



E2E Performance Tools: Internet2 Performance Architecture and Technologies Update

Eric L. Boyd

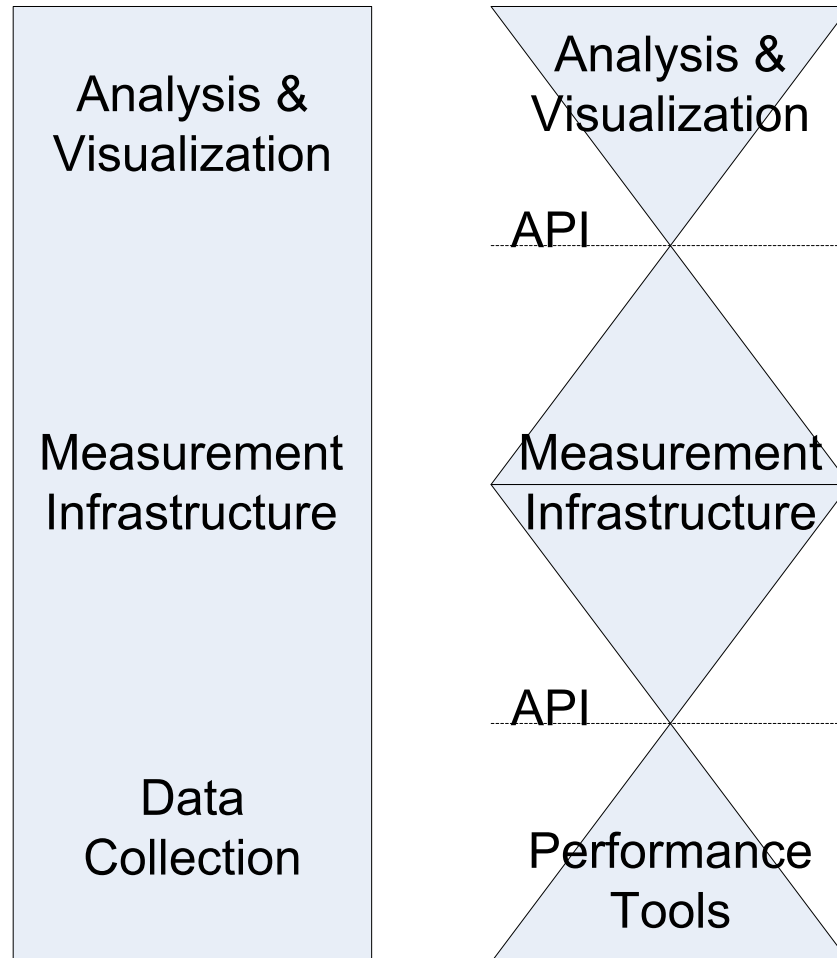
Director of Performance Architecture and
Technologies

Internet2

Performance Goals

- Enable end-users & network operators to:
 - determine E2E performance capabilities
 - locate E2E problems
 - contact the right person to get an E2E problem resolved.
- Enable remote initiation of partial path performance tests
- Make partial path performance data publicly available
- Be interoperable with other performance measurement frameworks
- Encompass needs of next generation infrastructures

