Optical Dynamic Intelligent Network Services (ODIN): An Experimental Control-Plane Architecture for High-Performance Distributed Environments Based on Dynamic Lightpath Provisioning

Joe Mambretti, David Lillethun, John Lange, and Jeremy Weinberger, Northwestern University

The authors are with the International Center for Advanced Internet Research (iCAIR) at Northwestern University. For more information please go to http://www.icair.org.

ABSTRACT

Currently, many large-scale, resource-intensive applications and services are being developed that can be supported only with high-performance, highly distributed, heterogeneous infrastructures, including Grids. This type of infrastructure is particularly effective for supporting applications and services that must quickly adjust to continuously changing conditions. Such processes require the flexibility of highly adaptive, dynamic, and deterministic resource provisioning. One such architecture is described here. To enhance the performance and flexibility of distributed environments, an experimental architecture for optical dynamic intelligent network (ODIN) services has been designed to enable core optical network capabilities to extend to edge processes, including applications. This architecture allows those processes to directly address and control core network resources, for example, individual lightpaths on demand. This approach supports flexible and deterministic communications by integrating signaled requirements with adjustable network resources. An experimental prototype of ODIN has been designed, developed, and implemented on several optical network testbeds.

INTRODUCTION

We present an experimental architecture for optical dynamic intelligent network (ODIN) services [1]. One motivation for the design and development of a more flexible optical control-plane architecture is to support the diverse infrastructure required for many current and emerging large-scale, data-intensive, distributed processes and applications [2]. For example, Grid environments provide capabilities for the sophisticated orchestration of heterogeneous resources at highly distributed facilities. Grids incorporate services that automatically discover, integrate, utilize, and continually reconfigure multiple resources, including those that are distributed at remote locations [3]. Traditional communication architectures incorporate characteristics that restrict capabilities required by Grids, especially support for large-scale, dynamically changing data flows. These architectures have been designed and developed through processes that anticipate fulfilling fairly static service and process requirements with a limited range of resource allocation options. The architectures on which these environments are based are derived from a set of relatively fixed assumptions about communication service characteristics, traffic volumes and classes, special security considerations, quality of services, sensitivity to latency, and other parameters.

New communication mechanisms must be created in order for highly distributed processes to be supported by dynamic environments. A new architecture is required that includes capabilities for managing real-time provisioning of multiple network resources, including lightpaths. A key component to this approach is an optical control plane that can provide more flexible and dynamic capabilities, especially for traditionally static (e.g., implementation times measured in years) and manual optical provisioning.

RELATED RESEARCH

Current infrastructure research projects, including those that are based on Grids, are experimenting with — and deploying as prototypes — novel methods for optimizing specific combina-

0163-6804/06/$20.00 © 2006 IEEE

tions of distributed resources (e.g., computational, memory, storage, communications, etc.) that can be continually and directly reconfigured and adjusted to meet the needs of dynamic highly distributed processes, services, and applications. Many of the driver applications behind these projects are extremely resource intensive — requiring high levels of compute time, bandwidth, memory, and storage as well as access to specialized distributed facilities and instruments, often at sites throughout the world [4]. To date, almost all of these environments have been based on traditional packet-routed networks. Although these routed data networks have been highly successful in providing scalable services that support millions of small data packets with bursty attributes, there is now a general recognition that they cannot optimally support many emerging applications, services, and other edge processes [5].

More recently, research studies have begun to investigate the potential for enhancing data flow performance within large-scale distributed environments through new architectures that provide options for supplementing, and in some cases bypassing, traditional routing processes by using dynamically provisioned lightpaths. A range of significant problems related to this type of service provisioning exist, especially for providing such channels within highly dynamic unpredictable environment. These projects propose architectures that provide support for data flows with responsive, high quality, high-performance point-to-point channels, including those based on dynamically switched lightpaths. This approach is not a substitute for traditional routed networks, but rather provides for complementary capabilities. These experimental architectures for dynamic provisioning are now being implemented on several advanced optical testbeds, and experimental research is being conducted to investigate their potential for supporting large-scale resource-intensive distributed processes.

One of these experimental projects is OptIPuter, a large-scale research project funded by the National Science Foundation, which is designing an advanced infrastructure that closely integrates optical networking, IP data communications, computer storage, high-performance computational processing, and visualization technologies [6–8]. Another example of this new architectural model is the user-controlled lightpath (UCLP) architecture, which has already been placed into production environments. (www.canarie.ca). Other research is examining methods that combine new L3-L4 protocol architectures with dynamic wavelength switching [9].

