What Can Network Performance Monitoring Do For You?

Jeremy Nowell, EPCC

Grid Operations Workshop, Stockholm

13-15 June 2007

jeremy@epcc.ed.ac.uk

www.eu-egee.org www.egee-npm.org

Information Society and Media

Overview

- **Introduction**
- **Why NPM?**
  - Real life examples
- **What's Available?**
  - Tools
  - Data
  - Diagnostic Tool
- **Deployment plans**
  - What is useful and what will be used?

Introduction

- **Network Performance Monitoring (NPM), formerly part of JRA4, is now part of SA1**
- **Tools have been developed to collect and provide access to a wide range of network data**
- **The challenge is now to start to gather data that is**
  - Useful for EGEE sites
  - Useful for EGEE Grid Operations
  - Presented in ways appropriate for consumers

Why NPM?

- **For Site and Grid operations**
  - Help diagnose performance problems ***between*** sites
    - This transfer is slow, what's broken? – the network, the server, the middleware…
    - I can't see site X, has the network gone down or just a particular service or machine?
    - My application's performance varies with time of day – is there a network bottleneck?
  - Help diagnose problems ***within*** sites
    - Most network problems, especially performance issues, are not backbone related, they are in the "last mile"
  - Help with planning and provisioning decisions

- **For Grid services and middleware**
  - I want to increase the performance of file transfers between sites
  - I want to know which compute site is "closest" to my data to submit a job to it

Why NPM? (2)

- **What's different about networks for the Grid?**
  - Without the network there is no Grid…
  - Large amounts of application data, often continuous
  - Multiple connections and streams
  - New technology – eg provisioned light paths
  - ***End-to-end*** performance crucial
    - What's the use of a 10 Gb/s dedicated connection if your application is only transferring data at 10 Mb/s?

Why NPM? (3)

Q: **Why don't we just throw some more bandwidth at the problem? - Upgrade the links.**

A: **Bandwidth is bad for you. It's like a narcotic…**

- **It's very addictive. You start off with a little, but that's not really doing it for you; it's not enough. You increase the dose, but it's never as good as you thought it would be.**
- **By analogy you can keep buying more and more bandwidth to make your network faster but it's never quite as good as you thought it would be.**
- **Why? Because simple over-provisioning is not sufficient**
- **Doesn't address the key issue of *end-to-end performance*:**
  - Network backbone in most cases is genuinely not the source of the problem.
  - Last mile (campus network→end-user system→your application) often cause of the problem: firewall, network wiring, hard disc, ***application*** and many more potential culprits.

This can get to be an expensive habit – dedicated high speed fibre is not cheap

Also, If simple over-provisioning was a total solution, there would not be so much other work going on, e.g. protocol research (high speed TCPs)

Network Performance Factors

- **End System Issues**
  - Network Interface Card and Driver and their configuration
  - TCP and its configuration
  - Operating System and its configuration
  - Disk System
  - Processor speed
  - Bus speed and capability
  - ***Application*** eg old versions of scp
- **Network Infrastructure Issues**
  - Obsolete network equipment
  - Configured bandwidth restrictions
  - Topology
  - Security restrictions (e.g., firewalls)
  - Sub-optimal routing
  - Transport Protocols
- **Network Capacity and the influence of Others!**
  - Many, many TCP connections
  - Congestion

How can NPM help?

- **Applications and sites can make operational decisions based on previous network performance.**
  - Wrong decisions will cause pain. Having the ‘right’ metrics available will allow ‘better’ decisions to be made.

- **NPM data let end users see the performance they should expect from their Grid applications**
  - Misleading to infer network performance from application performance.
  - Seldom the same as what they know (or think they know) about the specification of their network connections.

- **As the examples will show, faults and inefficiencies can be identified and solved if NPM data are available.**
  - Of benefit to the whole site, as well as the Grid in general.
  - Sometimes the data can show up strange configurations that even site network admins are not aware of.
  - Network admins will likely not investigate application problems without hard evidence.

Real Life Examples

Courtesy of Mark Leese (STFC Daresbury Lab - UK Gridmon/GridPP)

Real Life Examples (1)

Q: What if we share existing fibre, and use circuit-switched lightpaths? That is dedicated bandwidth, but without the cost of dedicated fibre.

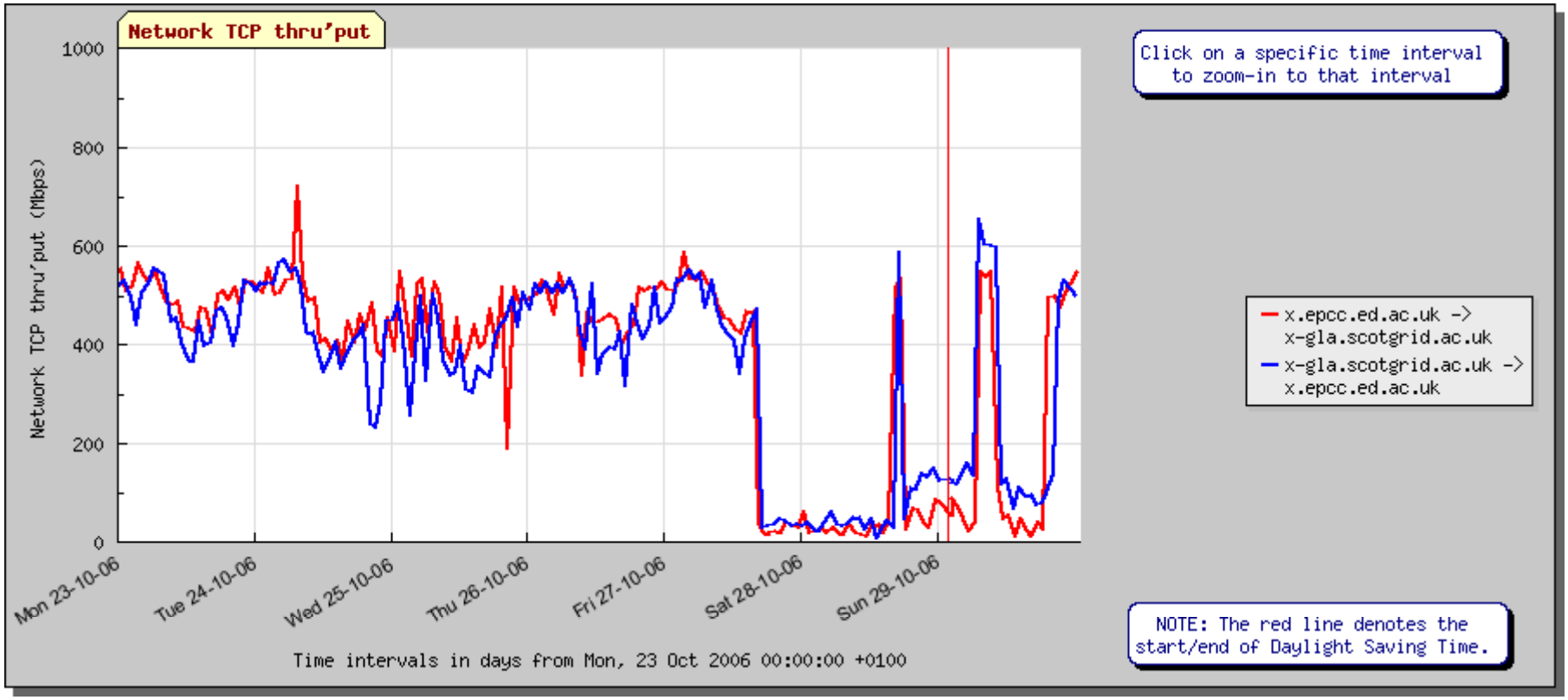
A: Good idea in theory, and we can see the benefits from a fibre infrastructure like *UKLight* via the *ESLEA project, but this still doesn't address the end-to-end issue. Take a *real-life* ESLEA example (thanks to ESLEA for the figures)...**

- UCL (London) wanted to transfer data from FermiLab (Chicago) for analysis, before returning the results
- datasets were 1-50TB
- 50TB would take > 6 mths on public network, or one week at 700Mbps
- 1 Gbps circuit-switched light path provisioned as a result
  - Still disc-to-disc transfers only came in at 250Mbps, just 1/4 of theoretical network maximum
- NPM data revealed an end-site problem

* Exploitation of Switched Lightpaths for e-Science Applications

Real Life Examples (2)

- Glasgow running transfer tests to Edinburgh
- Seeing poor rates (80Mb/s)
- 1st thing: despite transferring just 80Mb/s, residual TCP bandwidth drops by ≈ 400Mb/s
- Warning bells



Real Life Examples (2)

- **Traceroutes reveals suspect router…**

```
traceroute to gridmon.epcc.ed.ac.uk (129.215.175.71), 30 hops max,
    38 byte packets
1  194.36.1.1 (194.36.1.1)  0.941 ms  0.882 ms  0.815 ms
2  130.209.2.1 (130.209.2.1)  0.875 ms  0.831 ms  0.830 ms
3  130.209.2.118 (130.209.2.118)  60.415 ms  55.453 ms  31.327 ms
4  glasgowpop-ge1-2-glasgowuni-ge1-1-v152.clyde.net.uk
    (194.81.62.153)  32.420 ms  34.404 ms  29.424 ms
5  glasgow-bar.ja.net (146.97.40.57)  43.467 ms  52.298 ms  39.349
    ms
6  po9-0.glas-scr.ja.net (146.97.35.53)  45.856 ms  44.445 ms
    41.388 ms
7  po3-0.edin-scr.ja.net (146.97.33.62)  51.509 ms  63.493 ms
    31.435 ms
8  po0-0.edinburgh-bar.ja.net (146.97.35.62)  22.454 ms  25.412 ms
    31.381 ms
9  146.97.40.122 (146.97.40.122)  44.602 ms  42.494 ms  35.492 ms
10  gridmon.epcc.ed.ac.uk (129.215.175.71)  33.515 ms  34.623 ms
    37.694 ms
```

