Meeting Minutes from SMM 2009 in Arlington

Performance WG Face-to-Face Meeting

at Internet2 Spring Member Meeting
April 27, 2009

Attendees

Carla Hunt, MCNC (Chair)

Jeff Boote, Internet2

Tom Throckmorton, MCNC (via phone)

Chris Hawkinson - CENIC (via phone)

Brian Tierney - ESnet (via phone)

Peter O'Neil, MAX (Mid-Atlantic Crossroads)

Rich Carlson, Internet2

Linda Winkler, University of Chicago

Per Nihlen, NORDnet

Scott Colburn, U.S. Department of Commerce Boulder Labs

John Hicks, Indiana University

Don McLaughlin, Indiana University

John Streck, University of North Carolina at Chapel Hill

Charles Hollingsworth, Georgia State University

Martin Swany, University of Deleware

Maciej Strozyk, PSNC/PIONIER

Kazunori Konishi, APAN

Jon Dugan, ESnet

John Bartin, Washington University in St. Louis

Hans Wallberg, SUNET

Katsuhiro Sebayashi, Nippon Telegraph and Telephone Corp (NTT)

Hisao Uose, Nippon Telegraph and Telephone Corp (NTT)

Kenji Shimizu, Nippon Telegraph and Telephone Corp (NTT)

Takehito Suzuki, Nippon Telegraph and Telephone Corp (NTT)

Kazuto Noguchi, Nippon Telegraph and Telephone Corp (NTT)

Grant Miller, National Coordination Office, Computing, Information, and Communications

Bob Gerdes, Rutgers

John Stier, Stony Brook University, State University of New York

Andrea Blome, Internet2

Emily Eisbruch (scribe)

Discussion

Internet2 Update

Jeff Boote provided an update on the software releases available for Internet2 performance and measurement tools and a preview of the roadmap.

Release candidate for perfSONAR-PS 3.1 RC1 is available at http://software.internet2.edu.

REDDnet's Use of Performance Tools

REDDnet has disk depots that cache data placed throughout the U.S, and they have been experiencing challenges moving data between the depots.

Ezra Kissel, University of Deleware, presented on **REDDnet Performance Monitoring**

View Ezra Kissel's REDDnet presentation

Highlights:

- REDDnet provides "Working storage" to help manage the logistics of sharing, moving and staging large datasets across wide areas and distributed collaborations.
- Participating Institutions: Vanderbilt, Tennessee, Stephen F. Austin, NC State, Nevoa Networks, Delaware
- Host Sites: Caltech, Florida, Michigan, ORNL, SDSC, TACC, UC Santa Barbara (Stephen F. Austin, Tennessee, Vanderbilt)
- Tools used for performance monitoring inlcude:
    - OWAMP (3.1)
    - BWCTL (1.3)
    - NDT client (3.5)
    - perfSONAR-PS perfSONAR-BUOY (regular testing framework for BWCTL)
- Performance troubleshooting approach:
    - Ensure TCP is tuned on all hosts
    - Pick a set of hosts to investigate from the "worst offenders"
    - Divide and conquer testing:
        - Test from end-to-end path
        - Break up path into smaller segments
        - Narrow down the source of the problem by testing along the smaller segments and seeing which segments have the same symptoms as the end-to-end path
    - **Examples**:
        - REDDnet Umich and CHIC I2 POP
        - REDDnet Vanderbilt to Atlanta I2 POP

Internet2 and Cisco Telepresence

A behind-the-scenes look at the planning and setup for the Cisco Telepresence Demo shown at Wednesday's General Session.

Aaron Brown, of Internet2, presented on "**Regular Latency Monitoring Or: How I Learned to Start Worrying and Hate the Jitter**"

View Aaron Brown's slides on Latency Monitoring

Highlights

- Throughput testing is being done by LHC, ESNET and others
- But what about latency testing for latency sensitive applications such as Cisco Telepresence Demo?

- Cisco Telepresence Limits
    - 10 ms jitter
    - 160 ms delay
    - 0.05% loss
- Polycom Limits
    - 30-35 ms jitter
    - 300 ms delay
    - <1% loss

- Goals:
    - Measure delay/jitter/loss between points
    - Be able to fix any issues that come up
- Approach: Deployed measurement machines at the endpoints and a number of hosts in between and set up regular latency tests between the machines
    - Benefits: Shows end-to-end problems, and allows a "Divide-and-Conquer" approach to narrow down the source of the problem
    - Tools: OWAMP (Latency Tester) and perfSONAR-BUOY (Test Scheduling Framework)
    - Analysis software was written or modified to make it easy to view and understand the data.
- Results: Several potential performance issues, in both the network and the monitoring systems, were identified, and all were solved and verified through diagnostics and monitoring.

