A Flexible Advance Reservation Model for Multi-Domain WDM Optical Networks

Eric He*, Xi Wang, Jason Leigh
Electronic Visualization Laboratory
Department of Computer Science
University of Illinois at Chicago
* eric@evl.uic.edu

Abstract — **Advance reservation is a mechanism to guarantee the availability of resources when they are needed. In the context of LambdaGrid, this mechanism is used to provide data-intensive applications with the needed deterministic network Quality of Service to transport data between grid instruments, high-performance storage systems, compute clusters and visualization systems. Flexible scheduling affords users greater convenience while also improving resource utilization and acceptance rate. In this paper we propose a Flexible Advance Reservation Model (FARM) and describe how to implement this model in the meta-scheduling problem. Then we extend this methodology to the cross-domain lightpath reservation problem by incorporating Routing and Wavelength Assignment algorithms. Next, we present the architecture, implementation and APIs of a coordinated Interdomain and Intradomain optical control plane called AR-PIN/PDC, which is capable of flexible advance reservations. Our simulation results show that by relaxing the reservation time constraint, the acceptance rate and resource utilization can be improved dramatically. Through simulations, we also analyze the impact of advance reservations on immediate reservations and conclude that both AR and IR requests need admission control algorithms in order to let both types of reservations coexist and use resource properly.**

Keywords-LambdaGrid, Advance Reservation, Interdomain, Intradomain, RWA, Lightpath

I. INTRODUCTION

The OptIPuter [3] is a National Science Foundation funded project to interconnect distributed storage, computing and visualization resources using photonic networks at tens of gigabits per second. The main goal of the project is to exploit the trend that network capacity is increasing at a rate far exceeding processor speed, while at the same time plummeting in cost. This allows one to experiment with a new paradigm in distributed computing - where the photonic networks serve as the computer's system bus and compute clusters, taken as a whole, serve as the peripherals in a potentially planetary-scale computer. We consider photonic networks as all-optical networks comprised of optical fibers and 3D MEMS (Micro-Electro-Mechanical Systems) optical switching devices. There is no translation of photons to electrons in photonic networks, hence we can avoid electronic bottlenecks. MEMS optical switches are controlled by special control software called Photonic Domain Controller (PDC), that allows applications to request and acquire end-to-end lightpaths.

Increasingly, research organizations are buying dark fiber or wavelengths, and they want to share their resources with each other in a manner similar to how they might share computing resources in Grid environments. A collection of Grid computing resources interconnected by an application-configurable network of lightpaths is called a LambdaGrid [1]. This provides data-intensive applications with the necessary deterministic network bandwidth to transport data between grid instruments, high-performance storage systems, compute clusters and visualization systems, which is often needed for real-time interactive scientific exploration. An international virtual organization, GLIF, the Global Lambda Integrated Facility, was established to promote this paradigm [19].

Photonic Interdomain Controller (PIN) is software that allows applications to provision and control multi-domain lightpaths [4]. PIN specializes in the interdomain routing and signaling schemes over heterogeneous optical network domains. In a multi-domain environment, security management and policy administration are also critical. Our collaborator, at the University of Amsterdam, has done some pioneering research on Authorization, Authentication and Accounting (AAA) and we are leveraging it within PIN software [12].

Advance reservation is needed to guarantee the availability of network resources. The nature of resource reservations in Grid computing is quite different from those of telephone calls. For the latter, their durations are usually not known in advance and hence cannot be planned in advance. In contrast, resource allocations in Grid environments usually require a large number of different types of resources to be acquired simultaneously. Therefore, they have to be reserved in advance, in a manner similar to the reservation of hotels, airlines, and rental cars for vacation travel.

For customers, the major performance parameter of resource reservations is *acceptance rate* or *blocking rate*, which is defined as the ratio of accepted (blocked) reservation requests of all submitted requests. For network operators, the major performance parameter is *resource utilization*, which is related directly to their revenue. In comparison to immediate reservations, *advance reservations* usually degrade the resource utilization and the acceptance rate due to the resource fragmentation [2]. In order to improve the network performance, fragmentation must be avoided. Allowing flexibility in defining the advance reservations can result in better resource utilization while offering greater convenience to users. In this paper we will examine, through simulations, the degree by which flexibility affects performance.

Incorporating flexible advance reservation into PIN/PDC is not trivial. Because PIN/PDC is based on all-optical networks, one main problem that PIN/PDC has to solve is Routing and Wavelength Assignment (RWA). The RWA problem is a NP-hard problem. Usually it can be simplified by decoupling the problem into two sub problems: the routing problem and the wavelength assignment problem. The routing problem can be solved by Fixed Routing, Fixed Alternate Routing, or Adaptive Routing algorithms. Adaptive Routing is considered to be able to achieve the best performance by feeding the wavelength assignment status back to the routing algorithm [6]. The flexibility of advance reservations introduces a new temporal dimension into the resource allocation problem. The wavelength resources along the path have to maintain both wavelength and temporal consistency.

For interdomain distributed control, the addition of a temporal dimension makes the resource state of each domain too large to disseminate to other domains. Therefore, only the relatively static topology summary information of each domain is disseminated to other collaborating domains. The Grid community consists of many Virtual Organization (VO) based collaborations, which means that the resource of each domain is usually not open for all the world, instead, each domain wants to define their own collaborators and individual access policy. We think that the peer-to-peer publish/subscribe model is more effective in this regard and more scalable for interdomain topology exchange.

The multi-domain lightpath reservation problem is actually one type of meta-scheduling problem. Meta-scheduling can be defined as the act of locating and allocating resources for a job from a collection of distributed resources [17]. The key to meta-scheduling is that the user need not be aware of where the resources are, who owns the resources, or who administers the resources in order to use them. Therefore, the meta-scheduler has to be in charge of probing, selecting and reserving the best set of resources by communicating with a bunch of local schedulers. This same methodology can be applied to the cross-domain lightpath reservation problem.