Decouple 3 components of a Measurement Infrastructure



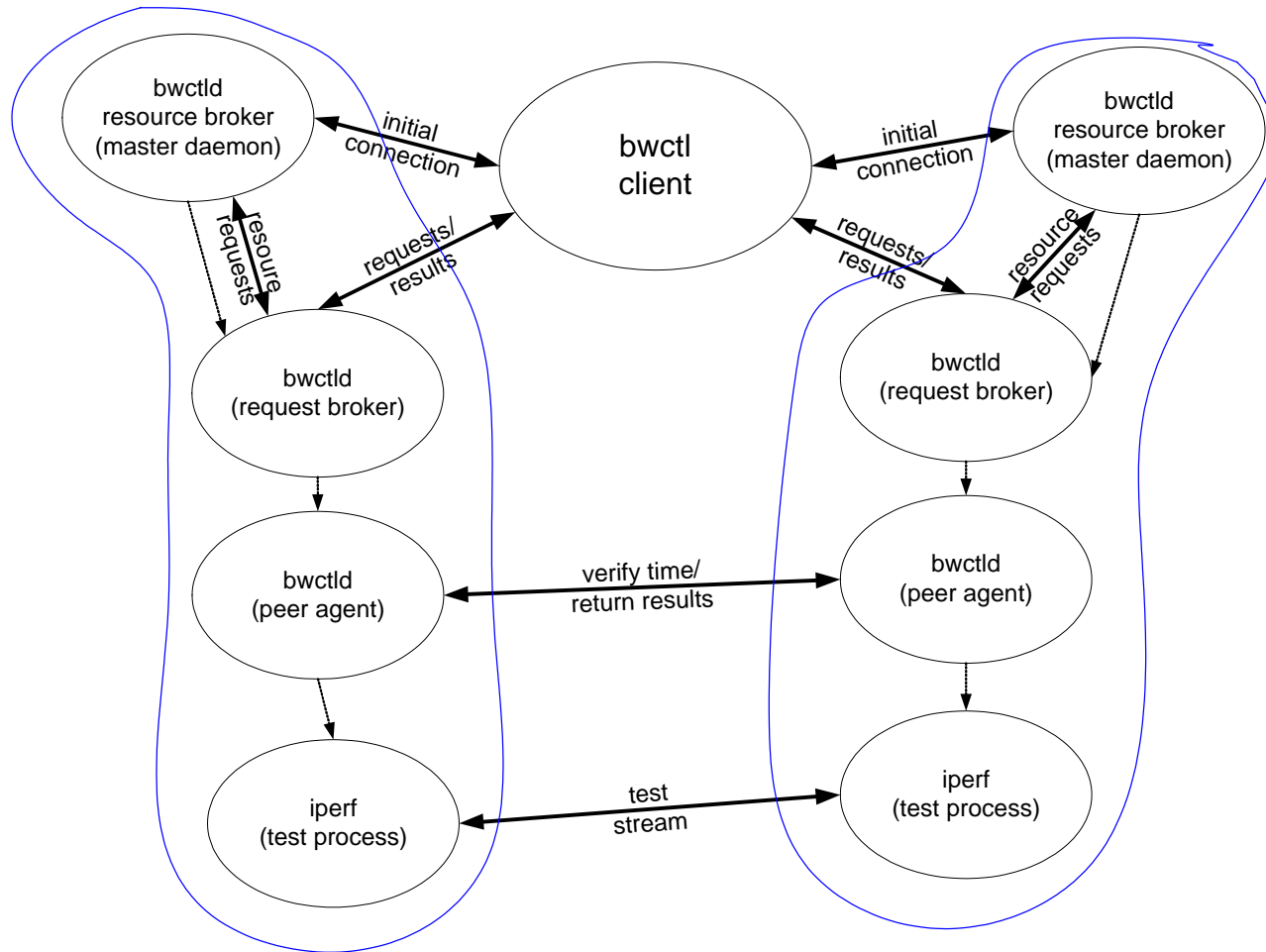
Current Projects

- Performance Tools
 - BWCTL
 - NDT
 - OWAMP
 - Thrulay
- Performance Measurement Framework
 - piPEs -> perfSONAR
 - GGF NMWG
- Member Outreach
 - Network Performance Measurement Workshops
 - Performance Tool Cookbooks
- Bulk Transport
 - Design Space
 - Prototype

BWCTL (Bandwidth Controller)

- What is it?
 - A resource allocation and scheduling daemon for arbitration of iperf tests
- Typical Solution
 - Run “iperf” or similar tool on two endpoints and hosts on intermediate paths
- Typical road blocks
 - Need permissions on all systems involved
 - Need to coordinate testing with others
 - Need to run software on both sides with specified test parameters

