pS Performance Toolkit – No More Excuses for Poor Network Performance
Overview

- Intro
- Special Thanks
- Historical Perspective
- ‘How’ and ‘Why’ Regarding Performance Problems
  - Don’t Be a Victim
- What’s in the Box?
  - Tools – New and Old
  - Location Knowledge and Discovery
  - Testing Configuration
- What’s Next?
- Performance WG Discussion – Testing Parameters
Intro

• pS Performance Toolkit (v3.1)
  – First RCs available July (Joint Techs)
  – 2 Months of testing (4 RCs, ~60 Bugs found and resolved!)
  – Final release was September 25th

• Platform
  – Bootable Linux CD (Knoppix Distribution)
  – x86 Architecture (64 Bit too – no special release)

• Purpose
  – Encourage the adoption of performance tools by lowering the bar
  – Installation time: ~5 minutes (Download/Burn a CD)
  – Configuration time: ~10 minutes (Console and Web interfaces)
  – No building or installing of individual software packages required
Special Thanks

- Release was only possible due to diligence of all testers
  - ESnet
  - Fermilab
  - Georgia Tech
  - Indiana University (GRNOC)
  - Lawrence Berkeley Laboratory
  - Mid Atlantic Crossroads
  - MCNC/NCREN
  - National Energy Research Scientific Computing Center (NERSC)
  - Pacific Northwest National Laboratory
  - The Pennsylvania State University
  - The REDDnet Project
  - SLAC
  - University of Delaware
  - USAtlas
    - Brookhaven National Laboratory
    - University of Michigan
    - Michigan State University
    - University of Oklahoma
    - Vanderbilt University (ACCRE)
Network Performance Workshop

- Presentations geared toward potential performance problems and troubleshooting; using useful tools
- Old Format
  - Spend ½ the time installing/configuring the tools
  - Remainder of time learning how/why to use

1st Generation: Network Performance Toolkit (NPToolkit)

- Needed an easier way (and more productive use of workshop time) to introduce tools
- Idea: Pass out a Live CD (Knoppix) with the tools pre-installed
- New Format:
  - ½ the learning about networking problems
  - ½ the time solving them on a ‘broken’ network.
Network Performance Problems

• Not new to this group, See Joe’s Slides for an excellent primer (e.g. ‘Soft Failures’)
  – Campus/Laboratory/Institution infrastructure is not capable of driving large flows – ‘Protect the Windows boxes’
  – Staff Mentality – WAN problems are not for the LAN team to fix
  – Configuration (Routers, Switches, etc.)

• Why does it matter?
  – Large science flows are no longer an exception to the rule
  – LHC of course - worry about other projects too...

• The pSPT is not the complete solution (but it’s a part of it)
  – Adoption and installation of the tools is key to solving end-to-end problems
    • ‘chicken in every pot’
  – Educating the staff: the tools really can help and they are out there
  – Education the users that *THEY SHOULD COMPLAIN*
Quick Case Study – REDDnet

- Distributed Data Storage (LHC Related)
- ~10 Locations (most at major institutions)
- Knew that data movement applications *could* drive the network hard – observed performance did not back this up
- Wanted to know performance between locations – turned to perfSONAR
  - Installed BWCTP, NDT, OWAMP, and perfSONAR-BUOY at each site
  - Let run to get a ‘baseline’
  - Started to see particular patterns – no coincidence that the tools validated the observed data movement behavior
  - Worked with Internt2 engineers to identify problems and possible solutions
It Could Happen to You ...

- Typical performance graph (Caveat – would you complain?):

TCP Throughput

Between Texas Advanced Computing Center BWCTL (Austin TX, USA) and University of Florida BWCTL (Gainsville FL, USA)
Another interesting pattern:

**TCP Throughput**

Between San Diego Supercomputer Center BWCTL (La Jolla CA, USA) and University of Florida BWCTL (Gainsville FL, USA)
It Could Happen to You ...

• Possible Problems
  – Infrastructure – Questionable equipment, interesting network decisions (i.e. don’t place sophisticated network equipment in a Departmental closet)
  – Protection – Shapers, Firewalls. Large science flows do not get along with these.
  – Configuration – or ‘you have some tiny buffers’
  – Host Tuning
  – Applications

• Solutions
  – Tune the hosts, configure (replace) the routers.
  – Work with Campus networking teams – segment by segment
  – Adjust application behavior (after it was shown the network could be fixed...)
  – Suggest ‘permanent’ deployment of tools and regular monitoring
It Could Happen to You ...