- **Graphs and traceroutes provide evidence for further investigation**

Real Life Examples (2)

- **Reverse route confirms. Traceroutes are normal until we hit suspect router…**

```
traceroute to gppmon-gla.scotgrid.ac.uk (194.36.1.56), 30 hops max,
    38 byte packets
1  vlan175.srif-kb1.net.ed.ac.uk (129.215.175.126)  0.435 ms  0.387
    ms  0.380 ms
2  edinburgh-bar.ja.net (146.97.40.121)  0.357 ms  0.329 ms  0.322
    ms
3  po9-0.edin-scr.ja.net (146.97.35.61)  0.564 ms  0.485 ms  0.485
    ms
4  po3-0.glas-scr.ja.net (146.97.33.61)  1.656 ms  1.511 ms  1.499
    ms
5  po0-0.glasgow-bar.ja.net (146.97.35.54)  1.850 ms  1.352 ms
    1.422 ms
6  146.97.40.58 (146.97.40.58)  1.679 ms  1.661 ms  1.569 ms
7  glasgowuni-ge1-1-glasgowpop-ge1-2-v152.clyde.net.uk
    (194.81.62.154)  1.796 ms  1.677 ms  1.646 ms
8  130.209.2.117 (130.209.2.117)  31.197 ms  34.615 ms  29.121 ms
9  130.209.2.2 (130.209.2.2)  32.814 ms  32.158 ms  32.145 ms
10  gppmon-gla.scotgrid.ac.uk (194.36.1.56)  41.634 ms  37.555 ms
    24.635 ms
```

Real Life Examples (2)

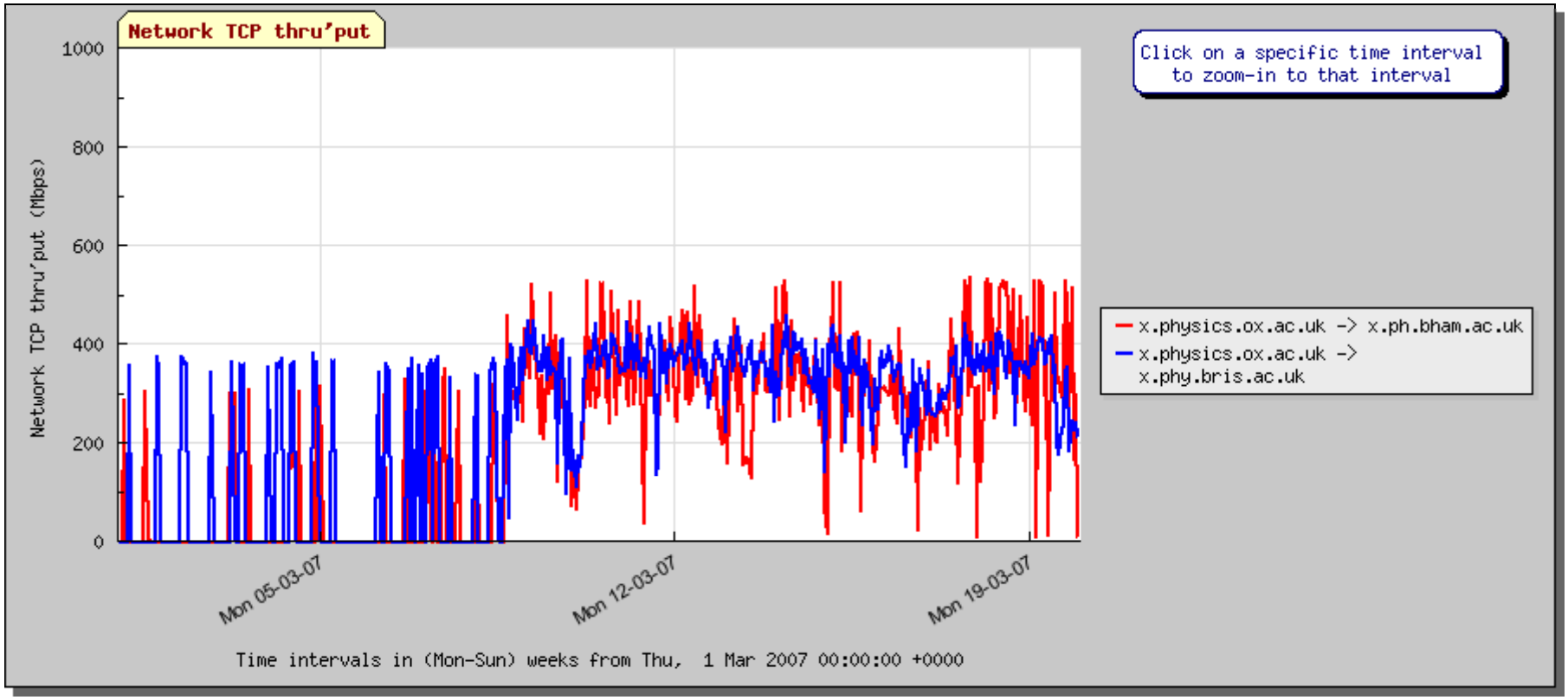
- Further investigation revealed that the router had exhausted it's CAM space and was essentially switching in software
- CAM = Content-Addressable Memory
- Hardware implementation of an associative area
- a data word is supplied (not a memory address) and the CAM searches its entire memory to see if the data word is stored. If the word is found, the CAM returns a list of one or more corresponding storage addresses, or other associated pieces of data
- CAM memory is used for switching and routing, e.g. switches store learned MAC addresses and their associated switch port in CAM

```
MAC Address   Located on Port
------------- ---------------
000039-0643f5 26
000089-01af9a 5
000102-162346 16
```

- A particular table lookup was not being hardware accelerated causing problems under certain flow conditions
- The CAM dynamic database was re-optimised and the unit began switching in hardware again

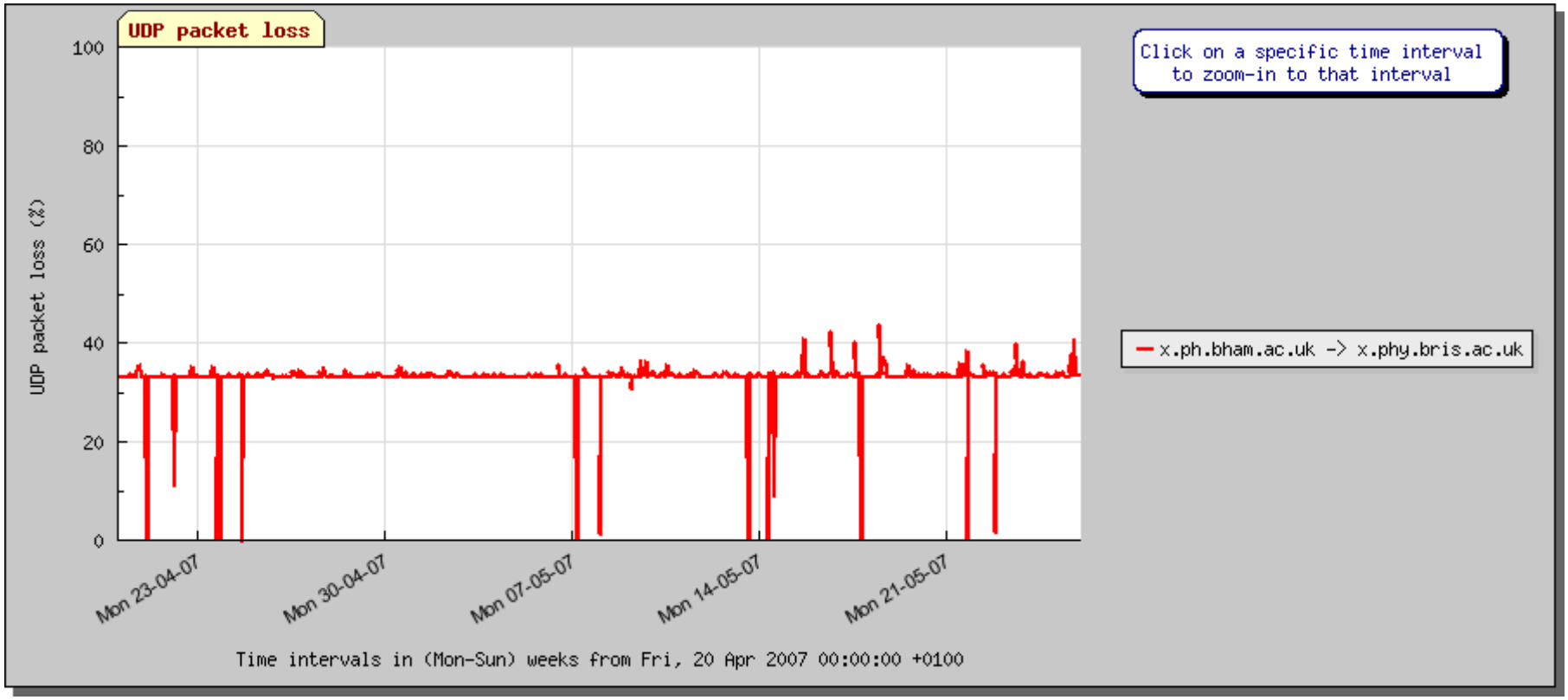
Real Life Examples (3)

- Local departmental firewall reconfigured to switch off strict checking of TCP sequence numbers
- Potential minefield: SACK etc.



