Report of the GENI Measurement Workshop

June 26, 2009

Paul Barford (U. Wisconsin-Madison), Joe Chabarek (U. Wisconsin-Madison),
Mark Crovella (Boston University), Joel Sommers (Colgate University),
Charles Thomas (U. Wisconsin-Madison), Dave Plonka (U. Wisconsin-Madison)

ABSTRACT

The GENI Measurement Workshop was held in Madison, WI on June 26, 2009. The goal of the workshop was to bring together experts in the area of network measurement as well as other GENI constituents to discuss the foundational components of the GENI measurement capabilities. This report summarizes the presentations and discussions held at the workshop.

Table of Contents

Participants

Aditya Akella - U Wisconsin
Paul Barford - U Wisconsin
Mike Blodgett - U Wisconsin
Joe Chabarek - U Wisconsin
Mark Crovella - Boston U
Heidi Picher Dempsey - GPO
Nick Duffield - AT&T
Chip Elliot - GPO
Nick Feamster - Georgia Tech
Franz Fidler - Columbia
Victor Frost - NSF
James Griffioen - U Kentucky
Deniz Gurkan - U Houston
Suzanne Iacono - NSF
Gianluca Iannaccone - Intel
Ron Kramer - U Wisconsin
Will Leland - BBN
Larry Peterson - Princeton
David Plonka - U Wisconsin
Elizabeth Nichols - PlexLogic
Ivan Seskar - Rutgers
Christopher Small - Indiana
Karen Sollins - MIT
Neil Spring - UMD
Robert Ricci - Utah
Vyas Sekar - CMU
Christopher Small - Indiana
Karen Sollins - MIT
Joel Sommers - Colgate U
Charles Thomas - U Wisconsin
Darryl Veitch - Melbourne
Matt Zekauskus - Internet2

Introduction

The Global Environment for Network Innovations (GENI) represents a compelling vision for the development and evaluation of next generation Internet technologies. Similar to other large-scale scientific infrastructure, comprehensive instrumentation, measurement and analysis capability will be critical to the GENI infrastructure's ultimate success. While initial development efforts on Spiral 1 have already begun, the scope and complexity of the infrastructure call for on-going discussions about its key components, and in particular its measurement capability.

Unfortunately, there is little that can be borrowed from the current Internet or other experimental network testbeds in terms of a blueprint for the instrumentation and measurement systems for GENI. What are the key architectural components of an instrumentation, measurement and analysis infrastructure for GENI? What kinds of instrumentation and sensors should be deployed in GENI in order to enable the broadest range of experiments? How can measurements be taken for simultaneous experiments across multiple layers of protocol stacks from widely distributed systems in an efficient and scalable fashion? How can the critical issues of data privacy and measurement integrity be addressed in an infrastructure that is owned and maintained by many different stakeholders? These and other important questions must be addressed in order to create a measurement capability for GENI that will enable its full vision to be realized.

To address these questions, the GENI measurement workshop was organized and held on June 26, 2009 at the Pyle Center on the campus of the University of Wisconsin-Madison. The goal of the workshop was to bring together experts in the area of network measurement and other GENI constituents to discuss the foundational components of the GENI measurement capabilities.

To that end, the technical focus of the workshop will include:

- Measurement architecture: key components required to realize GENI mission
- Instrumentation: what/where/how to deploy sensors in GENI
- Experiment specification: interfacing with control frameworks
- Data management: aggregating, labeling and archiving data

This report documents the workshop outcomes, presentations, and discussions. The workshop was divided into four sessions; each session consisted of presentations followed by question-and-answer time. This report presents the key points made by speakers in each session and the contents of the question and answer sessions.

