G E N I

Global Environment for Network Innovations

GENI Quarterly Status Report

Document ID: GENI-QR-ERM-Dec08

Dec 30, 2008

Prepared by:
F. Fidler, C.P. Lai, and K. Bergman
Dept. of Electrical Engineering, Columbia University New York,
500 W. 120th Street ,
New York City, NY 10027

under Project Nr. 1631
“Embedding real-time measurements for cross-layer communications”

Document Revision History

The following table provides the revision history for this document, summarizing the date at which it was revised, who revised it, and a brief summary of the changes. This list is maintained in chronological order so the earliest version comes first in the list.

Revision	Date	Revised By	Summary of Changes
1.0	30 Dec 08	K. Bergman	Initial draft
1.1	01 Feb 09	F. Fidler	Equipment update

Embedding real-time substrate measurements for cross-layer communications
GENI Quarterly Status Report
Project Nr.: 1631

P.I. Keren Bergman
Department of Electrical Engineering, Columbia University, New York

1. Major accomplishments
During this past quarter, our main accomplishments involved progress on all project milestones. As stated in the following document: in order to effectively implement a real-time measurement based cross-layer communications infrastructure, we must first assess the measurement capabilities offered by several control framework groups. This ongoing work, coupled with our simulation based investigations constitute our main accomplishments. We are continuing focused interactions with other GENI teams, mainly Renaissance Computing Institute (RENCI) and Duke University, as well as University of Houston, and making progress on all fronts regarding the design of a complementary and effective cross-layer infrastructure for these control frameworks.

2. Milestones
The following section discusses progress made on the agreed milestones as given in [geni08_1].

Milestone 1: GENI requirements for real-time measurements (03/01/2009)
The purpose of this milestone is to assess and evaluate GENI requirements for real-time user access to measurement data across a diverse set of technologies.

We are currently in the process of assessing the capabilities of GENI's future infrastructure with respect to real-time measurements. Our survey is based on the current version (12/10/2008) of GENI's "Spiral 1 substrate catalog" [geni08_2], on personal communication with the different working groups, as well as on information from equipment vendors.

While information about the main equipment available in each GENI node was derived from [geni08_2], information about the real-time measurement capabilities has to be gained directly from the equipment vendors or to some extend from the GENI working groups which control the nodes. General information about some measurement capabilities could be already derived from publicly available product datasheets, vendor webpages, and personal communication with different working groups (cf. Section 5). We are also working directly with several vendors to gain more specific information about how performance indicators are accessible on individual node equipment like routers or switches.

Milestone 2: Develop specifications and networking protocols (target 03/01/2009)
This milestone involves the development of a set of specifications and networking protocols based on GENI requirements for real-time user accessed cross-layer measurements.
To achieve this milestone, we first require the results of the measurement capability survey, as stated in 2.1. We must first be aware of and understand what metrics can be measured and obtained. We will then focus in on a specific set of control frameworks groups with which we will dynamically interact to design an appropriate cross-layer infrastructure that is complementary to the groups' endeavors and abilities. Once we have sufficiently determined the

measurement metrics that will be realistic within the scope of the GENI work, we will then develop the software and hardware specifications that must be met in order to have adequate access for real time measurements. For instance, a specification that must be achieved is that the minimum measurement interval should be a single (OTN) frame. We will then use these specifications to further develop networking protocols for the control framework(s) that will leverage the measured data to optimize networking functions within the project. The networking protocols' role is to instruct the control plane logic to incorporate the measurement data, use cross-layer signals to send the data upwards, make decisions on a real-time timescale on networking routing, and then reconfigure the network accordingly.

Our progress on this front is ongoing; as we discover the metrics that can be used for cross-layer communications, we are investigating the appropriate networking protocols that can be developed in a software platform. **It may be necessary to push back the delivery date on this milestone based on its dependency on milestone 1.**

Milestone 3: Perform discrete-event network simulations (03/01/2009)

We are performing discrete-event network simulations to quantitatively evaluate performance impact under several scenarios of cross-layer information exchange. Simulations are being developed into the ns-2 [ns08_1] open source network modeling environment.

