OPTICAL EDGE / ACCESS NETWORKS FOR GENI

Lokesh Mandvekar
lsm5@cse.buffalo.edu

Dr. Chunming Qiao
qiao@computer.org

Computer Science & Engineering
University at Buffalo

Dec. 23, 2008

Executive Summary

The project aims to study the feasibility of building an experimental facility for optical edge/access networks as a part of GENI. The facility considered is a campus/city wide Passive Optical Network (***PON***). This report describes the PON technology, an example setup including hardware and software, and ways to make the facility is programmable and virtualizable as per need of GENI and possible ways to integrate with GENI control frameworks. It also lists the major activities. This report serves as the base for a proposal to build such a campus/city wide optical access network for GENI.

1. Overview

The report is organized as follows. Sec. 2 gives a brief overview of PON technology. Sec. 3 describes the existing facilities (mainly the fiber layout) at the University at Buffalo (UB) – The State University of New York. Sec. 4 lists the potential sites for the system's nodes. Sec. 5 discusses the PON topologies that can be used for the testbed including the advantages of using a modified ring topology. Sec. 6 explains the major hardware and software blocks that would constitute a typical node. Sec. 7 briefly describes how a GENI user can access the system and the features available to him/her. Sec. 8 describes how a user can create and run his/her experiments. Sec. 9 enlists the various system parameters available to the GENI user, some of which are tunable. Sec. 10 states the integration plan with the PlanetLab control framework. Sec. 11 contains the references while Sec. 12 summarizes the general scope and outcome of this project.

2. Overview of PON

A PON, basically, consists of one Optical Line Terminal (OLT) (present at the Central Office (CO)) providing shared access to multiple Optical Network Terminals / Units (ONTs/ONUs) at the customer premises (access nodes). The various PON technologies available today are:

- ***Ethernet-PON (EPON)***: The EPON [1] standard was developed by *IEEE* (*802.3ah* technical committee). The underlying protocol is the Multipoint Control Protocol (***MPCP***) and data transmission is done using Ethernet frames. Currently, it supports data-rates of 1.25Gbps (symmetric). The EPON standard currently under development, called *10GigEPON* (expected in 2009), can support data rates of up to 10Gbps (symmetric).
- ***Gigabit-capable PON (GPON)***: The GPON standard was developed by the *ITU-T (G.984)* sponsored by FSAN. GPON allows multi-protocol transmission for e.g. ATM traffic is carried using ATM frames, TDMA / Ethernet traffic is encapsulated using *GEM* frames. Currently it supports data-rates of up to 2.5Gbps/1.25Gbps (asymmetric). GPON

is very well suited to supporting optical wireless integration (OWI) since it is time division multiplexed (TDM) based. The next step in the GPON will be increased bit-rate, 10Gbps/2.5Gbps (asymmetric) (*10G GPON*).

- ***Wavelength Division Multiplexing-PON (WDM-PON)***: The WDM-PON protocol essentially involves allocating dedicated wavelength(s) to each ONU (logical Point-to-Point). Thus, each ONU can communicate with the OLT without requiring synchronization/scheduling mechanisms as in the above two standards. Implementation of virtual-PONs (VPON) is also possible using WDM-PON [3], providing physical separation/isolation of each user signal.
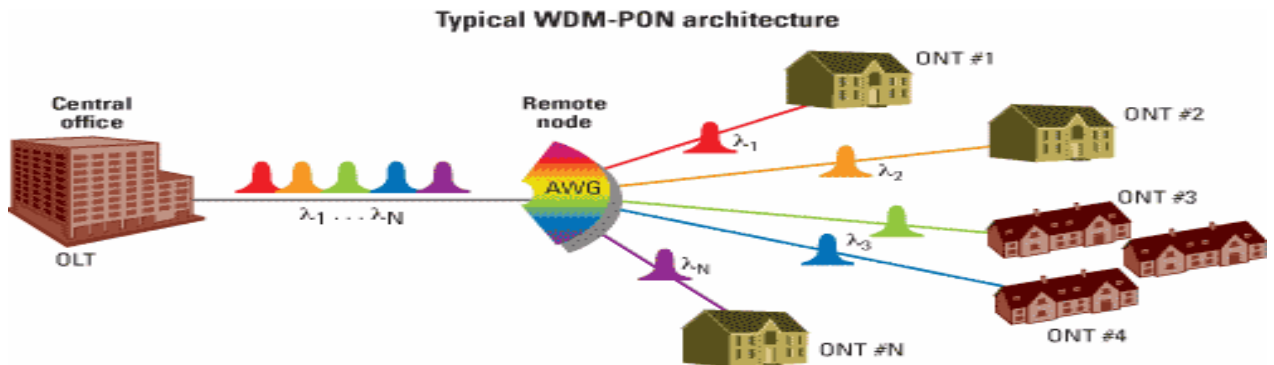

Figure 1: A Typical WDM-PON Structure

3. Existing Infrastructure

The University at Buffalo (UB) is spread over three campuses: South Campus (medicine and heath related departments), North Campus (school of engineering, natural sciences and law departments) and Downtown Campus (comprises of research centers). Fig. 2 shows a bird's eye view of the fiber map of Buffalo city and also the amount of fiber owned by UB (marked in black). Most of the fibers owned by UB are single-mode fibers (labeled with a single number which is the number of dark fibers available along that particular link).