- Better looking graph (closer to expectations):
It Could Happen to You ...

- Another

TCP Throughput

Between California Institute of Technology BMCTL (Pasadena CA, USA) and University of California, Santa Barbara BMCTL (Santa Barbara CA,
What’s In the Box?

• Performance Tools
  – BWCTL – Bandwidth Tester (Iperf/Nuttcp)
  – NDT/NPAD – Browser-based applications to diagnose performance
  – OWAMP (and JOWAMP) – One way ping, now with browser based client
  – perfSONAR PS
    • Lookup Service – Location and Discovery tool (more later…)
    • perfSONAR-BUOY – Framework for regular BWCTL and OWAMP tests
    • PingER – Round trip latency and statistics
    • SNMP MA (and Cacti) – SNMP monitoring
  – Reverse Traceroute/Ping – Traceroute/Ping from the toolkit to your browser.
• System Tools
  – Apache2
  – MySQL
  – Syslog-ng
What’s In the Box?

• Configuration Framework
  – Console-based ‘basic’ configuration
    • Answer questions regarding networking, storage
    • Or don’t configure at all! [Machine becomes a basic testing beacon]
  – Web-based ‘enhanced’ configuration
    • Personalized settings – Where the Toolkit is located and how to identify it (for Discovery)
    • Enabling Regular Testing
      – OWAMP, BWCTL, and PingER
      – Set Measurement Parameters (e.g. test length)
      – Set Hosts to Test with (Discover hosts or manually enter)
      – Step-by-step setup
    • Data Visualization – See the results of regular tests
What’s In the Box?

• Location and Discovery
  – Solving end-to-end problems relies on knowledge of the path
  – Testing with available tools on the path (when available)
  – perfSONAR-PS Solutions
    • Home Lookup Service (hLS)
    • Global Lookup Service (gLS)
    • Intelligent GUIs to take advantage of the two
What’s In (outside) the Box?

Global Lookup Service

- pSPT 1
  - Domain A:
    - OWAMP
    - BWCTP

- pSPT 2
  - Domain B:
    - SNMP

- SNMP Data for Domain B
- Contact pSPT 2

- GUI
Each node comes with an hLS
Each node registers information about itself
  – Where it Is, what domain it may answer for
  – Information on what tests it may be running (Type, Test Points)
Each hLS will ‘communicate’ with the Global Discovery Framework
Ability to search for
  – Measurement Services that match a specific Domain or IP Range
  – Data Type (e.g. throughput measurements)
  – Items that match a specific ‘keyword’
    • ‘What’s a Keyword and why is that important?’
• Communities = Web 2.0 Content Tagging
  – Think Flicker (tag your pictures with a category)
  – Think iTunes (tag your music with a genre)
• How does this help measurement discovery?
  – One more axis to search on
  – More human readable and understandable than IP address or hostnames
• Use as many (or as few) as required:
  – Networks (e.g. Campus, Regional, Network)
  – VO or Project (e.g. USATLAS, eVLBI, etc.)
  – Organization (DOE)
  – Other?
Example: Some VO is setting up monitoring.
- All sites want to test with each other
- Not everyone is coming online at once, and VO membership may be volatile.

Strategy 1:
- Central VO coordinator maintains a list of participants (and must update it often)
- All monitoring is manual: add/remove test hosts when the list changes

Strategy 2:
- VO recommends a tag for all new hosts
- All VO members search for test hosts (periodically) that share this tag – N.B. the GUIs on the disk can organize this automatically
- Screenshot from the toolkit (when setting up the host):

<table>
<thead>
<tr>
<th>Communities[^1] This Host Participates In</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet2</td>
</tr>
<tr>
<td>perfSONAR–PS</td>
</tr>
<tr>
<td>Add New Community</td>
</tr>
</tbody>
</table>

**Popular Communities As Of 2009-09-22 08:02 (Click To Join)**

| Atlas  CMS  DOE  DOE Sites  DOE–SC–LAB  ESnet  GRNOC  Internet2–CTP  KREONET  LHC  RNP Sites  USATLAS  Utah |