With the help of the event-driven network simulator it will be possible to:

- identify the key differences of several types of networks when enabled with cross-layer communications,
- evaluate the performance in terms of quantitative measurements such as packet loss or throughput,
- provide a comparison for different cross-layer based routing protocols and network control and management efforts,
- facilitate system architecture planning by reducing system design time.

The notion of a clean-slate network design with a cross-layer optimized protocol stack has been explored to some extent within the wireless networking domain; as such, some modules enabling cross-layer simulations have been previously implemented for wireless networks. We want to implement additional cross-layer enabling modules for the optical physical layer to show that these functionalities will provide significant performance enhancements.

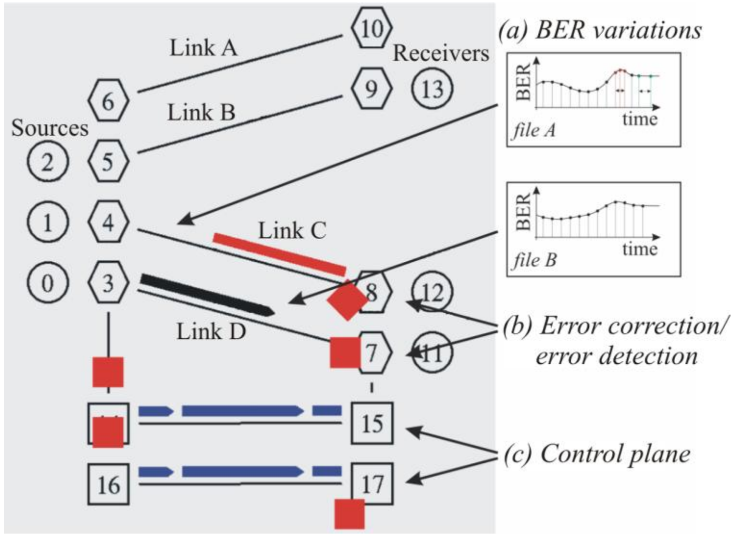

Fig. 1: New ns2 modules: (a) intra-packet BER variation for every link, (b) error correction and error detection, (c) control plane.

The open source software ns2 allows for integrating additional new modules based on the programming language C++. As a first step towards cross-layer simulations, we have implemented the following functionalities (cf. Fig. 1): intra-packet BER variations, a forward error correction (FEC) module, and a local control plane. The modules, which will be made downloadable from the GENI webpage in January 2009, are described in more detail in Section 5.

Milestone 4: Develop a software architecture (06/01/2009)

This milestone deals with the development of a software architecture based on GENI real-time measurement requirements and other developments/resources within the GENI prototyping activities. This will include interacting with the other prototype efforts to identify and leverage relevant activities or products.

The outcome of this milestone will strongly depend on the results from milestones 1-3. Cooperation to achieve the interaction with other working groups' efforts has been started with the Renaissance Computing Insititute (RENCI) and Duke University (Cluster D) [orcaben08_1], and University of Houston [gurkan08_1].

Milestone 5: Support the GPO in developing an experimental use-case (09/01/2009)

Within this milestone, we support the GPO in developing experimental designs for use-cases based on a measurement driven cross-layer communications.

Following successful simulation work, we plan to work closely with the GPO to further validate the cross-layer communication schemes in developing experimental use-case.

Milestone 6: Identify a candidate control framework (09/01/2009)

The aim of this milestone is to work with the GPO to identify a candidate control framework for integration of the SW module.

3. Deliverables made

- Input for GENI Wikipedia page on "Embedding real-time substrate measurements for cross-layer communications" (Sept. 2008)
- Poster and oral presentation at 3rd GENI Engineering Conference, Palo Alto, CA (Oct. 2008)
- GENI quarterly report (Dec. 2008)
- No milestones submitted yet. Next filling date is 03/01/2009 for milestones 1, 2, and 3.