Multi-mode fibers are also laid along some links (labeled as 'sm/mm' where sm and mm denote number of single and multi-mode fibers). As can be inferred from the map, there are plenty of dark fibers available for use (a minimum of 96sm between campuses). Intra-campus fiber layout and fiber types are shown in Figs. 3-4. A campus wide PON can be built using single-mode fibers.
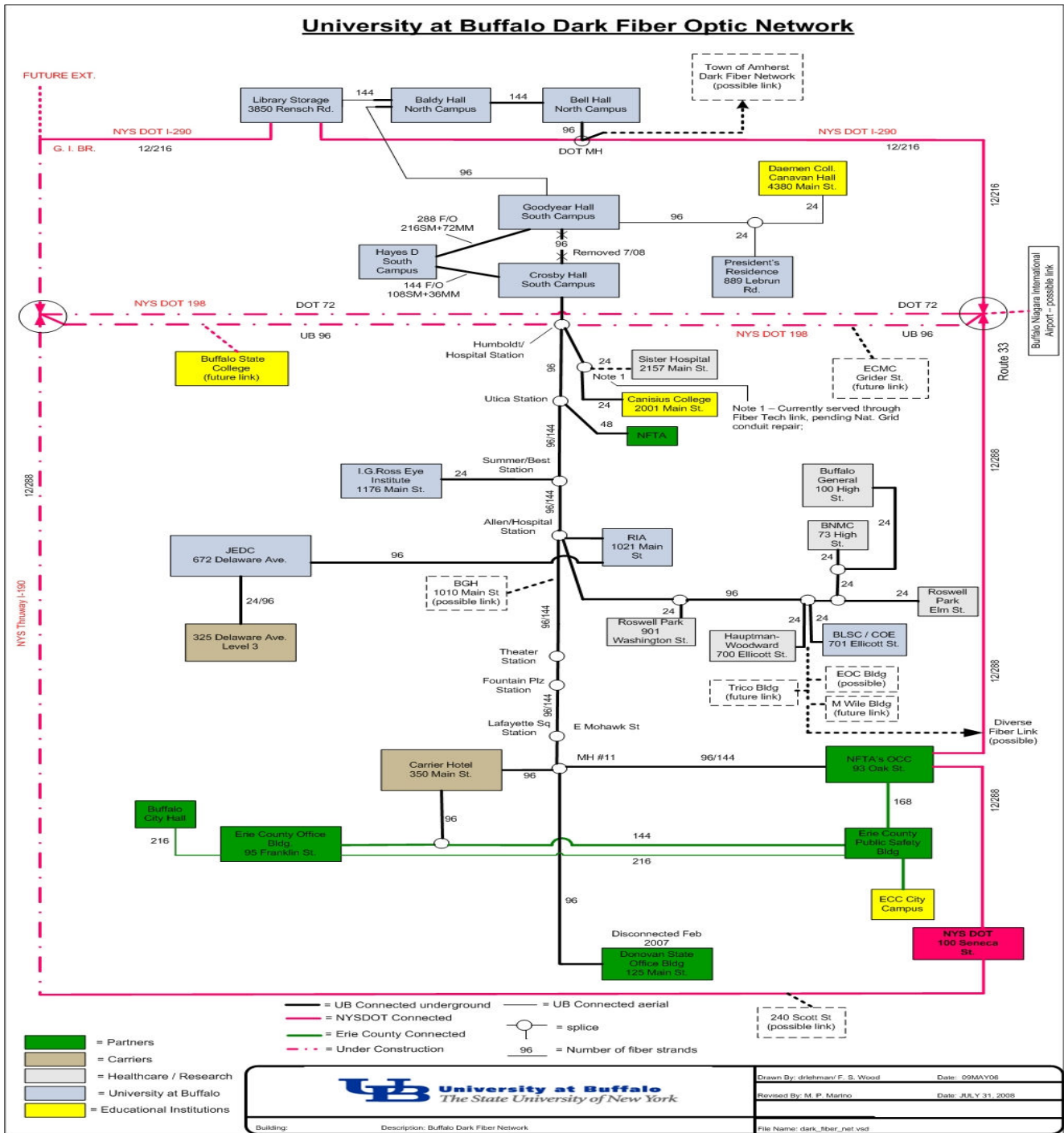
University at Buffalo Dark Fiber Optic Network

FUTURE EXT.

Town of Amherst Dark Fiber Network (possible link)

Library Storage 3850 Rensch Rd. —144— Baldy Hall North Campus —144— Bell Hall North Campus

NYS DOT I-290   96   NYS DOT I-290
G. I. BR.   12/216   DOT MH   12/216

12/216

Daemen Coll. Canavan Hall 4380 Main St.   24

96

Goodyear Hall South Campus —96—

288 F/O 216SM+72MM

96 Removed 7/08

24

President's Residence 889 Lebrun Rd.

Hayes D South Campus

144 F/O 108SM+36MM

Crosby Hall South Campus

NYS DOT 198   DOT 72   DOT 72   Route 33
UB 96   NYS DOT 198   UB 96

Buffalo Niagara International Airport – possible link

Buffalo State College (future link)

Humboldt/ Hospital Station   96   Sister Hospital 2157 Main St.   Note 1   ECMC Grider St. (future link)

24

24   Canisius College 2001 Main St.

Note 1 – Currently served through Fiber Tech link, pending Nat. Grid conduit repair;

Utica Station   96/144   48   NFTA

Summer/Best Station   24   I.G.Ross Eye Institute 1176 Main St.   Buffalo General 100 High St.

96/144

Allen/Hospital Station   RIA 1021 Main St   BNMC 73 High St.   24

JEDC 672 Delaware Ave.   96   24   24

BGH 1010 Main St (possible link)   96   Roswell Park Elm St.

24/96   96/144   Roswell Park 901 Washington St.   24   24   24   24

325 Delaware Ave. Level 3   Hauptman-Woodward 700 Ellicott St.   BLSC / COE 701 Ellicott St.

EOC Bldg (possible)

Theater Station   96/144   Trico Bldg (future link)   M Wile Bldg (future link)

Fountain Plz Station   Diverse Fiber Link (possible)

Lafayette Sq Station   96/144   E Mohawk St

MH #11   96/144   NFTA's OCC 93 Oak St.