The new version of PIN/PDC with flexible Advance Reservation (AR) capability is called AR-PIN/PDC. The remainder of the paper is organized as follows. In section II we describe related work. In section III we describe a unified Flexible Advance Reservation Model (FARM) and show how to implement the model for the meta-scheduling problem. In section IV we elaborate on the architecture of AR-PDC and AR-PIN, especially the interdomain routing and signaling processes and their application programming interfaces. In section V we provide simulation results to show how flexibility improves network performance and the impact of advance reservations on immediate reservations. Then the paper is concluded by section VI.

II. RELATED WORK

Significant research has been done on circuit-based intradomain or interdomain lightpath provisioning. User Controlled LightPath (UCLP) is a web service based software deployed over Canada's CA*net4 networks [7]. Bandwidth on Demand (BoD) concentrates on multi-domain policy-based access control [5, 12]. Circuit-switched High-speed End-to-End Transport arcHitecture (CHEETAH) provides end-to-end circuit connectivity by concatenating high-speed ethernet segments [8]. All of the aforementioned work assumes the physical network is based on SONET or Ethernet segments and therefore do not incorporate RWA algorithms. Our previous work assumes that border switches are OEO switches and there is no wavelength continuity constraint between domains, and therefore only considered RWA algorithms within domains [4]. This assumption is not necessary valid, we extend the wavelength continuity constraint beyond domain borders in this paper.

Advance reservation has been widely studied in networks other than all-optical networks. Zheng and Mouftah [9] studied the design of RWA algorithms for different types of advance reservations. Guerin and Orda [10] investigated the computational complexity of routing algorithms when supporting different models of advance reservations. Greenberg et al. [11] proposed a call admission control algorithm that occasionally allows a call in progress to be interrupted in order to efficiently share resources among book-ahead (BA) calls and non-BA calls. The Globus Architecture for Reservation and Allocation (GARA) is a toolkit used to implement advance reservations of grid resources in Globus software [13, 14]. The performance issues of applying advance reservations to meta-scheduling problem have been examined by Snell et al. [17].

The Dragon Project (Dynamic Resource Allocation via GMPLS Optical Networks) is trying to build an interdomain lightpath resource management system and leverage Generalized Multi-Protocol Label Switching (GMPLS) standard as the intradomain control plane [15]. It takes advance scheduling and AAA into consideration during the end-to-end path computation. In the process of state exchange, the topology or topology summary, Label Switching Path (LSP) reservation information, and AAA policy information of each domain will be disseminated to all other domains. This puts a huge amount of load on the control plane network which usually has relatively low bandwidth.

In this paper, we describe a coordinated intradomain and interdomain control plane, taking into account both cross-domain RWA and flexible advance reservation. We propose a publish/subscribe model and On-Demand Parallel Probe (ODPP) algorithm to achieve the scalability of interdomain information dissemination. The intradomain control plane can work on not only GMPLS-enabled switches, but also bare MEMS switches. Through simulations, we found that flexibility in advance reservations can improve performance dramatically. We also explored the impact of introducing advance reservations schemes into an immediate reservation system.

III. FLEXIBLE ADVANCE RESERVATION MODEL (FARM)

We assume that Immediate Reservation (IR) requests use resources immediately upon arrival if they are admitted, without announcing their holding times. In contrast, Advance Reservation (AR) requests specify clearly a start time and an end time (or a holding time). The AR request holding time is usually an estimate or a safe upper bound. The resource used by this request will be made available to other requests when the customer finishes his/her job or the holding time expires, whichever happens first.

The specification of an Advance Reservation (AR) request consists of two types of parameters: time-related parameters and resource-related parameters. For fixed advance reservations, time-related parameters include reservation start time t_{start} and reservation end time t_{stop}.

We believe that if we give some flexibility on the time parameters, the acceptance rate will be improved. This is because flexibility tends to aggregate the reservations together thereby reducing the effect of fragmentation, and in turn enhancing resource utilization. This hypothesis will be proven in the simulations in section IV-A. Zheng and Mouftah [9] classify advance reservations into three types: specified starting time and specified duration (STSD); specified starting time and unspecified duration (STUD); and unspecified starting time and specified duration (UTSD). Different wavelength assignment algorithms are used for each request type. There is however another possible condition where both starting time and duration are unspecified and only a time window is specified with an earliest time and latest time (UTUD). We wish to use this to express the notion of flexibility because all the other three reservation types can be expressed as UTUD with some constraints such as the earliest time or the longest time. It is possible the RWA algorithm will find that there are many candidate solutions. We want to put a limit on the maximum number of returned candidate solutions. Also we need to specify the criteria by which the resource agent can select the best candidate. The criteria could be the earliest or the longest.

The resource-related parameters are dependent on the type of the resource. In this paper, since we are focusing on layer 1 lightpaths, the parameters should include a source node and a destination node. In all-optical circuit switching networks, the bandwidth granularity is a wavelength. A request which needs multiple wavelengths can be decomposed into multiple requests wherein each request provisions a single wavelength. Therefore, in our scheme, we consider only single wavelength reservations.

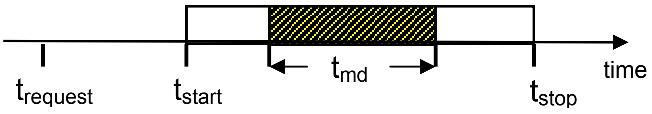

Figure 1. The specification of flexible advance reservations.

Therefore, a flexible advance reservation is defined as follows:

$$R = (s, d, t_{start}, t_{stop}, t_{md}, c) \quad (1)$$

where s is the source node, d is the destination node, t_{start} is the earliest time, t_{stop} is the latest time, t_{md} is minimum duration, and c the selection criteria. The time related parameters are shown in Figure 1.