All original presentation materials (slides) are available at

`http://groups.geni.net/geni/wiki/GENIMeasWS`

Organization

Session 1: Measurement Architecture
Moderator: Mark Crovella
Speakers: Will Leland (BBN), Deniz Gurkan (U Houston), Elizabeth Nichols (PlexLogic), Darryl Veitch (U. Melbourne)

Session 2: Instrumentation
Moderator: Paul Barford
Speakers: Aditya Akella (U Wisconsin), Neil Spring (UMD), Jim Griffioen (U. Kentucky), Nick Feamster (Georgia Tech), Franz Fiedler (Columbia)

Session 3: Experimental Specification
Moderator: Joel Sommers
Speakers: Larry Petersen (Princeton), Robert Ricci (U. Utah), Ivan Seskar (Rutgers), Matt Zekauskas (Internet2)

Session 4: Data Management
Moderator: Paul Barford
Speakers: Karen Sollins (MIT), Nick Duffield (AT&T Labs-Research), Gianluca Iannaccone (Intel), Christopher Small (Indiana U)

Session 5: GPO and NSF Responses / General Discussion

Session Summaries

Session 1: Measurement Architecture

Moderator: Mark Crovella

Will Leland - BBN

- Disruption Tolerant Networking (DTN).
- What are measurement requirements for doing DTN on GENI?
- What are benefits to GENI in doing these experiments?
- Destinations can be described, rather than addressed - e.g. "Deliver content to everyone within two miles of this site for the next two hours."
- Disruption in a federated environment goes beyond a node going down. It can mean an entire network going down for a wide variety of reasons.
- Disruptions can occur on all sorts of scales.
- DTN concepts: exchange bundles, no network end-to-end path necessary when network is responsible for delivery, how to measure delivery time, etc?

Q&A

Q: The issue of timestamps is very important because without coordinated common timebase, measurements lose their meaning.
A: May be able to do creative accounting to reconcile non-coordinated timestamps.
Q: Are people working on using this to deliver content faster?
A: Content-based networking example, if you put DTN and CBN together, there are interesting consquences for what to measure, etc.
Q: There are Meta-level questions: how does the meta data get distributed? What is the rendezvous mechanism for requests?
A: This is why GENI could be useful to study this there are lots of fun, hard problems. There needs to be some deep theoretic understanding and experience at large scale.

for geni - we need to look at a measurement architecture
(e.g., synchronized time is likely meaningless for DTN)
think about how to flexible record the information that matters
current GENI System Reqs doc doesn't really address things in terms of DTN
(e.g., and doesn't really define "components")

Q: topological/physical disruption may be quite different than policy disruption
(i.e, physical may be all or nothing, whereas the policy ones are selective)
thus there are consequences for understanding the measurements in these situations
A: Similar for security concerns

Q: The role of time and timestamps is mysterious how do we make sense of measurements without roughly synchronized timestamps?
A: Assumed coordination time base, have a project without assuming common time stamps. Want more time available than the experimental network sees

Deniz Gurkan - U Houston

- Data plane measurements. Embedded measurements. Suggested external measurements.
- Embedded measurements = available at the nodes. Sometimes restricted by vendor API.
- Examples: Ciena CN 4200, Adva FSP 3000, Infinera DTN:
  - ALL have different API's but have Ethernet Connectivity
- External measurements depend on the needs of the control framework.
- Wide variety of external measurement systems available for both wireless and optical networks.
- Basic attributes must be in place (networkable, configurable, programmable).
- API to interface with Aggregate Managers across GENI.
  - Remote access mechanisms.
  - Storage
- Must determine how to slice a measurement device for multiple user access.

Q & A

Q: Layer 1 and 2 measurements seem fuzzy re: which research questions. Do you have a set of target research questions for L1/L2 that would suggest a combination of embedded and external measurement components? i.e., what problems are people focused on and how can these lower layer measurements be brought to bear?
A: This is deloping since only recently has this data become available an example about link-level costs for routing algorithm also optical BERT (?) switching - might want to know physical-layer end-to-end measurements.

Elizabeth Nichols - PlexLogic

- Metrics Management.
- Instrumentation Frameworks - Some of these are mature technologies.
- Terminology is important
- Instrumentation - Generates raw observations
- Sensor - Listens to instruments, creates measurements
- Measurement - Quantitative value
- To the extent that dimension and metric management can be automated, it should be so as to avoid human error/laziness.
- Metric layer vs Instrumentation layer
- Metric layer enriches data obtained from instrumentation layer
- Important to filter metric data so as not to swamp users with useless information.