Carrier Hotel 350 Main St.   96

96

Buffalo City Hall   168

216   Erie County Office Bldg. 95 Franklin St.   144   Erie County Public Safety Bldg   12/288

96   216   ECC City Campus

96   NYS DOT 100 Seneca St.

NYS Thruway I-190   12/288

Disconnected Feb 2007

Donovan State Office Bldg 125 Main St.

240 Scott St (possible link)

= UB Connected underground   = UB Connected aerial
= NYSDOT Connected   = splice
= Erie County Connected
= Under Construction   96   = Number of fiber strands

= Partners
= Carriers
= Healthcare / Research
= University at Buffalo
= Educational Institutions

University at Buffalo
The State University of New York

Drawn By: drlehman/ F. S. Wood   Date: 09MAY06
Revised By: M. P. Marino   Date: JULY 31, 2008
Building: Description: Buffalo Dark Fiber Network   File Name: dark_fiber_net.vsd

Figure 2: Buffalo Fiber Map

North Campus

The North Campus occupies approx. 1192 acres of land and is home to most of the non-medical departments of UB. Fig. 3 shows the fiber map of North Campus.
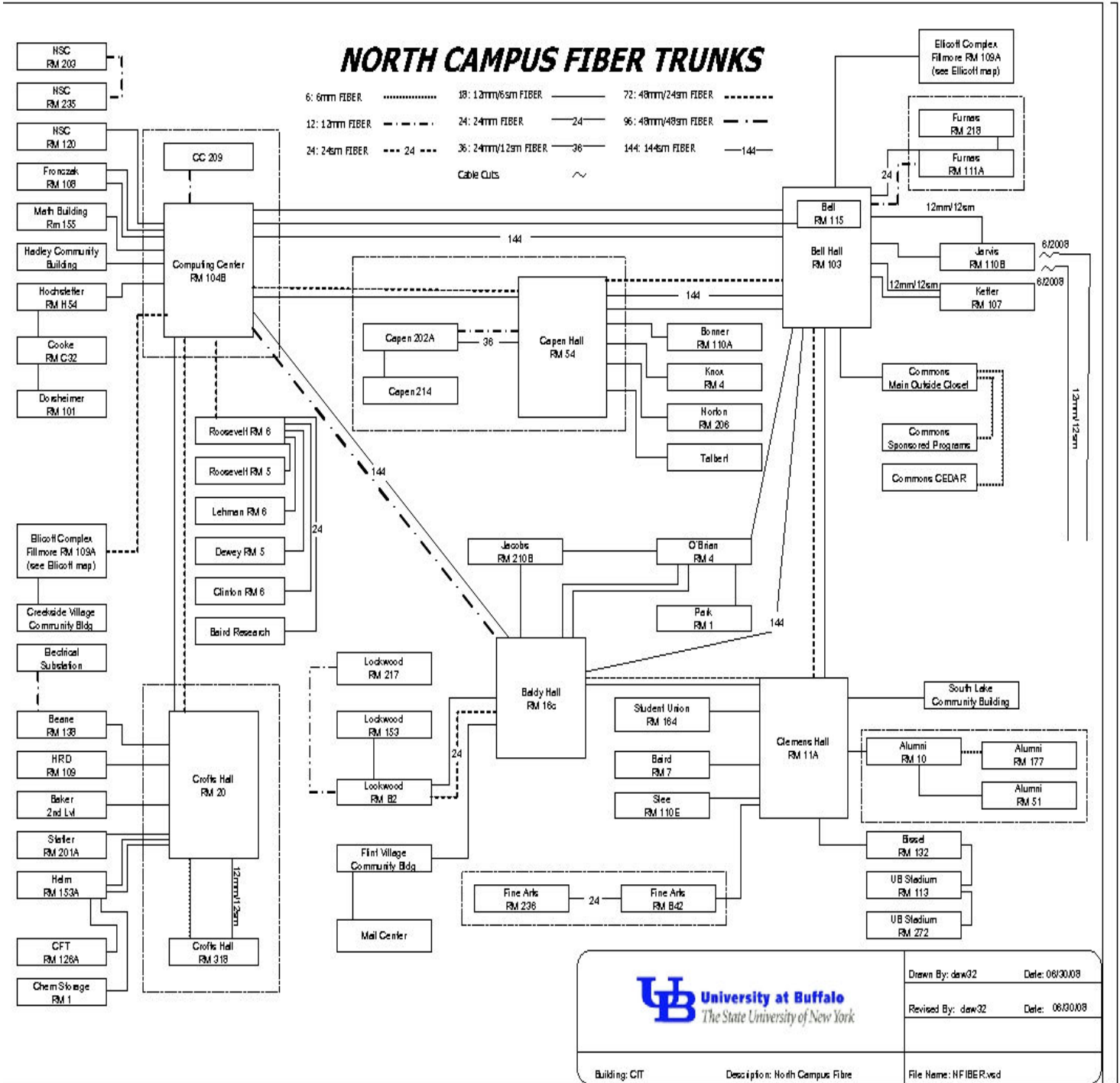