4. Description of work performed during last quarter

- Organizational work

Together with the GPO, i.e. John Jacob, detailed milestones were defined and information about our group's contribution to GENI was put on the GENI wikipedia page. The aim and methodology of our project was then presented at the 3rd GENI engineering conference by means of a poster and an oral presentation. A number of working groups in various clusters was contacted with the aim to establish possible cooperation for the investigation and implementation of real-time measurements in the GENI network. During this process we have established collaborations with the Renaissance Computing Insititute (RENCI), Duke University, and the University of Houston were established (cf. Section 9).

- Survey on equipment

Based on the actual version of the GENI "Spiral 1 substrate catalog" [geni08_2], equipment vendors' product datasheets, vendor webpages, and personal communication with different working groups a survey on available equipment within the GENI network was performed, with special focus on the potential use for embedded real-time measurements (cf. Section 5).

- Simulation

To assess the influence of real-time measurement based cross-layer communication within the GENI network we programmed new software modules for the open-source network simulator ns2. The following functionalities were added: intra-packet BER variations, a forward error correction (FEC) module, and a local control plane. The modules, which will be made downloadable from the GENI webpage in January 2009, are described in more detail in Section 5.

5. Activities and findings

5.1 Survey on equipment

Based on the actual version of the GENI "Spiral 1 substrate catalog" [geni08_2], equipment vendors' product datasheets, vendor webpages, and personal communication with different working groups a survey on available equipment within the GENI network was performed.

	Cluster B			*Cluster C*	*Cluster D*
	DRAGON	*GpENI*	*SSP Overlay*	*Proto-GENI*	*BEN*
HP ProCurve 5400 switch				***x***	
Juniper EX3200 switch					***x***
NetFPGA			***x***	***x***	
Polatis Fiber Switch					***x***
Ciena CN4200 switch		***x***			
Ciena CoreDirector switch		***x***			
Cisco 6509-E switch					***x***
Infinera DTN ROADM				***x***	***x***
Adva Optical Networking switch	***x***				

Table 1: Node equipment in candidate GENI networks as given in [geni08_2]

Table 1 shows an overview of the equipment available in the nodes of networks used by different clusters. Based on this information, a survey was performed to reveal the real-time measurement capabilities of different nodes (cf. Table 2).

HP ProCurve 5400 switch	*Juniper EX3200 switch*	*NetFPGA*	*Polatis Fiber Switch*	*Ciena CN4200 switch*	*Ciena CoreDirector switch*	*Cisco 6509-E switch*	*Infinera DTN ROADM*
-	Detection of link brakes	Freely programmable FPGA	Optical power monitor	FEC	-	-	Bit error counter

Table 2: Real-time measurement capabilities of GENI node equipment according to vendor's webpages and datasheets.

From Table 2 the most promising candidates to be applied in real-time measurement based cross-layer experiments are:

- **Infinera DTN:** The *Infinera DTN* is a remotely configurable optical add/drop multiplexer (ROADM) employed in a majority of nodes within the GENI network (BEN, Internet2, ...). One key advantage is its ability to provide bit-transparent services with digital performance monitoring [infinera08_1]. The DTN's PIC technology allows cost-effective O-E-O conversion of the optical signals as they transit the network, allowing Bit Error Rate (BER) monitoring from Forward Error Correction (FEC) and performance monitoring of

SONET/SDH and G.709 overhead at every digital node. The Infinera Digital Network Administrator (DNA), a comprehensive integrated element and network management system then provides users with a graphical interface with performance management capabilities.