- Top: Communities the host has chosen to associate with
- Bottom: ‘Popular’ communities
  - The word cloud is based on what we found in the GLS – the larger the word = the more people that are using this classification
Lookup and Discovery - Communities

- Adding hosts via the scheduled testing screen
  - Manually add
  - See what hosts are available for a particular community

Test Members
No Members In Test

Add New Host
Find Hosts To Test With

Communities This Host Participates In (Click To Find Community Hosts)
- Internet2 perfSONAR-PS

Popular Communities As Of 2009-09-22 08:02 (Click To Find Community Hosts)
- Atlas CMS DOE DOE Sites DOE-SC-LAB ESnet GRNOC Internet2 CTP KREONET LHC RNP Sites USATLAS Utah
We are interested in who is involved in the LHC community:

<table>
<thead>
<tr>
<th>Test Members</th>
<th>No Members In Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add New Host</td>
<td>Looking up hosts from community LHC...</td>
</tr>
<tr>
<td>Find Hosts To Test With</td>
<td>Communities This Host Participates In (Click To Find Community Hosts)</td>
</tr>
<tr>
<td>Popular Communities As Of 2009-09-22 08:02 (Click To Find Community Hosts)</td>
<td>Internet2 perfSONAR-PS</td>
</tr>
</tbody>
</table>

- Atlas
- CMS
- DOE
- DOE Sites
- DOE-SC-LAB
- ESnet
- GRNOC
- Internet2_CTP
- KREONET
- LHC
- RNP
- Sites
- USATLAS
- Utah
List of hosts from the LHC community:

<table>
<thead>
<tr>
<th>Members Of LHC Community As Of 2009-09-22 08:02</th>
</tr>
</thead>
<tbody>
<tr>
<td>BWCTL Server at HEP, University of Pennsylvania in Philadelphia, PA, USA</td>
</tr>
<tr>
<td>i2perf.hep.upenn.edu(128.91.45.144)</td>
</tr>
<tr>
<td>BWCTL Server at Internet2 in Washington, D.C., USA</td>
</tr>
<tr>
<td>nms-rthr2-eth2.wash.net.internet2.edu(64.57.16.22)</td>
</tr>
<tr>
<td>BWCTL Server at Internet2 in Kansas City, MO, USA</td>
</tr>
<tr>
<td>eth-2.nms-rthr2.kans.net.internet2.edu(64.57.16.214)</td>
</tr>
<tr>
<td>BWCTL Server at Internet2 in Houston, TX, USA</td>
</tr>
<tr>
<td>eth-3.nms-rthr2.hous.net.internet2.edu(64.57.16.131)</td>
</tr>
<tr>
<td>BWCTL Server at Internet2 in New York, NY, USA</td>
</tr>
<tr>
<td>nms-rthr2.newy32aoa.net.internet2.edu(64.57.17.66)</td>
</tr>
</tbody>
</table>
What’s Next?

• Security and Bugs will drive the next releases
• Migration to Fedora/RHEL Platform
  – 2010 Time Frame
  – Planned backwards compatibility
• Additional Tools
  – Smokeping
  – Nagios
  – ESxSNMP
• Customization
  – Add branding/style for your deployment
Discussion – Testing Parameters

- pSPT “defaults” heavily influenced by LHC Requirements
  - Throughput (perfSONAR-BUOY)
    - 20 Second tests, every 4 Hours
    - Tier2 to Tier2, Tier2 to Tier1 (USAtlas)
  - RT Latency (PingER)
    - 10 Packets with 1 Second inter-packet gap
    - Every 5 Minutes
  - OW Latency (perfSONAR-BUOY)
    - Continuous stream of 10 Packets Per Second
    - Tier2 to Tier2, Tier2 to Tier1 (USAtlas)
Discussion – Testing Parameters

• Problem
  – What is best for each? Should the pSPT use recommendations from the PerfWG?
  – If so, what should we base these on
    • Experience
    • Best performance vs. least impact on networks

• Open questions:
  – Throughput length vs frequency
  – Throughput value of UDP
  – Latency BCP (length, frequency)
  – Whom to test with (within the VO yes, but where on the outside)
pS Performance Toolkit – No More Excuses for Poor Network Performance

October 5, 2009, Internet2 Fall Member Meeting
J. Zurawski, Network Software Engineer

For more information, visit psp.s.perfsonar.net/toolkit