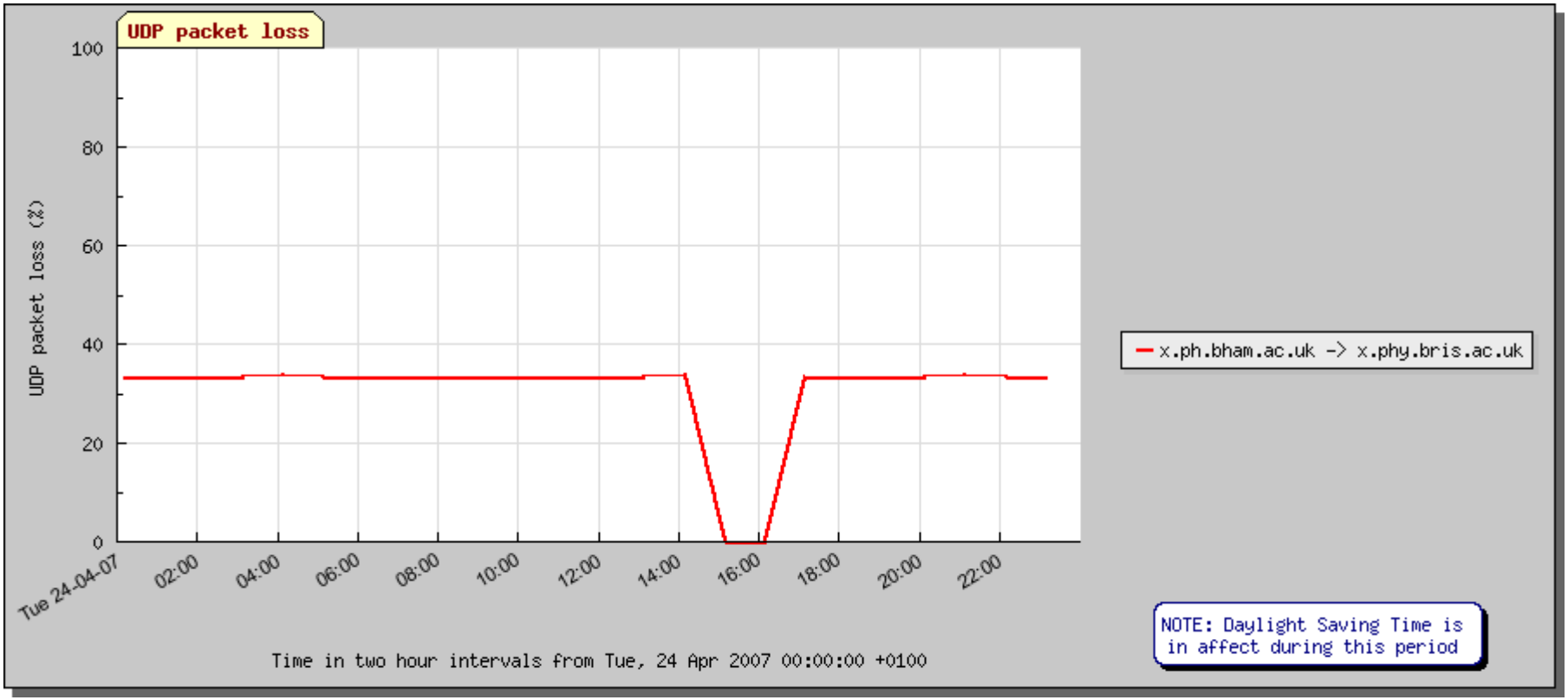
Real Life Examples (4)

- Almost constant 33% UDP packet loss
- Fatal to most/all apps using UDP
- Occassional dip to 0%



Real Life Examples (4)

- Zooming into particular day shows period of 0% loss
- Site firewall limits UDP to 1000 pps per endpoint pair
- Temporarily raised to 20,000 pps for Video Conferences



NPM General Requirements

- **Scale and heterogeneity of EGEE fabric poses a requirement to support diversity of all kinds**
  - Multitude of ways of collecting monitoring data
    - Different measurement types
      - *end-to-end*
        - *Appropriate to experience of user and application, eg TCP achievable bandwidth*
      - *Backbone*
        - *Lower level measurements, used to pin-point source of problems*
    - Different measurement tools
    - Different data formats
  - Many administrative domains
  - Different user groups

NPM User Requirements

Middleware

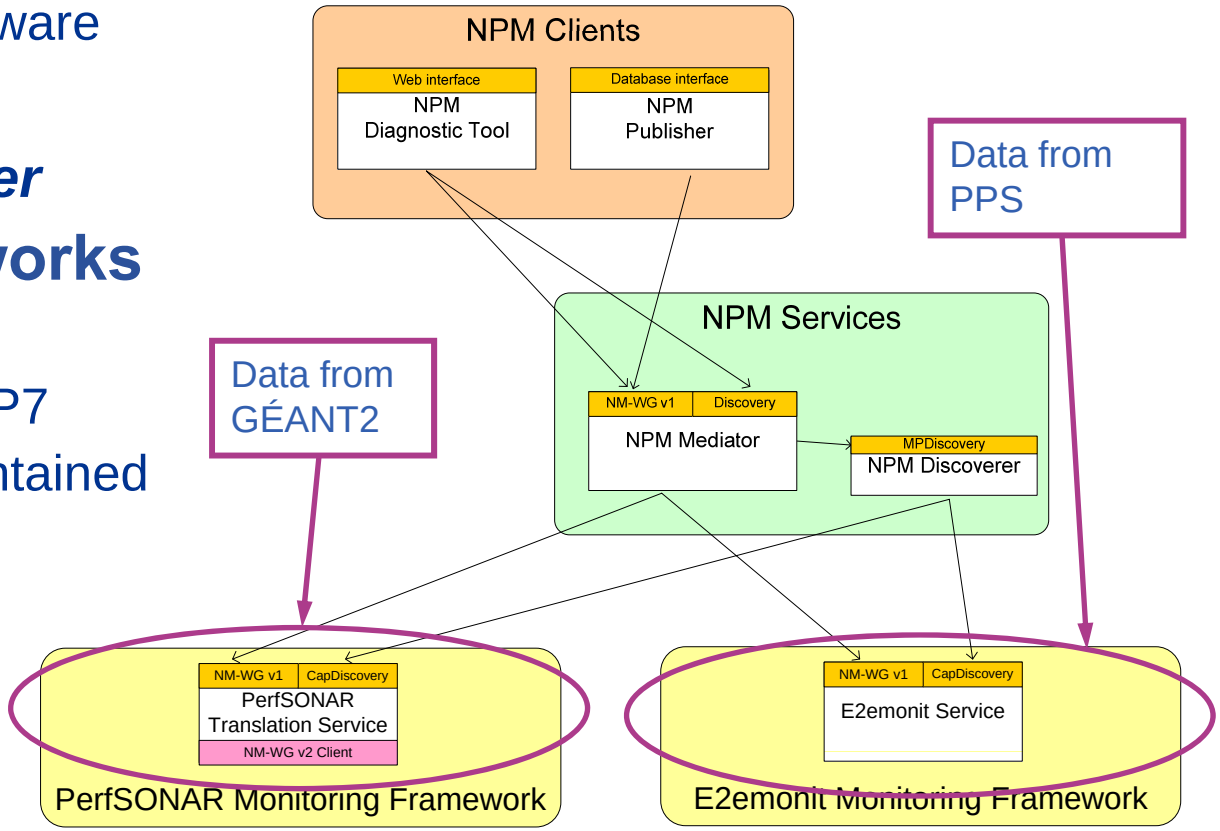
- **Programmatic interface**
  - Web service
  - Database
- **Info for 100 paths returned in 0.2s**
- **Relate Compute/Storage Element with NMP**
- **Raw, historical data for 24 hrs**
- **Mainly end-to-end data**

Operation Centres

- **NOCs and GOCs**
  - Web-based GUI
  - Interface to define alarms
  - On-demand & historical data
  - Backbone & end-to-end data
- **NOCs**
  - Display which tool gathered the results and how
  - Per hop data/ability to zoom in
- **GOCs**
  - High-level statistics

What's available - Software

- **Clients**
  - The ***Diagnostic Tool*** (DT)
    - For use by people
  - The ***Publisher***
    - For use by middleware
- **Middleware**
  - ***Mediator/Discoverer***
- **Monitoring Frameworks**
  - ***e2emonit***
    - Formerly EDG::WP7
    - Provided and maintained by NPM team
  - PerfSONAR
  - LHC-OPN
    - Soon?



What's available - Metrics

- **Metrics depend on which tools you use!**
  - We will allow access to any relevant data, provided it is available using a OGF NM-WG compliant interface
- **e2emonit**
  - ping
    - Connectivity
      - *Round trip time, packet loss*
  - iperf
    - Real life application performance
      - *TCP achievable bandwidth*
  - udpmon
    - Network health, congestion etc
      - *UDP achievable bandwidth, one-way delay, UDP packet loss*
- **PerfSONAR**
  - Developed by GÉANT, Internet2 and ESNet
  - Currently accessing utilisation data

Site Monitoring

- **A recent survey* carried out via the ROC managers gave some feedback on network monitoring tools used by sites**
  - Passive
    - RRDTool
    - MRTG
    - Cacti
    - Flow based tools
  - end-to-end
    - Smokeping
    - Gridmon (UK GridPP)

*https://twiki.cern.ch/twiki/bin/view/EGEE/SA1_Network_Monitoring

Data Federation

- **Our tools designed for data federation**
  - Currently
    - e2emonit from EGEE sites that deploy it
    - e2emonit from related projects – BalticGrid
    - PerfSONAR Measurement Archives
  - In the future
    - Gridmon (UK GridPP)
    - Other PerfSONAR components
      - *E2E layer 2 link status (relevant for LHC-OPN)*
      - *Measurement Archives through native interface*
      - *BWCTL, OWAMP Measurement Points*
    - Others – RRD based, MRTG, Smokeping etc?