Flexibility can also improve the user efficiency and satisfaction. For example, for fixed reservations, a user can only get the answer yes or no for a proposed reservation. When flexibility is introduced, the local resource manager will search a wider range of time slots when resources are available. This eliminates the need for the user to begin the process over again with another new proposed reservation.

The improvement is even more significant when a meta-scheduler wants to co-reserve multiple resources for the same period of time. For example, if the blocking rate of each resource is 0.05, then the blocking rate of meta-scheduling of 10 resources is $1-(0.95^{10}) = 0.401$. If we can improve the blocking rate of each resource to 0.01 through flexible advance reservations, then the blocking rate of meta-scheduling can be improved to $1 - (0.99^{10}) = 0.096$.

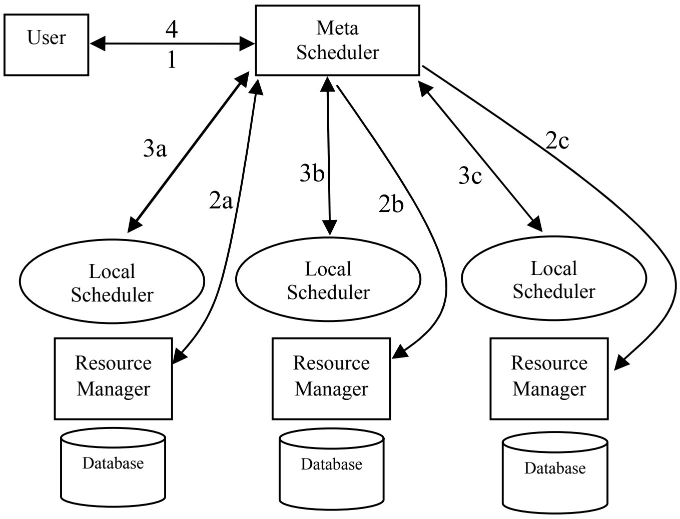

Figure 2. Apply FARM to meta-scheduling

Figure 2 illustrates the steps to apply flexible advance reservations to meta-scheduling. These are:

1). The user submits a flexible advance reservation request to the meta-scheduler. The meta-scheduler analyzes the request, computes a set of resources which can satisfy the user request.

2). The meta-scheduler decomposes the original request into sub-requests and send them to their local resource managers. The resource managers fetch the resource availability information from their scheduling database and send it back to the meta-scheduler.

3). The meta-scheduler collects all the resource availability information. It then determines the range of time that is common for all the resources to meet the original request. If there are many of them, the best one will be selected based on the user specified criteria. At this point, the original flexible request is fixed. Then the meta-scheduler sends the fixed advance reservation requests to all involved local schedulers to hold the needed resources. After all requested resources are held, the meta-scheduler sends commands to local schedulers to reserve them. This method is called a two phase commit reservation.

4). The meta-scheduler returns success to the user and sends back the reservation handle.

Cross-domain lightpath reservation is similar to meta-scheduling of multiple resources. However, the multiple domain all-optical lightpath reservation has one more constraint: wavelength continuity. Therefore, both meta-scheduler and local schedulers have to maintain time continuity as well as wavelength continuity.

Before we explain how we implement flexible advance reservations over cross-domain lightpaths, we need to describe the AR-PIN/PDC architecture in some detail.

IV. AR-PIN/PDC CONTROL PLANE DESIGN

AR-PIN and AR-PDC are interdomain and intradomain lightpath control software that work together to enable advance reservations for end-to-end interdomain lightpaths

The system architecture is shown in Figure 4. We use an example to show the sequence of interactions between users, AR-PIN and AR-PDC. The following steps will be executed when client A in domain 1 sends a reservation request to the AR-PIN/AR-PDC system:

Periodically, all the collaborating domains exchange topology summary with each other.

1. Client A sends a lightpath reservation request to its local interdomain agent AR-PIN1.
2. AR-PIN1 computes the domain-level paths.
3. The source domain queries resource availability from each AR-PDC on the domain-level path.
4. Each queried AR-PDC checks its own AAA policy, resource database, then returns the timeslot-wavelength availability matrix.
5. All the returned timeslot-wavelength availability matrices are intersected at AR-PIN1. Based on the result, the best switch-level path is selected. Then the reservations of all involved domains are performeds in parallel.
6. Within the reservation time window, the lightpath provisioning is triggered by committing the reservation. To do that, the device drivers send TL1 commands to switches to set up the end-to-end lightpath.

Next we will explain each major component of AR-PIN and AR-PDC in detail.

A. Interdomain Routing

A domain is an independently managed network cloud exposing a set of ingress and egress points and links with service specifications. Each link is controlled and managed by a single domain. The separation points between neighboring domains are switches. We call these switches as **border switches**. Ports on border switches can terminate links of multiple different domains. Every border switch needs a globally unique address or name for addressing purposes.

When a domain advertises its topology information to other collaborating domains, it is not necessary to include the details such as internal switches and internal links. Instead, it will just send out a topology summary of its own domain consisting of only border switches and abstracted links. For example, the advertisement from domain A will be:

Switch 1-2: wavelength w1, w2, w3, w4.

Switch 1-3: wavelength w1, w2, w3, w4, w5, w6, w7, w8.

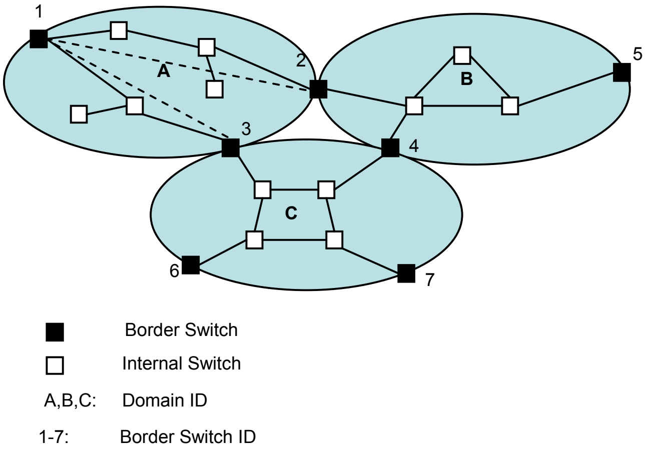

Figure 3. Structure of multiple photonic domains.