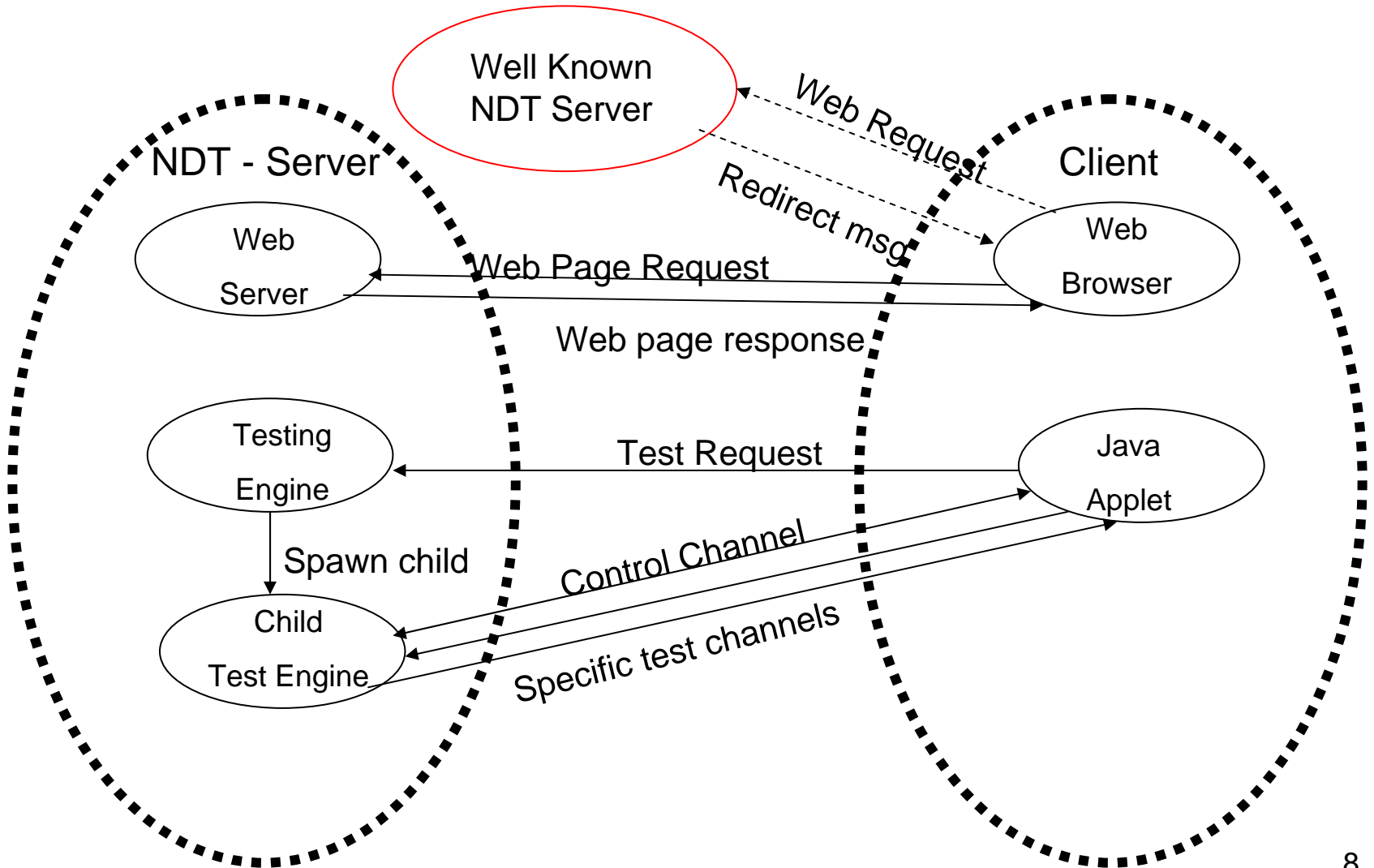
BWCTL: 3-Party Flow Diagram



NDT: Network Diagnostic Tool

- Web100 enhanced server handles testing and diagnostic services
- Java based and command line clients allows testing from any client (local or remote)
- Performance and configuration faults reported back to client
- Drill-down functions provide more details & error reporting capabilities
- Grant from NIH/NLM to explore duplex mismatch detection