- **Polatis 24port fiber switch:** The Polatis switch is employed within BEN and offers the possibility to remotely monitor the optical power [polatis08_1]. It also can be used as a variable optical attenuator, which could be used to compensate for power transients in a cross-layer based fashion.
- **Ciena CN4200 switch:** The installation of the Ciena CN4200 is planned within networks used by Cluster B [geni08_2]. Transponder modules (e.g. the F10-T_10G or FC4-T module) already incorporating forward-error correction (FEC) devices with the possibility of accessing pre-FEC BER are available [ciena08_1] but not yet installed.
- **NetFPGA:** The NetFPGA is a low-cost platform, primarily designed as a tool for teaching networking hardware and router design [netfpga08_1]. It has also proved to be a useful tool for networking researchers. Currently it is implemented in networks used by Cluster B and Cluster C. At a high level, the board contains four 1 Gigabit/second Ethernet (GigE) interfaces, a user programmable Field Programmable Gate Array (FPGA), and four banks of locally-attached Static and Dynamic Random Access Memory (SRAM and DRAM). It has a standard PCI interface allowing it to be connected to a desktop PC or server. The use of an FPGA allows implementing customized performance monitoring functionalities from scratch. However, the 1 Gb/s interfaces are rather slow. The next version of NetFPGAs is planned to incorporate 10Gb/s serial interfaces [netfpga08_2].

5.2 Simulation

This section gives an overview on the newly developed software modules for the discrete-event network simulator ns2, which should allow for the simulation of real-time measurement based cross-layer communication within the GENI network. The first modules will be made publicly available via the GENI wikipedia page in January 2009.

5.2.1 Intra-packet BER variations

The current version of ns2 (ns-2.33) does not allow for the simulation of individual bit-errors on a time scale shorter than the packet duration. The decision to drop a packet is based on a given packet loss ratio and its statistics. The number of bit errors per packet is not calculated or recorded.

Therefore, we have implemented a new software module that calculates the number of bit errors and burst errors per packet according to the variations of the BER with respect to time. These variations may be on a time scale which is less than the packet duration, i.e. the intra-packet variations of the BER can now be taken into account. The time course of the BER is accessible to the simulation environment by means of an external file. Therefore, any simulation tool that allows for the modeling of BER variations over time or even recorded real-time measurements of a physical link may be used for our simulations. In this way, individual files, i.e. BER variations, can be assigned to each network link in the ns2 network topology. As a packet traverses a link (Fig. 2), the number of bit errors and the number of burst errors that the packet incurs are calculated according to the timestamps of the transmit and receive times; these are based on the BER variations that occur during this time period.

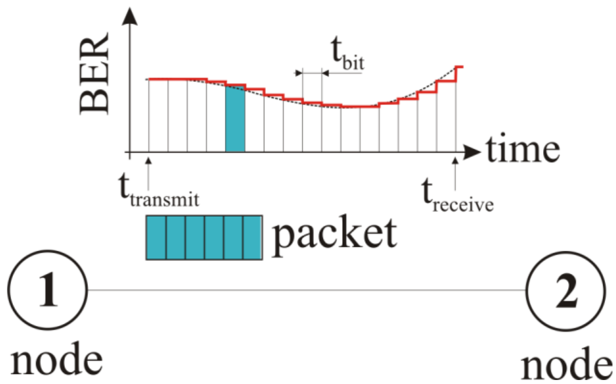

Fig. 2: Calculation of bit errors within a packet traversing a link based on BER variations stored in external file.

During simulation, the calculated bit errors and burst errors within each packet are recorded. Each receiving node can then decide whether these errors can be corrected by means of an FEC or whether the packet must be dropped. To allow for the BER variation functionality, an additional ns2 header was added to each packet that stores information about the transmission and receive time of each packet (two timestamps), the number of bit errors, the number of burst errors, and an error flag (Fig. 3**)**.

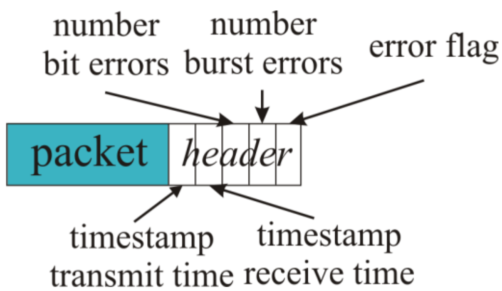

Fig. 3: New ns2 header added to packet to allow for simulation of intra-packet BER variations.

