Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®
Future Generation Computer Systems 19 (2003) 897–908
FGCS
FUTURE GENERATION COMPUTER SYSTEMS
www.elsevier.com/locate/future

The Photonic TeraStream: enabling next generation applications through intelligent optical networking at iGRID2002

Joe Mambretti*, Jeremy Weinberger, Jim Chen, Elizabeth Bacon, Fei Yeh, David Lillethun, Bob Grossman, Yunhong Gu, Marco Mazzucco

International Center for Advanced Internet Research, Northwestern University, Evanston, IL 60201, USA

Abstract

Traditionally, the design and implementation of network based applications, especially large-scale, high performance applications, have had to be compromised across multiple dimensions interfaces, services, performance, flexibility, protocols, architecture, technology, etc. These restrictions exist, in part, because the most widely deployed communications infrastructure was designed to optimize traditional communications, not high performance data communications. At iGRID2002, the International Center for Advanced Internet Research (iCAIR) and its research partners demonstrated "Photonic-Empowered Applications," based on next generation intelligent optical networking technology and dynamic data services provisioning. These demonstrations indicated the potential for creating next generation global applications when traditional barriers to network optimization at multiple levels are removed. These application demonstrations were based on high performance communications infrastructure utilizing novel techniques for managing globally distributed resources and extremely large volume data streams. The innovative advanced optical networking technologies being developed by these research organizations will allow for many new types of high performance global applications across multiple disciplines and industries.

Keywords: TeraStream; iGRID2002; Advanced optical networking; Optical control planes; IP-over-DWDM; Advanced photonic technology; Dynamically switched wavelengths; Lambda Grid

1. Introduction

Traditionally, the design and implementation of network-based applications, especially large-scale, high-performance applications, have had to be compromised across multiple dimensions—interfaces, services, performance, flexibility, protocols, architecture, technology, etc. These restrictions exist, in part, because the most widely deployed communications infrastructure was designed to optimize traditional communications, not high-performance data communications. At iGRID2002, the International Center for Advanced Internet Research (iCAIR) and its research partners demonstrated "Photonic-Empowered Applications", based on next generation intelligent optical networking technology and dynamic data services provisioning. These demonstrations indicated the potential for creating next generation global applications when traditional barriers to network optimization at multiple levels are removed. These application demonstrations were based on high-performance communications infrastructure utilizing novel techniques for managing globally distributed resources and extremely large volume data streams. The innovative networking technologies

* Corresponding author.
E-mail address: j-mambretti@northwestern.edu (J. Mambretti).

0167-739X/03/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0167-739X(03)00069-4

being developed by these research organizations will allow for many new types of high-performance global applications across multiple disciplines and industries.

2. Background

iCAIR and its partner research organizations have undertaken a number of projects directed at creating new advanced digital communication services based on intelligent next generation optical networks [1]. These initiatives have been undertaken primarily to respond to the needs of multiple large-scale advanced applications [2]. New data communication services have the potential to significantly enable many types of innovative, powerful global applications, such as those demonstrated at iGRID2002. Almost all of these applications are data and computationally intensive, especially those based on computational Grids. Grids are part of a next generation "cyberinfrastructure", often used for extremely large-scale, resource intensive applications [3]. Grids are "persistent environments that enable software applications to integrate instruments, displays, computational and information resources that are managed by diverse organizations in widespread locations [4]". Some emerging Grid infrastructures, such as the TeraGrid, use networks not for standard communications support but as backplanes for high-performance computational clusters, comprised of hundreds or thousands of individual compute nodes within widely distributed clusters [5]. Other emerging infrastructures envision global services based on data communications infrastructure that is primarily dependent on layer 2 transit end-to-end as opposed to routed paths [6]. The "Global Lambda Grid", a concept that is being formulated by the StarLight community, envisions supporting Grid infrastructure through world-wide wavelength-based data communications [7]. The global lambda Grid will support multiple advanced applications, including those related to high-performance computational scientific research, high energy physics, engineering, bioinformatics, computational genomics, high resolution medical imaging, materials sciences, data streaming, digital media, data mining, and astrophysics.

Many of the technologies supporting the photonic-empowered application demonstrations at iGRID2002 were developed on, or for, OMNInet, a metro area photonic networking testbed in Chicago. OMNInet has been established, in part, to create a reference model for next generation optical metro networks [8]. At iGRID2002, this testbed was extended to Amsterdam through StarLight and NetherLight in order to show that Photonic Enabled Applications are possible not only on next generation optical metro area networks (such as OMNInet), but can also be extended to global networks, e.g., through StarLight and NetherLight, as a basis for the global lambda Grid. StarLight is a next generation global optical networking exchange facility in Chicago [7], and NetherLight is a trans-Atlantic high-performance link connecting StarLight to SURFnet in Amsterdam [9]. These facilities are devoted to research, that is, experimental, not production, networking.