- "Top Down" - Let the definition of your questions define the data you need and let that define the metrics.
- www.metricscenter.org - cloud service
- GENI metrics layer as a "long-running experiment" managed by GENI O&M

Q&A

Q: People are bad at updating dimension information
A: This needs to be automated, needs an operation and mgmt framework (e.g., asset mgmt system). "Dimensions" come from other source (not the sensors) then a metric is a raw measurement enriched with dimensions or analytics then these metrics are objects that need to be managed for metrics, it's better to do this top-down (while instrumentation thus far is bottom-up). See http://www.metricscenter.org Supports more thought on metrics rather than instrumentation layer
Q: Is there some tension with the top-down piece? You don't know the questions in advance (for GENI research) so what context do you use?
A: Pick example questions: fault mgmt, configuration mgmt for instance.
Q: Maybe there's a role for both because you want to know what would drive particular instrumentation.
Q: One context that we can define tightly for experiments, i.e., metrics on the instrumentation itself

Darryl Veitch - Melbourne

- NTP System
- GENI can help research on timing, but results of timing research can also be implemented as a solution for GENI
- NTP has many limitations (accuracy, robustness, modularity, etc)
- NTP codebase difficult to understand and basically is only understood by one person
- NTP neglects important sync issues
- Difference clock offers very robust and stable measurements over absolute differences.
- Path asymmetry causes problems for absolute clock but not difference clock
- NTP doesn't provide a difference clock.
- Problems pronounced in multiple-server/multiple-client environments.
- Changes made to compensate for timing issues can cause more problems than it solves.
- System needs to support difference timing at all levels.
- Greater quality monitoring is required.
- Stratum 1 should be secured as an important part of the network
- RADClock project (Robust Absolute and Difference Clock). http://www.cubinlab.ee.unimelb.edu.au/radclock/

Q&A

Q: Can we run this on planetlab?
A: The client software is available so a radclock hierarchy could be put in planetlab there is no radclock stratum-1 yet though.
Q: Do we want the same stratum arrangement, server placement, etc.? Do we want better clocks?
A: We haven't thought a lot of that yet it would be possible to flatten the hierarchy, can come up with wan level broadcast that can flatten hierarchy. Connect to the highly accurate stratum 1.
Q: In GENI we have an opp'y to make an optimal infrastructure now that GPS is low cost, where should the stratum-1 infrastructure be, someone should use gps stratum 1 placement in geni to improve but, even if you have a very accurrate clock, there is still the issue of how do timestamps get applied (ie, kernel, user mode, etc.) In the end we care about how the timestamps get applied to events.
A: The timestamping is really a separate, and important, issue there are custom kernel timestamping sol'ns and other good sol'ns in the radclock package for stratum-1 it's difficult - b/c it must be able to see the satellites at all time so its useful to have a good stratum-2. Time accurate GPS is still expensive.
Q: CDMA takes timing from gps to 10us, doesn't work on GSM.
Q: There's a lot of good technology, incl. multicore, to use for better instrumentation.

Session 2: Instrumentation

Moderator: Paul Barford

Aditya Akella - U Wisc

- Role of memory and storage in supporting measurement.
- Indices are very large (> 100GB) and need to be very fast with reads/writes.
- B-tree, log structures not suitable. B-treeLookups are tool slow, logs have poor support for flexible fast garbage collection.
- Hash tables are optimal, but current solutions (e.g. DRAM) have issues (heat, cost).
  - Flash is an alternative to the suboptimal nature of current implementations.
  - Fast sequential writes
  - Insertions and expirations are difficult
- BufferHash - use DRAM as buffer, then commit to flash.
- Read and write latencies under 0.1ms for 99th percentile (100x better than disk, similar cost)

Q&A

Q: Is there one index? or one per
A: One index for all the data on that link, but there could be multiple indexes.
Q: There's a balance betw. reads and writes? why not one write, many reads?
A: Depends on application, (e.g.) balanced read/write for de-duplication
flash is good for lookups b/c random read is fast
however, random writes are expensive
Q: Java garbage collection effect on response time. Do you have to pause a real time process to collect garbage?
A: Can erase multiple blocks at a time, so there is a tradeoff between how often and at what cost.

Neil Spring - UMD

- Key problems in supporting measurement.
- Support for summarization and filtering
  - "Complete" information is intractable
  - Difficult to predict what's needed
- Precise, prompt, standardized timestamps
- Design timestamps on forwarding path.
- Time should be public.
- Compensation for virtualization
  - Need to understand when interference from other activities is occurring

- Could network error reports be descriptive enough to allow applications to fix errors?
- Expose errors at each layer to be uniformly propagated and handled?

