

March 10th 2011, OSG All Hands Workshop - Network Performance Jason Zurawski, Internet2

Performance Measurement Tools

Agenda

Tutorial Agenda:

- Network Performance Primer Why Should We Care? (15 Mins)
- Getting the Tools (10 Mins)
- Use of the BWCTL Server and Client (30 Mins)
- Use of the OWAMP Server and Client (30 Mins)
- Use of the NDT Server and Client (30 Mins)
- Diagnostics vs Regular Monitoring (30 Mins)
- Network Performance Exercises (1 hr 30 Mins)





Basic Premise

- Use of the network should meet user expectations.
 - See Examples from earlier
 - When things don't look right, time to explore 'why'
- If they don't you should complain!
 - Learn how to do so effectively
 - Produce results
 - Isolate problem
 - The network is slow is not how to complain effectively





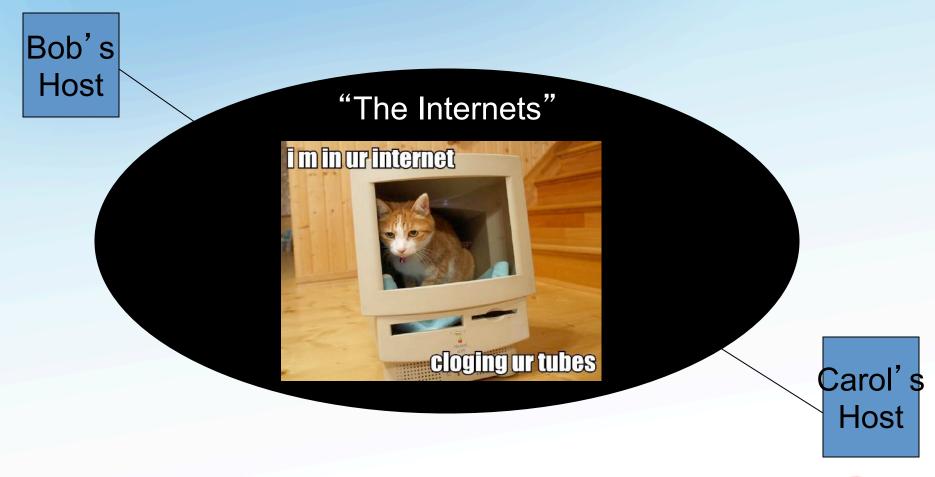
Underlying Assumption

- When problems exist, it's the networks fault!
 - Easy to blame a resource, but where else could a problem be when transferring large data sets?
 - Host (Disk, CPU, Kernel, NIC Drivers)
 - Network Interface Cards
 - Routers/Switches, Routing and Configuration
 - Physical Infrastructure
 - Protocols
- The network is viewed as a single resource in many cases
 - Reality complex series of components
 - Multiple vendors/technologies
 - Multiple configuration options
 - Crossing administrative domains





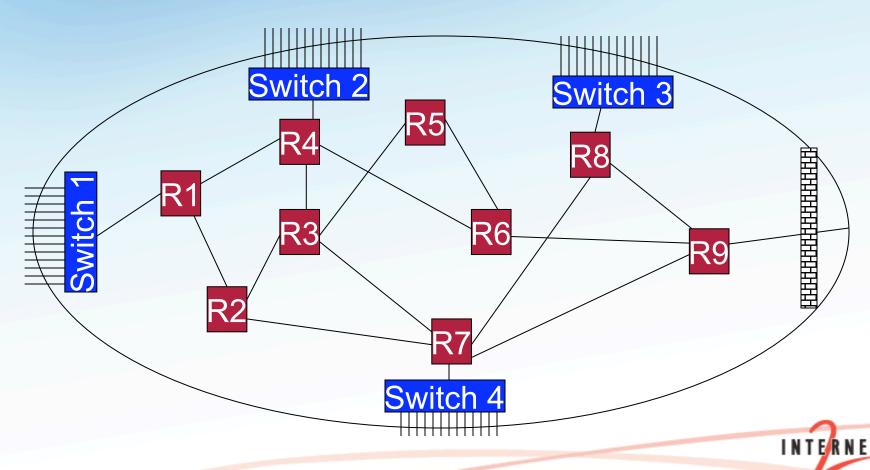
Network View (Layman's Terms)







Network View (Actual)





Addressing a Performance Discrepancy

- What are the first steps to address problems related to network performance?
 - Try a Tool
- What tools are out there
 - Numerous
 - Different metrics (measurements) available
 - How to interpret the results?