Figure 3: North Campus Fiber Map

South Campus

The South Campus (Fig. 4), located approximately 5.9 km (3.6 miles), is spread over an area of approximately 154 acres and houses the medicine, health and nursing departments.
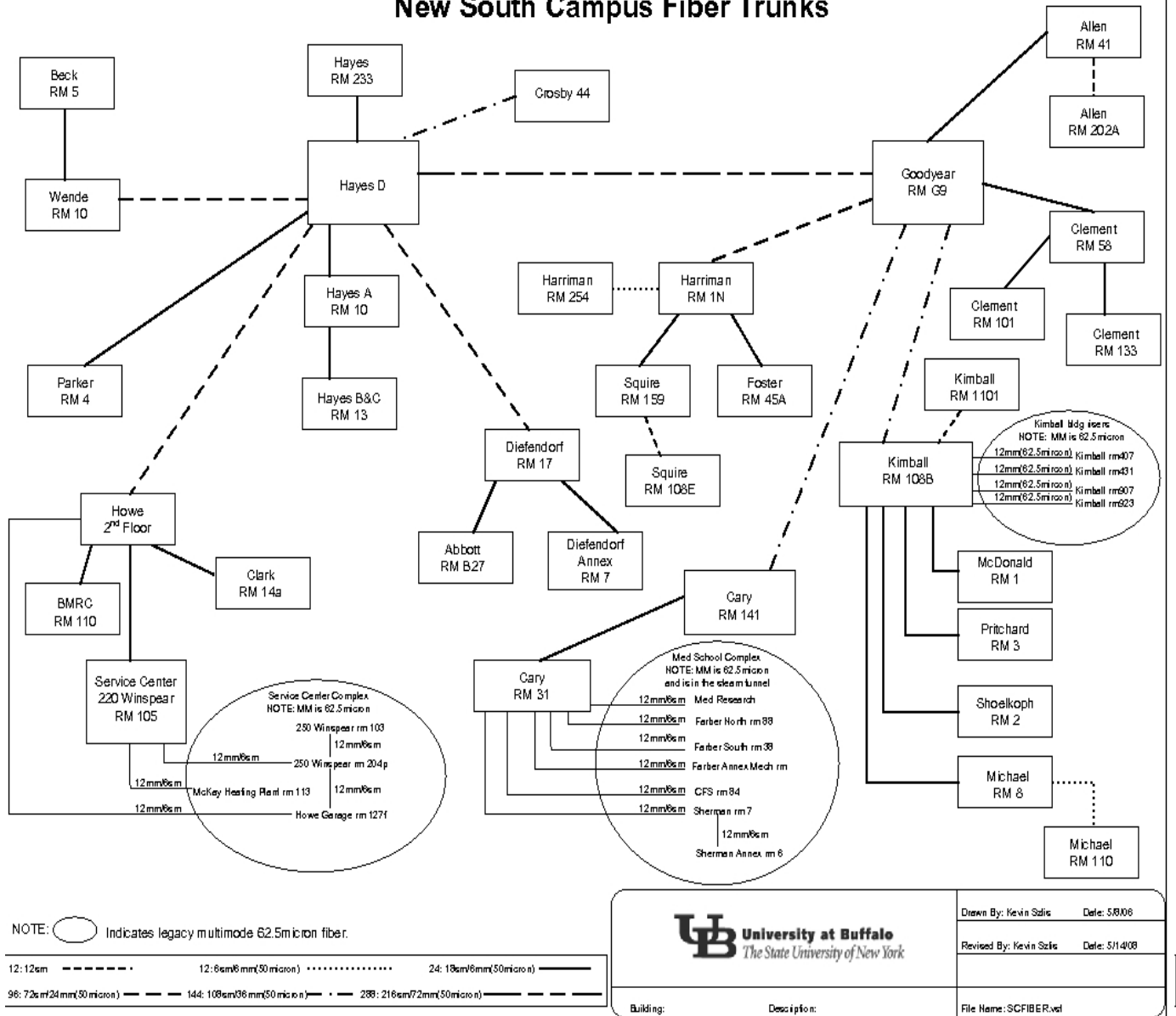
New South Campus Fiber Trunks

Figure 4: South Campus Fiber Map

Downtown Campus

The Downtown Campus (see Fig. 2) is located approximately 18 km from the North Campus, and consists of research centers like the Center for Excellence which works in many areas of research involving science, health and the general community.

4. An Example System Map
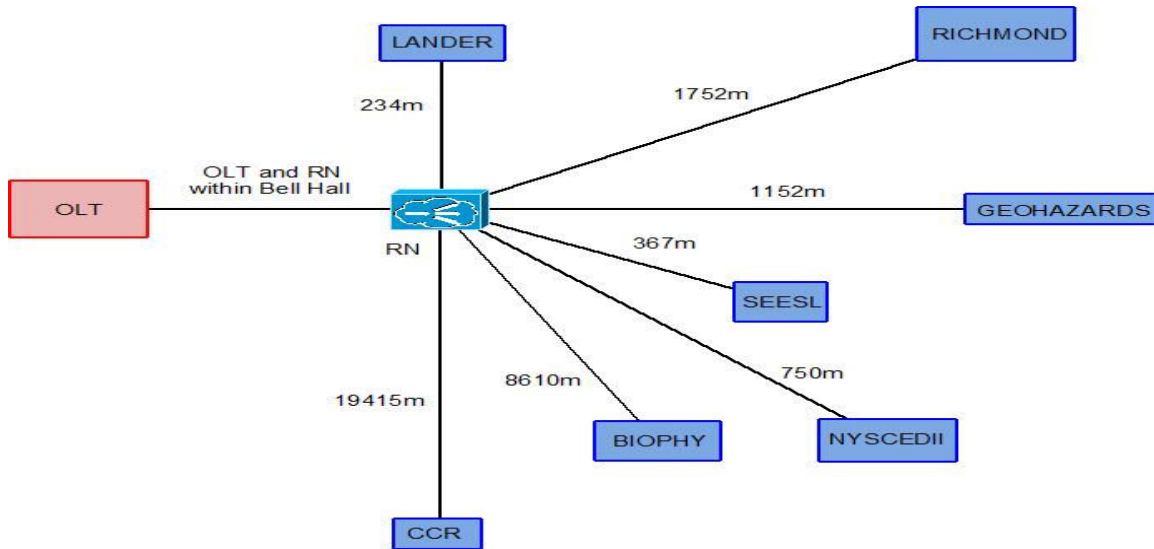
Figure 5: Logical Topology

As an example set up, the OLT can be located at Bell Hall in the North Campus. Also, five of the seven ONUs can be located in the North Campus. The system can also have one ONU in the South Campus and one in the Downtown Campus. The ONU in Downtown Campus is located approximately 19.4 km (12.1 miles) from our OLT at North Campus and hence, is within the OLT's reach.