Q&A

Q: Timestamp event rather than packet? Moment of composition or sending is more important than when packet hits the log?
A: Right, a packet trace is not precise.
Q: Getting extra data requires adding load to potentially overloaded machines.
Q: Exposing information in this way causes security problems?
A: We are examining things that go wrong with pkt., not sure what kinds of information need to be protected in that realm.

Jim Griffioen - Univ. Kentucky

- Instrumentation for ProtoGENI-based Edulabs.
- Enhancements to Emulab to automate instrumentation and publish results via the web.
- Capture of network state deemed more instructive than packet traces.
- SNMP daemons monitor experimental network and send to control network.
- Each experiment has its own out-of-band measurement controller (MC).
- Each MC is an extra sliver in the experiment.

Q&A

Q: For proto-GENI is it important to get an out-of-band measurement network?
A: This "measurement channel" could require more bandwidth than the experiment. Is instrumentation a feature of a resource or a resource?
Q: Will you need something more complicated than the student interface, or will your architecture work for GENI?
A: Don't rule out expanding current capabilities. Some additional info about the "global" state may be useful.
Q: Do you see only your own traffic, or is the view expanded?
A: We have to look at that, so they can see what other experiments and traffic are going on. (Right now they only see their own w/in a GRE tunnel) It's important to know about local traffic as it can skew the experiment. Therefore, the local view is very important in terms of understanding interference.
Q: Its important to know about cross traffic some measurement of unwanted interference.

Nick Feamster - Georgia Tech

- Virtual networks need BGP, but this is cumbersome particularly for transient experiments.
- Each time you change a dynamic experiment you don't want to reconfigure BGP.

- BGP Mux provides a separate BGP view for each upstream ISP.
- Downstream networks look as though they are directly-connected to upstream (mux is transparent).
- Upstreams are isolated from transient behavior.
- Proto-GENI compatible.
- Potentially scaling problems, but potential solutions (e.g. split views, hardware acceleration)
- Potential to do measurement using the Mux.

Q&A

Q: Advertisements flow in but what about advertisements that flow out of GENI. Demultiplex data traffic? NAT? Issues related to outbound traffic can support more than 8 experiments? What about unstable experiments?
A: Right now advertise /21, could chop it up (for multiple experiments). Can demux data traffic... multiplex in both directions.
Similar to NAT. Outbound traffic has other issues (> 8 experiments, unstable experiments, etc).

Franz Fiedler - Columbia

- Unified Measurement Framework for Embedded Real-Time Measurements
- Need to control and acquire measurement data in a unified way
- GENI needs to support embedded real-time measurements.
- Performance monitors already support at least one management protocol.
- Eventually XML based.
- Many monitors out there (optical power, bit counter, etc), we need to provide access to this information.
- Access control, abstraction of measurement capabilities, single point of access.
- SILO - For complex tasks, forms stacks of elemental functional blocks (service)
  - Flexible, promotes cross-layer communication
  - All services can be developed in isolation and interfaces can be exposed
- perfSONAR - Allows gather/distribution of measurement information
  - Measurement point service, archive service, lookup service, etc.

Q: Are you implying that perfSONAR access control would work for GENI - Architecture handles authorization and access control?
A: Also handled by control framework. Need to look into perfSONAR to see if it will work or may need to be augmented.
Want to give user the ability to decide on the framework they wish to use.

Session 3: Experimental Specification

Moderator: Joel Sommers

Larry Petersen - Princeton

Instrumenting PlanetLab.

- Co-Mon - diagnostics on per-slice/per-node basis.
- You monitor a PLATFORM not the slice
- Measurement capabilities need to be in RSpec.
- Trouble-shooting/auditing packets from other experiments that a user may have received.
- Analysis per user of CPU vs bandwidth, etc.
- RSpec->NetSpec->NodeSpec
- Demo.
- Planet Flow: Search, auditing for annoyed IT managers
- Visualization for cycle vs. transmit to weed out compute hogs.