Tools, Tools

- Ping
- Traceroute
- Iperf
- Tcpdump
- Tcptrace
- BWCTL
- NDT
- OWAMP
- AMP
- Advisor
- Thrulay
- Web100
- MonaLisa
- pathchar
- NPAD
- Pathdiag
- Surveyor

- Ethereal
- CoralReef
- MRTG
- Skitter
- Cflowd
- Cricket
- Net100
- Pathload
- Pathchrip
- MRTG
- Cacti
- Smokeping
- PingER
- FDT
- perfSONAR
- Nagios
- Ganglia
- Thurlay
- Etc. etc. etc.





Highlighting some Interesting Tools

- Focus on 3 Types of tools (for now)
 - Basic Diagnostics
 - Ping, Traceroute
 - Advanced User Tools
 - NDT
 - Network Admin Focused
 - OWAMP, BWCTL
- What about the others?
 - Try them out, learn how they work.
 - Most tools are designed to solve a specific problem and they may add value to your organization
- Integration of multiple solutions
 - Measurement frameworks integrate use of tools (operation, collecting results) along with analysis and presentation
 - perfSONAR





Software Availability

- BWCTL, OWAMP, NDT client applications are available in VDT
- Source Packages (Client and Server)
 - http://software.internet2.edu/sources/
 - Typical 'configure/make/make install'
- RPM Installation (CentOS 5.5 Supported):
 - Install our RPM package to enable the Internet2 Repository
 - See instructions here: http://software.internet2.edu/
 - Support for YUM and APT-RPM
- pS Performance Toolkit ISO
 - All tools, pre-installed and configured
 - More info: http://psps.perfsonar.net/toolkit/
- Others Notes:
 - Other RPM based distros (Fedora/RHEL) may work with packaged RPMs ... YMMV
 - To install on Debian, consider source. Alien conversions of RPMs may be problematic





Basic Diagnostic Tools

Ping

- Round Trip (e.g. source to destination, and back)
- Confirms that remote host is 'up'
- Some network operators block these packets
 - Play w/ command options to see if that will change anything

Traceroute

- Identifies the routers along the path
- Same blocking problem as above
- Routers treat TR packets with lower priority
 - See presentation from prior JTs:
 - http://www.internet2.edu/presentations/jt2009jul/20090722litvanyi.pdf





Ping Output

```
zurawski@latrobe:~ - ssh - ttys001 - 85×25
[zurawski@latrobe ~] $ ping -c 4 packrat.internet2.edu
PING packrat.internet2.edu (207.75.164.10) 56(84) bytes of data.
64 bytes from packrat.internet2.edu (207.75.164.10): icmp seq=1 ttl=57 time=16.2 ms
64 bytes from packrat.internet2.edu (207.75.164.10): icmp seq=2 ttl=57 time=16.4 ms
64 bytes from packrat.internet2.edu (207.75.164.10); icmp seq=3 ttl=57 time=16.5 ms
64 bytes from packrat.internet2.edu (207.75.164.10); icmp seg=4 ttl=57 time=16.4 ms
--- packrat.internet2.edu ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3000ms
rtt min/avg/max/mdev = 16.225/16.423/16.594/0.204 ms
[zurawski@latrobe ~]$
```





Ping Output

- Intro message
 - Identifies remote host name and IP address
 - States size of packets being sent
 - Setting larger sizes may reveal hidden problems
- Output lines
 - Who responded, and the RTT, maybe other details
- Summary lines
 - Number of packets sent/received/lost
 - RTT statistics min/average/max