RELATED STANDARDS BODIES ACTIVITIES

One context for this recent research consists of the architectural models that are being developed by standards bodies. For example, the Global Grid Forum (GGF) is defining the Open Grid Services Architecture (OGSA), an architecture allowing direct access to, control over, and utilization of multiple distributed resources (www.ggf.org). OGSA addresses both resource-intensive Grid applications as well as more general requirements. As part of that effort, the Grid High-Performance Networking Research Group (GHPN-RG) has been developing descriptions of the types of services that would enable Grid environments to take direct advantage of network resources, for example, through defined Grid Network Services (www.ggf.org) [10].

The IETF has undertaken a variety of efforts related to IP control-plane signaling within multiple working groups. The IETF Common Control and Management Plane (CCAMP) working group coordinates initiatives that are defining a common control plane and a separate common measurement plane. Related IETF efforts are designing the Link Management Protocol and the Generalized Multiprotocol Label Switching (GMPLS) Protocol, an important emerging standard [11]. Several optical control-plane research projects rely on IETF's authentication, authorization, and accounting (AAA) protocol standard for policy-based access management [12].

A related experimental protocol that is being developed as part of a signaling process for optical networks is the Simple Path Control Protocol (SPC) [13]. The SPC protocol allows an application to communicate its need for a specific path through the network to a server capable of establishing the path on lower-layer network elements. When this type of request is signaled, a receiving server is responsible for identifying the appropriate path through the controlled network topology and configuring it to fulfill the request. The server has direct access to state information, which is continuously updated, on the current network topology. SPC also provides a communication method for explicitly releasing resources, such as a path, after it is no longer needed. It also provides a means to make simple queries to core processes about the network state. Also, it can be integrated with policy servers to customize resource utilization and to provide for security controls.

ODIN ARCHITECTURE

The experimental control-plane architecture described here, optical dynamic intelligent network services (ODIN), was designed within the context of these emerging standards. However, it was also designed specifically to address the requirements of large-scale, resource-intensive processes within highly distributed, dynamic environments, including Grids. This architecture incorporates novel mechanisms that allow for dynamic and deterministic lightpath service provisioning as a capability complementary to data transport based on packet routing. This architecture was created to closely integrate edge services with dynamically allocated lightpaths as well as, via other processes, with extensions through non-optical environments, such as vLANs. Essentially, ODIN is an intermediary optical-network service layer, between high-performance distributed processes and lower-level network services and resources. ODIN is comprised of a set of modules, including those governed by policy-based access modules, which

The IETF has undertaken a variety of efforts related to IP control-plane signaling within multiple working groups. The IETF Common Control and Management Plane (CCAMP) working group coordinates initiatives that are defining a common control plane and a separate common measurement plane.

enable edge processes, such as network-aware applications, to signal for defined sets of network services and core optical resources.

ODIN DESIGN GOALS

This architecture is based on multiple design goals. One is to provide a higher degree of virtualization, or abstraction, between edge processes and network services while enabling those processes to directly manipulate core network resources. There are several benefits to this approach: it is complementary to the general industry trend of developing a transport network core with as few layers as possible, it enhances capabilities for providing flexible data-over-lightwave (IP-over-DWDM) service, and it can support network "services on demand," even the creation of ad hoc optical virtual private networks (OVPNs). It can also be used to correlate and integrate edge-process requirements with specific network resources, such as lightpaths with specifically defined characteristics. Another benefit is that it can be implemented within Grid environments to enable the use of lightpaths less as a traditional communication service and more as an extended backplane for distributed computational clusters.

This virtualization approach also allows dynamic optical provisioning capabilities to be integrated directly into applications — enabling them to utilize the full power of dynamic optical networking supported by optical switching and agile photonic components. For example, this architectural approach allows individual applications to provision and manipulate dynamically their own lightpaths without having to deal with the complexity of network provisioning. Also, this architectural approach can make provisioning more optimal or "intelligent," by enabling iterative communication processes between individual application and distributed network elements under changing conditions.

Another design objective was to create an architecture that could be implemented within a central or distributed model — as a centralized process with a single point of control (e.g., residing on a control server) or highly distributed across remote sites. In either case, it has a complete understanding of the topology and current resource allocations within a single domain or across multiple network domains.