2.1. Photonic-empowered applications

The term "photonic-empowered applications" refers to multiple, global, next generation applications that are being designed and developed as highly asymmetric, highly distributed (including those based on resources at sites world-wide), and resource intensive, e.g., computationally intensive, bandwidth intensive, storage system intensive, etc. However, in addition, they are distinguished also by their utilization of advanced data communications based on dynamic multiwavelength lightpath provisioning and supported by more flexible DWDM-based networking technology than that which is implemented in today's static point-to-point optical networks. They are also optical network "aware", that is, they have a capability for directly discovering and signaling for use of the networking resources that they require, including signaling for the provisioning of lightpaths. In addition, some of these types of applications may be highly periodic and transient (e.g., they may exist only for a few moments at different times throughout a month or throughout a day). Consequently, they may transition instantaneously from a state requiring little or no network utilization to one requiring enormous network resources for days, hours, minutes, or moments, or even milliseconds. Many of these types of applications require a much closer integration of such resources than are currently available through existing information technology infrastructure, which tends to distinctly segment system components. Within the emerging new infrastructure, the boundaries between applications, computers, and networks truly dissolve. This architectural direction is inherent within leading-edge information technology development projects, such as those focused on Grid computing, Globus middleware (http://www.globus.org), and new experimental (non-production) research networks.

Although these types of application characteristics may seem unusual, they actually reflect real world activities. If they seem extraordinary, it may be because applications are so frequently severely compromised by various restrictive digital infrastructure. One simple specific example of this type of restriction, of many that could be cited, is the issue of visualizing dynamic high definition 3D models and simulations, which is critically important in many organizations. Yet, very few public, or even private, venues exist that provide capabilities for displaying these visualizations. Consequently, addressing the resource needs of these next generation applications requires a new approach to data communications.

2.2. Photonic-empowered applications at iGRID2002

The Photonic TeraStream was designed to allow for experimentation with new techniques for provisioning for high-performance composite applications. The Photonic TeraStream application was developed as a prototype "composite application"—that could potentially integrate several component applications, including high-performance data transfer (based on GridFTP), digital media streaming (270 Mbps encoding, 600 with Forward Error Correction enabled), high-performance remote data access methods (based on iSCSI [10]), and dynamic provisioning. At iGRID2002, the prototype Photonic TeraStream demonstration illustrated the potential for supporting global applications with next generation wavelength-based networking, which includes allowing those applications to utilize directly the optical network control plane. Such new applications could be based on techniques for provisioning "Global Services-on-Demand", a method that allows applications to select services used. For example, iCAIR, in partnership with the Materials Sciences Research Center at Northwestern University is developing an International Virtual Institute for Materials Science (IVIMS), which is being funded by the National Science Foundation. The IVIMS requires a high-performance capabilities for instantaneously discovering, gathering, integrating, and presenting for a global set of users different sets of resources from throughout the world. These resources include large-scale data streams from experimental repositories at remote locations, scientific visualizations and digital media, and computational processes.

2.3. Photonic-empowered data mining, National Center for Data Mining at UIC and OMNInet

To provide for another set of reference application requirements for next generation network services, the National Center for Data Mining at UIC (http://www.ncdm.uic.edu) and iCAIR are developing methods for integrating high-performance data mining techniques with the methods for dynamic lightwave provisioning (Photonic Data Services). To prepare for iGRID2002, researchers at iCAIR and NCDM conducted a series of tests to ensure optimal performance of a variety of network components and protocols, such as TCP and UDP, including testing methods using services for parallel TCP striping (GridFTP). Researchers at the NCDM have been using OMNInet to test protocols that they developed to allow for the design of network-based applications with reliable end-to-end performance and speeds that scale to multiple Gbps. These protocols include PSockets and SABUL, which are open source libraries to build network applications with advanced functionality. NCDM's SABUL is an innovative protocol that uses UDP as a transit protocol but provides for reliability by using TCP as a control protocol. At iGRID2002, the NCDM and iCAIR presented a Photonic Data Services demonstration that set a new high-performance record for trans-Atlantic data transit [11,12].

2.4. Photonic-empowered e-science

Multiple large-scale e-science applications exist that can benefit from Photonic Data Services. These applications include many in high energy physics that utilize extremely data intensive traffic flows. The

1–100 TB data "transactions" for some high energy physics experimentation must be carried out within short time frames. Using Photonic Data Services, many transactions can be accomplished per day, using new methods for accessing, processing and analyzing data. Another application, in astrophysics, is realtime synchronized acquisition of data from large telescope arrays. Another is realtime data analysis in computational bioinformatics, especially with digital imaging.

2.5. Architectural considerations

The demonstrations shown at iGRID2002 envision large-scale, powerful, high-performance photonic-empowered applications liberated from the multiple restrictions inherent in today's traditional communication infrastructure. However, before applications can become "empowered", they must be "network aware" or "network intelligent"—they must have an understanding of available network resources, of how to access and use those resources, and of how to adjust to dynamic changes in those resources. Consequently, several projects have been established, at iCAIR and elsewhere, to focus on specialized signaling methods that will allow very large-scale, high-performance, distributed applications to directly manipulate a wide range of optical networking functions. Such signaling would allow applications to provision and control their own global, dynamically provisioned lightpaths, which could be implemented as global dynamic VPNs (GDVPNs), for example, as a basis for the global lambda Grid. The Photonic TeraStream demonstration was supported by an innovative architectural method that envisions closer integration between applications and advanced networks and that also anticipates providing applications with access to new types of intelligent network resource signaling methods, i.e., providing them with an "intelligence" through control capabilities provided by a series of optical networking service layers. Such "empowered" applications could use such service layers to access control plane and transport plane processes, including functions allowing for application-controlled dynamic lambda provisioning and switching.