Q&A

Q: Slice dying? (during demo) Can see what it did during lifetime.
Q: How does it decide what the qualifications are for getting a good time slice
A: Based on who is asking.
Q: Nice array of instrumentation. Assembled piece-meal. In hindsight are there elements that you would do differently to accomplish your goals?
A: Not really. We back-fit features from time to time. We separate the authoritative state from the dynamic state. We don't collect layer 1 and 2 stuff, mostly about interfaces and nodes.
Q: What do you do with GENI when you have aggregates that are not assumed to be part of planet lab? that's not in PlanetLab?
A: Enable a slice for CoMon to run in. Other solutions may come along. If you have a useful slice for monitoring, slice will be instantiated along aggregates. CoMon is a replacement for many earlier solutions.
Q: How do you bootstrap if you are not in planet lab?
A: Survival of the fittest.

Robert Ricci - Utah

- Measurement and Experiment Specification
- Filtering, tracing, etc available on links in Emulab (TCPDump) via transparent capture
- Measurement capabilities need to be in the RSpec
  - Help with component and link selection
  - Slice definition determines measurement capabilities
  - Measurement hardware capability needs to be defined
- Describing measurement devices is complex and non-standard (vs other components)
- Best level of abstraction?

- Informing user about devices they implicitly requested
  - Monitor link
  - Measurement control

Q&A

Q: Imagine running repeatable experiments. What is the appropriate mechanism to encapsulate experiments?
A: Pieces include exact RSpec of first experiment. Something to keep track of changes over time in components. Want to capture any executed or data files executed (closed world, capture all inputs and outputs). If you just want to compare result without rerunning, then there are a different set of things to keep.
Q: It could be treated statistically, run the experiment enough to get a set of results.
A: Sometimes detecting that there was a difference is just as good or better than being able to reproduce the experiment exactly.
Q: Measurement system itself runs completely dedicated to measurement task?
A: Yes, in Emulab. Management infrastructure is separate, out-of-band.
Q: System-wide considerations are important along with slice-based considerations, like Larry mentioned you measure the platform.
A: Yes.

Ivan Seskar - Rutgers

- Orbit - 400 node radio grid emulator + trial network. 1200 radios. Experimental infrastructure.
- Measurement infrastructure to collect real-time measurements.
- Real experimenters try crazy things, so need filters to manage behavior to keep it in reasonable bounds.
- Real experiments don't want to write code, they want applications ready to go (e.g. traffic generator, sink, common apps, routing protocols).
- Client filters
  - Average over a hundred packets
- Need to support disconnected operation.
- OML2: Now have proxy service running at each node.
- Originally non-virtualized. All-or-nothing provisioning of resources in 2-hour blocks.
- Need to be able to characterize environment to have any hope for reproducible results (both indoor and outdoor considerations).
- RF components change over time
- Building changes
- Must be able to observe wireless interference.
- Nodes are not the same, calibration necessary.
- Calibration can be expensive, it takes a few weeks.
- Need to be quite inventive to troubleshoot.

Q&A

Q: These considerations are similar to experiments on the wider Internet. Testing for quality on wireless, everything becomes critical.
The real-world environment is similarly complex, and there's a tolerance to the variation in quality that has to be built-in.
A: If you're trying to design a system, these considerations become very important.
Q: What sensitivity are you testing at?
A: [..]

Matt Zeukauskas - Internet2

- perfSonar - Performance monitoring framework
- Essentially a network middle-ware for measurement.
- Service oriented architecture
- Collection and visualization APIs - measurement archives
- SNMP, RRD, SQL, Circuit status measurement archive
- Visualization examples
  - E2e monview
- Tools available Java/Perl CPAN
- Need to harden packages
- Authenticate locally via trusted mechanisms (results in role not specific identity)

Q&A

Q: When you say wraps do you mean as a web service?
A: yes.
Q: Do you use the web service orchestration layer?
A: No, don't think so.
Q: Does this system extend into the computer world (monitor computer status)?
A: There have been thoughts about that by the grid people, some discussion of monitoring/scheduling computer centers.
Q: Is web services and SOA approach extendable, is it too heavy weight?
A: It's been very useful to do this, allows separate implementations to be developed. Need to be careful not to create enormous RSpec (XML) docs.
Optimization may be required.

