologies, SRRP may leads to an extension to RSVP or a new signaling protocol for the next-generation Internet.

2 QoS Provisioning in IPoW/IBS Networks: Problems and Solutions

In WBS or OBS networks a network node is required to have the capability to convert a burst in any wavelength into any other wavelength. This may lead to implementations that are prohibitively expensive. In this paper, we propose seveal other WDM-based burst switching techniques, including λ *Burst Switching* (λ*BS*), *Restricted-wavelength-conversion Burst Switching (RBS)*, and *Integrated Burst Switching (IBS)*, as aleternative switching formats that may considerably lower the implementation cost.

In λBS, a burst in one wavelength is not allowed to be converted to any other wavelength. λBS is the least expensive among the four burst switching techniques. In RBS, there are certain restrictions (that may be personalized for each network node or may be common to the entire network) imposed on the number or types of wavelength conversions that can be made at the same time. For example, the number of wavelength converters may be limited due to their high costs. If a network node with 4 links, each having 128 wavelengths, only has 16 wavelength converters, then at most 16 bursts can be converted at the same time. As another example, there may be restrictions on the output (or input) wavelength(s) for a wavelength converter. A burst may then be switched through a bipartite, multistage, or another type of network (within the node), until it reaches an appropriate wavelength converter that can convert it to the desirable wavelength. In IBS, a network may have heterogeneous nodes that uses WBS, RBS, and/or λBS. For example, bottleneck or heavily-utilized nodes may implement WBS, underutilized nodes may implement λBS, while other nodes in between may implement RBS. In order to mitigate the integrity problem, we may force all bursts to have equal length. The resultant switching techniques are referred to as *WDM-based slotted burst switching*, which can be further categorized into *slotted* λ*BS*, *slotted RBS*, *slotted WBS*, and *slotted IBS* according to the assumptions on the wavelength conversion capability of the network nodes.

In λBS, RBS, and IBS networks, connections may not share wavelengths as efficiently as in WBS networks, so "statistical multiplexing" is not effective enough to coordinate the usage of the wavelengths. As a result, priority-based techniques and reservation schemes on a burst-by-burst basis like the one proposed in [**?**] are in general insufficient to guarantee QoS, and resource reservations on a per connection basis may be necessary to provision QoS guarantees. When resource reservations on a connection-by-connection basis are allowed, however, prioritization based on larger reservation lag times [37], such as the approach proposed in [**?**] will fail, since a relatively low-priority connection that reserves resources in advance may be treated as having priority higher than any high-priority bursts that reserve bandwidth on a burst-by-burst basis. Therefore, a completely different signaling paradigm based on connection-by-connection reservations is mandatory for efficient QoS guarantees in such networking environments.

In most previous reservation protocols, only conventional resources such as network bandwidth and buffer space are reserved. In RBS and IBS networks, however, resource reservations become more complicated since wavelength converters and the associated paths (leading to appropriate wavelength converters through the multistage network within a node) may be resources that are more scarce than network bandwidth or wavelengths. Therefore, a signaling and resource reservation protocol for RBS and IBS networks has to reserve these unconventional resources as well and to schedule them appropriately according to the schedules for the reserved burst slots, as is done in SRRP.

Another limiting resources in optical networks are the available optical buffers. Since a burst-switched network may be bufferless or only have a relatively small amount of buffer space, and λBS, RBS, and IBS networks cannot utilize wavelengths as efficiently as WBS networks, the actual burst dropping rates may be considerably larger than the ones derived in [**?**] for WBS networks with optical buffers. As a result, to increase the network utilization and reduce call blocking rates, an adaptable reservation mechanism (see Subsection 3.3) and an appropriate admission control mechanism (see Sections **??**) should be employed.