The ONU locations have been identified on the basis of their need for heavy computational resources and high bandwidth connectivity. The ONUs in North Campus are at the following sites:

- *Laboratory for Advanced Network Design, Evaluation and Research* (*LANDER*): LANDER, located in the UB Commons, conducts research on advanced networking architectures, protocols, network control and management issues and performance evaluation. For more information visit http://www.cse.buffalo.edu/~qiao/index_files/Lander.html .
- *Structural Engineering and Earthquake Simulation Laboratory* (*SEESL*): SEESL, located in Ketter Hall, provides research services for Network for Earthquake Engineering Simulation (NEES) and Multidisciplinary Center for Earthquake Engineering Research (MCEER). For more information, visit http://seesl.buffalo.edu/ .
- *New York State Center for Engineering Design and Industrial Innovation* (*NYSCEDII*): NYSCEDII, located in Norton Hall, North Campus, focuses on the front-end high technology and information intensive design of complex products and the planning of manufacturing facilities, develops technologies and provides services that complement other economic development and training activities in the Region and State. For more information, visit http://www.nyscedii.buffalo.edu/ .
- *Center for GeoHazards Studies*: The Center for GeoHazards Studies, located in Natural Sciences Complex, North Campus, seeks to decrease harmful societal effects of natural

phenomena such as volcanic eruptions, landslides, mudflows and avalanches through research, service and education. For more information, visit http://www.geohazards.buffalo.edu/ .

- **Richmond Dormitory**: Richmond Dormitory (along with adjacent dormitories), located in the Ellicott Complex (adjacent to North Campus), is home to approximately 3200 UB students. Obviously, this site would need to cater to the demands of a huge number of students and staff, hence requiring high bandwidth connectivity.
- **Center for Single Molecule Biophysics**: The Center for Single Molecule Biophysics(located in Cary Hall, South Campus) is associated with the School of Medicine and Biomedical Sciences focuses on cell mechanics and channel and cell biophysics. For more information, visit http://www.sachslab.buffalo.edu/index.htm .
- **Center for Computational Research** (**CCR**): CCR, located at the Downtown Campus maintains high-performance, high-end visualization laboratories, and support staff with expertise in scientific, parallel and grid computing, software engineering, visualization, database design and networking. For more information, visit http://www.ccr.buffalo.edu .

5. Possible PON topologies

A PON consists of one Optical Line Terminal (OLT) (present at the edge node) providing shared access to multiple Optical Network Terminals / Units (ONTs / ONUs) (present at the access nodes). Such a network can be implemented in various topologies as shown.

- **Tree structure**: A PON is conventionally built in a tree structure. A trunk fiber runs from the OLT to a remote node, where its signal is split using an optical splitter and sent to ONUs. In the opposite direction, the splitter also acts as a combiner i.e. it multiplexes the signals from the ONUs headed toward the OLT.

A tree structure, however, has drawbacks. Since the ONUs cannot directly communicate with each other, it does not have the capability to support Local Area Network (LAN) functionality. Also, since only a single path exists from the OLT to any ONU, there is no provision to implement survivability/redundancy in the network.

- **Ring structure**: In a ring structure, all the ONUs are attached to a ring network. The trunk fiber from the OLT connects to the ring network using devices like circulator/isolator as in e.g. [2] in order to enable selective directional flow of signal and to block transmission in other directions. A ring structure can support LAN functionality and also provide survivability/redundancy mechanisms in the network.

Note that, given the current fiber layout at UB (Figs. 2-4), it is not possible to implement the ring topology. Nevertheless, we can use the physical topology(Fig. 6) either as a tree architecture or as a ring architecture depending on the needs of the user, i.e., the facility is quite flexible in supporting two major optical access topologies. For example, if a user requires a tree architecture, (s)he uses only one of the two active links between the concerned nodes. A *redundant* tree architecture could use one active link and either of the two protection links between the required nodes. The protection links run parallel to the active links shown in Fig. 6.

If the user requires a ring topology, the ring can easily be realized by appropriately configuring the optical switches (or coupler). A ring architecture requires two active links, one for incoming traffic and one for outgoing traffic, while the two protection links provide redundancy. Redundancy mechanisms can be implemented as in e.g. [2].