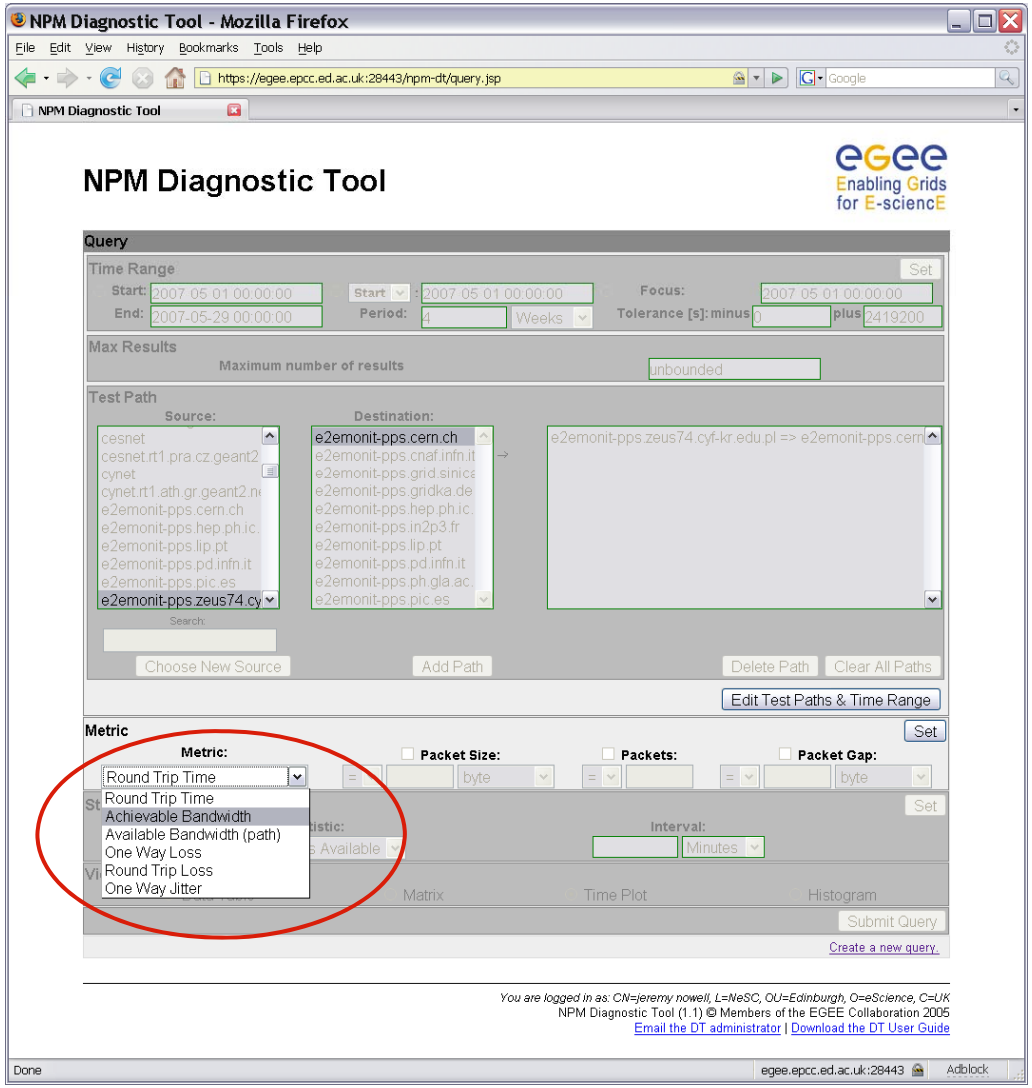
NPM Diagnostic Tool

– The Diagnostic Tool can be accessed using a standard web browser, which users are individually authorised to use.

- *Please mail us for access!*

– The intended user is a NOC/GOC/ROC operator, but anyone can use it to investigate problems

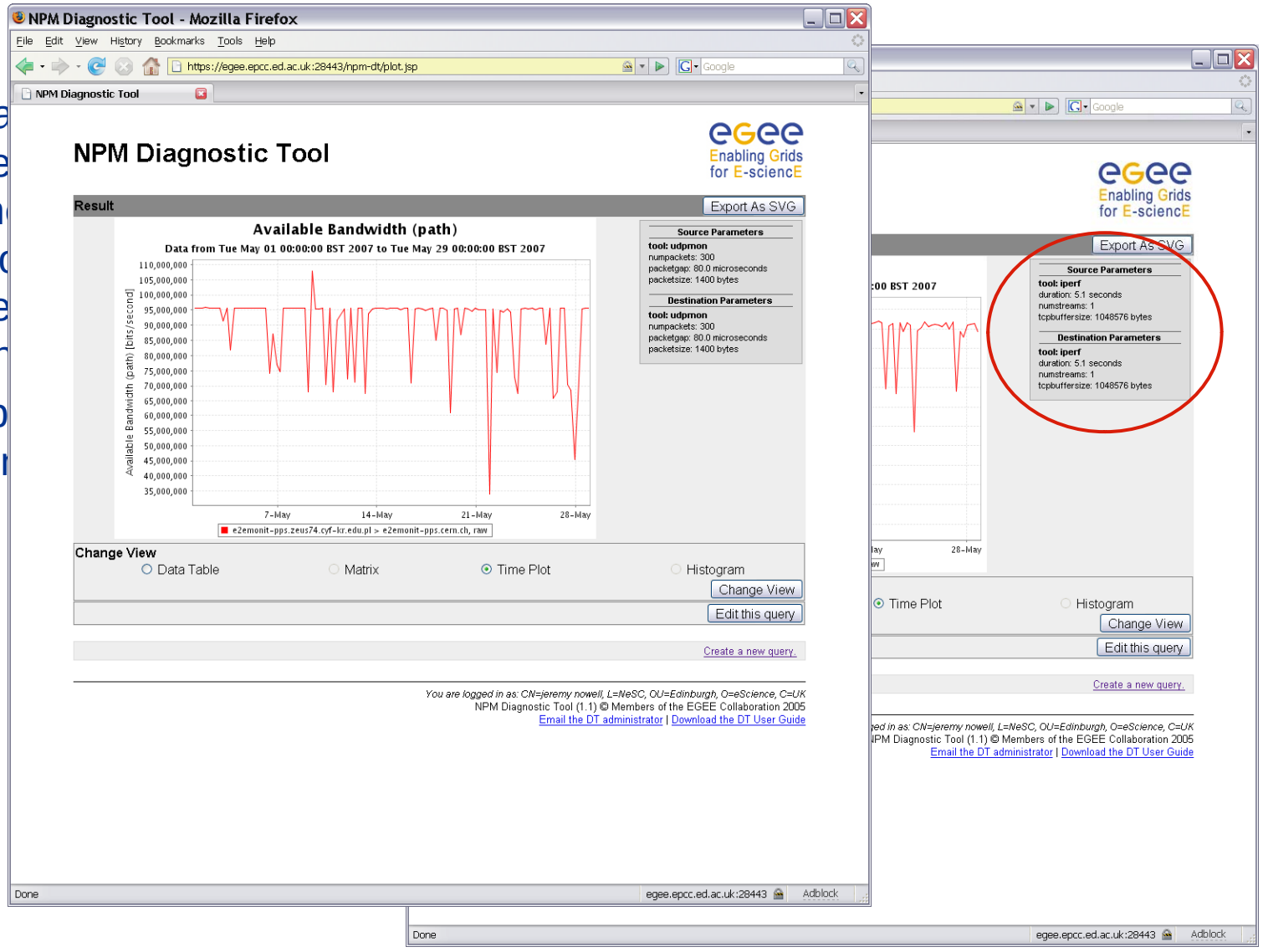
–The sites and metrics displayed depend on where and which measurement tool has been deployed



NPM Diagnostic Tool (2)