N.B. 1 msec RTT = \sim 50 miles of between hosts





Traceroute Output

```
zurawski@latrobe:~ - ssh - ttys001 - 88×24
[zurawski@latrobe ~] $ traceroute packrat.internet2.edu
traceroute to packrat.internet2.edu (207.75.164.10), 30 hops max, 40 byte packets
1 192.52.179.1 (192.52.179.1) 0.445 ms 0.427 ms 0.388 ms
   clpk-ucaid-gige.maxgigapop.net (206.196.177.29) 1.772 ms 1.737 ms 1.737 ms
  xe-7-2-0-0.lvl3-t640.maxgigapop.net (206.196.178.90) 2.954 ms 2.969 ms 2.936 ms
4 i2-lvl3.maxgigapop.net (206.196.178.46) 2.841 ms 2.812 ms 2.918 ms
5 xe-0-1-0x1004.wsu5.mich.net (192.122.183.9) 14.723 ms 14.567 ms 14.623 ms
6 tenge0-0-0-0x22.aa2.mich.net (198.108.23.49) 16.617 ms tenge0-0-0-0x76.aa2.mich.net
 (198.108.23.10) 16.617 ms tenge0-0-0-0x22.aa2.mich.net (198.108.23.49) 16.724 ms
7 mam-77.merit.edu (192.122.200.77) 16.492 ms 16.497 ms 16.603 ms
8 packrat.internet2.edu (207.75.164.10) 16.324 ms 16.292 ms 16.282 ms
[zurawski@latrobe ~1$
```





Traceroute Output

- Intro messages
 - Name and address of remote host
 - Maximum number of link before giving up
- Status messages
 - One line per router in path
 - '*' indicates router didn't respond
 - Routers usually rate limit replies
 - No name indicates DNS entry is missing
 - Hops required to reach remote host or max number from above





Advanced User Tool - NDT

- NDT = Network Diagnostics Tool
- Measure performance to users desktop
- Identify real problems for real users
 - Network infrastructure could be the problem
 - Host tuning issues could be the problem
- Make tool simple to use and understand
 - Presentation in a method almost all users can access: web browser
- Make tool useful for users and network administrators
 - Variables for many aspects of host, protocol, and network performance





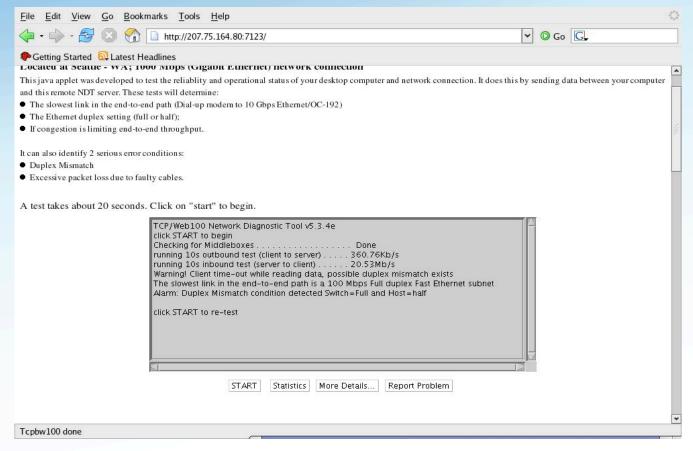
NDT User Interface

- Web-based JAVA applet allows testing from any browser
 - One Click testing
 - Option to dig deep into available results
 - Send report of results to network administrators
- Command-line client allows testing from remote login shell
 - Same options available
 - Client software can be build independent of server software





NDT Results







Finding Results of Interest

Duplex Mismatch

 This is a serious error and <u>nothing</u> will work right. Reported on main page, on Statistics page, and mismatch: on More Details page

Packet Arrival Order

- Inferred value based on TCP operation. Reported on Statistics page, (with loss statistics) and order: value on More Details page
- Packet Loss Rates
 - Calculated value based on TCP operation. Reported on Statistics page, (with out-of-order statistics) and loss: value on More Details page
- Path Bottleneck Capacity
 - Measured value based on TCP operation. Reported on main page