These research project directions are consistent with the general data communication architectural trend directed at removing hierarchical layers in networks, not only to eliminate cost and complexity but also to advance the goal of providing digital services with transparency. As an example of this trend, service providers have been replacing traditional infrastructure based on ATM-over-SONET/SDH-over fiber with packet-over-SONET (POS). Although SONET is highly reliable and provides for efficient, well-known management and traffic engineering (TE) techniques, it is neither optimal nor cost-effective as a support foundation for digital services. In addition, SONET does not allow for rapid deployment of new and enhanced services and capacity.

Consequently, the research described here is developing techniques that remove the SONET layer and that place data services directly on lightpaths, with dynamic wavelength provisioning. This architecture takes advantage of the capabilities of a fairly mature set of technologies, the dense wavelength division multiplexing (DWDM) and related technologies that have been used in long haul networks since the late 1980s [13]. DWDM-based services have now been deployed in regions, metro areas, and are even moving into the enterprise. DWDM divides a beam of light into multiple "colors", lightwaves—or "lambdas", so that multiple optical signals can be transmitted through fiber in parallel—extremely high volumes of data can be communicated through each strand of optical fiber. Newer technology provides capabilities for dividing the light into many dozens of channels on each fiber length and for sending tens of gigabits or more through each channel. Single fibers are capable of transmitting many terabits of data, which will lead to ubiquitous broadband and significantly lower cost for digital services. However, almost all of the current deployments are fairly static, supporting primarily point-to-point links. Although various efforts are underway to develop techniques for what has been described as "point-and-click" provisioning, iCAIR and its research partners are attempting to move beyond this type of manual intervention and provide mechanisms that allow for applications to directly signal for and utilize network resources that they require. Key prototype technologies that comprised the iGRID2002 demonstration were: (a) intelligent application signaling, (b) dynamic lambda provisioning, and (c) extensions to lightpaths through dynamically provisioned L2 and L3 configurations, in part, to allow for access to multiple types of edge resources.

2.6. Intelligent application signaling and TeraAPI—a UNI for photonic-empowered applications

The Photonic TeraStream prototype application is based on new signaling methods, not only for dynamic lightpath provisioning, which could be used to create optical virtual private networks (OVPNs), but also for extending those lightpaths to edge resources through other types of dynamically provisioned L2 gigabit Ethernet links, including complete Ethernet vLANs. As an initial step toward accomplishing these goals, a simple preliminary application protocol for making requests of low level network service layers has been developed, the Simple Lightpath Control Protocol Specification (SLCP) [14]. This type of mechanism is required because these applications must be much more "network aware" than most current applications especially about changing network and edge resource dynamics. At iGRID2002, Photonic TeraStream application interacted with a client (odin_cli) based on a preliminary version of this architecture, that served as the signaling method for requests from lower-layer network service layers. The odin_cli also served for demonstration purposes as an application proxy. In addition, for demonstration purposes, a preliminary architectural model, the TeraAPI was designed and developed (in C) as a prototype API. (The term "Tera" is used to indicate that the architecture is being designed to scale to TeraScale infrastructure, e.g., in a networking context, terabits per second.) This architectural approach is being designed and developed as a mechanism to experiment with methods for network resource discovery and high level (i.e., application to network service layer) intelligent application signaling for core resources, e.g., dynamic lightpath management, and for adaptive application responses to dynamic network conditions.

The TeraAPI served as the interface between the application, e.g., Photonic TeraStream, and low level optical networking resources, through additional service intermediaries currently in development at iCAIR, such as: (a) the optical dynamic intelligent network (ODIN) service layer, (b) the TeraScale high-performance optical resource-regulator (THOR), which manages the optical network control plane and resource provisioning, including dynamic provisioning, deletion, and attribute setting of lightpaths, and (c) the dynamic ethernet intelligent transit interface (DEITI), which has a capability extending lightpaths beyond wavelengths to other L2 links, currently GE links, for example, to allow the Photonic TeraStream to access edge resources such as compute clusters and data storage repositories. These links can be provisioned as GE vLANs. In addition, the TeraAPI must interact with a range of network middleware functions, described below. There are multiple ways that such an interface can be utilized. For example, two general models are: (a) the application requirement parameters can be signaled into the network at each instantiation requiring recourses, setting off a sequence of resource discovery and utilization processes or (b) the application can signal a specific call including a set of known parameters that would evoke the utilization of a pre-set "package" of resources, including all required resources and linkpaths that allow access to those resources.

2.7. ODIN services

The Photonic TeraStream utilizes the TeraAPI to access the ODIN services module [15]. The ODIN service layer is software being designed and developed at iCAIR as an intermediary between high-performance distributed global applications and lower level network service layers. Collectively, these service layers allow for dynamic, integrated coordination between the photonic-empowered application that may reside on a client network with various processes and resources at the optical network layer. The ODIN services layer enables high-performance applications while providing for network transparency, allowing the applications to utilize the full power of photonic networks without having to deal with their complexity. ODIN is a powerful mechanism that can be used to establish "services on demand" and, as noted, not only dynamically allocated lightpaths, but also dynamically allocated transient (or permanent) OVPNs.