COMPONENTS

ODIN consists of several modular components, which are defined within a larger services-architecture context. Other elements within this wider context include an OGSA services model. Initial implementations of ODIN, based on the Open Grid Services Infrastructure (OGSI) implementation, are now being migrated to the Web-Services Resource Framework (WS-RF) as a standard information-exchange process implementation. This technique is compatible with standard communication mechanisms such as the Extensible Markup Language (XML) and the Simple Object Access Protocol (SOAP). Access policy and process registration is based on the IETF AAA architecture, as instantiated through a particular implementation that is being developed for distributed environments [12].

Initial implementations of ODIN, based on the Open Grid Services Infrastructure implementation, are now being migrated to the Web-Services Resource Framework as a standard information-exchange process implementation.

PROVISIONING

Provisioning can be accomplished by individual applications using the ODIN service layer and a signaled overlay control plane for resources (e.g., lightpath allocation and provisioning). Process components consist of server software that can

- Accept requests from clients for resources (the client requests a resource, i.e., implying a request for a path to the resource — the specific path need not be known to the client)
- Determine an available path — possibly an optimal path if there are multiple available paths
- Create the mechanisms required to route the data traffic over the defined optimal path (virtual network)
- Notify the client and the target resource to configure themselves for the configured virtual network

Signaling is a particularly important component of this architecture. By creating a virtualization service layer, this approach can be integrated with existing specialized signaling for provisioning capabilities for discovering, utilizing, and discarding dedicated and individually addressable wavelength-based channels. Various models exist for signaling. For example, an overlay model has different control-plane functions at each level of a layered network without interactions among the layers. A signaled overlay model provides for a limited set of interactions. A peer model enables network nodes of various types to fully interact while maintaining separate control-plane functionality. The ultimate model may be a unified control plane that provides ubiquitous functionality across all network elements. By providing for network service as a virtualization layer, this architecture allows for an insight into traffic behavior within such an integrated environment without fully implementing that model. For example, an implementation of this architecture has been integrated with GMPLS, which is based on the signaled overlay model.

Providing for virtualization through a service layer also provides enhanced capabilities for dynamic requirement changes, which are especially useful for unpredictable environments. The service layer can be the basis for flexibly integrating many different forms of signaling. Given that many edge processes are dependent on requirements within continually changing environments, signaling mechanisms must exist that can provide high-performance support for ongoing cycles of signal-acceptance allocations, monitoring, analysis, and adjustment. The signal processing required includes those related to many functions, such as those for signaling from edge processes (clients), for example, applications, systems, devices, from the management plane (for provisioning and restoration), and from resource management processes (including discovery, scheduling, and reservations). Others