Installation

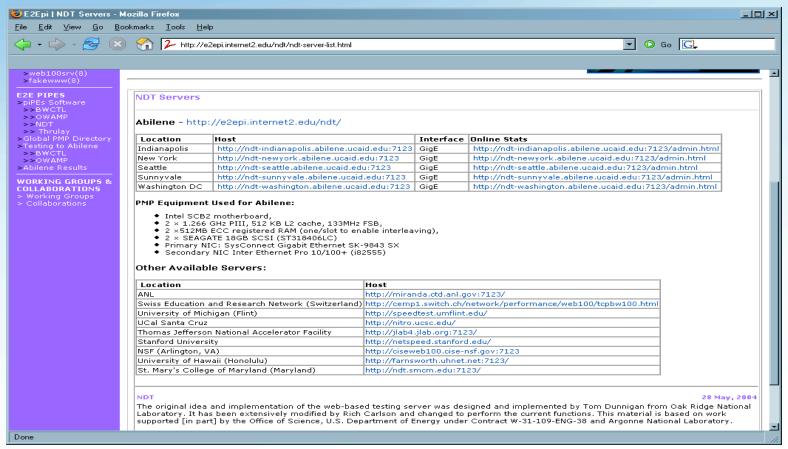
- Software Dependencies
 - Web100 Patched Kernel
 - Traditionally not a 'simple' operation to maintain
 - Questions on security when patching a kernel also may not cleanly apply to a 'vendor' kernel
 - NDT Client/Server Package
- Software available as:
 - Source
 - RPM for Web100 Kernel and NDT
 - pS Performance Toolkit
 - Web100 patched kernel, all measurement tools on a Live Linux CD





Finding a Server – The Old Way

Static List of servers – doesn't scale

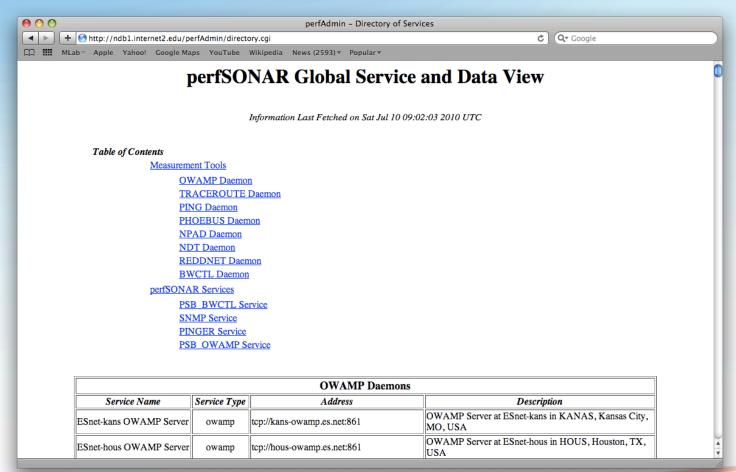






Finding a Server – The New Way

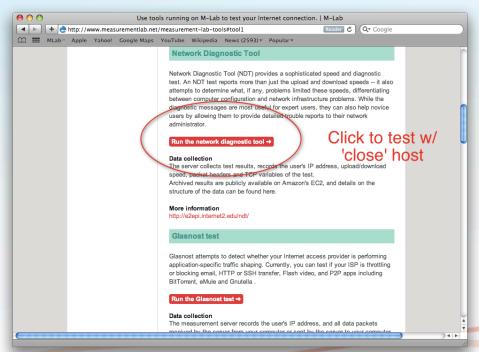
perfSONAR Infrastructure – automatically search for instances





Finding a Server – MLab

- Measurement Lab
 - Joint Project between several partners
 - More Info Here: http://www.measurementlab.net/
- Locate a 'close' NDT server using DONAR (http://donardns.org/)