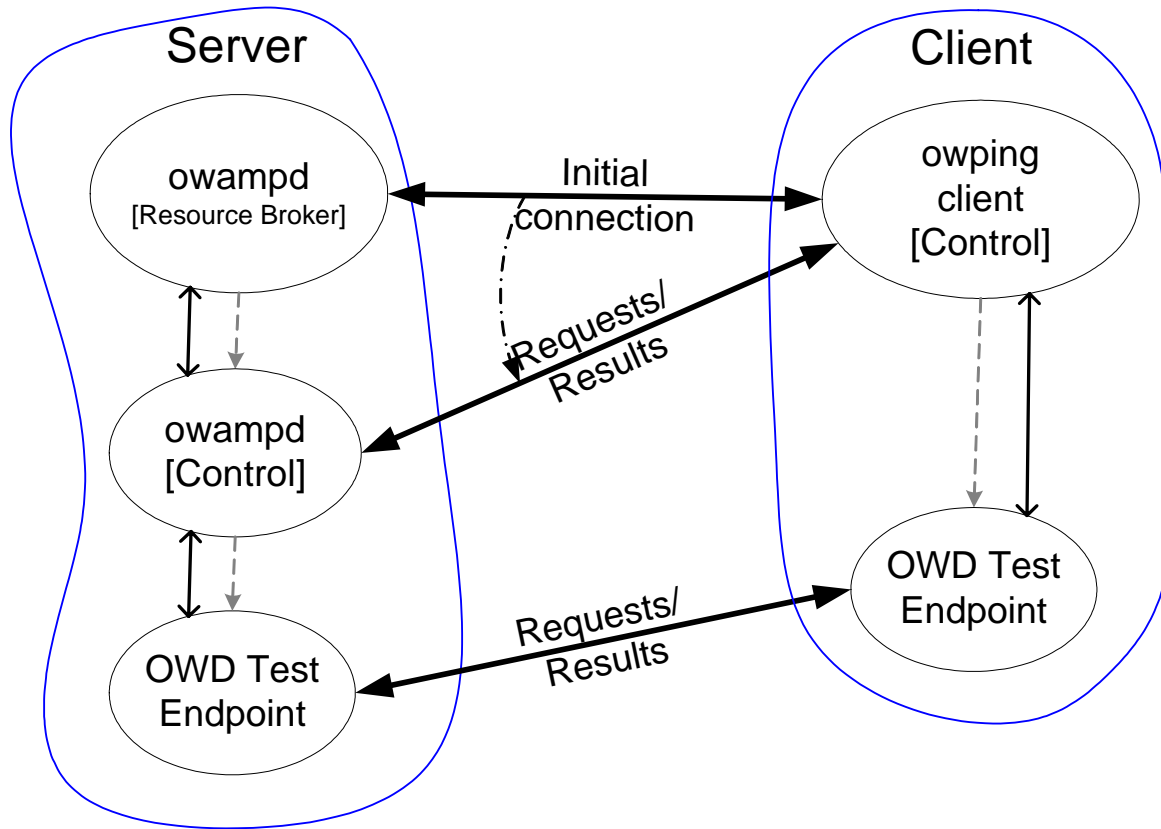
NDT Flow Diagram



OWAMP: One-Way Active Measurement Protocol

- What is it?
 - Measures one-way latency: 1-way ping
 - Control connection used to broker test request based upon policy restrictions and available resources. (Bandwidth/disk limits)
- Specification
 - <http://tools.ietf.org/wg/ippm/draft-ietf-ippm-owdp/draft-ietf-ippm-owdp-14.txt>