5.2.2 Forward error correction module

As mentioned above, the decision to drop a received packet depends on the FEC scheme used in the receiver. We have implemented a new ns2 module that allows for the specification of whether an FEC for error correction is used in the receiving node, whether a CRC check is performed for error detection, if bit errors should be ignored, or if the packet should be dropped whenever it contains bit errors. The error correction and error detection capability can be enabled or disabled independently by means of a flag. The ns2 user can also specify the maximum number of correctable errors, the maximum number of detectable errors, and the maximum number of consecutive erroneous bits within a packet which can still be corrected (i.e. the maximum burst error length). If any of these values is exceeded by the number of errors and burst errors stored in the ns2 packet header, the packet is dropped and/or marked with an error flag.

5.2.3 Local control plane

The newly implemented "local control plane" is a module which aims to monitor the QoS requirements for each packet, in addition to the performance on a link, and takes appropriate action if the signal quality degrades, i.e. if the actual BER exceeds a certain, predefined target BER. The control plane module constitutes an essential entity for future cross-layer simulations in the ns2 network simulator and is currently in the development stage.

6. Project participants

P.I. Keren Bergman, Columbia University New York, *bergman@ee.columbia.edu*
Caroline P. Lai, Columbia University New York, *cpl2114@columbia.edu*
Franz Fidler, Columbia University New York, *ff2213@columbia.edu*

7. Publications (individual and organizational)

- C.P. Lai, H. Wang, K. Bergman, "Cross-Layer Communication With an Optical Packet Switched Network via a Message Injection Control Interface," IEEE Photon. Technol. Lett. 20 (12) 967-969 (Jun 2008)
- C. P. Lai, K. Bergman, “Demonstration of Programmable Broadband Packet Multicasting in an Optical Switching Fabric Test-bed,” Accepted at OFC 2009, Paper OTuA5, San Diego, CA, Mar 2009
- C.P. Lai, H. Wang, A. Shacham, K. Bergman, “Priority Encoding Scheme for Contention Resolution in Optical Packet-Switched Networks,” ECOC 2008, P.5.7, Sep 2008
- K. Bergman, “Embedding Real-Time Substrate Measurements for Cross-Layer Communications”, Presentation on 3rd GENI Engineering Conference, Palo Alto, CA, 29. Oct. 2008
- C. P. Lai, F. Fidler, K. Bergman, “Embedding real-time substrate measurements for cross-layer communications”, Poster Presentation on 3rd GENI Engineering Conference, Palo Alto, CA, 29. Oct. 2008
- C. P. Lai, F. Fidler, K. Bergman, “GENI Quarterly Report”, Dec. 2008

8. Outreach activities

none

9. Collaborations

- **Cluster D:** Renaissance Computing Insititute (RENCI) and Duke University, Ilia Baldin

We contacted Cluster D to discuss and assess the possibilities of using (or inserting) real-time measurement capabilities for cross-layer communications into the “Breakable Experimental Network” (BEN). Some basic questions on BEN's resources, especially on its bit error (BER) monitoring capabilities, were discussed via e-mail.

It was found, that BEN incorporates the *Infinera DTN* in its nodes which should offer the possibility of BER monitoring, as well as *Polatis fiber switches* with built-in power measurement capabilities. BEN management network is accessed through a VPN. Detailed information about how to access the BER and power measurements still has to be provided by RENCI (but might require an NDA with Infinera). RENCI is also interested in placing other measurement devices in BEN and make this information available as part of slice provisioning by the ORCA control framework.

- **GENI Substrate Working Group:** University of Houston, Deniz Gurkan

We are currently in active contact with Prof. Gurkan and her group at the Univ. of Houston regarding their current GENI project goals of evaluating the embedded measurement capabilities of the current GENI prototypes. The results of their findings will be crucial in achieving our milestone of knowledge of the physical layer measurement abilities. Although they will release the results of their survey according to their milestones, they have agreed to maintain a close collaboration with us.

