On Enabling Real-Time Large-Scale Network Simulation in GENI: The PrimoGENI Approach (Poster Abstract)

Miguel A. Erazo and Jason Liu

School of Computer and Information Sciences
Florida International University
Miami, Florida 33199

{meraz001,liux}@cis.fiu.edu

ABSTRACT

The Global Environment for Network Innovations (GENI) is a community-driven research and development effort to build a collaborative and exploratory network experimentation platform, a "virtual laboratory" for the design, implementation and evaluation of future Internets. In this paper, we present an overview of PrimoGENI, a GENI project with the goal of extending the GENI suite of interoperable infrastructure to allow network experiments at scale, involving physical, simulated and emulated network entities.

Categories and Subject Descriptors

I.6.3 [**Simulation and Modeling**]: [Applications]

General Terms

Real-time network simulation

Keywords

real-time simulation, emulation, virtualization

1. INTRODUCTION

Research on network testbeds has made steady progress. Physical testbeds, such as PlanetLab [3], provide realistic network environments (including live traffic) for testing network protocols and distributed services. Emulation testbeds, such as EmuLab [7] and ModelNet [6], allow more flexible and controlled network experiments. GENI [2] capitalizes on the success of these previous efforts and aims to provide an overarching technology that brings all different network testbeds together as a single collaborative and exploratory platform for implementing and testing new network designs and technologies fundamental to future Internet research. GENI is built on four premises: *i) programmability*: researchers shall be able to upload and execute software on GENI nodes deep inside the network hierarchy; *ii) resource sharing*: GENI shall be shared among multiple users and their experiments, possibly through virtualization; *iii) federation*: GENI shall support interoperability among various types of resources provided by different organizations; and *iv) slice-based experimentation*: each user shall be provided with a "slice" (or a subset of resources) of the GENI infrastructure, and network experiments shall be conducted independently on reserved resources within slices.

The PrimoGENI project aims to incorporate real-time network simulation capabilities into the GENI "ecosystem". Simulation is effective for studying the behavior of large complex systems that are otherwise intractable to close-form mathematical/analytical solutions. Real-time network simulation refers to simulation of potentially large-scale networks in real time so that the virtual network can interact with real implementations of network protocols, network services, and distributed applications. PrimoGENI will integrate our existing real-time network simulator, called the Parallel Real-time Immersive network Modeling Environment (PRIME) [4], into the GENI federation. PRIME is a real-time extension to the high-performance SSFNet simulator for large-scale networks [1]. We plan to augment PRIME with the GENI aggregate interface, through which PRIME will inter-operate with the rest GENI infrastructure to support large-scale experiments involving physical, simulated and emulated network components.

In this paper, we present our design for integrating the real-time network simulator into the GENI environment. Specifically, we provide the design of the PrimoGENI aggregate in the context of current GENI control framework. The design decisions will direct our development and prototyping activities with the objective of substantially broadening GENI's prospect in supporting realistic, scalable, and flexible experimental networking studies.

2. THE PRIMOGENI APPROACH

The current GENI design consists of three main types of entities: clearinghouses, aggregates, and principals. A clearinghouse is the central location for management of GENI resources for experimenters and administrators; specifically, it provides registry services for principals, slices and aggregates, and authentication services for accessing the resources. An aggregate represents a group of components encapsulating the GENI sharable resources (including computation, communication, measurement, and storage). When an experimenter from a research organization (i.e., a principal) decides to conduct a GENI experiment, she will negotiate with the clearinghouse and the associated aggregate managers through an elaborate resource discovery and allo-

SIMUTools 2010 March 15–19, Torremolinos, Malaga, Spain.

Digital Object Identifier: 10.4108/ICST.SIMUTOOLS2010.8636
http://dx.doi.org/10.4108/ICST.SIMUTOOLS2010.8636

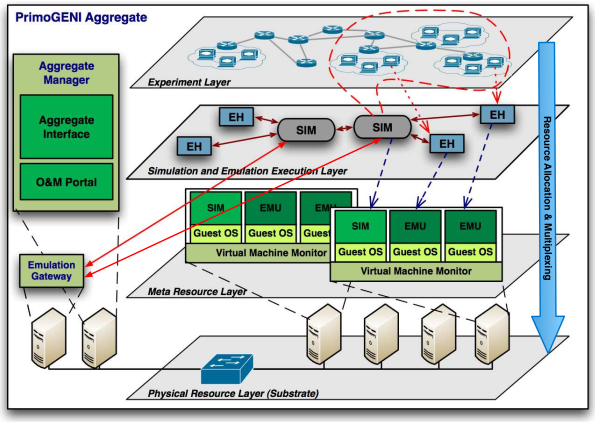

Figure 1: PrimoGENI as a layered system

cation process. In response to the experimenter's request, each participating aggregate will provide a set of requested resources, which are called *slivers*. Jointly, these slivers form a *slice*, in which the experimenter conducts experiments, with the help of GENI experiment support services.