These two abstracted links are shown as dotted lines in the Figure 1. The abstract link is actually an abstraction of a bunch of consecutive physical links in the same domain. The topology summary can be generated manually or automatically from the intradomain topology database. The topology summary generation is a maximum-flow problem and it can be solved by the Ford-Fulkerson method [16].

AR-PIN runs a peer-to-peer publish/subscribe based routing protocol to exchange topology summaries among different domains. The peer-to-peer exchange mode is more suitable than blind flooding because it is possible that a domain may want to selectively advertise different sets of resources to different domains. The information exchange is based on the nature of the subscription. Every domain maintains a list of collaborating domains (subscriber). The information exchange is triggered by any change of the interdomain topology of the domain. In other words, whenever the topology changes in a domain, the topology summary will be regenerated, the AR-PIN in this domain will update the new topology summary to all subscribed domains (push model) or just send a change notification and let the domains to request the update by themselves (pull model). Of course, the pull mode should always be supported to handle newly started domains or out-of-sync domains. After receiving the topology summaries from all collaborating domains, each domain can compose its own

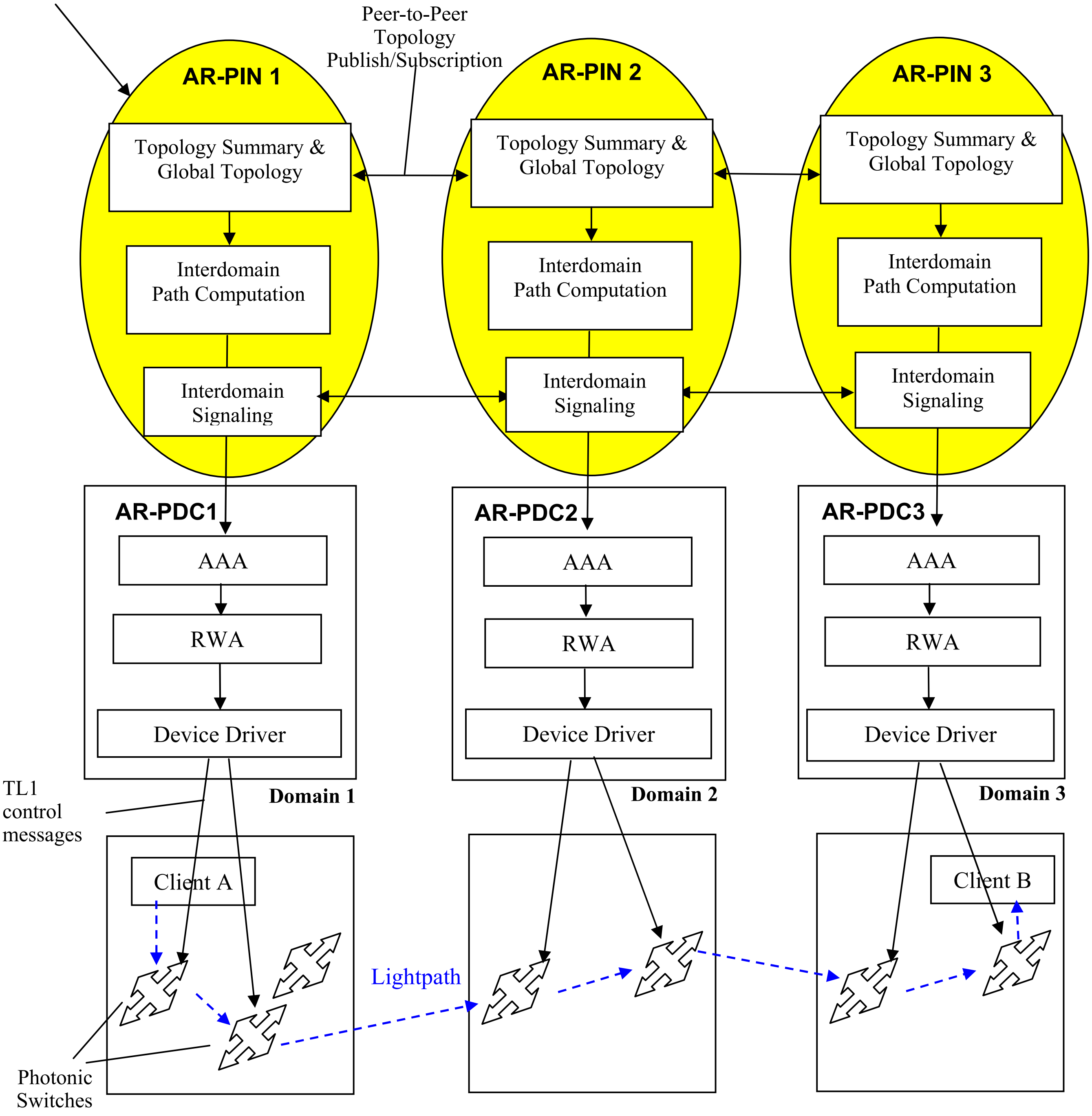

Figure 4. AR-PIN/AR-PDC system architecture.

global topology. Because each domain gets different topology summaries from different collaborating domains, every domain has its own unique global topology database. In this global view, each node is a border switch, each link is an abstract link managed by a domain.

When a lightpath reservation request arrives, the local domain will compute a domain-level path based on its own view of global topology. This path includes only border switches. Source routing will be used to compute the path. There are several possible path computation algorithms such as Shortest Path First, Fixed Alternate, Least-Load-Path, etc. The detailed discussion of these algorithms is beyond the scope of this paper.