ODIN provides a single point of control for a defined set of network service requests within a single administrative domain. This point of control is incorporated within a process that resides on a control server. The process has a complete "understanding" of the topology and current resource allocations within the administrative domain. ODIN accepts requests for resource allocations from applications over the network. The process listens on a TCP socket for requests from applications, and responds to those requests over a connected session linked to that application. When resources are allocated to fulfill those requests it communications with the requisite network switches to configure them to meet those application requests. These switches can be optical-domain DWDM switches, Ethernet switches and/or IP routers. In sum, ODIN is server software that is comprised of components that: (a) 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), (b) determines an available path—possibly an optimal path if there are multiple available paths, (c) creates the mechanisms required to route the data traffic over the defined optimal path (virtual network), (d) notifies the client and the target resource to configure themselves for the configured virtual network (ODIN returns a new IP and subnet mask in response to a resource request).

There are four primary architectural models for optical network control planes (with many hybrids and variations), overlay, signaled overlay, peering, and integrated. The overlay model has separate control planes, for example, for both a client network, e.g., an IP network, and for its underlying supporting optical network, which is connected to the client network via a UNI. OMNInet is based on signaled overlay architecture, which allows for client networks to signal into the optical network to allow for dynamic services and resources provisioning. On the OMNInet testbed, this type of provisioning is being accomplished directly by applications through an optical route allocation and management system, an interface between the ODIN service layer and the OMNInet control plane that does the primary work of the dynamic lambda provisioning required by applications such as the Photonic TeraStream. This interface is the THOR.

2.8. THOR

THOR is a process that establishes and deletes lambda-switched paths, i.e., lightpaths, based on an understanding of application requirements, physical optical network topology, potential capabilities for resource allocations within that topology, and performance optimization. THOR components include mechanisms for receiving requests, fulfilling requests, such as allocating and managing network resources (e.g., routes), and for monitoring state information such as route configuration data. After receiving a client (e.g., application) request(s) through ODIN, THOR determines the state of the lambda-based lightpaths (lambda-switched paths) in the optical network, determines the most optimal lightpath(s) for a particular request, creates a lightpath, which could be an OVPN, by configuring the photonic node switches, notifies the client and the target resource to configure themselves (e.g., for use of the OVPN), and reallocates optical resources when they are no longer being used. THOR has an understanding of the network configuration to such a degree that it can allocate lambda-switched paths (resources at the level of multiple Gbps), but also provide for lambda resource sharing (i.e., multiple paths on a single lambda). THOR controls the DWDM layer by direct calls to a UNI API that Nortel Research Labs developed for OMNInet.

2.9. OMNInet, GMPLS, and standardization activities

This research is being conducted within the context of multiple emerging architectures being developed within standard bodies. For example, key related IETF efforts are those that relate to the link management protocol and the control plane efforts related to the development of the generalized multiprotocol label switching (GMPLS) protocol, an emerging IETF standard. The photonic layer control tools used on OMNInet are based on a GMPLS implementation, designed and developed by Nortel Research Labs, which has additional optimization components. The GMPLS architecture [16–21] has gained significant momentum. GMPLS can be used for resource discovery, link provisioning, label switched path creation, deletion, and property definition, TE, routing, channel signaling, and path protection and recovery. GMPLS is the generalized extension of multiprotocol label switching (MPLS) [22–25], a signaling protocol with a flexible framework which specifies separating forwarding information from IP header information, allowing for forwarding through label swapping and various routing options. The MPLS architecture assumes a forwarded plane that recognizes packet or cell boundaries and provides processes based on packet or cell headers.

GMPLS provides for extensions that include forwarding planes that are not capable of recognizing such boundaries, such as traditional devices based on time-division multiplexing (e.g., SONET ADMs) and newer devices, based on wavelengths (optical lambdas) and spatial switches (e.g., in flow port or fiber to out flow port or fiber) [26]. Consequently, GMPLS allows for forwarding decisions to be based on time slots, wavelengths, or ports. Path determination and optimization is based on labeled switched path (LSP) creation. This process gathers the information required to establish a lightpath and determines characteristics, including descriptive information (address identifiers, reachability, etc.) [27]. This type of IP control plane provides 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, etc. General functions include: (a) specifying and updating lightpath addressing, (b) employing unique identification of path end points, (c) determining lightpath availability and reachability, (d) dynamically provisioning lightpaths through lambda processing (discovery, add, delete, switch, change, restore, etc.), (e) multiservice and multiprotocol supports, including IP, GE, 10 GE, etc., (f) TE, (g) performance monitoring and analysis, and (h) survival and protection.

Extensions to GMPLS provide for a range of extensions for routing, OSPF, and general TE functions [28–30]. GMPLS TE extensions include those that allow for CR-LDP specific formats and mechanisms and for RSVP-TE signaling [31,32]. Path protection that ensures the protection of existing paths is a key requirement, and requires continual high-performance monitoring of state information [33]. Detecting and locating faults at both the IP and optical layers and rapid responses are also high priority functions. Currently, GMPLS development is focused on intradomain networks. Other efforts with the IETF are attempting to develop techniques that address interdomain networking while avoiding link-state protocols and their related complexity [34]. Currently, the efforts described here are primarily focused on Intradomain issues. However, future work will explore the challenges of interdomain provisioning. In this regard, some innovations such as optical BGP (OBGP) seem promising [35].