Session 4: Data Management

Moderator: Paul Barford

Karen Sollins - MIT

- Experiment specification. Beyond configuration, but how the experiment should behave.
- Defining what is and is not acceptable is important (failures may be a success).
- Understand what to do in case of unacceptable behavior.
- Must specify experimental behavior in terms of
  - Network architecture and design - comparisons, tuning, reproduce-ability
  - Security - potential threats to other users/outside world (e.g. releasing a virus)
  - Education - constraints, refinement of constraints, demonstration of understanding
- Specifications are important for all aspects of GENI, not only experiments.
- Direct measurement vs inferred measurement
- Specification should include expectations
- Metadata important for sharing and improving performance

Q&A

Q: One could despair looking at this... is there anything like an ESpec that would be a starting point, even for a subset of conditions?
A: Yes, we wrote a DHS paper that talks about this, takes some first steps (deterlab.net).

Nick Duffield - AT&T Labs-Research

- Scaling Data Collection and Analysis in Measurement Infrastructures
- Attempt to anticipate what is going to happen in the future
- We are in a mature phase of internet measurement - too much data!
- Trends -
  - increasing scale and rates,
  - increasing demand for detail
  - increasing availability of embedded measurements
- Demand for detail
  - Endpoints, application, customer demand
  - Performance measurements
  - Infrastructure measurements
    - Direct instrumentation of equipment and protocol operation
  - Detailed temporal granularity
- Equipment and protocol support
  - Netflow, ipfix,psamp

        - Performance measurements
    - Protocols are in place for flow and performance measurements.
    - Constraints for measurements
        - Consider analyzer
            - Increased detail means greater than liner growth of measurement infrastructure
        - Can't solve only with more aggressive sampling, aggregation
            - would lose detailed view of data
    - Challenges for distributed infrastructure
        - Data reduction within infrastructure v. ability to correlate
        - Push analysis functionality out closer to data sources
    - Distributed measurement infrastructure
        - Embedded and external instrumentation generating raw data
        - Multiple measurement collection/analysis locations
        - Dynamic reconfiguration
        - Distributed control
        - Protocol support for transmission of control information, data summaries, correlation and combitation of data summaries
    - What does this mean for current measurement projects
        - ex. Net tom.
            - too little data
            - too much data
            - claim problem 1 goes away problem 2 gets worse
    - To manage scale make measurement infrastructure dynamically configurable.
    - Need protocol support for creation, transmission and correlation of data summaries.

Q&A

Q: We had hierarchical network managers previously. We looked at a peer-peer grid for management to execute a series of steps to do pin-point correlation via grid requests. Have you looked at approaches like this that don't rely on a hierarchy (since hierarchical solutions may not be flexible enough)?
A: I don't see the need to enforce strict child-parent relationships, so not locked into strict hierarchy.
Q: How do we address different legal domains (europe vs US, etc) for data retention, etc ?
A: Have not thought about it specifically. Clearly needs to be considered.
Q: You mentioned network tomography. It's easy to get measurements between IPs, but not hosts. How do you deal with multiple IPs per host (interfaces)?
A: I don't make a claim to network tomography. There is always a question of determining where elements start and end. That's an issue that may go away when direct measurements become more ubiquitous.
Q: Are they becoming more ubiquitous.

A: I think so, yes.
Q: Network tomography, how is the router interface aliasing issue being dealt with?
A: How do you determine where one network element begins and another starts? I think this problem will go away with more embedded measurements

Gianluca Iannaccone - Intel

- Data Management - What data? What about management?
- Requirements: fast, safe, easy.
  - Data stream processing main focus of past efforts
  - Started with specialized hardware (DAG) now within reach with general purpose servers
  - Safe: privacy/data anon, resource management
- Fast: today = 20Gbps
  - IO bottleneck not really a huge problem
  - CPU bottleneck
- Safe: Users cannot compromise each other, monitor machine is not overwhelmed.
  - EU Onelab2 approach.
    - Using como platform and API
    - Queries go to data, raw data never leaves systems
    - Monitoring systems accessed only via proxy
    - Users submit their own monitoring apps
      - Run in a slow safer environment
      - Or review process
      - Cannot access network traffic of other users
    - Load shedding alg. (usenix '07) enforces fair access to system resources (CPU only right now)
- Easy: Hard part. Under-specified and unresolved. Force API, or force new declarative language. Problems with both.
- Support for C creates lots of safety problems
- Declarative languages are nice but not easy
- Support for data annotations is very important
  - Needed for "write once, deploy everywhere"
  - COMO allows to attach any metadata string to any packet
  - Hard to specify in a way that is generic
- http://como.sourceforge.net

Q&A

Q: Sending code to data (like como) is a good way to get characterization, who's doing the como stuff these days?
A: Intel project... it's in the project group... a company has acquired funding to support it. Like PlanetLab.