Interoperability Testing with Dante - Update

Tom Throckmorton, of MCNC, presented an update on the Multi-Vendor 10 Gigabit Testing that Matt Zekauskus and Tom discussed at the Performance WG at the Feb. 2009 Joint Techs in College Station. The goal is to determine 10GE vendor interoperability and higher speed circuits work between differing vendor hardware over long distance.

Tom reported that interoperability testing on a 10GE transatlantic circuit connecting Internet2 and DANTE has been ongoing over the past year, which was reported on in February 2009. At that point, we had reached limitations in the use of commodity systems as test endpoints. Since the Feb 2009 update, DANTE, Internet2 and MCNC had the opportunity to pursue simultaneous product evaluation of network test equipment from Xena Networks, a new company out of Denmark. This has been an opportunity to quickly complete the interop testing using suitable test equipment before turning the circuit over to production. These test units are FPGA-based, and priced about a tenth of the cost of similar testers.

DANTE had received testers in Jan 2009; Internet2 received testers towards the end of Feb 2009. We had set an aggressive timeframe for completing testing, based on anticipated turnover of the circuit for production. Having equipment on hand allowed us to complete the testing in timely fashion and also to complete tests with a higher degree of confidence than w/ using commodity PCs. There had also been issues around being able to drive circuits beyond certain rates with the commodity systems; with this test equipment in place, we were able to drive the circuit almost to full capacity. Did suite of scripted tests at a range of packet sizes (64-9000 bytes). We were able to iterate through the same set of tests independently and get the same results consistently, leading to high confidence in the numbers.

One issue emerged as a result of this testing. In one direction, we observed some throughput dropoff as the frame size approached 64 bytes. After a number of back-to-back tests, and repeated tests to be sure we got numbers accurately, we surmised a packet processing rate limitation on one side of the connection. Based on the anticipated use of this circuit, this is not a problem interoptablility-wise.

Overall we got excellent results from this gear -- more consistent than out of PCs. Another positive was interaction with Xena -- they were eager to please and responsive to issues we raised with them. Made corrections for us based on feedback we had given them. Will provide a general product evaluation around end of May and an interop test report will be delivered at end of June. Dante is pursuing the purchase of these testers for some use; neither Internet2 nor MCNC are pursuing a purchase at this time.

Some ideas surfaced on how we could make improvements in using commodity systems to do testing, which will be useful for future similar test scenarios. The underlying circuit was turned over for production in mid April and it's been carved up in different ways to serve connections between Dante and a couple of points in the U.S.

If some one wants to learn more, contact Tom Throckmorton or Matt Zekauskus.

Assembling a Performance Enhancement and Response Team (PERT) team in the U.S.

Discussion of establishing a team of network engineers representing each of the RONs that would be available on a rotating basis to troubleshoot complex, multi-domain issues.

This is the link to the Geant PERT team:

http://www.geant2.net/server/show/conWebDoc.1602

Jeff Boote is on the PERT team. The team undertakes operational debugging efforts for multi-domain paths. It's labor intensive. Knowledgeable engineers interpret the data. Need an operational component of some sort. Important points:

- To start a U.S. team, we can't do things exactly the same way that Geant does things. They have a more hierarchical org structure, and they are more centrally funded.
- However, we can work together and get a rotating on call person to help with multi domain issues.
- Lesson learned from the PERT team: At first, they had the responsibilities rotating thru member countries. That was not a successful model. Need a system where group that opens the ticket sticks with it.
- Possibliity of getting NSF or DOE funding.
- There are not a lot of people with experience for analyzing the longer latency paths.
- Physics organizations already have people on staffs dealing with this. ESnet has about 3-4 engineers focusing on performance problems. Smaller scientific groups don't have the experience. We could address needs there.
- Much of the community doesn't realize the bad performance is not acceptable. We should get folks educated on expectations and get them to complain if they don't get it.

Q: Is PERT team only engaged for troubleshooting once already in operation? Or are they also involved in design?

A: in Europe, PERT focuses on troubleshooting, but we can design the U.S. team how we want.

Advantage of developing a U.S. Team: Spread the knowledge around on how to do this to more of the community. It's not just about solving the problem it's also about spreading the knowledge. Another goal is to put the knowledge into a process, make checklists, create a knowledge base, use techniques to make the effort scale. GEANT is making a knowledge base. Also, we can help end users gather info necessary to start analyzing problems.

Anyone interested in working on defining this team, please send Jeff or Carla an email.

WG Charter

Carla presented the draft WG Charter, and invited comments. Carla would like volunteers to serve with her as a co-chair of the working group.

Next Performance Working Group Call

The next call will be scheduled for June, 2009. Stay tuned for details.