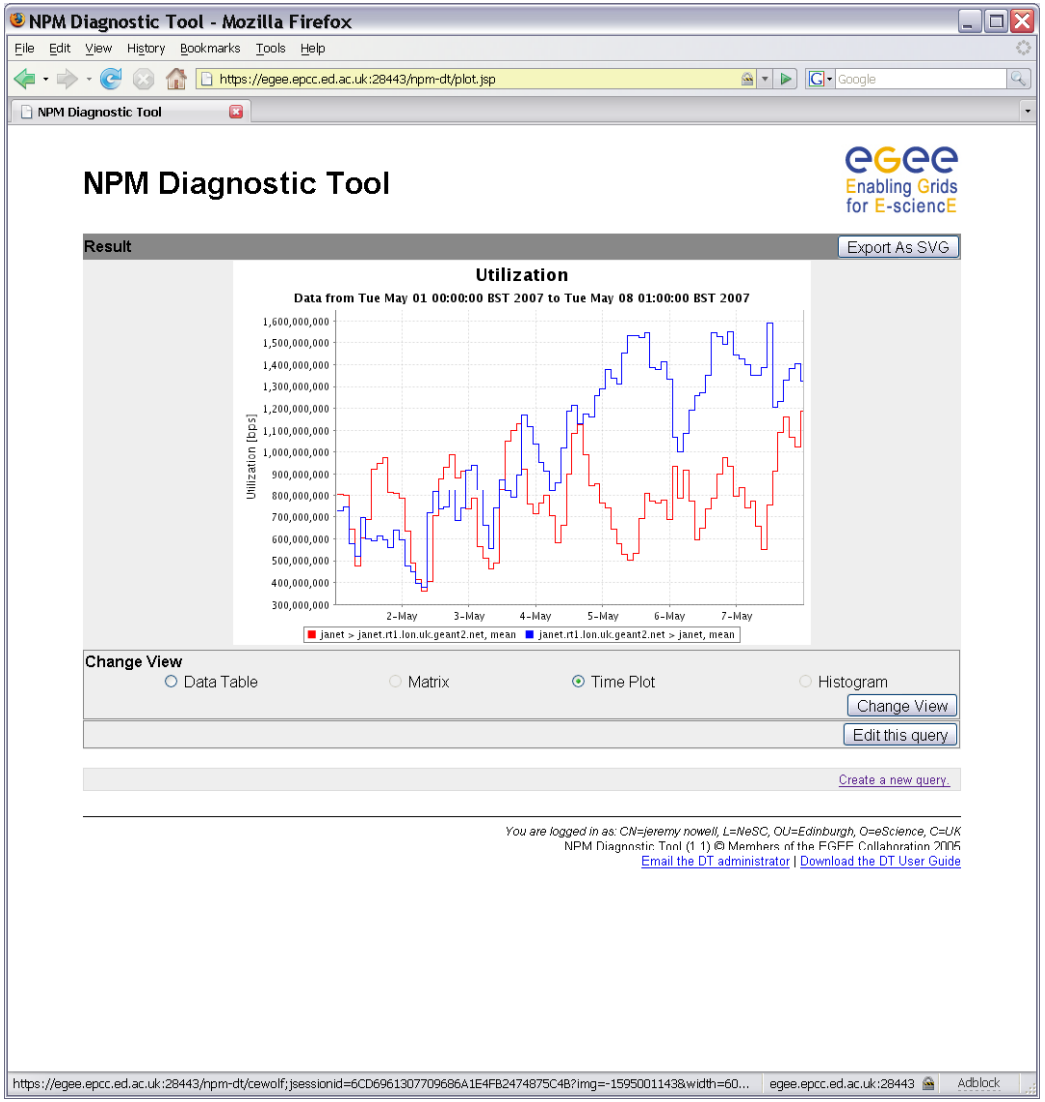
– The para
gather me
shown - h
the iperf t
gather the
bandwidth

– These p
useful in i
results.



NPM Diagnostic Tool (3)

- Information from multiple paths may be plotted at the same time.
- Here utilisation data for the GÉANT2/JANET router is plotted for both inbound and outbound traffic over the course of one week, obtained from the GÉANT2 PerfSONAR Measurement Archive.



Deployment Issues (1)

- **The usefulness of all this depends critically on the data that is available**
  - The plan was always to use measurement data that is already available
    - Probably not sufficiently deployed across sites
      - *e2emonit could be an option, but not the only one*
    - Ideally federations or VOs make deployment decisions
      - *E.g. GridPP or BalticGrid*
    - We can help with network monitoring topology, based on application requirements.
  - The monitoring tools questionnaire suggests NPM data are already collected in some ROCs/sites
    - RRD/Smokeping and RRD/Flow in particular
      - *We aim to write a Webservice that makes some of these data widely available*
      - *Will you deploy it on top of your suite?*

Deployment Issues (2)

- Is the Diagnostic Tool useful or useable?
- Are alarms more immediately useful for site and service administrators?
- Firewall Issues – eg ICMP for ping

Deployment Plans (1)

- **Provide a general "network availability" test for sites**
  - Ping like connectivity test without using ping
  - In conjunction with SA2
    - attempt connection to BD-II port and make data available to ENOC
  - Big, conscious assumption that non-availability of service suggests non-availability of network
  - Data to be made available through Grid Monitoring Data Exchange Standard being developed by Grid Service Monitoring Working Group

However:

- **This *is* a service test**
- **If the network really is so broken you'll probably already know**
- **You could do 10^6 pings before noticing a dropped packet, but such an error rate would be critical for TCP throughput**
  - (ping is still useful for RTT measurement and absolute connectivity)

Deployment Plans (2)

- **Deploy end-to-end monitoring tools on selected sites decided by their usefulness**
  - eg Tier0-Tier1 FTS
    - Probably UDPmon and iperf (e2emonit)
    - Could provide a more useful alarm based on UDP packet loss
      - *Congestion issues, number of streams etc.*
    - Historical data available for mining and diagnosis of problems, service provisioning and planning, as well as middleware and applications
      - *Possibly lead to even more useful alarms based on historical TCP performance data*
  - New node profile for e2emonit that can be deployed on any box that sites choose (rather than necessarily MON)
    - The closer the box is to the end-to-end service being monitored, the better

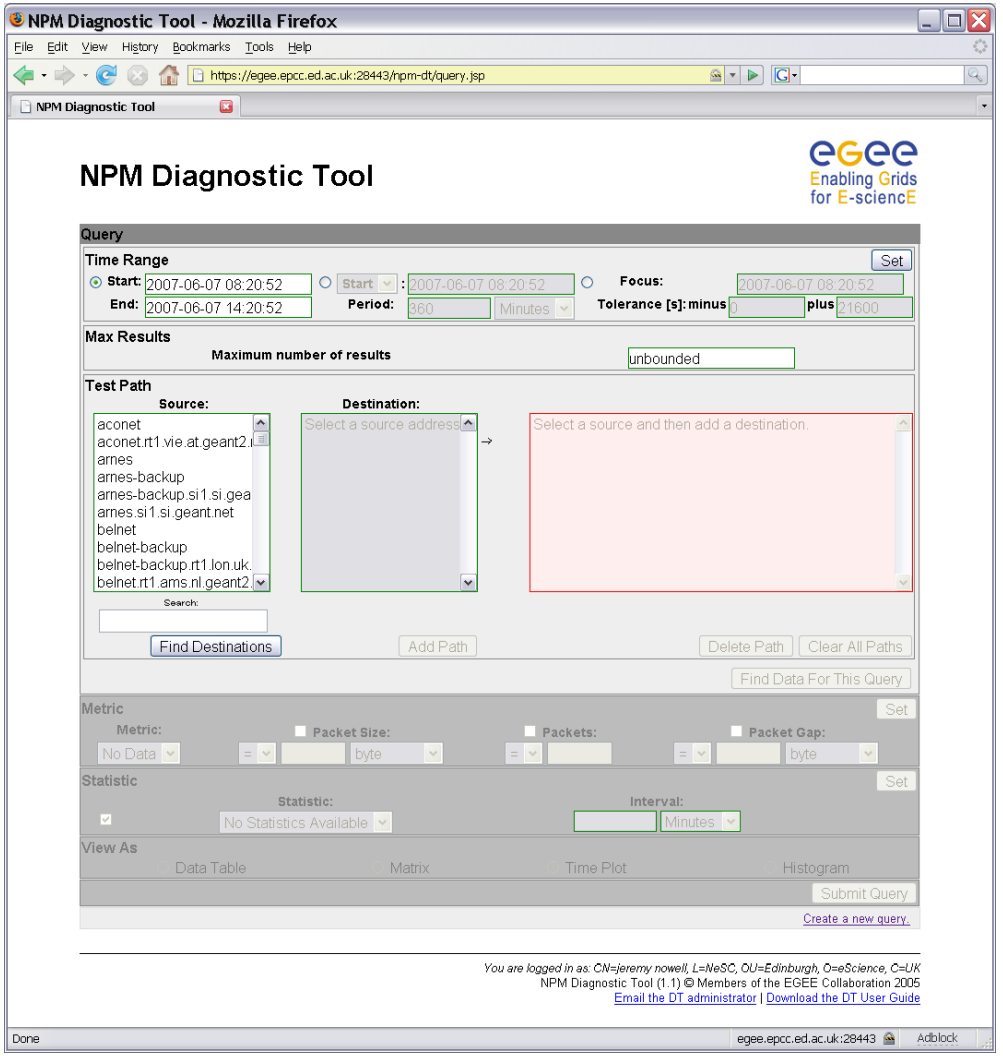
Summary

- **Network performance data (especially *end-to-end*) is crucial for site, grid and service operations**
  - Clear idea of site network performance at any given time
  - Historical data to inform operational decisions
  - Site end-user and application support
- **There is a deployment challenge to be faced to gather useful data**
- **We are ready to face this challenge, and will help you by**
  - Providing network alarms for use by site and service admins
  - Pushing forward the deployment of end-to-end monitoring tools to collect useful data for important services and paths
- **Please contact us if you would like to deploy e2emonit or talk about network monitoring at your site.**
  - http://www.egee-npm.org/
  - jeremy@epcc.ed.ac.uk, egee@epcc.ed.ac.uk