Real-time simulation complements the GENI concept of federating global resources as a shared experimental network infrastructure. Our immersive large-scale network simulator, PRIME, supports experiments potentially with millions of simulated network entities (hosts, routers, and links) and thousands of emulated elements running unmodified network protocols and applications. PRIME can be readily extended to support GENI experiments. In order to interact with other GENI facilities, PRIME will function as a GENI aggregate, so that experimenters can use a well-defined interface to remotely control and realize network experiments consisting of both physical, simulated and emulated network entities exchanging real network traffic. In addition, PrimoGENI will provide experiment support services to facilitate large-scale network experiments involving simulated and emulated components; these services will include model configuration, resource specification, simulation deployment and execution, online monitoring and control, data collection, inspection, visualization and analysis.

PrimoGENI will use the ProtoGENI control framework [5] to manage, control and access the underlying resources. PrimoGENI makes the distinction between meta and virtual resources. *Meta* resources include compute nodes and network connectivity between the compute nodes. We call these resources "meta resources" to distinguish them from the physical resources (also known as the substrate), since they could be virtual machines and virtual network tunnels. Meta resources are managed by and accessible within the ProtoGENI/EmuLab suite. *Virtual* resources are elements of the virtual network instantiated by PRIME, which include simulated hosts, routers, links, protocols, and emulated hosts. We call these resources "virtual resources" as they represent the target (virtual) computing and network environment for the GENI experiments; they encompass both simulated network entities and emulated hosts (which are run on the virtual machines). PrimoGENI will export an aggregate interface as defined by the ProtoGENI control framework, and provide mechanisms for instantiating the virtual network onto the ProtoGENI/EmuLab facilities as configured and allocated on behalf of the experimenter.

The PrimoGENI aggregate can be viewed as a layered system, as shown in Figure 1. At the lowest layer is the *physical resources (substrate) layer*, which is composed of cluster nodes, switches, and other resources that constitute the EmuLab suite. These resources will be made known to the clearinghouse(s) and can be queried by researchers during the resource discovery process. In addition, two servers are set up to run the aggregate manager (for exporting an aggregate interface to researchers and clearinghouses) and the emulation gateway (for communicating with other slivers on other aggregates), respectively.

A *meta resources layer* is created upon resource assignment in a sliver. PrimoGENI uses the EmuLab suite to allocate the meta resources (including a subset of cluster nodes, VLAN connectivity among the nodes, and possible GRE channels created for communicating with resources off site). Each physical cluster node is viewed by PrimoGENI as an independent scaling unit loaded with an operating system image that supports virtual machines (e.g., OpenVZ). Multiple virtual machines may be created on the same physical machine to run the PRIME instance and the emulated hosts, respectively. In particular, the simulator will run on a (privileged) virtual machine, and the emulated hosts will be mapped to separate virtual machines so that they can run unmodified applications.

A *simulation and emulation execution layer* is created according to the virtual network specification of a sliver. The PRIME instances and the emulated hosts are mapped to the meta resources at the layer below. High-performance inter-VM communication channels are established between the emulated hosts and the corresponding real-time simulator instance, so that traffic generated by the emulated hosts are captured by the real-time simulator and conducted on the simulated network with appropriate delays and losses according to the simulated network conditions. Each real-time simulator instance handles a sub-partition of the virtual network; they communicate through VLAN channels created by the EmuLab suite at the meta resources layer. They also establish connections to the emulation gateway for traffic to and from slivers on other aggregates.

Once the slivers are created and the slice is operational, experimenters can conduct experiments on the *experiment (logical) layer*. She will be able to log into individual emulated hosts, upload software, and launch it. Traffic between the emulated hosts will be conducted on the virtual network. Traffic originated from or destined to other physical network entities will be redirected through the emulation gateway. Experiment data will be collected and viewed in real time on demand through the measurement facility.

3. REFERENCES

[1] J. Cowie, D. Nicol, and A. Ogielski. Modeling the global Internet. *Computing in Science and Engineering*, 1(1):42–50, 1999.

[2] GENI. `http://groups.geni.net/`.

[3] L. Peterson et al. A blueprint for introducing disruptive technology into the internet. *HotNets-I*, 2002.

[4] The PRIME Research. `http://www.primessf.net/`.

[5] ProtoGENI. `http://www.protogeni.net/`.

[6] A. Vahdat et al. Scalability and accuracy in a large scale network emulator. *OSDI*, 2002.

[7] B. White et al. An integrated experimental environment for distributed systems and networks. *OSDI*, 2002.

Digital Object Identifier: 10.4108/ICST.SIMUTOOLS2010.8636
http://dx.doi.org/10.4108/ICST.SIMUTOOLS2010.8636