Network Administration Tools

- BWCTL Bandwidth Control
 - Allows single person operation over wide area testing environment
 - Runs NLANR 'iperf' program
 - Support for Thrulay, nuttcp
- OWAMP One way Delay Measurement
 - Advanced 'ping' command
 - One way vs round trip
 - Allows single person operation over wide area testing environment





BWCTL Sample Results

```
👼 nmsx.internet2 - SecureCRT
                                                                                       _ | _ | ×
<u>File Edit View Options Transfer Script Tools Window Help</u>
[rcarlson@nmsx-aami rcarlson]$ bwctl -L90 -i2 -t20 -w8388608 -A AE AESKEY rcarlson .aeskey
 -s nms1-sttl.abilene.ucaid.edu -c nms1-ipls.abilene.ucaid.edu
bwctl: 107 seconds until test results available
RECEIVER START
3349536592.521172: /ami/bin/iperf -B 198.32.8.174 -P 1 -s -f b -m -p 5002 -w 8388608 -t 20
Server listening on TCP port 5002
Binding to local address 198.32.8.174
TCP window size: 16777216 Byte (WARNING: requested 8388608 Byte)
[ 14] local 198.32.8.174 port 5002 connected with 198.32.8.190 port 5002
      0.0-2.0 sec 203240632 Butes 812962528 bits/sec
      2.0- 4.0 sec 247453364 Butes 989813456 bits/sec
Г 147
      4.0-6.0 sec 247453276 Bytes 989813104 bits/sec
  147
      6.0-8.0 sec 247392040 Bytes 989568160 bits/sec
  147
      8.0-10.0 sec 247454640 Bytes 989818560 bits/sec
  14]
 14] 10.0-12.0 sec 247451404 Bytes 989805616 bits/sec
[ 14] 12.0-14.0 sec 247447240 Butes 989788960 bits/sec
[ 14] 14.0-16.0 sec 247459968 Bytes 989839872 bits/sec
[ 14] 16.0-18.0 sec 247445468 Bytes 989781872 bits/sec
[ 14] 18.0-20.0 sec 247453536 Bytes 989814144 bits/sec
[ 14] 0.0-20.0 sec 2435407872 Bytes 972137651 bits/sec
[ 14] MSS size 8948 bytes (MTU 8988 bytes, unknown interface)
RECEIVER END
[rcarlson@nmsx-aami rcarlson]$
Ready
                                                        ssh2: AES-12E 26, 32 28 Rows, 90 Cols VT100
```



BWCTL Highlights

- You must pre-install BWCTL software package
 - Server/Client required for ends of a test
- Internet2 default permission: TCP test from any member
 - UDP available on request
- Sites can restrict access to 'known' remote users
 - Rich set of permissions and limitations





BWCTL Commands

bwctl -L90 -i2 -t20 -c bwctl.losa.net.internet2.edu

bwctl -L90 -i2 -t20 -s bwctl.newy.net.internet2.edu

- Bwctl = name of program
- L90 = wait up to 90 seconds for a test
- i2 = report intermediate results every 2 seconds
- t20 run test for 20 seconds
- s name = remote end will send data to you
- c name = you will send data to the remote host