2.10. OMNI and OMNInet

The optical metro network initiative (OMNI) OMNInet is a metropolitan optical network testbed currently being deployed in Chicago and Evanston, Illinois. This project is a joint partnership among SBC, Nortel, iCAIR, the Electronic Visualization Laboratory (EVL) of University of Illinois at Chicago (UIC), the Math and Computer Science Division at Argonne National Laboratory, and CANARIE, the Canadian national research network. The OMNInet testbed is being designed to optimize for metro area data services, which has several implications. One of these implications of this optimization approach is that there are no SONET components. (SONET-based networks are optimized for traditional communication services not data communications.) Another is that the types of networking technologies being designed and implemented for this testbed are those that specifically address metro area vs. long haul network requirements. Also, this testbed has been designed with a focus on optimizing for highly asymmetric, high-performance data communication services. In part, these initiatives are addressing issues such as optimization for network resources and high-performance scalability protocols as traffic flows transition from multiple Gbps, to tens of Gbps, to 40 Gbps and above. The development of such protocols is complementary to developments in methods for dynamic optical and packet path provisioning. Consequently, investigating and developing the integration of, and complementary roles of, packet and wavelength switching is important research topic for the near term.

OMNInet is an integrated network system with three primary functions: (1) unified IP routing, using, as noted, an implementation of an IETF GMPLS, (2) an optical networking system, based on microelectro-mechanical systems (MEMS), that supports lambda switching, and (3) a cross-layer network intelligence, including at the physical lightpath (DWDM) level. The physical core of OMNInet is a metro area dark fiber fabric that has been qualified for multiwavelength-based services. OMNInet's initial testbed is a metropolitan-area, four-node, optical network linking a core node on Northwestern University's (NU) Chicago campus (just northeast of Chicago's downtown "Loop" area), with nodes at the UIC (west of the Loop), at the CANARIE CA*net3/4 node at its Chicago Point of Presence (PoP) (just south of the Loop), and at a node on NU's Evanston campus (20 miles north of the Loop). This deployment constitutes a very wide area testbed sufficient to conduct advanced experimentation in next generation phonic-based networking. Each node includes an MEMS-based Wave Division Multiplex (WDM) photonic switch, an optical fiber amplifier (OFA), optical transponders/receivers (OTRs), and high-performance L2/L3 router/switches. These nodes were designed and developed by Nortel Research Labs. (The core nodes are not products but rather innovative research concepts, custom built by Nortel Research Labs.) The photonic switches are supported by Optera 5200 OFAs to compensate for link and switch dB loss with eight ports capable of supporting 10 Gbps optics. Application cluster and compute node access is being provided at each location by Passport 8600 L2/L3 switches, which are provisioned with 10/100/1000 Ethernet user ports, and 1 GE 1550XD trunks (i.e., dedicated fiber that supports 1550 nm cross-connect trunks).

2.11. Out-of-band management layer and physical layer monitoring and adjustment

Currently, a separate OMNInet control plane has been provisioned out-of-band using completely separate fiber, provided through the Metropolitan Research and Education Network (MREN). At some point in the future, this control plane will be migrated to a supervisory wavelength. This control plane enables User-to-Network Interface (UNI) control signaling via a UNI interface to the optical transport network and bi-directional signaling to the connection control plane. 10 GE trunk interfaces, using true 1550 nm 10 GE, have been implemented, with a specialized set of protocols that allows for enhanced optical network intelligence, including a wavelength signaling protocol, a wavelength routing protocol, and an optical link management protocol. Optical wavelengths are being used to implement virtual lightpaths to transport data streams (relevant to scientific applications) with specific service levels. In part, control plane software is being used to segment aggregated and individual virtual paths in accordance with the characteristics and requirements of specific applications, including those related to traffic classes, security requirements, latency sensitivity, etc.

OMNInet is supported at L1 by state-of-the-art lasers, amplifiers, and filters, which are part of subsystems and systems developed by Nortel Research Labs. To provide for reliability and optimal L1 performance, OMNInet is provisioned with a wide range of sophisticated pre-fault detection mechanisms, also developed by Nortel Research Labs, which monitor network conditions and adjust resources in response to specific detected characteristics. These types of critical components are key elements to next generation photonic-based networks. The photonic core of OMNInet is interconnected by multiple fiber strands, dedicated to the testbed, fully qualified for multiwavelength-based services. The fiber provisioning for OMNInet is being undertaken by SBC, as part of the technology trial.

2.12. DEITI

Applications such as the Photonic TeraStream require access to resources that are not linked to wavelengths. This research initiative recognizes that it is difficult to deploy lightpaths over wavelengths to all areas where resources must be accessed, especially at the edge of the network. However, to optimize performance, it is necessary to provide a means to instantaneously provision L2 transit paths while also avoiding the performance degradation that would result by introducing unnecessary L3 packet inspection processes. One potential solution would be to implement high-performance MPLS. However, currently few edge routers are MPLS enabled. Another potential solution is to utilize dynamic L2 provisioning through Ethernet vLANs. Consequently, iCAIR has established a project to design an L2 extension to its ODIN service layer beyond lightpath provisioning. This technology—DEITI allows for extending optical resource provisioning to dynamic vLANs. To some degree, this capability can be considered a proxy or substitute for MPLS. However, it may prove useful over the longer term in many areas where MPLS is not implemented.