In a burst-switched network, the bandwidth of a wavelength has to be allocated to a certain burst completely at any time instance, in contrast to virtual circuit networks where partial bandwidth can be reserved and allocated to a connection. As a result, new reservation formats such as the one used in the proposed *scheduled reservation mechanism* (see Subsection 3.1) as well as accompanying traffic shaping mechanisms are required for this type of reservations. Also, when connection-by-connection reservation is employed, scalability issues have to be tackled (e.g., by employing the *slotted reservation* introduced in Subsection 3.2), since the number of connections over a core router may be extremely large, even after aggregation. Other important issues include QoS guarantees in the presence of faults, which can continue to be provided based on protection techniques. The proposed *geographical area protection* and *logical area protection* with *virtual reservation* are new protection techniques that can considerably reduce the overheads introduced by signaling and the wasted resources due to idling.

In the following sections, we present several techniques and mechanisms of SRRP that are designed to enable efficient QoS guarantees in IPoW networks. Several techniques such as the adaptable reservation mechanism also make SRRP suitable for other networking technologies such as wireless cellular and ad hoc networks.

3 SRRP Reservation Mechanisms for IPoW/IBS networks

When SRRP is applied to WBS and λBS networks, the reservation and allocation of network resources are separated. More precisely, a fraction of the capacity of a link and other resources, such as buffer space, are reserved for certain intervals by each QoS connection (possibly after aggregation) or QoS class using the link. The purpose of reservations is bookkeeping, while allocation explicitly assigns the resources to a session and forbids other sessions from using the allocated capacity. In this section, we present SRRP reservation mechanisms and in Section **??** we propose SRRP allocation mechanisms for WBS and λBS networks.

3.1 The Scheduled Reservation Mechanism

In *scheduled reservation*, the capacity is not always reserved right after an intermediate node receives a setup packet. Each node maintains a reservation profile, which records the intervals during which capacity is reserved for each connection and QoS class. The reservation profile contains the curve for the reserved capacity with respect to time. The simplest form of the curve of a particular connection or QoS class is a straight line as in ordinary reservation protocols; the second form is a step function where the curve starts at zero, rises at the time the reservation starts, and remains at that value; the third form is a step function where the curve rises at the starting time of the reserved interval and go back to zero at the ending time of the reserved interval; the fourth form is a more complicated function, where the curve changes its value whenever applicable, according to the estimated required capacity at the time. Note that the fourth form may also be a step function with multiple steps. The curve of the reservation profile of a link has to be equal to, or larger than, the sum of theses functions for all the QoS connections or QoS classes using the link.

When the processing time and queueing delay for the setup packet and/or other control packets is large, we may use the *admit-and-forward discipline* to speed-up connection setup. Under this discipline, an intermediate node uses the minimum possible time to determine whether it will likely admit the request, and then forwards the setup packet to the appropriate outgoing link. The setup packet can continue to be processed at the node for the reservation of the bandwidth, buffers etc. The processing for admission and forwarding should be given higher priority so that several setup packets may be partially processed and forwarded first, postponing more time-consuming processing for a later time. The admit-and-forward discipline may considerably reduce the connection establishment/renegotiation time since the request may be processed in multiple nodes in parallel and the processing time does not accumulate along the path it traverses.

3.2 The Slotted Reservation Mechanism

Recording the starting and ending times for the interval and/or subintervals of each connection is prohibitively expensive when the number of connections is not small. In addition to the memory requirements, a node has to check a large part of its *reservation profile* for admission control. The number of additions and subtractions required to determine the acceptance or rejection of a new request is proportional to the number of QoS connections, which is extremely expensive in terms of the processing time.

To achieve better scalability, SRRP use *slotted reservation* [37], where the reserved interval or subintervals are rounded to the minimum set of slots that cover the entire interval/subintervals. Each slot only needs several counters to record the aggregate bandwidth reserved for each of the QoS classes during that slot. Admission control can then be performed in a small number of clock cycles, independent of the numbers of micro/macroflows or the number of slots, by performing several addition operations in parallel using multiple adders (whose number is upper bounded by the number of slots). Note that the durations of slots in slotted reservation are not necessarily equal to the transmission time of a burst – some of them may be the same, but some others may be considerably larger.

3.3 The Penalty-based Adaptable Reservation Mechanism

The reservation made by an SRRP connection may contain more than one set of QoS requirements, one for the normal mode and others for degraded modes. A node may then adapt to the traffic conditions and provide acceptable service quality to all the connections.