OWAMP Flow Diagram



Thrulay Overview

- Network capacity and delay tester
- Same class of tools as iperf, netperf, nettest, nuttcp, ttcp, etc.
- Unique features not found in other tools:
 - TCP: measures round-trip delay along with goodput
 - UDP: measures:
 - One-way delay, with quantiles
 - Packet loss
 - Packet duplication
 - Reordering
 - UDP: ability to send precisely positioned true Poisson streams (microsecond errors in sending times)
 - Human and machine-readable (ready to be fed to gnuplot)

Thrulay Update

- New release v0.8
- Tests with multiple TCP streams
- Set DSCP (a.k.a. first 6 bits of the TOS byte)
- Report MTU and/or MSS (whichever the OS makes available)
- More UDP statistics: duplication, reordering, quantiles of delay
- SPARC/Solaris support
- Mac OS X support
- IPv6 support
- Non-busy-waiting UDP mode (less precise, but can run more concurrent tests)
- Documentation: manual pages have been added
- Basic client authorization based on IP address
- Integration of TSC timekeeping projects for faster and more precise timestamping

Performance Tools in 2006 (1)

- BWCTL: Bandwidth Controller (all “nice-to-haves”)
 - Resource allocation broker for Iperf
 - Integrate Thrulay and/or library version of Iperf
 - Integrate “generic” programs (e.g. nutcp)
 - Schedule pre-circuit testing
- OWAMP: IETF Protocol
 - One-Way Active Measurement Protocol (1-way Ping)
 - Get it out of IESG (i.e. proposed standard)
 - Get an IANA port number
- OWAMP: The Tool
 - One-Way Active Measurement Protocol Reference Implementation
 - Keep it up to date with IETF Protocol
 - Release SC-inspired changes
 - Release ITEC-inspired changes
 - Improve tool statistics

Performance Tools in 2006 (2)

- NDT: Network Diagnostic Tool
 - Latest version released
 - Deployment results -> development feedback
 - Broaden number and depth of tests
- Thrulay
 - Network Capacity Tester
 - Continue to improve packaging (and ease of installation) and portability
 - Would be great to find funding to have interns help
 - Integrate into Network Performance Workshops

Current Projects

- Performance Tools
 - BWCTL
 - NDT
 - OWAMP
 - Thrulay
- Performance Measurement Framework
 - piPEs -> perfSONAR
 - GGF NMWG
- Member Outreach
 - Network Performance Measurement Workshops
 - Performance Tool Cookbooks
- Bulk Transport
 - Design Space
 - Prototype

Network Performance Measurement Workshops

- Example Course Materials:
 - <http://e2epi.internet2.edu/net-perf-wkshp/max/presentations.html>

Goals:

- Grow installed base of BWCTL/Iperf, OWAMP, and NDT at GigaPoP and regional campuses.
 - <http://e2epi.internet2.edu/pipes/pmp/pmp-dir.html>
- Begin integration into IT support processes.
- Create an installed base for perfSONAR deployment.
- Give each participant tool-specific cookbooks.

Network Performance Measurement Workshop Dates / Locations

- **Completed**
 - SOX / GaTech (03/05)
 - CENIC / UCLA (06/05)
 - JT – Vancouver (07/05)
 - OARNet / OSU (09/05)
 - MAGPI / FMM (09/05)
 - MAX / College Park (12/05)
- **Planned**
 - APAN (01/06)
 - JT - Albuquerque (02/06)
 - MERIT (02/06)
- **Under Consideration**
 - Wisconsin, Alaska, ...