2.13. (I)AAA and dynamic vLANs

Given that these research projects assume providing applications with access to key network resources, it is important that some consideration be given to middleware access policy and control. Middleware components include APIs incorporating a policy framework, such as authentication, authorization, and accounting (AAA), directories, resource management, networked information discovery and retrieval services, quality of service, security, and operational tools. Such considerations should also include specific separate considerations of identification, thus, (I)AAA. To address these issues, a partnership has formed with researchers from University of Amsterdam (UvA) who have been creating and experimenting with new techniques to in these areas.

At iGRID2002, the UvA research contingent on this project demonstrated a technology that includes multiple considerations of access control to network resources. This type of middleware can serve as useful context to identify various components that require integration into an application signaling architecture, including signaling methods, access/admission controls, and a series of defined services and related resources, management of service levels and priority attributes, scheduling, service attribute setting functions, feedback mechanisms for notifying applications or systems about performance variations, etc. These types of mechanisms imply various capabilities for: (1) an interaction with some type of policy implementation and enforcement, (2) dynamic assessment of available network resources, (3) policy monitoring, (4) service guarantees, (5) conflict resolution, and (6) restitution for lack of performance and/or fulfillment. The UvA researchers demonstrated an Ethernet-based switch setup allowing a 802.1Q VLAN configuration. For this demonstration, modular switches were interconnected using a 1000BaseSX connection. Each switch supported 16 FE connections into the GbE uplink. Four stations (each with dual FE interfaces) were connected in pairs to the two switches. One interface of a station was connected to a switch, the other to a common network. A fifth station was used as AAA server that controls both switches using SNMP and the 802.1Q bridge mib extensions. With this implementation, a station could send an XML/SOAP-based connection request to the AAA server, that then obtains a policy that determines if the request can be honored or not. If the request was granted, a bypass connection was created via the VLAN switches using the second ethernet interface for a specified amount of time. The iGrid2002 demonstration showed the setup, the request, the policy and its execution. This principle could be used to create connections on demand that bypasses the regular Internet for high-performance data intensive applications (that would justify this kind of implementation). This demonstration was good example of the application of the basic principles of the IETF Generic AAA architecture [36].

2.14. The optical Grid

Integration of Grid architectural concepts with those emerging from next generation optical networking, such as Photonic Data Services, has just begun. The concept of optical network aware (Photonically Empowered) applications signaling requests to an intelligent, dynamically configured wavelength-based network complements the overall objectives and direction of Grid architecture. This type of integration would provide Grid communities with powerful new global data path capabilities. Also, this integration could provide operational benefits. For example, given that these networks will require a global real-time monitoring system, e.g., to manage requests and ensure fulfillment, it would be useful if these functions could be integrated with related Grid intelligent system capabilities.

2.15. Future research: application of intelligent signaling and control of dynamically switched networks

Research with new techniques for next generation optical networking based on dynamic lambda provisioning is continuing on the OMNInet testbed. A number of these research projects are also part of an on-going project in “Application of Intelligent Signaling and Control of Dynamically Switched Networks”, which is being funded by the National Science Foundation. This project was established as a joint research partnership of the EVL of the UIC (http://www.evl.uic.edu) and iCAIR [37]. This project is experimenting with, and developing novel methods for, application level dynamic control of resource discovery, allocation and adjustment to allow more flexibility in service provisioning, infrastructure deployment and service resource management. The research components include: (a) research into the behavior of advanced science applications, not just on an extremely high-performance optical network, but on one that can be dynamically adjusted on a granular level, (b) identification of application-level network requirements, (c) investigation of management techniques for optical networks and of new service provisioning models, (d) research into new methods for application signaling, (e) investigation of interconnections between application signaling and IP-based control planes, including with GMPLS, (f) test deployment of these techniques on an advanced optical testbed, (g) analysis of results, (h) experimentation with multiservice provisioning to ensure gateways to traditional networks and protocols, (i) development of a system for performance metrics, monitoring and analysis, and (j) creation of a testbed for StarLight, the next generation, optically-based Science Technology and Research Transit Access Point (STAR TAP) and for other advanced research networks.

3. Summary

Traditionally, the design and implementation of many applications have to be compromised because of the technical limitations of infrastructure. iGRID2002 provided a unique opportunity to demonstrate the potential for global next generation applications with many of these restrictions eliminated. The photonic-empowered applications are one example, based on next generation intelligent optical networking technology, dynamic data services provisioning, and utilization of novel techniques for managing globally distributed resources and extremely large volume data streams. The further development types of innovative networking technologies will allow for the creation of high-performance global applications across many disciplines and economic sectors.

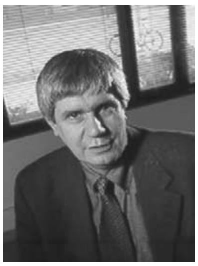
Acknowledgements

Hal Edwards, Paul Daspit (Nortel, http://www.nortel.com), Teresa Elliott, Rachel Alarcon (SBC/Ameritech, http://www.sbc.com), Kees Negers, SURFnet (http://www.surnet.nl), The Netherlands, Linda Winkler, Bill Nickless, Caren Litvanyi, Argonne National Laboratory.