The topology shown in Fig. 6 uses a maximum of four fibers on those links. Since, a minimum of twelve single-mode dark fibers are available on any of the concerned links, utilizing four fibers on each of the links is feasible and leaves a sufficient number of fibers free for other use.
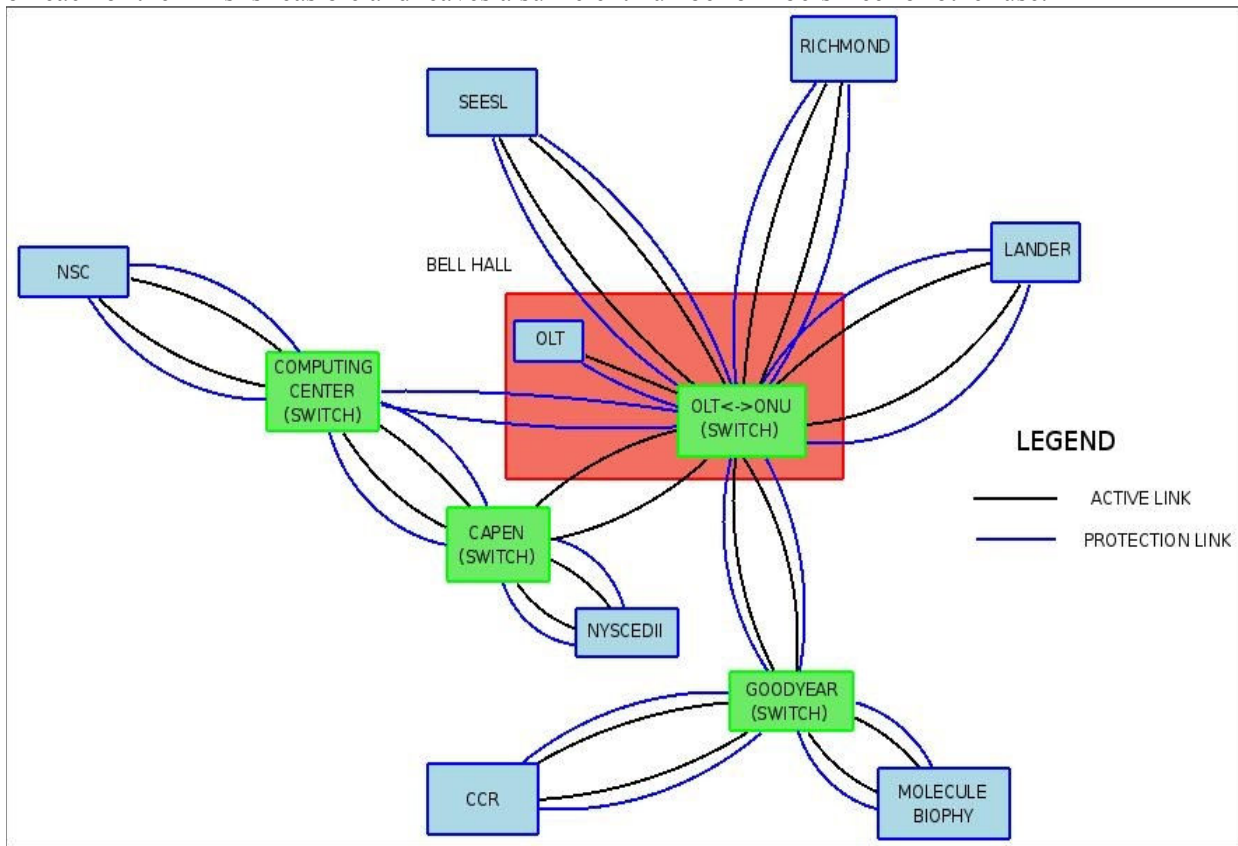

Figure 6: Physical Topology

6. Platform description

The PON testbed implementation provides end-to-end slice virtualization (each user is completely isolated and thus secured from other users in terms of node as well as network resources). Node slice isolation is provided by using the Xen virtualization system which gives near-native performance for guest OSes and also ensures isolation between them. Network resource isolation is provided using the concept of Virtual PON (**VPON**) [3] implemented within a WDM PON which dedicates unique wavelengths for different PON protocols. Thus, different PON protocols can be implemented simultaneously with guaranteed isolation from other PON protocols.

The GENI software is mainly a set of APIs which abstracts the lower layers of the network and makes it suitable for users to create and run their experiments without worrying about the lower layer structure/topology as well as provides other required features related to running experiments. This software could be packaged and made readily available to the GENI users in the form of operating system images, which are readily bootable on their node slices. Standalone GENI software source code could also be made available for users to download.

The Xen virtualization system runs on the nodes (OLT as well as ONUs) as shown in Fig. 7. The testbed administrators control the Xen hypervisor (Xen runs directly on top of system hardware) and the Dom0 OS (the privileged guest OS) for security, administration and arbitration purposes (in co-ordination with the Aggregate Manager), while the DomU OSes (unprivileged guests) are under the complete control of the concerned GENI users. NetBSD is a good candidate for host operating system (and hence Dom0 OS) for two major reasons: mature support for Xen, and high portability across various processor architectures (a system could be ported/migrated from one architecture to another with minimum possible effort as compared to other operating systems). Thus, scalability and portability can become a lot easier than otherwise.

First, the control and management protocols for GENI user specific applications are setup on-demand by the virtualization systems with Dom0 OS and DomU OSes. Then the Dom0 OS downloads the management/control protocols to the traffic management module. The dynamic bandwidth allocation and other OAM (operation, management & maintenance) functions can be implemented in an FPGA. Meanwhile, metro aggregation traffic is sent through a XAUI or GE interface on the backplane at the OLT while at the ONU, user traffic is sent through a number of Ethernet interfaces. Localized traffic inside the access network is sent to the MAC layer buffer with appropriate ONU id.

PON MAC functionalities, including framing and optical line supervision functions can also be implemented with FPGA. This PON MAC module supports GPON, EPON, WDM PON and other programmable framing formats with variable preamble lengths (for burst mode upstream, minimum preamble length is 44 bits due to timing limits in clock and data recovery circuitry; for WDM PON and other continuous transmission scheme, no preamble is needed). At the same time, power monitoring, signal detection, and other physical layer supervision functions are implemented in the PON MAC module; various alarm signals are generated and sent to Traffic Management module and CPU for further processing, and various protection signals such as APS (automatic protection switching) are activated through FPGA.