BWCTL 3rd Party Testing

bwctl -L90 -i2 -t20

- -c bwctl.salt.net.internet2.edu
- -s bwctl.atla.net.internet2.edu

User can run a test between 2 remote hosts





Installation

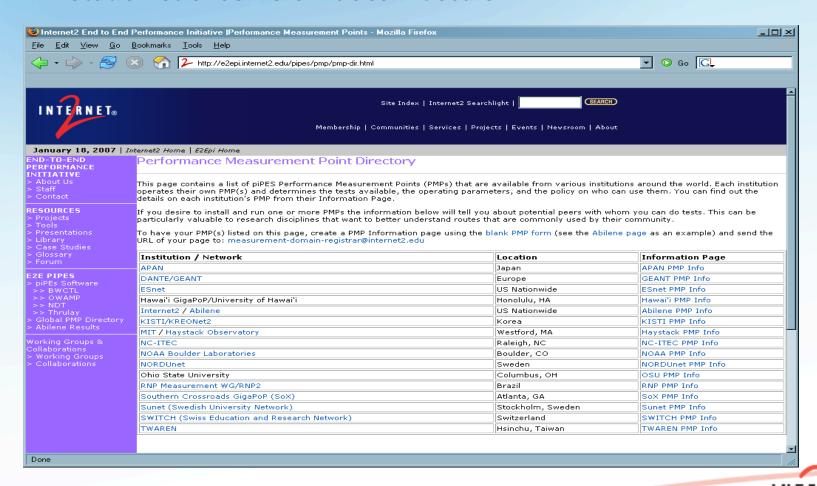
- Software available as:
 - Source
 - RPM Installable through YUM/APT-RPM
 - pS Performance Toolkit
 - Web100 patched kernel, all measurement tools on a Live Linux CD





Finding a Server – The Old Way

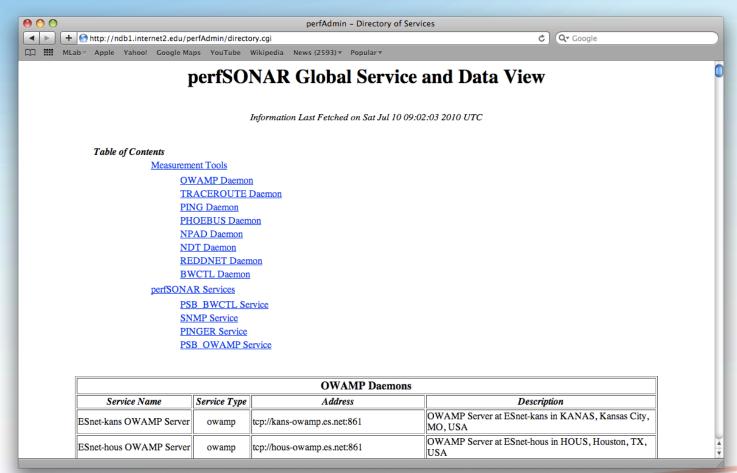
Static List of servers – doesn't scale





Finding a Server – The New Way

perfSONAR Infrastructure – automatically search for instances





OWPing Results

```
👼 nmsy.internet2 - SecureCRT
                                                                                             _ | _ | × |
File Edit View Options Transfer Script Tools Window Help
rcarlson@nmsy-aami$ ./owping -F /tmp/from.owamp -T /tmp/to.owamp nms4-ipls.abilene.ucaid.edu 🗖
Approximately 12.8 seconds until results available
--- owping statistics from [207.75.164.84]:61960 to [nms4-ipls.abilene.ucaid.edu]:57501 ---
SID: c6200ae4c7a5e3dc5ffc50773bbed85a
100 sent, 0 lost (0.0%), 0 duplicates
one-way delay min/median/max = 5.04/5.1/5.2 ms, (err=5.2 ms)
one-way jitter = 0.1 ms (P95-P50)
Hops = 5 (consistently)
no reordering
--- owping statistics from [nms4-ipls.abilene.ucaid.edu]:63467 to [207.75.164.84]:61887
SID: cf4ba454c7a5e3dc67124c14947f2e5d
100 sent, 0 lost (0.0%), 0 duplicates
one-way delay min/median/max = 5.14/5.2/5.35 ms, (err=5.35 ms)
one-way jitter = 0.1 \text{ ms} (P95-P50)
Hops = 5 (consistently)
no reordering
rcarlson@nmsy-aami$ ./owstats -a99 /tmp/to.owamp
--- owping statistics from [nmsy-aami-abilene.internet2.edu]:61960 to [nms4-ipls.abilene.uca
id.edul:57501 ---
SID: c6200ae4c7a5e3dc5ffc50773bbed85a
100 sent, 0 lost (0.0%), 0 duplicates
one-way delay min/median/max = 5.04/5.1/5.2 ms, (err=5.2 ms)
one-way jitter = 0.1 \text{ ms} (P95-P50)
Percentiles:
        99.0: 5.2 ms
Hops = 5 (consistently)
                                                             ssh2: AES-128 31, 21 31 Rows, 93 Cols VT100
Ready
```