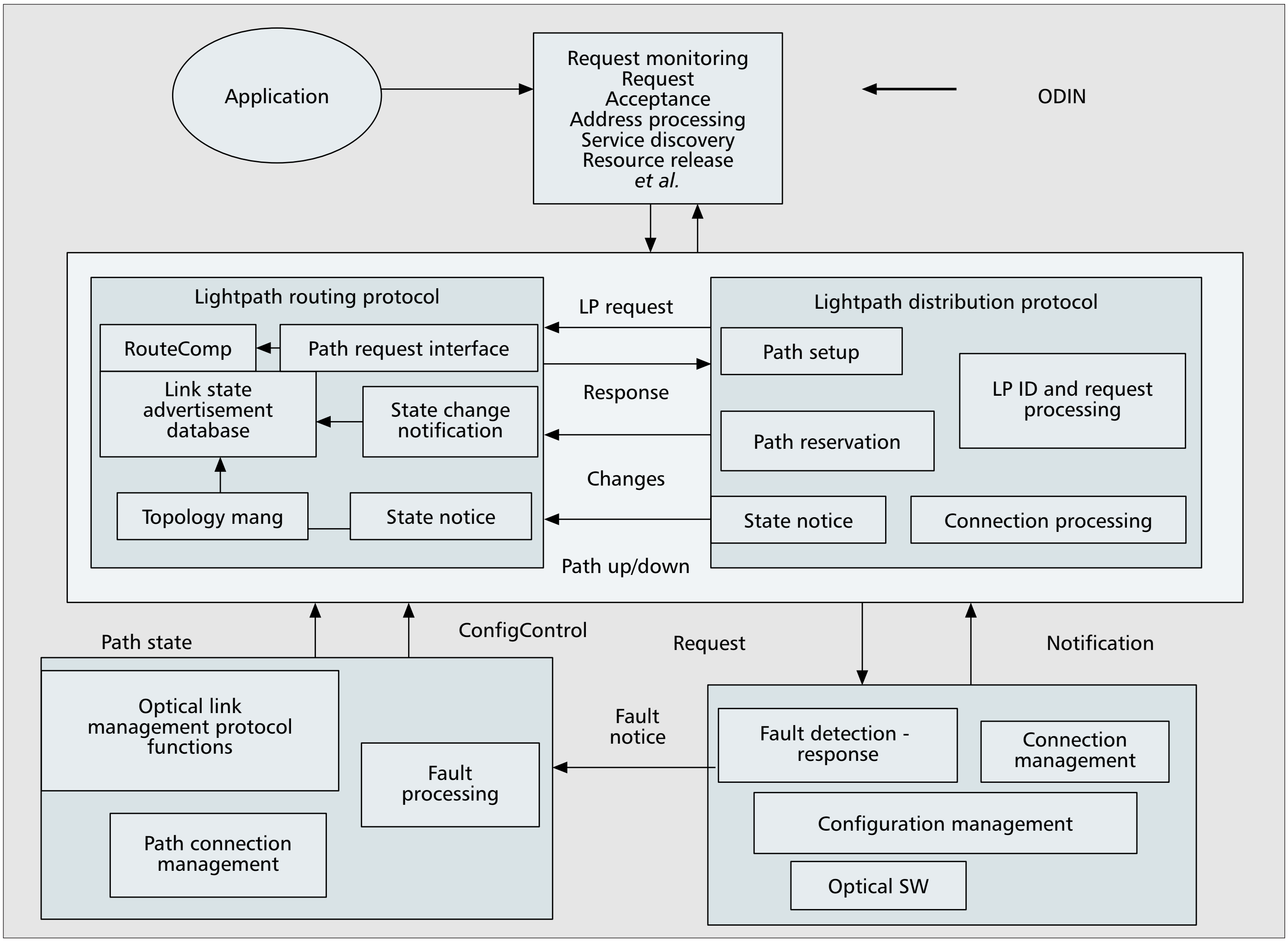

Figure 1. *The ODIN architecture.*

are related to resource contention resolution, lightpath provisioning, deletion, characterization, wavelength assignment, resource advertisements, wavelength routing, device interfaces, optical-switch configuration, data-plane management systems, and optimization signaling through extensions for OSPF for optical routing (Fig. 1).

The categories of services provided by this architecture include those that are oriented to the client side (e.g., edge processes), those that are oriented to the network resource side, those that integrate service requests with available resources, and those that fulfill requests and claim resources when the requests are fulfilled. These processes are highly iterative. This architecture incorporates an overlay signaling model, in that it interoperates with other signaling processes and it is not integrated with them.

PROCESSES

On the client interface, a process monitors a TCP socket for requests from clients (i.e., applications or other edge processes) and responds to those requests over a connected session linked to that application. The process accepts the requests based on policy determinations, and interprets the requests. ODIN returns a new IP and subnet mask in response to a resource request. Requests for services can be simple or highly complex, for example, incorporating many levels of specific service descriptions and implementation characterization, for example, to enable multiservice and multiprotocol support. Simple forms of signaling make the network more transparent. As this signaling increases in complexity, the more “network aware” the edge process or application becomes. On the network side, ODIN continually monitors current network status. This process can be implemented as an active process (e.g., sending inquiries) or a passive process (e.g., continually receiving updates on network state). This architecture provides for a separate process to assist in this task. Such a process is specifically designed as a resource manager integrated with a discovery process that can access topology information and state data such as current resource allocations within an administrative domain.

The basic processes are those that interpret the signaled request in terms of appropriate lightpath routing through dynamic path computation based on a topology discovery of a pool of available links and reachability. However, optimization methods, integrating considerations of constraint-based routing, traffic engineering, and best (versus shortest) path, and associated protection links can be highly complex.

When resources are selected to fulfill the requests, the process communicates with the requisite network switches to configure them to meet those application requirements. These switches can include optical-domain dense wavedivision multiplex (DWDM) switches and Ethernet switches. After the optimal paths are determined, a process implements them, that is, allocates lightpaths, by a signaling process that could include attribute designations (e.g., unidirectional versus bidirectional, protection levels, performance and quality attributes, etc.). This implementation process can also create an ongoing monitoring process in order to ensure that the requested resources are being provided, signal for any required adjustments, and de-allocate the resources when they are no longer required. This core process dynamically provisions lightpaths through processes for discovery, allocation, de-allocation, swap, change attribute, restore if there are faults, and so forth.