References

[1] http://www.icair.org.
[2] J. Mambretti, A. Schmidt, Next Generation Internet: Creating Advanced Networks and Services, Wiley, New York, 1999. ISBN 0-471-32762-X.
[3] I. Foster, C. Kesselman (Eds.), The Grid: Blueprint for a Future Computing Infrastructure, Morgan Kaufmann, Los Altos, CA, 1999.
[4] http://www.globalgridforum.org.
[5] http://www.teragrid.org.
[6] Information on the OptIPupter project. http://www.calit2.org and http://www.evl.uic.edu.
[7] http://www.startap.net/starlight.
[8] http://www.icair.org/omninet.
[9] http://www.uva.nl.
[10] J. Satran, K. Meth, iSCSI, September 2002. draft-ietf-ips-iscsi-18.txt.
[11] R. Grossman, Y. Gu, D. Hanley, X. Hong, J. Levera, M. Mazzucco, D. Lillethun, J. Mambrett, J. Weinberger, Photonic data services: integrating path, network and data services to support next generation data mining applications, in: Proceedings of the NGDM'02, November 2002, in press.
[12] R. Grossman, Y. Gu, D. Hanley, X. Hong, J. Levera, M. Mazzucco, D. Lillethun, J. Mambrett, J. Weinberger, A case for global access to large distributed data sets using data webs employing photonic data services, in preparation.
[13] T. Stern, K. Bala, Multiwavelength Optical Networks: A Layered Approach, Addison-Wesley, Reading, MA, 1999.
[14] D. Lillethun, J. Weinberger, Simple lightpath control protocol specification, in preparation.
[15] D. Lillethun, J. Weinberger, J. Mambretti, ODIN: path services for optical networks, in preparation.
[16] E. Mannie, et al., GMPLS extensions for SONET and SDH control. draft-ietf-ccamp-gmpls-sonet-sdh-01.txt.
[17] E. Mannie, GMPLS signaling extension to control the conversion between contiguous and virtual concatenation for SONET and SDH. draft-mannie-ccamp-gmpls-concatenation-conversion-00.txt.
[18] E. Mannie, et al., Generalized multi-protocol label switching (GMPLS) architecture. draft-ietf-ccamp-gmpls-architecture-00.txt.
[19] A. Bellato, G.709 optical transport networks GMPLS control framework. draft-bellato-ccamp-g709-framework-00.txt.
[20] A. Bellato, GMPLS signaling extensions for G.709 optical transport networks control. draft-fontana-ccamp-gmpls-g709-00.txt.
[21] O. Aboul-Magd, A framework for generalized multi-protocol label switching (GMPLS). draft-many-ccamp-gmpls-framework-00.txt.

[22] R. Callon, et al., A framework for multiprotocol label switching, 1999. draft-ietf-mpls-framework-03.txt.
[23] E. Rosen, et al., Multiprotocol label switching architecture, 1999. draft-ietf-mpls-arch-05.txt.
[24] E. Rosen, A. Viswannthan, R. Callon, Multiprotocol label switching architecture, IETF RFC-3031, January 2001.
[25] D. Awduche, Y. Rekhter, Multi-protocol lambda switching: combining MPLS traffic engineering control with optical crossconnects, IEEE Commun. Mag. 39 (March 2001) 111–116.
[26] B. Davie, P. Doolan, Y. Rekhter, Switching in IP Networks: IP Switching, Tag Switching, and Related Technologies, The Morgan Kaufmann Series in Networking, Academic Press, New York, 1998.
[27] J. Lang, et al., Generalized MPLS recovery mechanisms. draft-lang-ccamp-recovery-01.txt, draft-mannie-ccamp-gmpls-concatenation-conversion-00.txt.
[28] K. Kompella, et al., OSPF extensions in support of generalized MPLS. draft-ietf-ccamp-ospf-gmpls-extensions-00.txt.
[29] K. Kompella, Routing extensions in support of generalized MPLS. draft-ietf-ccamp-gmpls-routing-00.txt.
[30] A. Banerjeeet, et al., Generalized multiprotocol label switching: an overview of routing and management enhancements, IEEE Commun. Mag. 39 (March 2001) 144–150.
[31] P. Ashwood-Smith, et al., Generalized MPLS signaling: CR-LDP extensions, IETF Network Working Group Draft Report, May 2001. http://www.ietf.org/internet-drafts/draft-ietf-mpls-generalized-cr-ldp-03.txt.
[32] P. Ashwood-Smith, et al., Generalized MPLS signaling: RSVP-TE extensions, IETF Network Working Group Draft Report, May 2001. http://www.ietf.org/internet-drafts/draft-ietf-mpls-generalized-rsvp-te-03.txt.
[33] J. Lang, et al., Link management protocol (LMP). draft-ietf-ccamp-lmp-00.txt.
[34] Y. Xu, A BGP/GMPLS solution for inter-domain optical networking. draft-xu-bgp-gmpls 01.txt.
[35] M. Blanchet, F. Parent, B.St. Arnaud, Optical BGP (OBGP): InterAS lightpath provisioning, IETF Network Working Group Draft Report, March 2001. http://search.ietf.org/internet-drafts/draft-parent-obgp-01.txt.
[36] J. Vollbrecht, et al., AAA Authorization Framework, RFC2904, August 2000.
[37] Investigators on this project include Tom DeFanti, Maxine Brown, Jason Leigh, Oliver Yu, and Mitchell Theys, Electronic Visualization Laboratory, University of Illinois at Chicago (http://www.evl.uic.edu), National Science Foundation (http://www.nsf.gov, #ANI-0123399 et al.).