OWAMP Results

- You must pre-install OWAMP software package
- User program is called 'owping'
- Internet2 default permissions: basic test from any member
 - Other options (e.g. longer test, more data sent) available on request
- Sites can restrict access to 'known' remote users
 - Similar rich set of permissions to BWCTL





Using OWPing

- owping owamp.salt.net.internet2.edu
 - owping = program name
 - owamp... = name of server
- Output results
 - Separate statistics for both directions
 - Number of packets sent and lost
 - One-way delay statistics min/median/max
 - Number of IP hops in path
 - Number of packets that arrives out-of-order





Installation

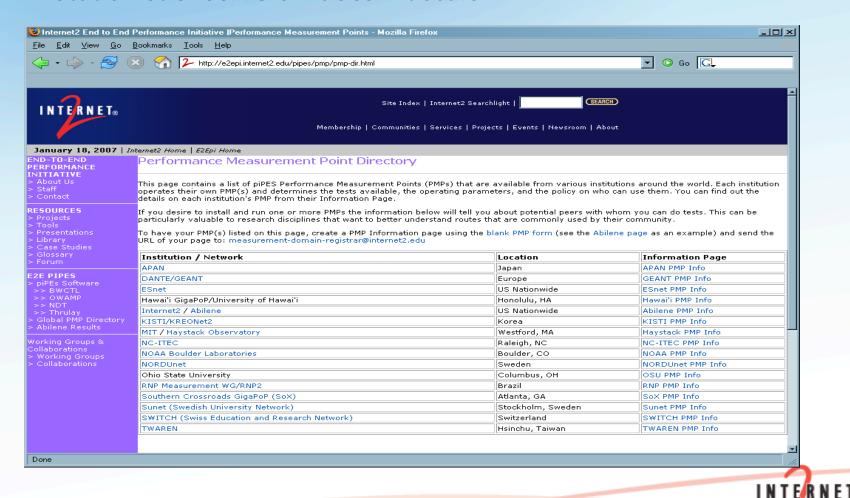
- Software available as:
 - Source
 - RPM Installable through YUM/APT-RPM
 - pS Performance Toolkit
 - Web100 patched kernel, all measurement tools on a Live Linux CD





Finding a Server – The Old Way

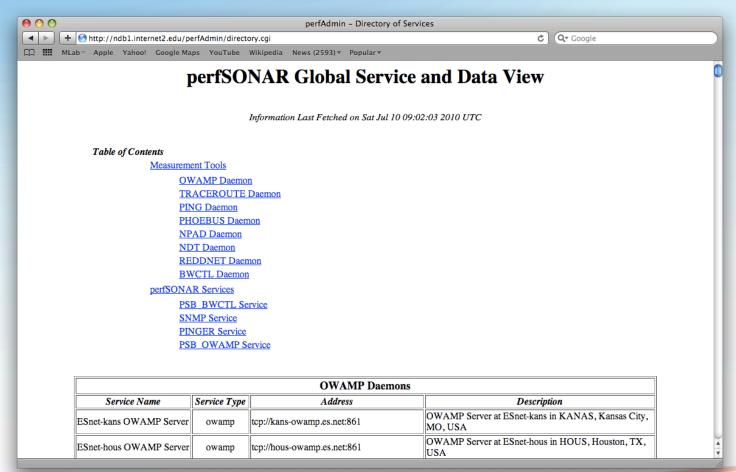
Static List of servers – doesn't scale





Finding a Server – The New Way

perfSONAR Infrastructure – automatically search for instances





Conclusions

- Primary tools still useful
 - Know the tool
- Advanced tools are being developed all the time
- Installation using simple methods (e.g. VDT, RPM based package managers)







Performance Measurement Tools

March 10th 2011, OSG All Hands Workshop – Network Performance Jason Zurawski – Internet2

For more information, visit http://www.internet2.edu/workshops/npw