When the adaptable reservation mechanism [37] is used for a connection (possibly after aggregation) or a QoS class, its setup packet is required to carry with it the resource requirements under the degraded modes, and the associated session(s) has to be able to cope with such a de-

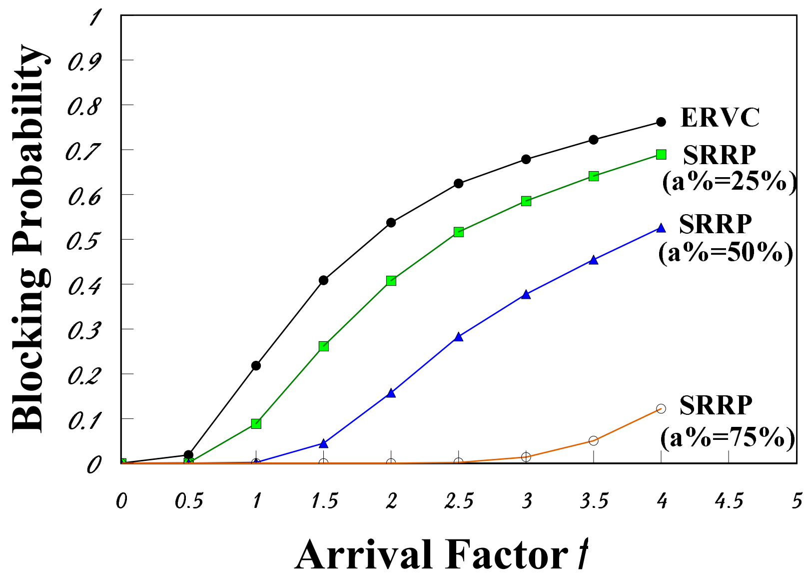

Figure 1: Comparisons of the blocking rates (based on simulation) for SRRP and previous protocols without adaptable reservation, varying $a\%$, the percentage of traffic volume that is adaptable. Note that the maximum allowed arrival factors for SRRP are considerably larger than those of previous protocols for given blocking rates.

graded bandwidth enforced by the network when there are no sufficient network resources. Network nodes with such an advanced feature have to record the requirements of each connection or QoS class, including the maximum tolerable delay and delay variation, the peak, average, and/or minimum bandwidths without degradation, and the bandwidth requirements (e.g., a fixed value or a window) under degradation level 1, level 2, and so on, along with the priority of the session or QoS class and the penalty for each degradation level (e.g., a fixed value for that level or a function of the assigned bandwidth within the associated window as well as the traffic conditions and resources consumed).

When there is sufficient bandwidth available, a network node allocates the required bandwidths without degradation to all the connections/QoS class, as in ordinary reservation protocols; when there is not sufficient bandwidth left, a network node may first hold a certain amount of traffic that is time-noncritical (which typically has the smallest penalty). If this is still insufficient, the network node reallocates the required bandwidth under degradation to some connections/QoS classes that have lower priority and smaller penalty. The goal of bandwidth allocation and reallocation is to minimize the aggregate penalty and thus maximizing user satisfaction, revenue, fairness, or other criteria.

4 SRRP Transmission Disciplines for IPoW/IBS networks

In the conventional transmission discipline used in most reservation protocols, a source node is not allowed to transmit any data until it receives an acknowledgement from the destination. Similarly, the *conservative transmission discipline* of SRRP does not allow transmissions until the ingress router receives an acknowledgement from the egress router and until its local time is at least equal to the transmission starting time of its scheduled reservation.

In the *aggressive transmission discipline* of SRRP, an ingress router can start transmitting data packets before it receives an acknowledgement for the establishment of the connection. This transmission discipline can reduce latency for high-priority connections/classes and/or reduce the buffer space required by low-priority connections/classes at the ingress routers. When the high-priority data bursts of a high-priority connection/class arrive at a node that is overutilized, some other connections/classes will be degraded to make room for the data bursts; if there is still not sufficient capacity, the data burst can be buffered or deflected. When the low-priority data bursts of a high-priority connection/class or the data bursts of a low-priority connection/class arrive at a node without successful reservation or before the reserved capacity is available, a certain action, such as deflection, buffering, or dropping, is taken to deal with these bursts.