B. AR-PDC: Intradomain Control Plane

AR-PDC provisions intradomain lightpaths. Reservation requests may come from local domain users or its interdomain control plane AR-PIN. During the ARPIN resource probing process, it relies on AR-PDC to extend the domain-level path into a switch-level path and check the wavelength availability status.

a) Authorization, Authentication and Accounting (AAA)

When a reservation request comes from foreign domains, they need to go through the AAA mechanism to ensure the foreign user is authenticated. Then according to the identity of the user and the local access policy, the network resources will be filtered and a virtual topology will be generated and it will be used in the following Routing and Wavelength Assignment (RWA) operation.

b) Routing and Wavelength Assignment (RWA)

AR-PDC does the RWA job at the switch level. We also divide the RWA problem into two sub problems: routing and wavelength assignment. For the routing problem, AR-PDC can use Fixed, Fixed Alternate or Adaptive algorithms – same as interdomain path computation. For the wavelength assignment problem, we execute an intersection operation on all hops from the ingress switch to the egress switch and return the resulting timeslot-wavelength availability matrix to PIN. When we use the Fixed Alternate algorithm, we can return the matrices of all paths to PIN and let PIN choose the best one according to the intersection result with the matrix of the explored part of the path.

c) Device Driver

If a request gets reserved successfully, the user needs to commit the request when he/she wants to activate the reservation. Then each domain along the path will send TL1 commands to the MEMS switches to set up cross connects. At the present time we have built device drivers for Calient DiamondWave PXC and Glimmerglass Reflexion 3D MEMS switches. PDC software has unified interface to different types of MEM switches.

C. Apply FARM to AR-PIN/PDC

We described how to implement flexible advance reservation in meta-scheduling in section III. The same principle can be applied to AR-PIN/PDC. In the context of cross-domain lightpath reservations, the meta-scheduler is implemented in AR-PIN, the local scheduler and the resource manager resides in AR-PDC.

We apply the FARM model to AR-PIN/PDC and reiterate the four steps shown in Figure 2. We call this algorithm On-Demand Parallel Probe (ODPP) because AR-PIN probes wavelength resources in each domain in parallel.

1). The user submits a flexible advance reservation request to AR-PIN, AR-PIN computes domain level path based on its own global topology view, which has been described in section IV-A.

2.1). AR-PIN decomposes the original lightpath request into sub-requests and sends them to their local AR-PDCs. Each AR-PDC then computes its own local switch level path. Next the AR-PDC resource manager will operate a two-dimensional JOIN operation over the computed path, which is shown as Figure 5. The parameter t_{start} and t_{stop} from the original request specification (see equation 1) specifies the time range. The purpose of the JOIN is to remove unusable resources which cannot satisfy the wavelength and time continuities. Therefore the two dimensions are time slot and wavelength.

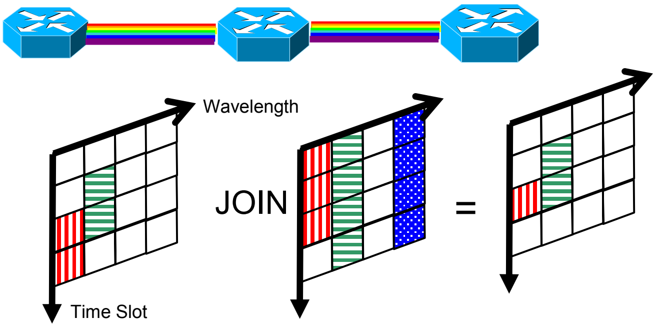

Figure 5. JOIN operation in AR-PDC resource manager.

2.2). After the JOIN operation, the available time slots and wavelengths are found. The next operation is FILTER. The t_{md} in equation (1), minimum duration, is used to filter out those small time fragments whose duration is smaller than t_{md}. For example, if the t_{md} in Figure 5 is 2 time slots, then the red block in the right 2D plane will be filtered because its duration is only one slot. The final resulting 2D matrix is then sent back to the source domain's AR-PIN.

3.1). The AR-PIN collects timeslot-wavelength 2D matrices from all involved domains. It will operate JOIN and FILTER one more time in order to maintain the wavelength and time continuities of the end-to-end cross-domain lightpath. If there exists more than one candidate, the best one will be selected based on user specified criteria, parameter c in the equation (1).

3.2). The AR-PIN needs to populate the selected timeslot-wavelength combination back to all involved domains. The resources may become unavailable due to other reservations. Two phase commit is adopted to make sure the reservations of all domains are all successful or all properly rolled back.

4). The AR-PIN will send back the reservation handle to the user if the two phase commit succeeds.

The *reserve* algorithm for AR-PIN and the *probe* algorithm for AR-PDC are described as follows:

```
$reserve(s,d,t_{start},t_{stop},t_{md},c)$
    $init(matrix);$
    $compute\text{-}domain\text{-}path(s,d);$
    $for\ each\ domain\ d\ on\ the\ path$
        $find\ ingress\ and\ egress\ for\ domain_d;$
        $matrix_d=probe(domain_d,t_{start},t_{stop});$
        $matrix = join(matrix, matrix_d);$
    $filter(matrix,t_{md});$
    $result = select(matrix, c);$
```

```
        return result;

probe(ingress,egress,$t_{start}$,$t_{stop}$)
        init(matrix);
        compute-switch-path(ingress, egress);
        for each link l on the path
                get 2D matrix from database;
                matrix = join(matrix, $matrix_l$);
        filter(matrix,$t_{md}$);
        return matrix;
```

D. ARPIN/PDC Application Programming Interface (API)

AR-PIN/PDC provides advance and immediate reservation service for applications or higher layer resource management systems. Their APIs should be defined clearly. As described earlier AR-PIN accepts interdomain lightpath reservations while AR-PDC accepts intradomain lightpath reservations. Both in fact have similar APIs. The APIs include:

- **Reserve**: This function allows the application to submit a reservation with a specification described by equation (1) to the reservation system. If the reservation succeeds, the system will reply with a unique reservation handle. This handle will be used for other operations such as modification and cancellation.
- **Cancel**: Before the reservation is bound, it can be cancelled.
- **Modify**: Before the reservation is bound, it can also be modified. For example, one can extend or shorten the reservation duration. If the modification request failed because part of the resources cannot be reserved, the original reservation should keep intact.
- **Bind**: When the application is ready to use a reservation, the resource manager may need to do some special processing for the application, or provide some run-time information to the application. For instance, in lightpath reservation systems, the control plane needs to set up the end-to-end lightpath for the application by make proper cross-connects in photonic switches. Also, the control plane may need to provide the IP addresses of end-points to application. This process is known as binding a reservation.
- **Unbind**: When a session of resource usage ends, the reservation should be unbound. After unbinding, the resource is still in reserved status and cannot be used by others. Therefore, if the application will not use the resource any more and the original reservation end time is in the future, it should cancel the reservation so that the resource can be returned to the pool of available resources.
- **Terminate**: This operation should be used when the reservation is in bound status. In fact Terminate is implemented as a combination of executing unbind and followed by cancel.
- **Query Reservation Status**: The client can discover the status of a reservation by polling it. The status includes whether the start of the reservation has begun and whether the reservation has been committed.
- **Query Reservation Attributes**: The client can discover the attributes associated with an existing reservation. These attributes include time-related or resource-related.
- **Subscribe Notification**: The client can subscribe to certain topics so that the resource manager can send messages when the status of the reservation changes or the reservation manager wishes to provide extra information to the application.

V. SIMULATIONS

The simulation work we conducted in the paper mainly consists of two parts. In the first part we validate that how the flexibility in advance reservations can improve acceptance rate and resource utilization. In the real world, immediate reservations co-exist with advance reservations. Therefore, in the second part we analyze the impact of advance reservations on immediate reservations and conclude that both AR and IR requests need admission control algorithms in order to let both types of reservations live together and use the resources properly.

We ran simulations on the NSFNET topology with 14 nodes as shown in Figure 6. We assumed that each link is a single bi-directional fiber with 8 wavelengths. The entire topology was fully-optical without any wavelength converter. In the workload, the starting time of both advance and immediate reservations is a Poisson distribution and the reservation duration is a negative exponential distribution. All these distributions are mutually independent. For advance reservations, the book ahead time, $t_{start} - t_{reserve}$, is randomly selected. All requests try to reserve a lightpath with exactly one wavelength. We ran the simulations on 5 different generated workloads and took the average. The network load is decided by call inter-arrival time and call holding time. In all simulations, the path computation method employed is an adaptive routing algorithm using FIXED wavelength search since it can reach a reasonable combination of performance and complexity [18]. It searches through wavelengths in a fixed order until the available path is found. A standard shortest path algorithm is used to find a path on the effective topology.

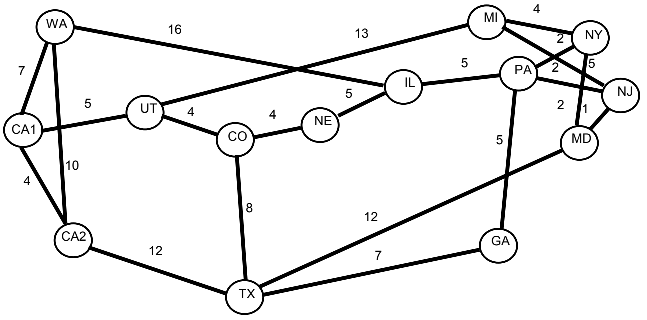

Figure 6. 14 node NSFNET topology.

For flexible advance reservations, the degree of flexibility is defined as:

$$flex = (t_{stop} - t_{start}) / t_{md} \qquad (2)$$

A. Flexibility Improves Both Accepence Rate and Resource Utilization

The goal of the first set of experiments is to evaluate how flexibility affects the blocking rate and resource utilization of advanced reservations. We changed the flexibility of the starting time of reservations from 0, 1, …, to 10. Figure 7 shows how the blocking rate varies with network load. The different curves represent the different degrees of flexibility. We can see that simply introducing 1 or 2 units of flexibility improves the performance considerably, but more flexibility does not help as much. For example, when the blocking rate is 5%, the system load is improved from 1100 requests to 1450 by introducing 1 unit of flexibility, and to 1720 by introducing 2 units of flexibility. Figure 8 shows how the relation of resource utilization vs. network load is affected by different flexibility. The maximum resource utilization can be improved from 47% to 57% by just introducing 1 unit of flexibility.

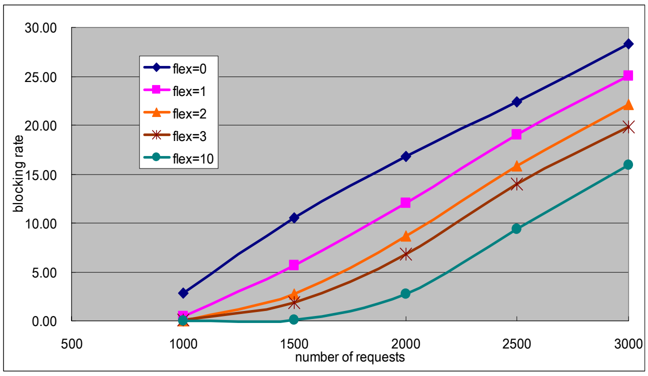

Figure 7. Blocking rate under different flexibilities

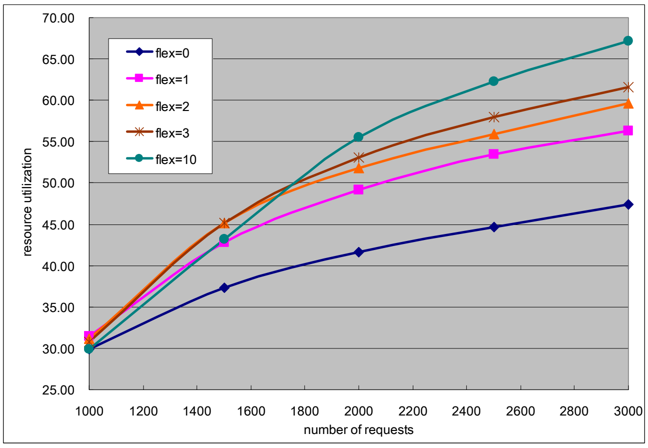

Figure 8. Resource utilization under different flexibilities

B. Impact of Advance Reservations on Immediate Reservations

Certain extemporaneous activities can not or need not be planned ahead. As such our system should be able to take immediate reservations as well. There are three ways to share the wavelength resources between AR and IR requests: full sharing, partial sharing and strict partitioning.