This provisioning process interacts with multiple other processes, including those for resource discovery and management, dynamic lightpath-state information management (with capabilities for rapid revisions and updating through iterative processes), and equipment configuration information. Many of these processes require high-performance databases with network element information, including addresses (especially for lightpath identification) and configuration information. Also required are separate processes for performance monitoring and analysis, and processes that ensure survival, protection, and restoration. This architecture can be integrated with instrumentation capabilities that provide for the procedures and metrics for monitoring and analyzing the performance of dynamic large-scale lightpath processes. This architectural model addresses end-to-end provisioning. Consequently, it incorporates a process that enables edge processes not only to control optical paths but also extensions to those paths so as to allow them to be terminated on edge devices through L2 paths, for example, through dynamic vLANs.

EXPERIMENTAL RESEARCH

A prototype of this architecture was designed, developed, and implemented on several wide-area distributed Grid environments, and interconnected by advanced optical testbeds within a metro area, within a statewide optical network, and across an international testbed. One testbed was implemented with completely separate fiber for the control plane and the data plane. These testbed implementations made possible a series of experiments that could be conducted using particularly data-intensive distributed processes, made up of sets of actual Grid application codes and data.

EXPERIMENTAL METHOD

The basic experimental methods consisted of conducting a series of tests using data-intensive processes on distributed Grid clusters. These series of experiments employed an actual applications suite rather than artificially generated data flows. Multiple Grid applications were used in these tests. They included digital media streaming, data mining, distributed data services, and computational astrophysics simulation using adaptive mesh refinement (AMR), and scientific visualization. These applications were highly parallel, based on multiple Grid computational and data clusters running GNU/Linux, and MPICH (1.2.5) and Lam-MPI (7.0.4) libraries were used for messaging. Interprocess communication was significantly enhanced because of the low latency within the network. These clusters were interlinked through several optical testbeds, including the distributed optical testbed, which is an overlay distributed Grid environment. The interfaces between the clusters and the optical networks were high-performance L2 switches, which were directly connected to optical equipment supporting both LAN PHY and WAN PHY on individual wavelengths.

These applications were given access to an API that enabled them to make requests of low-level network service layers, based on the Simple Path Control Protocol (SPC). The ODIN services layer provided an interface to core processes, enabling various sites and applications to provision and manage their own optical paths. ODIN segmented and aggregated individual virtual paths in accordance with the needs of applications. To investigate specifically the traffic characteristics of Grid applications on L1 and L2 channels, no routers were used in the network core or, for most tests, at the edge.

ODIN segmented and aggregated individual virtual paths in accordance with the needs of applications. To investigate specifically the traffic characteristics of Grid applications on L1 and L2 channels, no routers were used in the network core or, for most tests, at the edge.

OPTICAL TESTBED GMPLS IMPLEMENTATION

The ODIN architecture can be integrated with multiple optical-provisioning tools and processes. For these experiments, an implementation of this architecture was integrated with an implementation of GMPLS, which was used as a tool for resource discovery, link provisioning, label-switched path creation, deletion, property definition, traffic engineering, routing, channel signaling, and path protection and recovery. This process gathered the information required to establish lightpaths and determined characteristics, including descriptive information (address identifiers, reachability, etc.). This implementation of an IP-signaled control plane provided for extremely high-performance capabilities for a variety of functions, such as optical-node identification, service-level descriptions (e.g., request characterizations), managing link-state data, especially for rapid revisions, allocating and re-allocating resources, establishing and revising optimal lightpath routes, determining responses to fault conditions, and so on. General functions include:

- Specifying and updating lightpath addressing
- Employing unique identification of path end points
- Determining lightpath availability and reachability
- Dynamically provisioning lightpaths through lambda processing (discovery, add, delete, switch, change, restore, etc.)
- Multiservice and multiprotocol supports (including IP, GE, 10GE, etc.)
- Traffic engineering

- Performance monitoring and analysis
- Survival and protection (Fig 1)

In these experiments, the Constraint-Based Routing Label Distributed Protocol CR-LDP was used for TE signaling [14].