Backup Slides

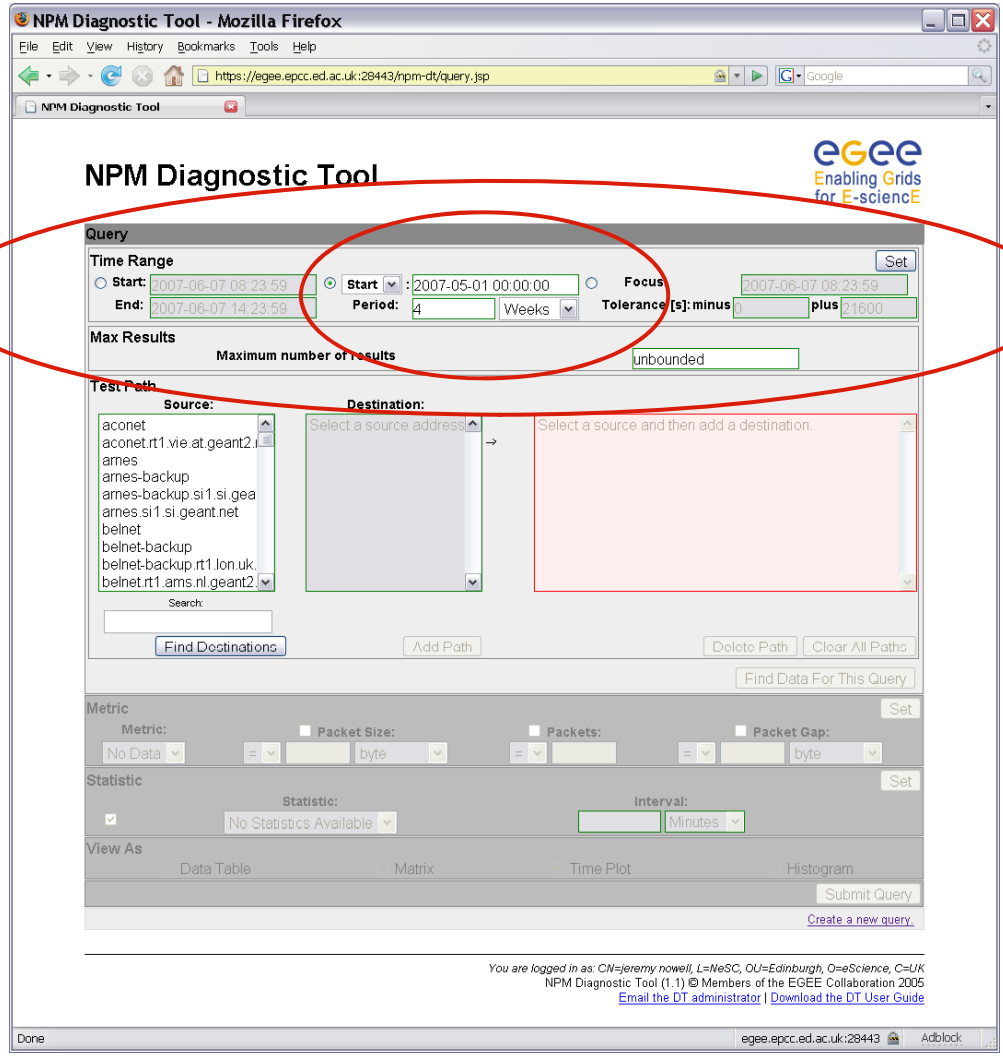
DT Usage (1)

- Step 1: Access the NPM Diagnostic Tool.
  - The Diagnostic Tool can be accessed using a standard web browser, which users are individually authorised to use.
    - In the future, we plan to use VOMS for authorisation.
    - *Please mail us for access!*
  - The intended user is a NOC/GOC/ROC operator



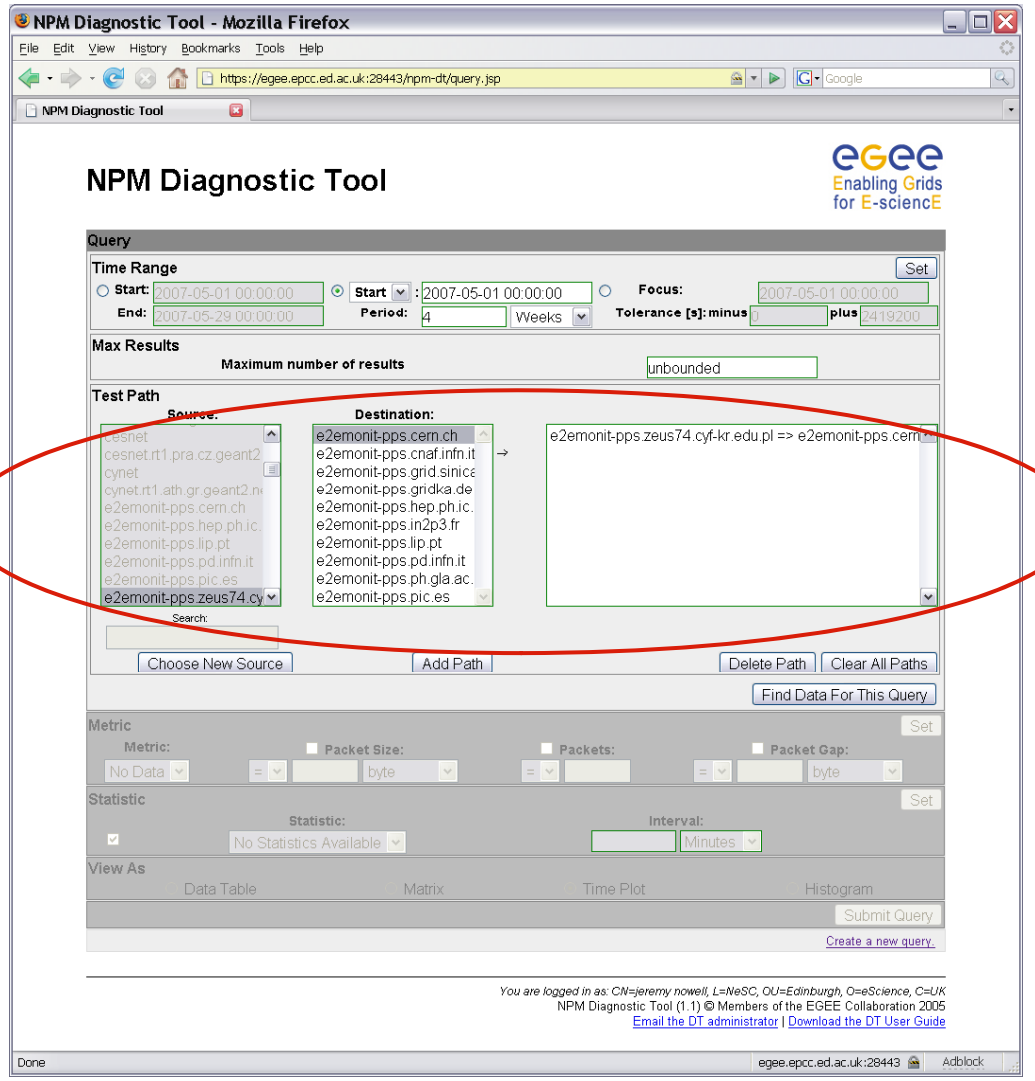
DT Usage (2)

- Step 2: Select a Time.
  - The end-user does not have a specific time, but wants to see the performance for the past four weeks.
  - The user enters the appropriate time range, specifying a Start date/time of 2007-05-01 00:00:00 and a period of 4 weeks.
  - The user presses the Set button to confirm and the alternate time range representations update.



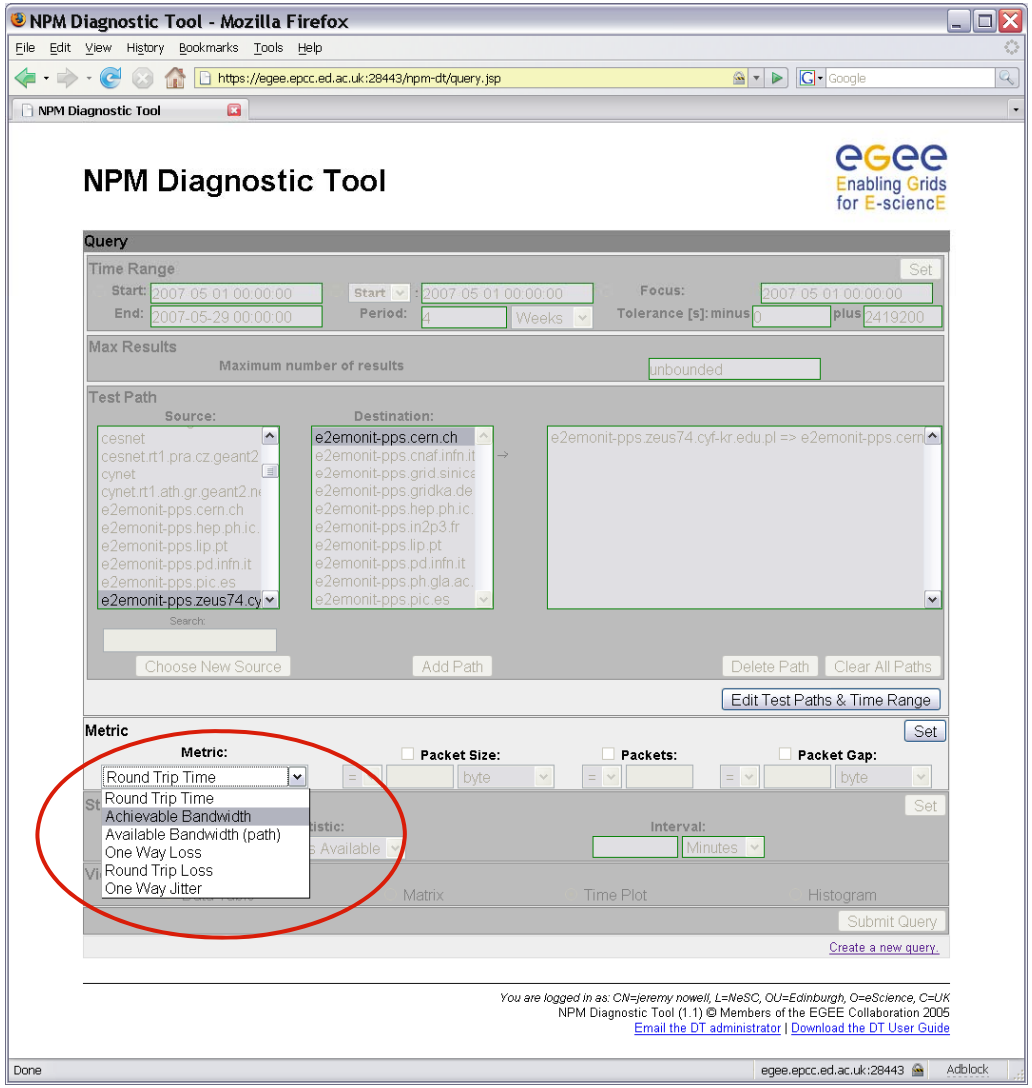
DT Usage (3)