Network Performance Measurement Workshops in 2006

- Goal is 9-12 workshops
- Develop / schedule “Application Measurement Workshops”
 - Target specific application communities
 - Focus on use of tools, not installation
 - Generate second source of pressure on network engineers to deploy measurement infrastructures
 - Create more “have” (vs. “have-not”) communities
- Integrate Thrulay
- Plan for Advanced Network Performance Workshops (i.e. perfSONAR workshops)

Current Projects

- Performance Tools
 - BWCTL
 - NDT
 - OWAMP
 - Thrulay
- Performance Measurement Framework
 - piPEs -> perfSONAR
 - GGF NMWG
- Member Outreach
 - Network Performance Measurement Workshops
 - Performance Tool Cookbooks
- Bulk Transport
 - Design Space
 - Prototype

Bulk Transport Tool (1)

- Killer App for High Performance Networks (i.e. why else do we need fat pipes)
- Remedies for TCP's maladies
 - Tuning: buffers, window scaling, timestamps, SACK
 - Use multiple streams
 - **Something Else**
 - Replace the kernel and use different congestion control
 - Replace all the routers and kernels

Bulk Transport Tool (2)

- Many alternative TCP/IP congestion control algorithms
- Modified kernels are incompatible with regular kernel security patches
- Get the benefits of kernel-level modifications to TCP/IP congestion control algorithms in a user level tool, avoiding security issue with alternate kernels

Bulk Transport Tool (3)

- Design Space Document
- Early Stage Prototype

Google Summer of Code

- Google “sponsored” many students to work on open-source projects summer ‘05.
 - Internet2 mentored 10 students.
<http://transport.internet2.edu/student-projects.html>.
 - One student continues as an intern
- Projects:
 - Timekeeping using TSC register - timestamp fetching without a context switch and relating the TSC value to UTC.
 - Noise calibration - data analysis of noise in delays for packet measurements and development of filtering algorithms.
 - Thrulay enhancements
 - Bulk Transport API over UDT
 - Rich Presence Project



ESnet, G2 JRA1, Internet2: perfSONARUpdate

Eric L. Boyd (Internet2) with Loukik
Kudarimoti. (DANTE / GEANT2 JRA1)
and Joe Metzger (ESnet)

Current Projects

- Performance Tools
 - BWCTL
 - NDT
 - OWAMP
 - Thrulay
- Performance Measurement Framework
 - piPEs -> perfSONAR
 - GGF NMWG
- Member Outreach
 - Network Performance Measurement Workshops
 - Performance Tool Cookbooks
- Bulk Transport
 - Design Space
 - Prototype

- Joint effort of ESnet, GÉANT2 JRA1 and Internet2
- Webservices network performance framework
 - Network measurement tools
 - Network measurement archives
 - Distributed scheduling/authorization
 - Multi-domain policy