In addition to FPGA implementation, APIs for software implementation of the above functions(for custom PON protocols, especially) are also provided (packaged within the operating system image, and as standalone). If customer defined PON protocols are too complicated to be implemented in FPGA, then the firmware/hardware version of PON MAC in FPGA are bypassed and data streams are sent directly to the DomU OS, where a software version of PON MAC is implemented. In such cases, PON MAC protocols handled by software implementation are slower, and it might not be able to provide real-time processing. Such a platform becomes more like a software controlled simulator, but all the user defined protocols are fully functional. Hence, GENI experiments can be run on top of this simulator, and network performance and other metrics can still be evaluated with this platform.

For robustness, the OLT/ONU can be designed with dual sets of hot-swappable line cards, as shown in Fig. 8. The OLT rack includes two power supplies, two metro upstream line cards with 10G Ethernet interface, two CPUs for control/management, and two sets of OLT line cards. The two power supplies and two metro upstream line cards are identical – one active and one for backup. The OLT line cards can be of any PON protocol, be it GPON, EPON or WDM PON; and the slots on the rack can support any type of PON card. Furthermore, each transceiver on the line cards is pluggable; and any PON transceiver can be plugged into an OLT/ONU line card. The ONU rack configuration is similar, but there are no metro upstream line cards and the number of line cards is smaller but expandable, depending on the actual user bandwidth needs. APS

function can be achieved on electrical domain by hot-swapping two identical line cards or 2×2 optical switches with 50ms switching time. A typical OLT/ONU rack is about 4U in height, 19-in in width, weighs under 50 lb and consumes about 100 watts (with a typical number of line cards).
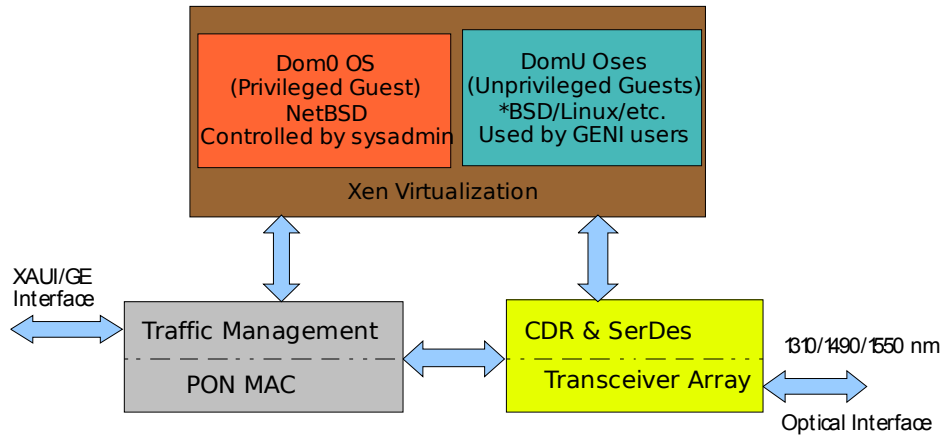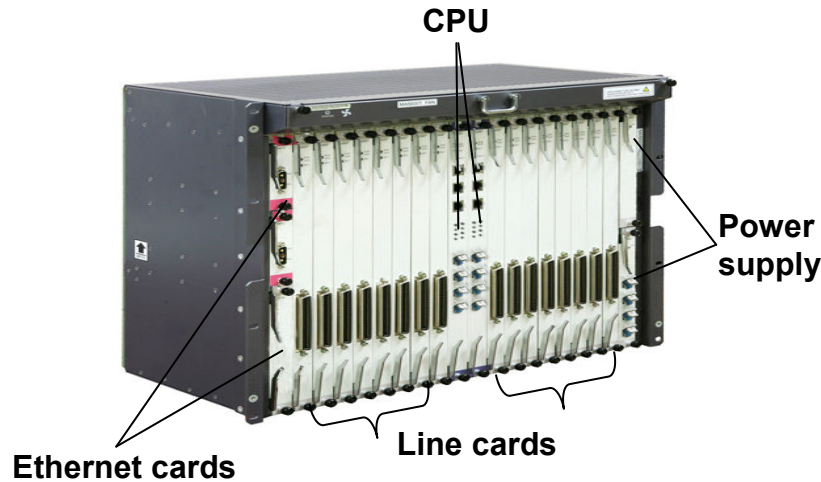

Figure 7: Structure of OLT / ONU

Figure 8. OLT / ONU rack configuration

7. GENI User Access

A GENI user is provided with a complete and fully featured operating system running on his/her nodes, (Fig. 9) which can be accessed via a secure protocol (for instance SSH). The user interface could either be a UNIX Terminal and/or a GUI depending on the user's demand and/or purpose. This particular operating system is actually the DomU OS running on the node (Fig. 7). The GENI user has complete control over all features of this OS and can modify them to suit his/her needs. The home directory of the user is NFS mounted to enable data transfer back and forth from any node to any other node used by him/her. The GENI user's home folder is physically located on a different machine(s) dedicated for data storage. In addition, RAID technologies are suitably implemented to provide data redundancy. The GENI user's OS contains GENI-specific software for various functionality like instrumentation and data collection, slice resource requests (for instance, increase or decrease of slice size), as well as functionality to modify the MAC layer
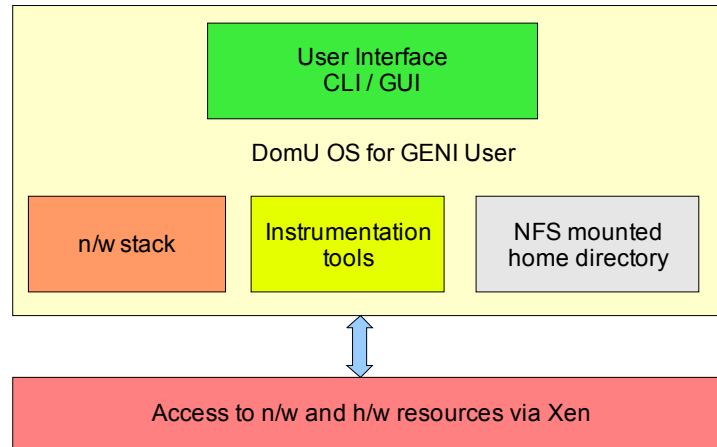
implemented in FPGA.