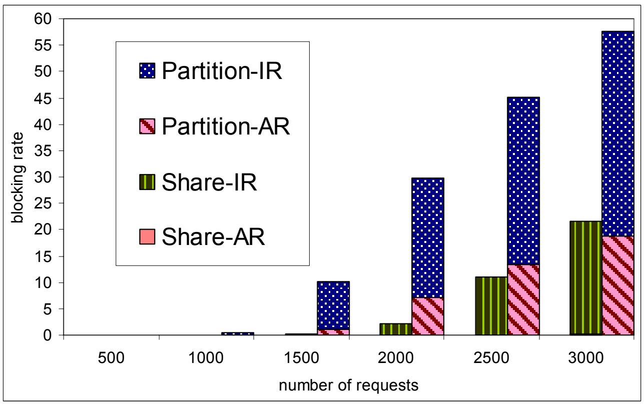

Figure 9. Comparison of wavelength sharing between ARs and IRs.

In this simulation, there are eight wavelengths in the WDM network. We consider two situations: all AR and IR requests share all eight wavelengths (Share), or AR requests use four wavelengths and IR requests use the other four wavelengths (Partition). All AR and IR requests have independent identically-distributed Poisson distribution and occupy 50% of the entire load respectively. From Figure 9, we can see that the Share case has much lower blocking rate than the Partition case. When the number of request is 3000, the blocking rate is 58% for strict-partitioning and only 21% for full-sharing case. The blocking rate of Share-AR is almost zero because it has time advantage over Share-IR.

C. The Dropping Problem and IR Admission Control

Even though sharing brings about greater blocking rate performance, it also introduces a new problem: IR dropping. An admitted IR request may be dropped when the IR request conflicts with a reserved AR request. High and unpredictable dropping degrades the service satisfaction dramatically. We have two means to improve the user experience. Firstly, we could introduce one more parameter for IR requests: Minimum Duration (MD). The IR admission control algorithm will scan the future time slots to make sure the needed resources of this IR request are vacant within the Minimum Duration. Another measure is to notify the user when he/she is possible to be dropped in advance. When the IR request is admitted, we can continue search the future time slot table to find the next conflict point. After the conflicting point, when the AR customer claims the AR request, the conflicting IR request will be dropped out after a short period. During the short period, the IR user can have time to gracefully stop his/her application. We can imagine that if we specifying a larger MD increases the probability of blocking. This is confirmed in the simulation results shown in Figure 10.

In this simulation, the IR profile is fixed. With an increase in AR load, the blocking rate of IRs increases because more resources are pre-occupied by ARs. At the same time, the blocking probability is higher if the user specifies higher Minimum Duration for IRs. The average duration of IRs is 1800 seconds.

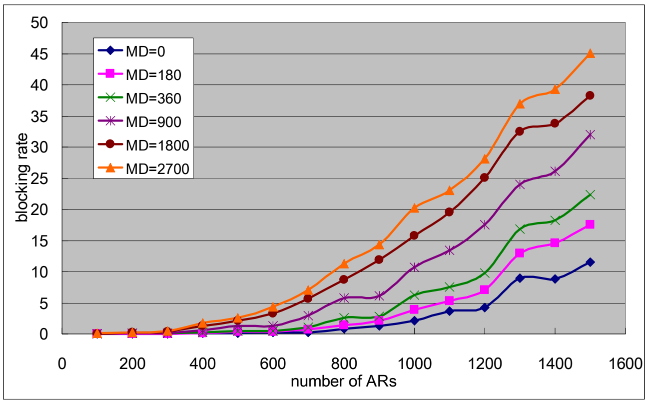

Figure 10. Blocking rate of IRs for different minimum durations.

D. *AR Admission Control*

Since AR requests book reservations relatively far ahead, this gives AR requests priority over IR requests. If there is no admission control for AR requests, it is possible that AR requests occupy most of the resources, which causes a high blocking rate of IR requests or even starvation. In order to provide a certain level of service guarantees to IR requests, it is necessary to put an upper limit on admitted AR requests. From an economic perspective, the charge of IR requests is usually more than AR requests. For example, the ticket fare is usually less expensive if you book earlier in airline reservation systems. At the same time, there are always some impromptu circumstances which cannot be anticipated. We need to keep this type of resource requests from starving.

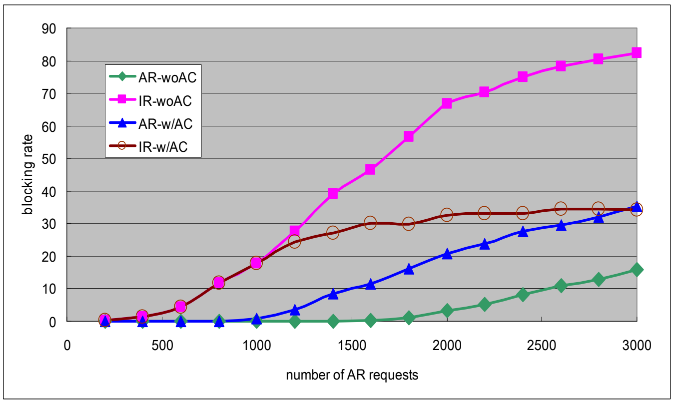

Figure 11. Effect of AR admission control.