Note that a typical and appropriately-operating IP/SRRP network should have some connections using conservative transmission if they are not sensitive to the latency for connection establishment, while some other connections using aggressive transmission if they are latency sensitive, have higher priority, and/or are droppable.

In the *optimistic transmission discipline*, an ingress router (or intermediate node) transmits (or forward) after it receives an acknowledgement (and the associated binding information) from the next node in the path and after its local clock has reached its transmission (or reservation) starting time. This transmission discipline may have to be used to replace the aggressive transmission discipline due to technology or cost constraints.

5 SRRP Admission Control Mechanisms for IPoW/IBS

Under the differentiated admission control discipline [**?**] of SRRP, different requests are treated differently according to their attributes. More precisely, a class-r request is admitted by a node or bandwidth broker only when the nonadaptable part (see Subsection 3.3) of the total bandwidth requirement over the link does not exceed the threshold t_r of its maximum capacity, where $0 < t_r \leq 1$. If class r_1 has lower priority than class r_2 (i.e., $r_1 > r_2$), then we have $t_{r_1} \leq t_{r_2}$. In this way, higher-priority traffic can be protected against occasionally excessive requests from lower-priority traffic. Our performance evaluation study [**?**] indicates that differentiated admission control can considerably reduce the call-blocking rates for higher-priority traffic, while only increasing the call-blocking rates for lower-priority traffic moderately. Note that the thresholds t_r are adaptive in SRRP. A node or resource broker can monitor the performance of a link and dynamical assign appropriate values to t_r.

6 SRRP Protection Mechanisms for IPoW/IBS Networks

6.1 Link Protection with Virtual Reservation

In order to guarantee the provision of QoS to existing connections with minimum disruption when a link fault occurs in MPLS networks [23], one or several protection *label switched paths* [23] with *virtual reservation* are established for each link that carries traffic with QoS guarantees. There are two major differences between our approach and previous protocols. The first difference is that we may use more than one path to protect a single link, in contrast to a single protection path in previous protocols. This way the burden for accommodating traffic of a broken link can be shared by multiple paths. The second difference is that the virtual reservations of the protection paths are utilizing the adaptable part (i.e., see Subsection 3.3) of the link bandwidth along the path. Note that these protection paths do not need to be link or node disjoint. Note also that the links on the protection paths can still be utilized to 100% under normal operation; when there is a fault occurs, the virtual reservation is activated and some of the *forwarding equivalence class (FEC)* over the link are degraded, but still with an acceptable service quality. Since link capacity does not need to stay idle unnecessarily for the protection purpose, while QoS can still be guaranteed when faults occur once the associated virtual reservations are activated, our approach leads to efficient fault recovery without wasting expensive network resources as in previous protocols.

When a new FEC is included in the forwarding table for the protected link, the associated node multicasts the information to all the nodes along the protection paths, which in turn record it in their local memories. When there are no appropriate paths with sufficient bandwidth to provide protection for all the QoS guaranteed traffic with separate virtual reservations, different links are also allowed to share a common virtual reservation for their protection. In such *shared protection*, a single link fault can still be tolerated without affecting the QoS guarantees, but multiple faults may cause some QoS sessions to be disrupted in the worse case. Such sessions will then have to resort to restoration by signaling an alternative path.

Similar strategy can be applied to tolerate channel or node faults based on channel, node, subpath, and path protection. The details will be reported in the near future. In the following subsection, we propose a new protection technique called *geographical area protection (GAP)*.

6.2 Area Protection with Virtual Reservation

In this subsection, we propose *geographical area protection (GAP)* and *logical area protection (LAP)* as new paradigms for protection in IPoW networks.