OPTICAL TESTBED ASON IMPLEMENTATION

This testbed implemented an experimental prototype of the automatically switched optical network (ASON) architecture, which was integrated with the ODIN service layer. Through GMPLS out-of-band signaling, top-level processes, including applications, were allowed to dynamically provision lightpaths over the optical core network. ODIN was used to extend the ASON reference architecture through a service layer that could provide for a wide range of capabilities. This implementation employed an interface based on the OIF Optical UNI (O-UNI) standard between edge devices and optical switches.

OMNINET TESTBED

The core testbed used for these experiments, OMNInet, was designed to optimize traffic flow for highly asymmetric, high-performance data communication services in support of Grid applications. OMNinet is a wide-area metro photonic testbed in the Chicago metro area, supporting multiple 10 GE optical channels among four core node sites (www.icair.org/omninet). Each of the core nodes includes a DWDM photonic switch based on micro-electro-mechanical systems (MEMS), an optical fiber amplifier (OFA) to compensate for link and switch dB loss, optical transponders–receivers (OTRs), and high-performance L2/L3 router/switches, providing connections to high-performance Grid clusters at multiple sites. Each photonic node is connected to each of the other nodes with 4 * 10 Gb/s optical channels, that is, each node supports 12 * 10 Gb/s channels (lightpaths or lambdas) on dedicated wavelength-service-qualified fiber, which are dynamically provisioned. In total, OMNInet supports 24 individually addressable wavelengths, 24 * 10 Gb/s channels, allowing individual applications to have 250 Gb/s of core capacity. There are no SONET components or routers in the testbed network. Resiliency and redundancy are addressed through mesh lightpath provisioning. Core photonic nodes are not commercial products but unique assemblies of advanced components customized for this testbed. The 10GigE trunk interfaces, which use true 1550 nm 10 GigE, have been implemented with a specialized set of protocols that allow for enhanced, cross-layer, optical-network intelligence, including at the physical lightpath level, through a wavelength-signaling protocol, a wavelength-routing protocol, and an optical-link management protocol.

Application cluster and compute node access is provided at each location by high-performance L2/L3 switches, provisioned with 10/100/1000 Ethernet user ports, and 1 GigE 1550XD cross-connected trunks. To provide for reliability and optimal L1 performance, OMNInet depends on sophisticated prefault detection and responsive adjustment mechanisms. The photonic core of OMNInet is interconnected by multiple dedicated fiber strands, fully qualified for multiwavelength-based services.

The ODIN processes were implemented on a separate OMNInet control plane that was provisioned out-of-band using completely separate fiber. Such control-plane implementations could also be provisioned on supervisory wavelengths. This control plane enabled user-to-network interface (UNI) control signaling through an OIF UNI compliant interface to the optical transport network and bidirectional signaling to the connection control plane.

A separate management plane was implemented to control network elements, employing standard functions (e.g., fault management, configuration management), including specific resource allocation/de-allocation, performance management, security management, and resource-usage audits. For a number of experiments, including many related to scalability, the OMNInet testbed was extended through a distributed optical overlay testbed, provisioned on OMNInet and I-WIRE, consisting of a meshed L2 implementation provisioned within configurable lightpaths connecting compute clusters at seven geographically separate sites throughout the state of Illinois.

The ODIN processes were implemented on a separate OMNInet control plane that was provisioned out-of-band using completely separate fiber. Such control-plane implementations could also be provisioned on supervisory wavelengths.

EXPERIMENTAL RESULTS

Multiple experiments were conducted on these testbeds using sequential versions of implementations of the ODIN architecture, including many that involved data-intensive Grid applications. In general, the experimental results indicate that this architecture can successfully address some of the current barriers that limit the flexibility and performance of traditional distributed processes based only on routed packets.

These experiments demonstrated that this architecture provides capabilities for exceptional distributed process performance by significantly enhancing the flexibility of distributed environments. Through the virtualization of core optical resources provided by an integrated service layer, this approach enabled capabilities that allowed for rapid control-information exchange, a high degree of responsiveness, direct interoperability between application and dynamic lightpaths, and exceptional data-flow performance among multiple highly distributed sites. The virtualization of network core capabilities enabled edge processes to benefit from flexible resource allocations through capabilities for direct signaling for network resources such as lightpaths (as required), even under changing conditions. Dynamic lightpath allocation and switching was particularly beneficial for high-performance, data-intensive Grid applications.