- **GPO:** John Jacob

We contacted John Jacob to request the current version of the GENI "Spiral 1 substrate catalog" to perform a survey on the real-time measurement capabilities within the GENI network. The document was sent to us but it was pointed out that the current version is not final and the quality of information will be improved in following versions. However, the document provided us with a good overview on some of the equipment already installed in GENI's future network infrastructure. Specific detailed information, e.g. on *Infinera DTN's* BER monitoring capabilities, was not available. It was pointed out that performance monitoring data should be available over a craft interface and a network management interface at time intervals down to 1 second.

- **Cluster B:** Mid-Atlantic Crossroads (DRAGON), Chris Tracy, Jarda Flidr

At the 3rd GENI Engineering Conference, Mid-Atlantic Crossroads provided us with the information that bit error counters are implemented within *Infinera's ROADMs* available at every node within the DRAGON network. Detailed information about how to access this BER information would require an NDA with Infinera.

- **Vendor:** Infinera, Drew Perkins

We contacted Infinera with the request for information on the BER monitoring capabilities of their DTN nodes (e.g. the *Infinera DNA* user guide, *Infinera DTN* hardware guide, *Infinera DTN* system description guide, material from the *Infinera DTN* OAM&P two day course). To this day we have not received any information yet.

10. Other Contributions

none

11. Bibliography

[1] [geni08_1] Global Environment for Network Innovations – Wikipedia (2008, December) [Online]. Available: http://groups.geni.net/geni/wiki/Embedded%20Real-Time%20Measurements
[2] [geni08_2] GENI Project Office, "Spiral 1 substrate catalog (draft version)", 10. December 2008
[3] [ns08_1] The network simulator ns2 - Wikipedia (2008, December) [Online]. Available: http://nsnam.isi.edu/nsnam/index.php/Main_Page
[4] [orcaben08_1] RENCI, Duke University, "A Prototype GENI Control Plane (ORCA) for a Metro-Scale Optical Testbed (BEN)" (2008, December) [Online]. Available: http://groups.geni.net/geni/wiki/ORCABEN
[5] [gurkan08_1] D. Gurkan, "Data Plane Measurements" (2008, December) [Online]. Available: http://groups.geni.net/geni/wiki/Data%20Plane%20Measurements
[6] [infinera08_1] Infinera, "The DTN in detail" (2008, December) [Online]. Available: http://www.infinera.com/products/dtn_detail.html
[7] [polatis08_1] Polatis, "VOA and Power Meter Switch Tray VST" (2008, December) [Online]. Available: www.polatis.com
[8] [ciena08_1] Ciena, "CN 4200 FlexSelect Advanced Services Platform Family" (2008, December) [Online]. Available: http://www.ciena.com/products/products_cn4200_overview.htm
[9] [netfpga08_1] (2008, December) [Online]. Available: http://www.netfpga.org
[10] [netfpga08_1] J.W. Lockwood, "NetFPGA - A quick update for the GEC3 conference" (2008, October) [Online]. Available: http://groups.geni.net/geni/attachment/wiki/presentations/NetFPGA_Status_Update_at_GEC3.ppt?format=raw
[11] [lai08_ecoc] C.P. Lai, H. Wang, A. Shacham, K. Bergman, "Priority Encoding Scheme for Contention Resolution in Optical Packet-Switched Networks," presented at ECOC 2008, P.5.7, Sep 2008
[12] [shacham07_micro] A. Shacham, K. Bergman, "Building Ultralow-Latency Interconnection Networks Using Photonic Integration," IEEE Micro 27 (4) 6-20 (Jul/Aug 2007)
[13] [lai08_ptl] C.P. Lai, H. Wang, K. Bergman, "Cross-Layer Communication With an Optical Packet Switched Network via a Message Injection Control Interface," IEEE Photon. Technol. Lett. 20 (12) 967-969 (Jun 2008)
[14] [lai09_ofc] C. P. Lai, K. Bergman, "Demonstration of Programmable Broadband Packet Multicasting in an Optical Switching Fabric Test-bed," Accepted at OFC 2009, Paper OTuA5, San Diego, CA, Mar 2009