To protect a geographical area, each working subpath with QoS traffic can be protected by using one or several protection subpaths that bypass the area. As long as sufficient virtual reservations are made along the protection subpaths, QoS of existing connections can continue to be guaranteed against any combination of faults within the protected area. To reduce overheads in signaling, we can establish one or several protection rings surrounding the protected area. When a connection is broken due to some faults within the area, the connection detours along a protection ring to get around the area, until the other end of the area is reached. Note that for different switching techniques, the the resources reserved are very different. For example, in a λBS network, SRRP may reserve a sufficient number of rings for each wavelength that carry QoS traffic, where each of the rings is a circular path with the same wavelength. To tolerate any combination of faults within that area, the total virtual reservations over the rings have to be sufficient to carry the nonadaptable part of all detour QoS traffic to the other end of the area; to tolerate i faulty nodes and j faulty links, the total virtual reservations over the rings have to be sufficient to carry the nonadaptable part of all detour QoS traffic caused by i faulty nodes and j faulty links in the worst case. This rule can be easily generalized to other assumptions on tolerable faults. Topologies other than rings may also be used to protect the area as long as they surround the area.

In LAP, the area protected can consist of any nodes selected from the network. Any link connecting a selected node and a node that is not selected becomes an incident link to the logical area, and has to be protected if it carries QoS traffic. A protection ring (or another topology) can then be established to connect all neighboring nodes of the logical area to protect against any combination of faults within the logical area, given that the protection ring (or the protection topology) does not use any link within or incident to the logical area.

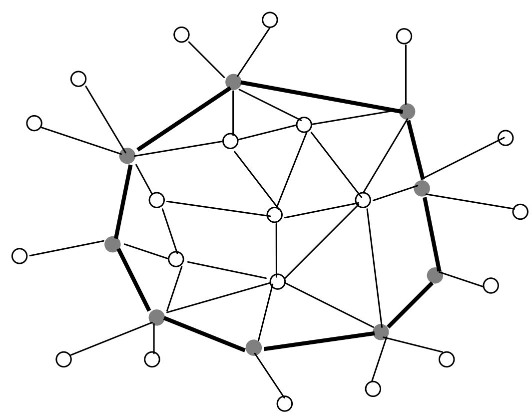

Figure 2: An example for the circumscribed protection technique. QoS traffic can detour along the protection ring when a fault within the area affects its transmissions. Any combination of faults within the area can be tolerated. Resources can be reserved for all QoS traffic going through the protected area; they may also be reserved only for the worst-case scenario for a certain number of faults or fault pattern. Multiple rings (possibly with overlapping links) surrounding the protected area can be established if required.

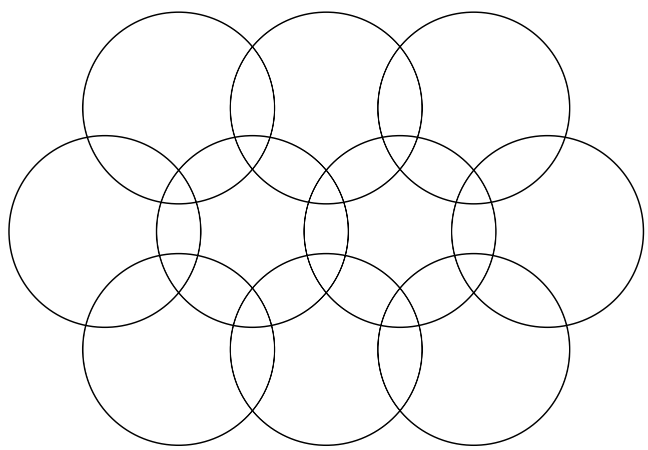

Figure 3: The entire (physical) space occupied by a WDM network should be covered by protected areas. Note that boundary nodes are protected using the other side of the protection ring.

7 Conclusions

In this paper, we proposed λBS, RBS, and IBS as cost effective candidates for IP over WDM. We propose to use the *scheduled reservation mechanism*, PARIS, and GAP and present a variety of SRRP mechanisms to provide QoS guarantees in IPoW networks. The mechanisms proposed for SRRP effectively support each other so that high throughput, low latency, scalability, adaptability, and reliability can be achieved for signaling and resource management across heterogeneous networks including WDM networks and wireless mobile networks.