Figure 9: GENI User Operating System Architecture (Guest OS on the nodes)

8. Operation Principle

This proposal extends virtualization (viz. programmability and resource sliceability) to optical access networks. The system architecture and middleware enables deployment of multiple coexisting PON systems isolated from one another on the virtualized optical access network (WDM-based). The general system architecture is shown in Fig. 10. The PON testbed is largely targeted at two types of users: **(a)** users who require the testbed resources to run their experiments without having to modify the supported PON protocols (such users may include the opt-in users), and **(b)** users who need to modify the existing PON protocols for their experiments. It provides, basically, three types of services: 1) programmable optical access; 2) sliceable optical access; 3) PON template for future PON (with open source optical access hardware/software).

The general working of the system is described as follows:
Login/Access: To begin using the system, a user must have an account created on the system. The nature of the account may vary depending on his/her credentials and may also depend on the experiment i.e. user could either be PI, research student, guest user or opt-in end-user of a particular experiment.

Experiment Creation: First, a user acquires a globally unique slice ID from the Clearinghouse upon request. The user then contacts the aggregate manager for available resources and the aggregate manager, in turn, runs suitable resource allocation algorithms to allocate resources (either immediately or reserves resources for later use). The aggregate manager allocates access to the OLT depending on the nature of the experiment. If the experiment requires modification of the underlying system/network protocols, the user is allocated a slice on the OLT along with required slices on the ONUs, else, the user only has access to his/her concerned ONUs while no access is given to the OLT. This is a required step in order to keep the OLT as lightly loaded as possible since the OLT is critical for a PON.

Experiment Execution/Storage of Results: Once the user can access his/her slice (nodes as well as network resources), (s)he can modify his/her slice as per need/liking. Since the underlying PON protocol is WDM-based, (s)he is allocated a fixed (dedicated) set of wavelengths

guaranteeing complete isolation in the network. Also, on the nodes, (s)he runs a virtualized operating system, thus isolating his/her system from those of others. Thus, (s)he is guaranteed complete isolation and security from other users. Appropriate mechanisms for data storage and redundancy are implemented as mentioned earlier to ensure protection of data and results.
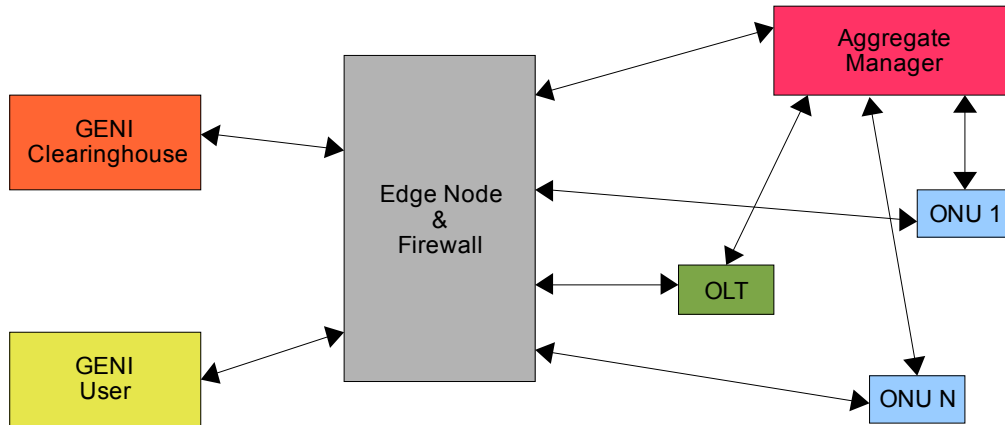

Figure 10: System Architecture – Major Blocks

9. System Parameters

The following table lists various parameters of the testbed, some of which are programmable/tunable.

PHY Layer	Wavelength	GPON/EPON: 1480-1500nm (down), 1260-1360nm (upstream) WDM PON: C or L band, Video overlay: 1550 nm
	Distance	0 – 20 km
	Overhead	Guard band / preamble / delimiter are all programmable
	Transmitted power	0 – 5 dBm
	Extinction ratio	$\geq$ 6 dB
	Receiver sensitivity	≤ -24dBm
	Jitter tolerance	0.25 UI
MAC Layer	Service	Supporting full services (Ether, TDM, arbitrary data format)
	Frame	GEM frame / Ethernet frame / programmable format.
	Resource management	Programmable (e.g. MPCP)
	Openflow	Supporting openflow
	Bandwidth	1G-10G, adjustable.

10. Integration with other control frameworks

The PON testbed can be integrated with PlanetLab's GENI Clearinghouse implementation. The integration process mainly involves incorporating PlanetLab's Geniwrapper functionality on our testbed nodes. Integration with Emulab's GENI implementation as well as integration with other control frameworks are also possible. Additional steps will become clear and defined once the base testbed system has been realized.

11. References

[1] Sudhir Dixit, **"IP OVER WDM"**, ISBN 0-471-21248-2

[2] ASM Delowar Hosain, Antonis Hadjiantonis, Georgios Ellinas, **"A Simple Self-Healing Ring-Based Local Access PON Architecture For Supporting Private Networking Capability"**

[3] Wei Wei, Ting Wang, Chunming Qiao, **"Resource Provisioning for Orthogonal Frequency Division Multiple Access (OFDMA)-based Virtual Passive Optical Networks (VPON)"**, in OFC 2008