- Step 3: Select a Path.
  - The end-user experienced the problem between Cyfronet in Krakow and CERN.
  - The user selects e2emonit sites at Cyfronet and CERN, adds the path and then selects “Find Data For This Query”



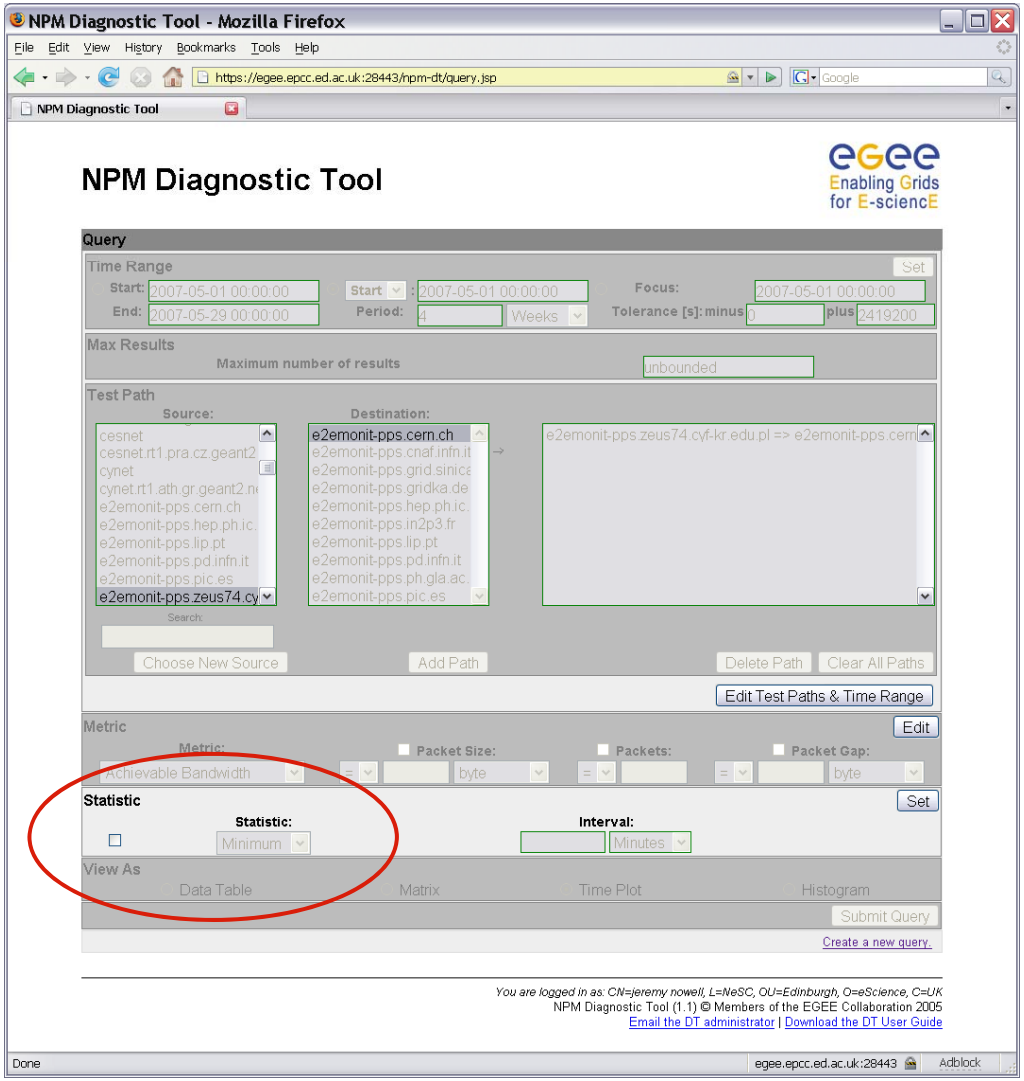
DT Usage (4)

- Step 4: Select a Metric.
  - The end-user experienced throughput problems.
  - Although there are several possibly relevant metrics to choose from (and only those measured are available to select from), the user decides to look at the Achievable Bandwidth on the path.
  - Achievable Bandwidth is selected from the Metrics box and the Set button pressed to confirm.



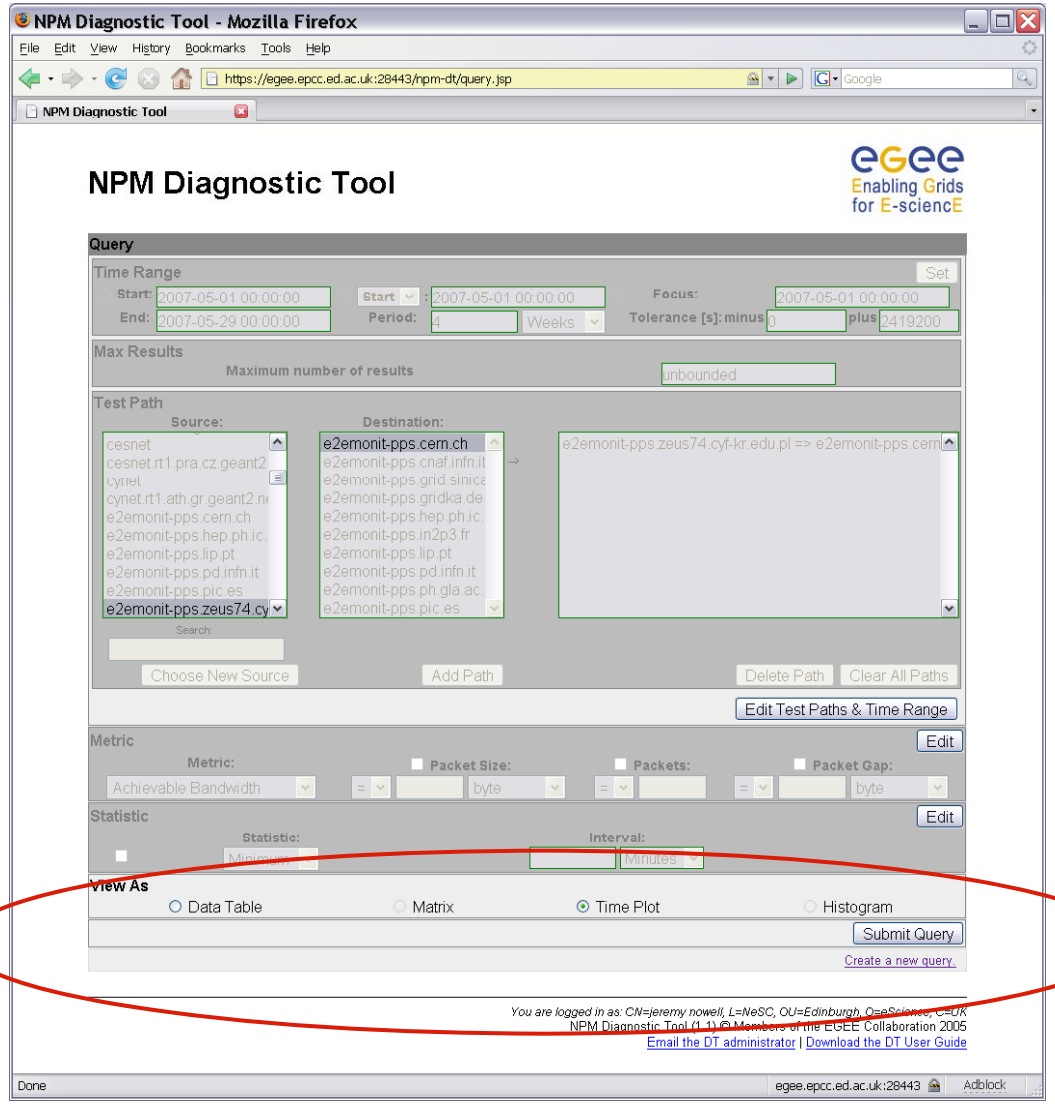
DT Usage (5)

- Step 5: Select a Statistic.
  - Several types of statistical data are available, such as Minimum, Maximum, Mean.
  - A particular interval can be applied to each, to provide, for example, an hourly mean over the past two days.
  - The user just wants a general overview of measurements and elects to retrieve raw data (Statistic check-box not checked).