perfSONAR: Credits

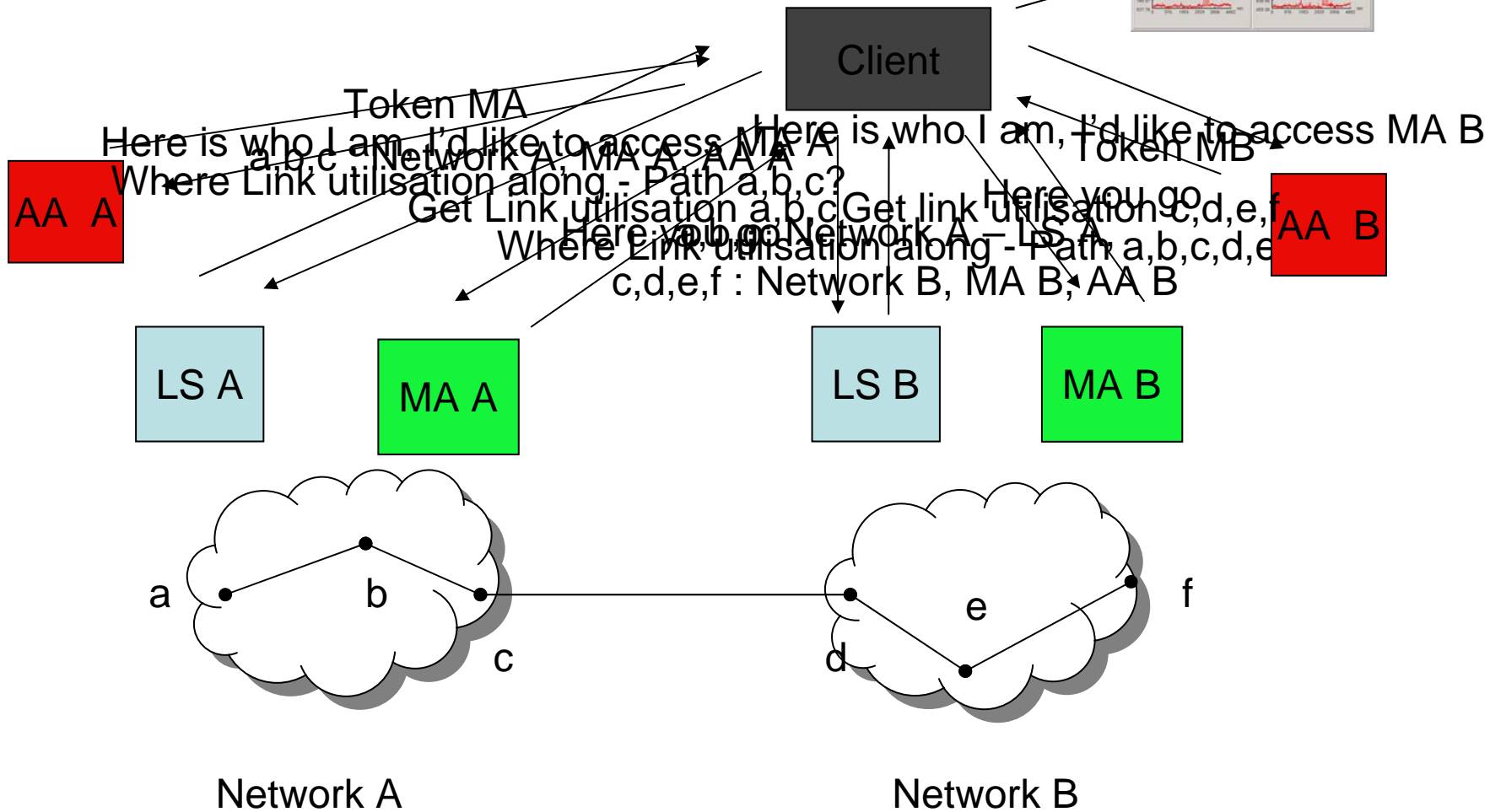
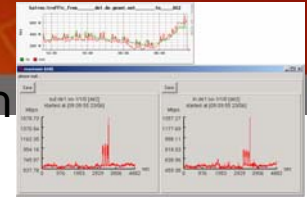
- perfSONAR is a joint effort
- Major Partners: ESnet, GEANT2 JRA1, and Internet2.
 - RNP may join
- Internet2 includes:
 - Internet2 staff
 - University of Delaware
- My apologies if I have overlooked someone (still working on the credits process)
- GEANT2 JRA1 includes:
 - Arnes
 - Belnet
 - Carnet
 - Cesnet
 - DANTE
 - DFN
 - FCCN
 - GRNet
 - ISTF
 - PSNC
 - Nordunet (Uninett)
 - Renater
 - RedIRIS
 - Surfnets
 - SWITCH

perfSONAR: Services (1)

- Lookup Service
 - Allows the client to discover the existing services and other LS services.
 - Dynamic: services registration themselves to the LS and mention their capabilities, they can also leave or be removed if a service gets down.
- Authentication Service
 - Internet2 MAT, GN2-JRA5
 - Authentication functionality for the framework
 - Users can have several roles, the authorisation is done based on the user role.
 - Trust relationship between networks

- Transformation Service
 - Transform the data (aggregation, concatenation, correlation, translation, etc).
- Topology Service
 - Make the network topology information available to the framework.
 - Find the closest MP, provide topology information for visualisation tools
- Resource protector
 - Arbitrate the consumption of limited resources.