References

[1] Awduche, D.O., L. Berger, D.-H. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: extensions to RSVP for LSP tunnels," Internet Draft draft-ietf-mpls-rsvp-lsp-tunnel-08.txt, Aug. 2001.

[2] Baker, F., C. Iturralde, and F. Le Faucheur, and B. Davie "Aggregation of RSVP for IPv4 and IPv6 reservations," Internet Draft draft-ietf-issll-rsvp-aggr-03.txt, Feb. 2001.

[3] Berger, L., D.-H. Gan, G. Swallow, P. Pan, and S. Molendini, "RSVP refresh overhead reduction extensions," Internet Draft draft-ietf-rsvp-refresh-reduct-05.txt, Jun. 2000.

[4] Bergkvist, J., I. Cselenyi, and D. Ahlard, "Boomerang – a simple resource reservation framework for IP," Internet Draft draft-bergkvist-boomerang-framework-00.txt, Nov. 2000.

[5] Bertsekas, D.P. and R. Gallager, *Data Networks*, Prentice Hall, Englewood Cliffs, N.J., 1992.

[6] Blake, S., D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," RFC 2475, Dec. 1998.

[7] Bolliger, J. and T. Gross, "A framework-based approach to the development of network-aware applications," *IEEE Trans. Software Engineering,* May 1998, Vol. 24, pp. 376.

[8] Braden, R., Ed., L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) – Version 1 functional specification," RFC 2205, Sep. 1997

[9] Chen, T.M. and T.H. Oh, "Reliable services in MPLS," *IEEE Communications Magazine*, Vol. 37, no. 12, Dec. 1999, pp. 58-62.

[10] Du, H., H.S. Hassanein, and C.-H. Yeh, "Fixed-zone-based routing protocols for wireless ad hoc networks," *Proc. IEEE Workshop Wireless Local Area Networks*, 2001, to appear.

[11] Guo, D., L. Zhang, J. Fu, and R. Cheung, "Extensions to RSVP-TE for bi-directional optical path setup," Internet Draft draft-sorrento-rsvp-bi-osp-00.txt, 2000.

[12] Hellstrand, F., "Extensions to CR-LDP and RSVP-TE for setup of pre-established recovery tunnels," Internet Draft draft-hellstrand-mpls-recovery-merge-01.txt, Nov. 2000.

[13] Hjalmtysson, G. and K.K. Ramakrishnan, "UNITE–an architecture for lightweight signaling in ATM networks," *Proc. IEEE INFOCOM'98*, 1998, pp. 832-840.

[14] Huang, J., Y. Wang, and F. Cao, "On developing distributed middleware services for QoS- and criticality-based resource negotiation and adaptation," *Real-Time Systems,* Vol. 16, no. 2-3, May 1999, pp. 187-221.

[15] Karn, P., "MACA – A new channel access method for packet radio," *Proc. Computer Networking Conf.*, 1990, pp. 134-140.

[16] Kim, Y., W. Tsai, M. Iyer, and J. Ros, "Minimum rate guarantee without per-flow information," *Int'l Conf. Network Protocols,* 1999.

[17] Kleinrock, L., *Queueing Systems, Vol. II: Computer Applications,* John Wiley & Sons, New York, 1976.

[18] Malamos, A., T. Varvarigou, E. Malamas, and C.-H. Yeh, "MEQA^3 – A multi-end QoS application adaptation architecture," *Int'l J. for Computer Research*, to appear.

[19] Owens, K., S. Makam, V. Sharma, B. Mack-Crane, and C. Huang, "A path protection/restoration mechanism for MPLS networks," Internet Draft draft-chang-mpls-path-protection-02.txt, Nov. 2000.

[20] Pan, P. and H. Schulzrinne, "YESSIR: a simple reservation mechanism for the Internet," IBM Research, Hawthorne, New York, no. 20697, Sep. 1997.

[21] Pan, P., E. Hahne, and H. Schulzrinne, "BGRP: a framework for scalable resource reservation," Internet Draft draft-pan-bgrp-framework-00.txt, Jan. 2000.