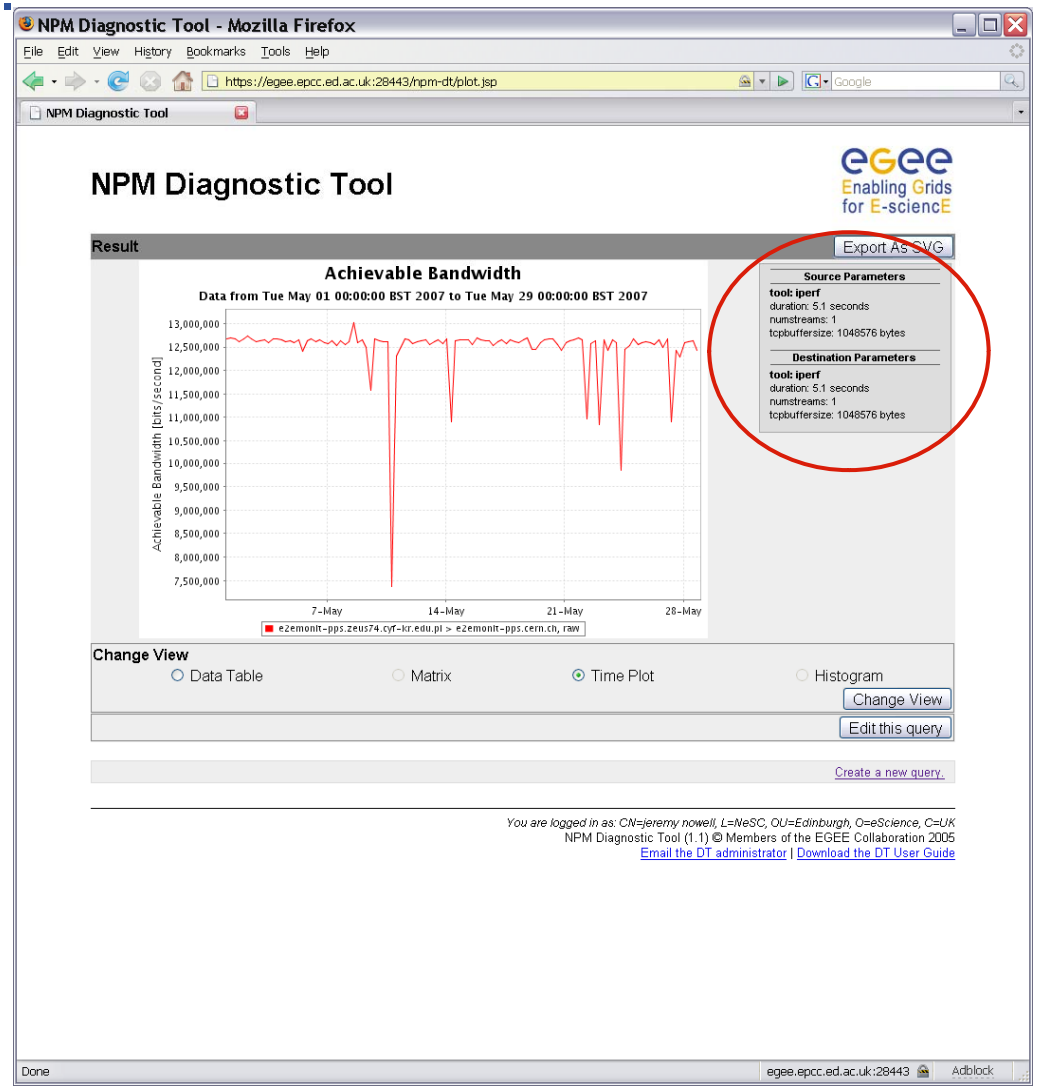
DT Usage (6)

- Step 6: Select a View.
  - Currently Data Table and Time Plot views are available.
  - The user wants an overview of how the Achievable Bandwidth has changed over time, so selects the Time Plot.
  - The Query entry is complete, and the user selects Submit Query.



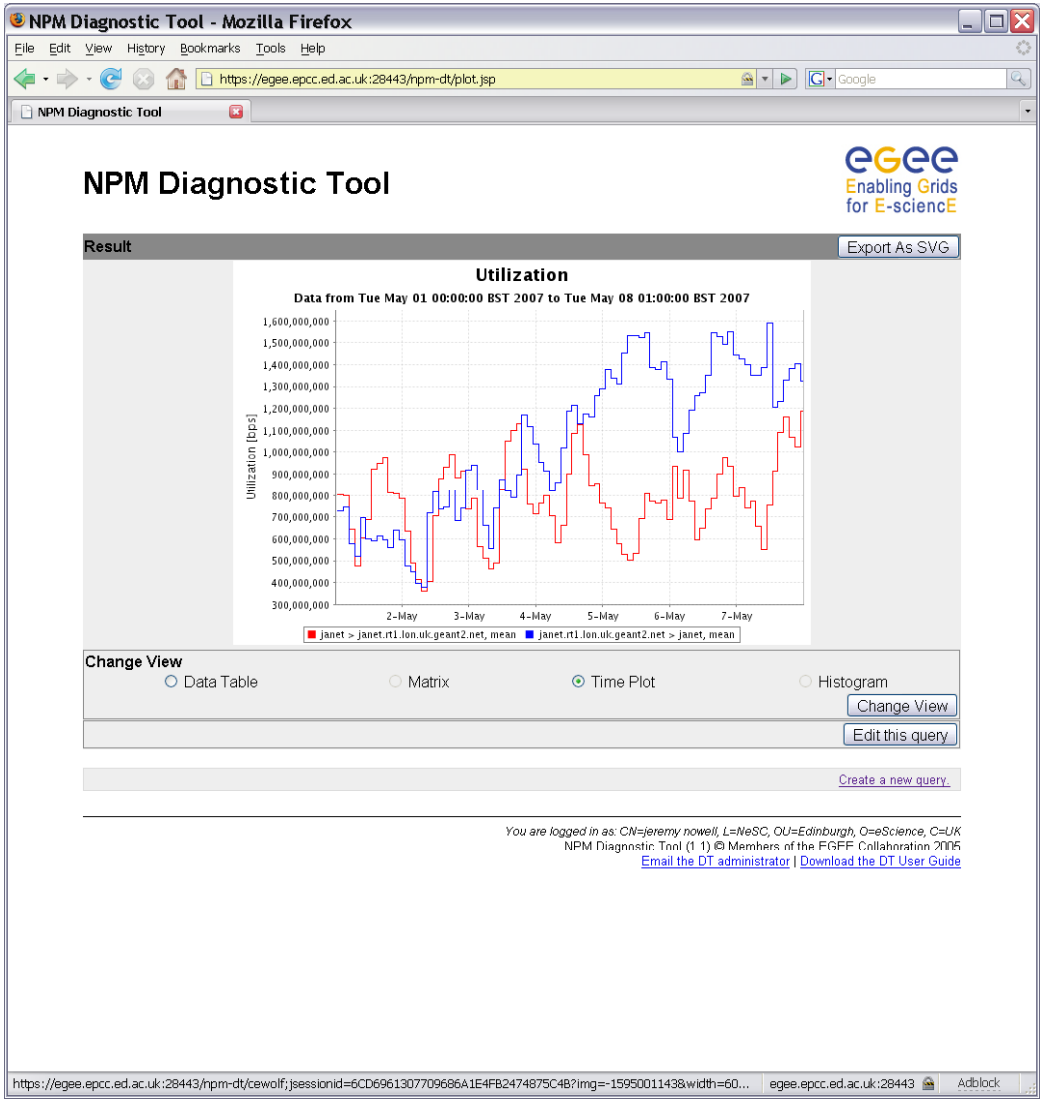
DT Usage (7)

- Step 7: Examine results.
  - The results are plotted, with Time on the x-axis and Achievable Bandwidth on the y-axis.
  - The parameters used to gather measurements are shown - here, showing that the iperf tool was used to gather the achievable bandwidth information.
  - These parameters can be useful in interpreting the results.



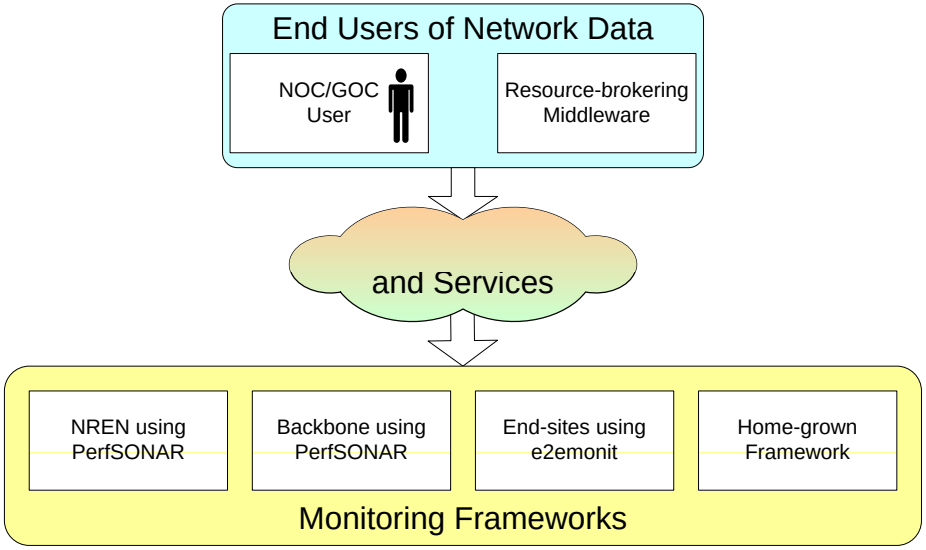
DT Usage (8)

- Information from multiple paths may be plotted at the same time.
- Here utilisation data for the GÉANT2/JANET router is plotted for both inbound and outbound traffic over the course of one week, obtained from the GÉANT Measurement Archive.



NPM Strategy

- **Aim to standardise access to NPM data across different domains and frameworks**
  - Note – we are not building measurement tools, but rather facilitating access to data collected by them
- **Interoperability pursued through use of OGF NM-WG schema**
  - EGEE should not and cannot aim to enforce the uptake of a specific NPM framework across the diverse EGEE fabric or the associated networks
  - Use NM-WG interfaces where they have been adopted; facilitate their use elsewhere.



NPM Architecture