Christopher Small - Indiana U

- Experiences with Measurement Data Collection on R&E Networks
- Global Research NOC
- GENI Meta Operations Center (GMOC) will need inputs to understand the state of the network.
- Researcher questions
  - What is the state of the network?
  - Where is the data (location, type, state)?
  - Why can't I get the data (privacy, cost, technology concerns, disruption of network)?
- Hard to describe changes in hardware.
- Varying levels of applicability to GENI.
- Streamlining (metadata attached to measurement data)
- Slice creation in GENI - Data discovery mechanisms and how they fit with the GENI slice.
- Possible relevant paradigms (Grid, HEP, astronomy, Emulab, DATCAT, etc)

Q&A

Q: Do you already use workflow languages?
A: We're examining it, but are not currently using one.
Q: Would you look at a GENI workflow language?
A: We would use it to understand a GENI experiment from an operational and management standpoint.
NSD107 paper describes a workflow language for EmuLab.
Q: An interface is up for quite a few years. Do you keep statistics about who accesses online data and what they do with it?
A: We do keep logs of all access. Most of the data has no registration, so we know who looked but not how they used it. A handle system might help with this, could go back to the last 10 years of sigcomm/infocom.
Q: Do you have a high-level sense for the impact?
A: The one set of data that has a registration process is I2 flow data (stored by I2). Good question and we can talk about flow data access, but since it's publicly-available it's hard to know what's going on with it. It would be difficult to figure out a good way to collect this kind of data.
Q: Our own experience has a voluntary registration page. Most people put something in there. Gives a ballpark for how it's being used. Would be useful to track the impact of this data.
A: We are also looking for data from GENI experiments for examination and to see how it fits with our setup. GENI needs to get credit where credit is due.

Session 5: GPO and NSF Responses / General Discussion

Observations (Chip Elliott)

If we're aiming toward measurement architecture, what is the minimal set?
i.e., the invariants -- we've heard a lot of options ...

Is there a desire to converge to a common software measurement platform?
(or do different groups do different things)

- GPO has no desire for it to converge to a common platform

What's the learning curve required to set up an experiment, collect measurements?

- if there isn't a common way to do it, is this an obstacle

What do we do to make it possible for students to use GENI easily/quickly?

- this is a major concern
- and what to be done in the short term (so that people who aren't on these teams can use GENI)

What's the minimum set that we know everyone is going to want, and get that agreed on

- remember to get the people that want to experiment involved in the conversation, especially security people, open up the discussion more
- Also, see if your own network research would benefit

As a non-GENI person, what does one do to go about doing an experiment?

We need people to start using these control frameworks and see what works

- others here should start designing and building GENI

Is there a specific measurement focused follow-up?

- There will be a report from this workshop
- as more of GENI comes together and we know what the infrastructure looks like, there will likely be a need for another workshop perhaps in 1 year
- lets do some things and get some experience with this

I have some software/components that could be used as GENI components… is this one way to get started? (e.g., catalog with meta-data)

- There is lot of documentation on GENI
- to get started in GENI, one way is to respond to CFPs (for funded projects) or talk to people about partnering with a funded participant
- communicate with the GPO to participate in the engineering conferences

Get to work. If you have expertise, make it so.

- I see the need for another such meeting in a year or so.

- We need vendor participation also to have embedded systems in network devices.
- We'll be getting real-world experience with GENI over the next six to twelve months.

Workshop Summary

The GENI measurement workshop was organized and held on June 26, 2009 on the campus of the University of Wisconsin-Madison. The goal of the workshop was to bring together experts in the area of network measurement and other GENI constituents to discuss the foundational components of the GENI measurement capabilities.