All applications demonstrated particularly high-performance results, even when they were dependent on data-intensive traffic flows of multiple Gb/s. For highly parallel applications, such as computational astrophysics modeling, performance was key both to data flow and to rapid message passing among distributed processes, which avoided delays in sequenced tasks. Low

> *All applications demonstrated particularly high performance results, even when they were dependent on data-intensive traffic flows of multiple Gb/s. For highly parallel applications, performance was key both to data flow and to rapid message passing among distributed processes, which avoided delays in sequenced tasks.*

latency enabled responsive interprocess communications among multiple node clusters distributed at sites statewide. Applications that require large-scale data flows can often experience suboptimal performance because of issues related to edge nodes, e.g. TCP widows not adjusted, systems tuned for sending data but not for receiving, low-performance NICs. For these experiments, all edge systems were tuned to ensure support for high-performance flows.

Exceptional application performance was primarily a reflection of the performance of the optical network. Almost all of the traffic flows consisted of Gb/s streams or multiple Gb/s streams on dedicated lightpaths. All data flows were provided with maximum network resources whenever possible through dedicated, unshared 10 Gb/s optical circuits. Issues related to fairness allocation for shared channels and to more granulated allocations were not addressed.

Because these experiments used only L1 and L2 transport, the traditional measures of delay-bandwidth product formulated for L3 flows were not meaningful. Another measure of performance is packet loss and loss probability. Quality of service was exceptional, most flows had fewer than 10 packets lost per million, and many had no packet loss at all, even for large-scale long duration (hours) data flows. Many sustained flows exceeded 960 Mp/s with no packet loss, and virtually 0 loss probability. The lightpaths used had a maximum packet loss of fewer than 10 per million under the highest utilization levels (better than $\sim 10^{-10}$ BER) and none with low utilization. For artificially stressed lightpaths, the packet loss was approximately 50 packets per million ($\sim 10^{-10}$ BER).

Other metrics were those that measured the total application process performance. In part, these measurements were made by disaggregating application processing time from that used for network resource signaling and resource provisioning. The control-plane signaling methods and responses were particularly important in enabling services to be established between the applications and optical network resources. These experiments did not attempt to optimize these processes, for example, through recoding. A primary measure of process performance for this series of experiments was time-to-completion for the process, measured in real time. Total performance (TP) was measured as initial request signaling (IRS) + lightpath services processing (LSP) + services acknowledgment (SA) + lightpath instantiation (LI) + process instantiation time (PIT) + processing time (PT) + data flow (DF). The services acknowledgment time and the process instantiation time (from initial signal to completion) were both virtually instantaneous. The path-allocation request time was measured in subseconds (~.5 s), and ODIN server processing time ranged from 3.6 s (allocation) to 11 s (de-allocation). Lightpath Ids were returned in subseconds. MEMs provisioning takes place on the order of milliseconds. However, the propagation of state information among the optical control nodes can take as long as 17–20 s (TP from start of process to initiation of data flow), in part, because each node maintains a separate database that must be opened and closed, and related data files must be written and verified.

Because of these timings, these techniques are best applied in support of large-scale, long-duration data flows. Through optimization, these processing times can be significantly reduced. However, even with significant improvements, challenges remain in integrating high-layer network mirocsecond provisioning processes with the slower timings within optical domains.

The timings for a few of these processes could be addressed by standardizing repeated cycles. For example, one potential method that could be developed for complex but repeated requests would be to create standard "packages" that relate certain types of signals to specific network resources. Many large-scale applications have identical resource-requirement parameters that can be signaled into the network with each initiation, setting off a sequence of resource discovery and utilization processes, or such applications can signal for specifically defined set of predefined resources through an identifier that would evoke the utilization of a preset "package" of all required resources, including the paths that allow access to those resources.

Scalability of this implementation was demonstrated by extending these control-plane signaling and processing techniques from the metro-area testbed, to a statewide testbed, to an international testbed, through optical OC-192 channels dedicated to research between the StarLight international exchange facility in Chicago and the SARA science center in Amsterdam, through the NetherLight facility (www.surfnet.nl).

Path protection that ensures link reliability and maintains the integrity of optical equipment is important. Detecting and locating faults at both the IP and optical layers and rapid responses were also high-priority functions, as was continual high-performance monitoring of state information to avoid fault conditions. Throughout these experiments, no optical components failed and there were very few fault conditions. Well over 1,000 lightpaths were successfully setup and de-allocated with no optical equipment failures.

SUMMARY

The architecture for a virtual optical-network service layer presented here has a significant potential to support large-scale, resource-intensive applications and services that are being created on high-performance, distributed, heterogeneous infrastructures, especially Grid environments. This architecture is especially effective for providing communications support for applications and services that must adjust to constantly changing conditions. These processes must be supported through the flexibility of highly adaptive, dynamic, and deterministic resource provisioning.