The method we employed is to reserve partial wavelengths for IR requests only. For example, AR requests can only use the first five of the total eight wavelengths. From Figure 11, we can see that the AR admission control brings down the blocking rate of IR requests from 82% to 34% when the number of requests is 3000, while it increases the blocking rate of AR requests at the same time because of reduced available wavelength resources. Therefore we achieved a much better balance between AR and IR requests. The percentage of resources specially left to IR requests can be adjusted in run-time by the network administrator.

VI. CONSLUSION AND FUTURE WORK

In this paper we described an architecture of intradomain and interdomain control planes for photonic networks which are capable of flexible advance reservations. A peer-to-peer based publish/subscribe topology model is used to avoid the problem of state flooding. The On-Demand Parallel Probe algorithm renders the periodic dissemination of time-based resource availability information unnecessary and hence makes the system more scalable. Through simulations, we found that by introducing some flexibility on the time parameters of advance reservations, the network performance can be improved dramatically. Also we decided that both AR and IR admission control are necessary in order to maintain a well-balanced AR/IR mixed environment. At the time of writing this paper, some modules of AR-PIN/PDC have been implemented. We will finish the entire AR-PIN/PDC software and deploy them in real testbeds and measure critical parameters such as resource utilization, end-to-end latency, etc. With greater adoption of GMPLS protocols as an intradomain optical control plane, we will explore the possibility of inter-work AR-PIN and GMPLS protocols.

REFERENCES

[1] T. DeFanti, M. Brown, J. Leigh, O. Yu, E. He, J. Mambretti, D. Lillethun, J. Weinberger, "Optical Switching Middleware for the OptIPuter", IEICE Transactions on Communications, invited paper in special issue on Photonic IP Network Technologies for Next Generation Broadband Access. Vol. E86-B, No. 8, pp. 2263

[2] Lars-Olof Burchard, Hans-Ulrich Heiss, Cesar A. F. De Rose, "Performance Issues of Bandwidth Reservations for Grid Computing", Proc. Of the 15th Sym. On Computer Architecture and High Performance Computing (SBAC-PAD'03), 2003.

[3] Larry L. Smarr, Andrew A. Chien, Tom DeFanti, Jason Leigh, Philip M. Papadopoulos, "The OptIPuter", Communications of the ACM, Volume 46, Number 11 (2003), Pages 58-67.

[4] O. T. Yu, T. A. DeFanti, "Collaborative User-Centric Lambda-Grid over Wavelength-Routed Network", In Proceedings of the 2004 ACM/IEEE Conference on Supercomputing (Nov 06 - 12, 2004), Washington, DC.

[5] L. Gommans, C. de Laat, B. van Oudenaarde, and A. Taal, "Authorization of a QoS path based on generic AAA", Future Gener. Comput. Syst. 19, 6 (Aug. 2003), 1009-1016.

[6] Hui Zang, Jason P. Jue, and Biswanath Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," SPIE Optical Networks Magazine, vol. 1, no. 1, Jan. 2000.

[7] Raouf Boutaba, Wojciech Golab, Youssef Iraqi, and Bill St. Arnaud, "Lightpaths on Demand: A Web-Services-Based Management System", IEEE Communications magazine, July 2004, pp 2-9.

[8] M Veeraraghavan, X Zheng, H Lee, M Gardner, W Feng, "CHEETAH: Circuit-switched High-speed End-to-End Transport ArcHitecture", Proc. of Opticomm 2003, 2003.

[9] Jun Zheng and Hussein T. Mouftah, "Routing and Wavelength Assignment for Advance Reservation in Wavelength-Routed WDM Optical Networks", IEEE International Conference on Communications (ICC), 2002.

[10] R Guerin, A Orda, "Networks with Advance Reservations: The Routing Perspective", IEEE INFOCOM 2000, Tel-Aviv, Israel, March 26-30, 2000.

[11] AG Greenberg, R Srikant, W Whitt, "Resource sharing for book-ahead and instantaneous-request calls", IEEE/ACM Transactions on Networking, Vol. 7, No. 1, Feb 1999.

[12] S Van Oudenaarde, Z Hendrikse, F Dijkstra, L. Gommans, C. de Laat, R. Meijer, "Dynamic paths in multi-domain optical networks for grids",

Future Generation Computer Systems, Vol. 21, No. 4, Page 539-548, Apr. 2005.

[13] I. Foster, C. Kesselman, C. Lee, R. Lindell, K. Nahrstedt, A. Roy, “A Distributed Resource Management Architecture that Supports Advance Reservations and Co-Allocation**”,** Int'l Workshop on Quality of Service, 1999.

[14] C Curti, T Ferrari, L Gommans, S Van Oudenaarde, et al, “On advance reservation of heterogeneous network paths”, Future Generation Computer Systems, Vol. 21, No. 4, Page 525-538, Apr. 2005.

[15] Xi Yang, Tom Lehman, Chris Tracy, Jerry Sobieski, Payam Torab, Shujia Gong, Bijan Jabbari, “Policy-Based Resource Management and Service Provisioning in GMPLS Networks”, Adaptive Policy-Based Management workshop, Barcelona, Spain, April 28, 2006.

[16] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, “Introduction to Algorithms”, Second Edition, The MIT Press, 2001.

[17] Quinn Snell, Mark Clement, David Jackson, Chad Gragory, “The Performance Impact of Advance Reservation Meta-Scheduling”, IPDPS 2000 Workshop, JSSPP 2000, Cancun, Mexico, May 2000.

[18] Ahmed Mokhtar, Murat Azizoglu, “Adaptive Wavelength Routing in All-Optical Networks”, IEEE/ACM Trans. Networking, Vol. 6, No. 2, pp. 197-206, April, 1998.

[19] http://www.glif.is.