The workshop was a first meeting of interested parties in the domain of measurement for GENI. At the outset, the questions to be addressed included: What are the key architectural components of an instrumentation, measurement and analysis infrastructure for GENI? What kinds of instrumentation and sensors should be deployed in GENI in order to enable the broadest range of experiments? How can measurements be taken for simultaneous experiments across multiple layers of protocol stacks from widely distributed systems in an efficient and scalable fashion? How can the critical issues of data privacy and measurement integrity be addressed in an infrastructure that is owned and maintained by many different stakeholders?

The major themes of the workshop were: (1) measurement architecture, (2) instrumentation, (3) experiment specification, and (4) data management.

In the area of measurement architecture, Will Leland presented an overview of Disruption Tolerant Networking, and argued that some architectural principles from DTNs apply to GENI. In particular, both DTNs and GENI have difficulties in establishing a unified timebase for measurements collected in different locations (or slices). Some of the techniques used for unifying timestamps in the DTN world may be useful for GENI. In another take on the timing problem, Darryl Veitch presented methods that support precise time measurement (both absolute time and time differences). He argued for a precise stratum-1 clock as an integral part of GENI, and for the underappreciated utility of differential-time measurement. Taking a broader look at architecture, Deniz Gurkan presented the results of her survey of measurement methods and tools at low levels (layers 1 and 2).

In the area of instrumentation, Aditya Akella presented work showing the utility of new data structures for use with flash memory. In particular, he argued for combining the general flexibility of hash tables with the advantages of flash memory. Neil Spring argued for a number of general principles: summarization and filtering to deal with measurement data size, a well-designed timestamping mechanism, and a transparent error reporting system. Jim Griffioen described his work providing support for instrumenting proto-GENI experiments for users that are not measurement experts. His tools automate certain network measurements and publish them on the Web. Nick Feamster laid out the challenges in using virtual networks in conjunction with BGP. He described the design of BGP Mux. Finally, Franz Fiedler described a framework for real-time measurements.

On the topic of Experimental Specification, Larry Peterson described and demonstrated the CoMon measurement system for PlanetLab. The CoMon system is able to monitor on a per-slice basis, a key to disambiguating measurements from a virtualized platform. Robert Ricci

described RSpecs that capture all necessary information to run repeatable experiments. Discusison brought out the notion that an RSpec that defines an experiment can be metadata that describes measurements. Ivan Seskar described the measurement framework used in Orbit, a large radio grid emulator. Here again the need to precisely characterize the experiment (including the experimental environment) is key to properly interpreting measurement results. Finally, Matt Zekauskas described the perfSonar measurement system used in Internet2, which is a set of tools (middleware) for generally-available network measurements. The general organization of perfSonar is based on a web service oriented architecture.

In the session on Data Management, Karen Sollins described the notion of an experiment specification – something going beyond just experimental configuration to define a broader environment, including what sort of behavior is and is not acceptable for the experiment. Nick Duffield laid out the challenges associated with increasing scale, rates, and detail of network measurements. He describe the potential for the use of distributed measurement infrastructure, network tomography, and correlated data summaries in addressing these problems. Gianluca Iannaccone describe the CoMo system, a system allowing multiple measurement experiments to be performed simultaneously, without compromising the isolation of separate experiments being performed by different users. Finally, Chris Small described experiences with measurement data collection on the Internet2 network, including the difficulty of communicating to users operational questions such as the state of the network.

Despite the early nature of many of the systems and technologies discussed, there were a number of points raised in the workshop that should impact GENI measurement design. Some of the most salient observations from workshop presentations and discussion include:

1. The recognition that the distributed and virtualized architecture of GENI presents special challenges for time measurement and timestamp unification.
2. Measurement tools will need to consider support for experimenters who are not measurement experts.
3. Virtualization of certain network capabilities (eg, BGP) will require additional meta-data collection.
4. Some good examples of virtualization-aware measurement frameworks already exist in the PlanetLab and Emulab environments and in the CoMo system.
5. The ability to define repeatable experiments allows one to provide essential metadata for storage with experimental measurements.

In end of the day discussion there was consensus that the conversations needed to support the design of effective GENI measurement had just begun. Plans were discussed for an annual GENI measurement workshop.