Joe Mambretti is the Director of the International Center for Advanced Internet Research (iCAIR, http://www.icair.org) at Northwestern University, the Director of the Metropolitan Research and Education Network (MREN, http://www.mren.org), member of the StarLight/STAR TAP partnership (http://www.startap.net), a member of the I-WIRE consortium, and a principal researcher on the OMNInet project. The mission of iCAIR is to accelerate leading-edge innovation and enhanced digital global communications through advanced Internet technologies, in partnership with the international community. iCAIR accomplishes that mission by undertaking large-scale (e.g., global, national, region, metro) projects in several key areas, high-performance resource intensive applications, advanced middleware and metasystems, and large-scale optical network infrastructure. Virtually, all sectors of the national economy require new types of network-based applications supported by an advanced information infrastructure—an internetworking fabric capable of providing high-performance, reliable, high capacity communication services that can be rapidly scaled and readily managed. iCAIR researchers are computer scientists and engineers.

Jeremy Weinberger is a Research Associate at iCAIR. He has focused on the design and development of new types of systems and control plane methods for dynamic provisioning of resources on optical networks, such as lightpaths provisioning on wavelength switched networks. He is a Manager of Operations for OMNInet, principal researcher, architect, and developer for ODIN, THOR, and DEITI and related protocols, and investigator of new methods for intelligent application signaling for optical networks. He has also been involved in experimentation and development of IETF DiffServ QoS methods, and developed experiments for a regional Science Grid. He also developed several successful international DiffServ experiments with CERN and various Asia Pacific research centers.

Jim Chen is an Associate Director of iCAIR, leads multiple projects focused on designing and developing advanced systems capable of providing high-performance, reliable, high capacity services that can be rapidly scaled and readily managed. He is a Project Manager for the IVIMS, Project Manager for high-performance computational clusters, and a specialist in large-scale advanced applications designed for next generation infrastructure. He has special expertise in high-performance large-scale digital media. For iGRID98 at SC98 in Orlando, Jim developed an innovative high-performance digital media demonstration. For iGRID2000 in Yokohama, Jim was a principal partner in the Global Internet Digital Video Network (GiDVN), with a high-performance digital media demonstration that involved almost 40 researchers from 10 countries.

Elizabeth Bacon is a Research Associate at iCAIR, Project Manager for the Digital Video Portal initiative, including the development of the C-SPAN multicast channel, and specialist in management methods for high-performance digital media intended for distribution over national and international networks. Elizabeth has designed and developed several new methods for managing and interlinking digital media types: capturing, storing, automatic indexing and metadata management, cataloging, interlinking, scheduling, searching, retrieving, streaming, reserving, retiring, etc. The Digital Video Portal integrates a number of key technologies including an access mechanism with multiple channels for distributed media assets (regional, national or international).

Fei Yeh is a Research Associate at iCAIR, Project Manager for Grid Computing, Project Manager for Distributed Optical Systems, and Assistant Manager of Operations for OMNInet. Fei Yeh is currently designing new methods for an experimentally distributed optical testbed that will be used to develop new techniques for integrating high-performance Grid computing and next generation optical networks, such as through new methods for dynamic provisioning of lightpaths across large-scale optical networks.

David Lillethun is a Research Associate at iCAIR and a principal architect and software engineer and protocol developer for ODIN, THOR, and DEITI, and investigator of methods for intelligent application signaling. David designed and developed a number of successful experiments related to innovative intelligent application signaling techniques on the OMNInet testbed. David also has an interest in developing new techniques for medical imaging. Previously, he was a designer and developer for a major software corporation.

Bob Grossman is the Director of the Laboratory for Advanced Computing and the National Center for Data Mining at the University of Illinois at Chicago, where he has been a Faculty Member since 1988. He is also the spokesperson for the Data Mining Group (DMG), an industry consortium responsible for the Predictive Model Markup Language (PMML), an XML language for data mining and predictive modeling. He is the President of the Two Cultures Group, which provides consulting and outsourced services focused on data. He has published over 75 papers in refereed journals and proceedings on internet computing, data mining, high performance networking, business intelligence, and related areas, and lectured extensively at conferences and workshops.

Yunhong Gu is a PhD candidate in Computer Science at UIC and a Research Assistant at the Laboratory for Advanced Computing. His research interests are data mining and high-speed networks.

Marco Mazzucco is a Postdoctoral Fellow at the University of Wales, Swansea. He is currently working on an ESPRERC funded project in theoretic computer science under the direction of Dr. Martin Otto. He also does research and consulting for the National Center for Data Mining. He received his PhD in Mathematics at UIC in 2000.