[22] Radhakrishna, C., "Use of RSVP for differentiated services signaling and admission control," Internet Draft draft-rk-diffserv-rsvp-sig-00.txt Jun. 2000.

[23] Rosen, E.C., A. Viswanathan, R. Callon "Multiprotocol label switching architecture," RFC 3031, Jan. 2001.

[24] Sharma, V., S. Makam, K. Owens, B. Mack-Crane, C. Huang, and B. Akyol, "Extensions to RSVP-TE for MPLS path protection," Internet Draft draft-chang-mpls-rsvpte-path-protection-ext-01.txt, Nov. 2000.

[25] Sharma, V., B.-M. Crane, S. Makam, K. Owens, C. Huang, and F. Hellstrand, J. Weil, L. Andersson, B. Jamoussi, B. Cain, S. Civanlar, and A. Chiu, "Framework for MPLS-based recovery," Internet Draft draft-ietf-mpls-recovery-frmwrk-02.txt, Mar. 2001.

[26] Suraev, V, "Global path recovery enhancement using notify reverse LSP," Internet Draft draft-suraev-mpls-globl-recov-enhm00.txt, Apr. 2001.

[27] Tanenbaum, A.S., *Computer Networks, 3rd Edition,* Prentice Hall, N.J., 1996.

[28] Vandalore, B., W. Feng, R. Jain, and S. Fahmy, "A survey of application layer techniques for adaptive streaming of multimedia," *J. Real Time Systems*, Jan. 2000, to appear.

[29] Varvarigos, E.A., and V. Sharma, "The ready-to-go virtual circuit protocol: a loss-free protocol for multigigabit networks," *IEEE/ACM Trans. Networking,* Vol. 5, no. 5, Oct. 1997, pp. 705-718.

[30] Varvarigos, E.A. and V. Sharma, "An efficient reservation connection control protocol for Gigabit networks," *Computer Networks and ISDN Systems,* Vol. 30, no. 12, Jul. 1998, pp. 1135-1156.

[31] Varvarigos, E. A. and J.P. Lang, "A virtual circuit deflection protocol," *IEEE/ACM Trans. Networking,* Vol. 7, no. 3, Jun. 1999, pp. 335-349.

[32] Wan, J.-W. and X.-C. Lu, "QoS level-based bandwidth split-level adaptation, *J. Software,* Vol. 11, no. 10, Oct. 2000, pp. 1375-1381.

[33] Wang, X. and H. Schulzrinne, "An integrated resource negotiation, pricing, and QoS adaptation framework for multimedia applications," *IEEE J. Selected Areas in Communications,* Vol. 18, no. 12, Dec. 2000, pp. 2514-2529.

[34] Yeh, C.-H. and O.Y. Chan, "A QoS-guaranteed MAC protocol for wireless/mobile communications," *Proc. Int'l Conf. Information Technology and Communication at the Dawn of the New Millennium,* Aug. 2000, pp. 411-420.

[35] Yeh, C.-H. and E.A. Varvarigos, "The scalable networking scheme for high-speed networks," *IEEE Int'l Conf. Communications (ICC'2000)*, 2000, pp. 1335-1342.

[36] Yeh, C.-H., E.A. Varvarigos, D.P. Bertsekas, and H.T. Mouftah, "Reservation-based session routing for broadband communication networks with strict QoS requirements," *Proc. Int'l Conf. Information Networking*, 2001, pp. 593-600.

[37] Yeh, C.-H., "Scalable QoS supports for multimedia applications in the next-generation Internet," *Proc. IEEE Real Time Technologies and Applications Symp.*, May/Jun. 2001, pp. 39-50.

[38] Yeh, C.-H., "Medium access control with differentiated adaptation for QoS management in wireless networks," *Proc. IEEE Int'l Conf. Mobile and Wireless Communication Networks*, Aug. 2001, pp. 208-219.

[39] Yeh, C.-H. and E.A. Varvarigos, "QoS guarantees for IP over WDM burst switching," *Proc. 39th Allerton Conf. Communication, Control, and Computing*, Univ. Illinois, Urbana, IL, USA, 2001, to appear.