ACKNOWLEDGMENTS

Contributors to the research described here include Jim Chen, Fei Yei, Aaron Johnson, Rachel Gold, Tom DeFanti, Maxine Brown, Jason Leigh, Oliver Yu, Bob Grosssman, Paul Daspit, and Rod Wilson. The research presented here was supporting by the National Science Foundation (#ANI-0123399 et al.), DARPA, Nortel Research Labs, and SBC.

REFERENCES

[1] Special issue of *Journal of Future Generation Computer Systems*, Elsevier Science Press, vol. 19, no. 6, Aug 2003, available at www.iGRID2005

[2] I. Foster and C. Kesselman, *The Grid: Blueprint for New Computing Infrastructure*, Morgan Kaufmann, 2003; I. Foster *et al.*, "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration," Open Grid Service Infrastructure WG, Global Grid Forum, June 2002, available at http://www.globus.org/research/papers/ogsa.pdf

[3] T. DeFanti *et al.*, "TransLight: A Global Scale Lambda Grid for E-Science," Special Issue on "Blueprint for the Future of High Performance Networking," *Commun. ACM*, Nov. 2003, vol. 46, no. 11, pp. 34–41.

[4] M. Goutelle *et al.*, "A Survey of Transport Protocols Other Than Standard TCP," Data Transport Research Group, Global Grid Forum, May 10, 2005.

[5] L. Smarr *et al.*, "The OptIPuter," special issue: Blueprint for the Future of High-Performance Networking, *Commun. ACM*, vol. 46, no. 11 Nov 2003, pp. 58–67.

[6] T. DeFanti *et al.*, "Optical Switching Middleware For the OptIPuter," special issue on Photonic IP Network Technologies for Next-Generation Broadband Access, IEICE Trans. Commun. E86-B, 8, Aug. 2003, pp. 2263–72.

[7] J. Mambretti, "Progress on TransLight and OptIPuter and Future Trends Towards Lambda Grids," *Proc. 8th Int'l. Symp. Contemporary Photonics Technology (CPT2005)*, Tokyo, Jan. 12–14, 2005.

[8] R. Grossman *et al.*, "Experimental Studies Using Photonic Data Services at iGrid 2002," *J. Future Comp. Systems*, Elsevier Press, Aug. 2003, pp. 945–56.

[9] D. Simeonidou *et al.*, "Optical Network Infrastructure for Grid," *Grid High-Performance Networking Research Group*, Global Grid Forum, Mar. 2004.

[10] A. Banerjeeet *et al.*, "Generalized Multiprotocol Label Switching: An Overview of Routing and Management Enhancements," *IEEE Commun. Mag.*, Mar. 2001, pp. 144–50.

[11] C. T. A. M. de Laat *et al.*, Generic AAA Architecture, Aug. 2000,IETF — RFC2903.

[12] D. Lillethun, J. Lange, and J. Weinberger, "Simple Path Control Protocol Specification," IETF draft-lillethun-spc-protocol-00.txt

[13] P. Ashwood-Smith *et al.*, "Generalized MPLS Signaling: CR-LDP Extensions," IETF Network Working Group Draft Report, May, 2001.

BIOGRAPHIES

JOE MAMBRETTI (j-mambretti@northwestern.edu) is director of the International Center for Advanced Internet Research (icair.org). His research interests include high-performance computational grids, signaling architecture, control plane architecture, dynamic optical networking, optical transport, switching, and photonics.

DAVID LILLETHUN (davel@cc.gatech.edu) is a Ph.D. student in computer science at Georgia Tech. His research interests include operating systems, signaling architecture, distributed systems, and advanced networking.

JOHN LANGE (jarusl@cs.northwestern.edu) is a Ph.D. student in computer science at Northwestern University. His research interests include grid computing, signaling architecture, operating systems, distributed systems, virtual machines, and advanced networking.

JEREMY WEINBERGER (jeremy@cs.nyu.edu) is a Ph.D. student in computer science at New York University. His research interests include computational music theory, digital media, operating systems, distributed systems, and advanced networking.

The architecture for a virtual optical-network service layer presented here has a significant potential to support large-scale, resource-intensive applications and services that are being created on high-performance, distributed, heterogeneous infrastructures, especially Grid environments.