Useful graph



- Architecture document (Fall '04)
- Detailed Design document (Spring '05)
- Workshops in Brussels (09/04), Zurich (04/05), Ann Arbor (05/05), Poznan (08/05), Cambridge (01/06)
- Communications:
 - E2EMON submission (03/05)
 - TNC05 paper / presentation (06/05)
 - ICSOC05 paper (12/05)

perfSONAR: Current Status (2)

- Measurement Point Service (MP)
 - Alpha release for ping available (Java)
 - In progress work on BWCTL (Perl) and IPPM (Perl)
 - In progress work on Scheduler
- Measurement Archive Service (MA)
 - Beta release for RRD data available (Java, Python)
 - Alpha release for SNMP data available end of month (Java)
- Look-up Service (LS)
 - Alpha release for monolithic look-up service available (Java)
 - Research (Target: SC paper) into P2P implementation issues
- Authentication Service (AS)
 - We are working closely with JRA5. We have explained our requirements to them and they are working on the first draft of the API.

- Transformation Service (TS)
 - Not started
- Topology Service (ToS)
 - Not started
- Resource Protector Service (RP)
 - Not started
 - Some discussions concerning interaction with scheduler
- User Interface
 - Joe Metzger demo
 - Various examples in progress by the Bulgarians, the Germans, the Greeks

perfSONAR: Demos

- GGF, EGEE, Supercomputing demos (10-11/05)
- Participants who have deployed infrastructure over RRD files:
 - Abilene
 - ESnet
 - Geant
 - Other NRENs (GRnet, PSNC, SWITCH)
 - University of Delaware

perfSONAR: Demo

- <http://monstera.man.poznan.pl/jra1-wiki/index.php/PerfsonarUI>
- <https://performance.es.net/cgi-bin/level0/perfsonar-trace.cgi>
- Cue Joe Metzger ...

perfSONAR: Issues (1)

- Consensus building is slow
 - Distributed management
 - Difficult to partition responsibility into subtasks
 - Competing responsibilities
- Consensus building is fruitful
 - Requires thorough exploration of issues
 - Forces consideration of alternate deployment environments
- Proposed Solutions
 - Keep the lines of communication open
 - Flexibility and good humor

perfSONAR: Issues (2)

- The reference implementation is in Java:
 - Java implementation is “harder than some thought.”
 - Java implementation is “more complicated to modify than we hoped.”
 - Java implementation is “more complicated to deploy / configure than we hoped.”
- Other partial implementations in perl, python
 - Bad: May imply frustrations with Java or it may imply consensus building is slow.
 - Good: Diversity of implementations extends viability / adoptability and exposes design flaws
- Proposed Solutions
 - Code simplification (Jeff Boote)
 - Better documentation (Jerome Durand / Susan Evett)

perfSONAR Issues (3)

- Development is coming along “just in time” (or pretty close to that)
- Not enough time between development deliverables and first adopters
- Not enough developers
- Proposed Solution:
 - Request more development resources / more efficient use of developers
 - Better documentation targeted at:
 - Adopters / deployers
 - Tool integrators
 - Developers

- Version 1 of the schema “all but done”
 - Employed by piPEs, Advisor, AMP, MonALISA, and SLAC
- Version 2 of the schema continues to evolve
 - perfSONAR work benefits from and informs this project

perfSONAR: Conclusion

- Many tools are becoming interoperable using the common GGF schema language (and specifically the pSonar dialect).
- Good working relationships with the principles across multiple continents (Brazil is even becoming involved!)
- The pace of development is increasing.
- Real tools (i.e. analysis tools building on the minimal framework) are actually in development now.



INTERNET[®]

www.internet2.edu