High Performance Networking for e-Science in the UK

Saleem Bhatti
Networks Research Group
Department of Computer Science
University College London
http://www.cs.ucl.ac.uk/staff/S.Bhatti/

Networks Research Group (NRG) at UCL

- http://nrg.cs.ucl.ac.uk/
- Internet:
  - Architecture
  - Protocols
  - Application
  - Evolution
- Multimedia Networking
- Mobile Networked Systems
- **High-speed networking**

e-Science and Grid in the UK

- e-Science Core Programme:
  - http://www.rcuk.ac.uk/escience/
- Funding for 'bringing distributed computing to the science community' (my interpretation!):
  - ~£520m over ~5 years
  - (some more funding recently announced)
- e-Science application users:
  - compute intensive
  - data intensive
  - both compute and data intensive
  - demands on networking

UK e-Science/Grid examples

- RealityGrid http://www.realitygrid.org/ :
  - simulation and discovery of new molecules
  - simulation checkpoint files of ~0.5TB+ each
- e-Diamond http://www.ediamond.ox.ac.uk/ :
  - mammogram analysis
  - image groups of ~80GB+ (for each 1 hour session per patient)
- Integrative Biology http://www.integrativebiology.ox.ac.uk/ :
  - cardiac analysis and tumour growth modelling
  - many data files of several 10s of GB each
- MERLIN http://www.merlin.ac.uk/ :
  - very long baseline interferometery (VLBI):
  - continuous flows of 500Mb/s+, between multiple sites

Differing user requirements

- Network requirements:
  - **all** require high data rates
  - other diverse requirements
- Diverse network requirements:
  - real-time vs. non-real-time
  - reliability (packet level)
  - **deadline based scheduling for non-real-time**
- Two general requirements:
  - high data rates
  - **flow protection (QoS)**

High-speed and QoS: Challenges

- High-speed (multi-Gb/s) operation:
  - **network** engineering
  - **end-system** tuning and engineering
- QoS at high speed within the core network
- QoS for individual application flows:
  - **state management** and **site-to-site signalling**
- Operation of protocols:
  - e.g. various TCP congestion control proposals
  - newer protocols such as DCCP and SCTP

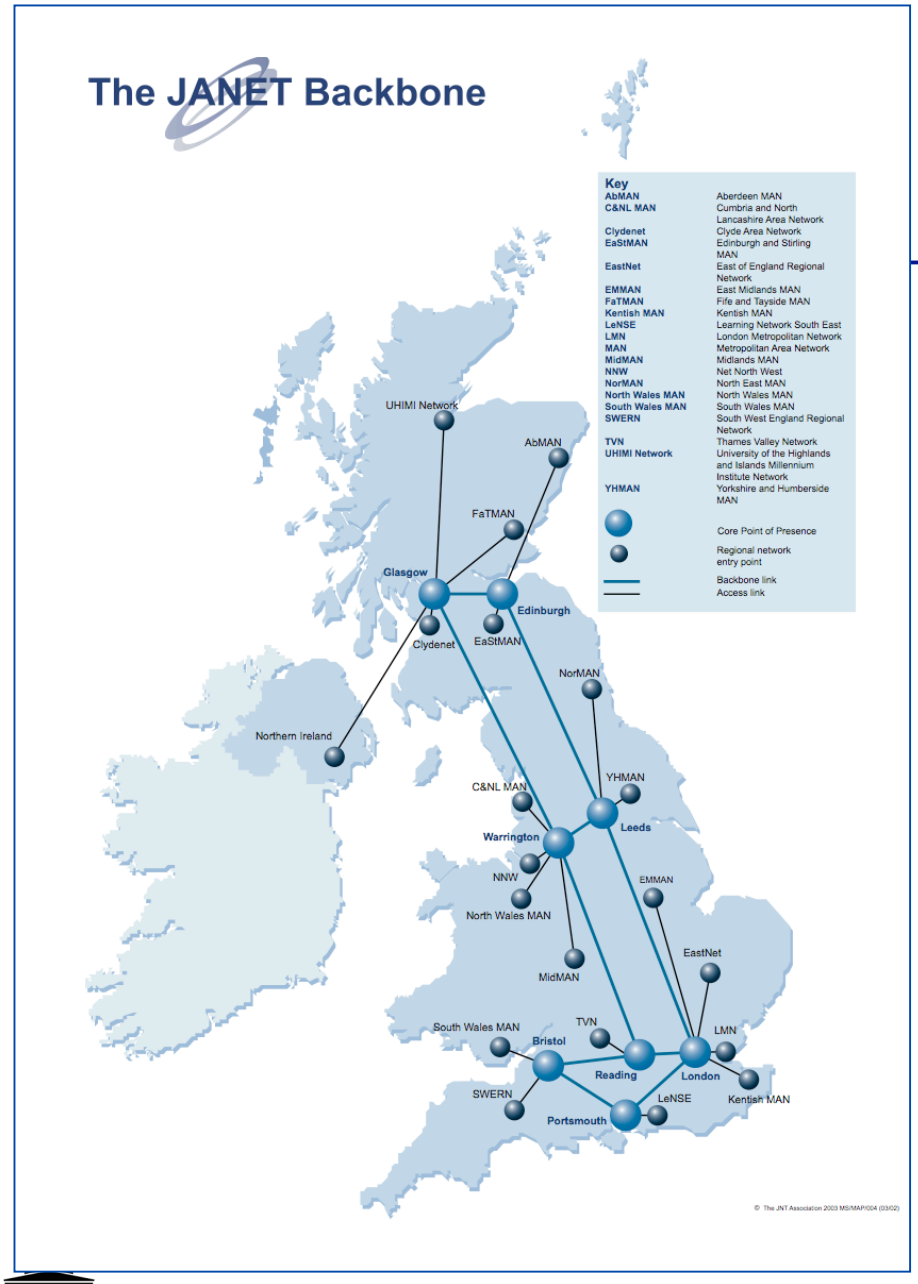
Control at the core and edge network

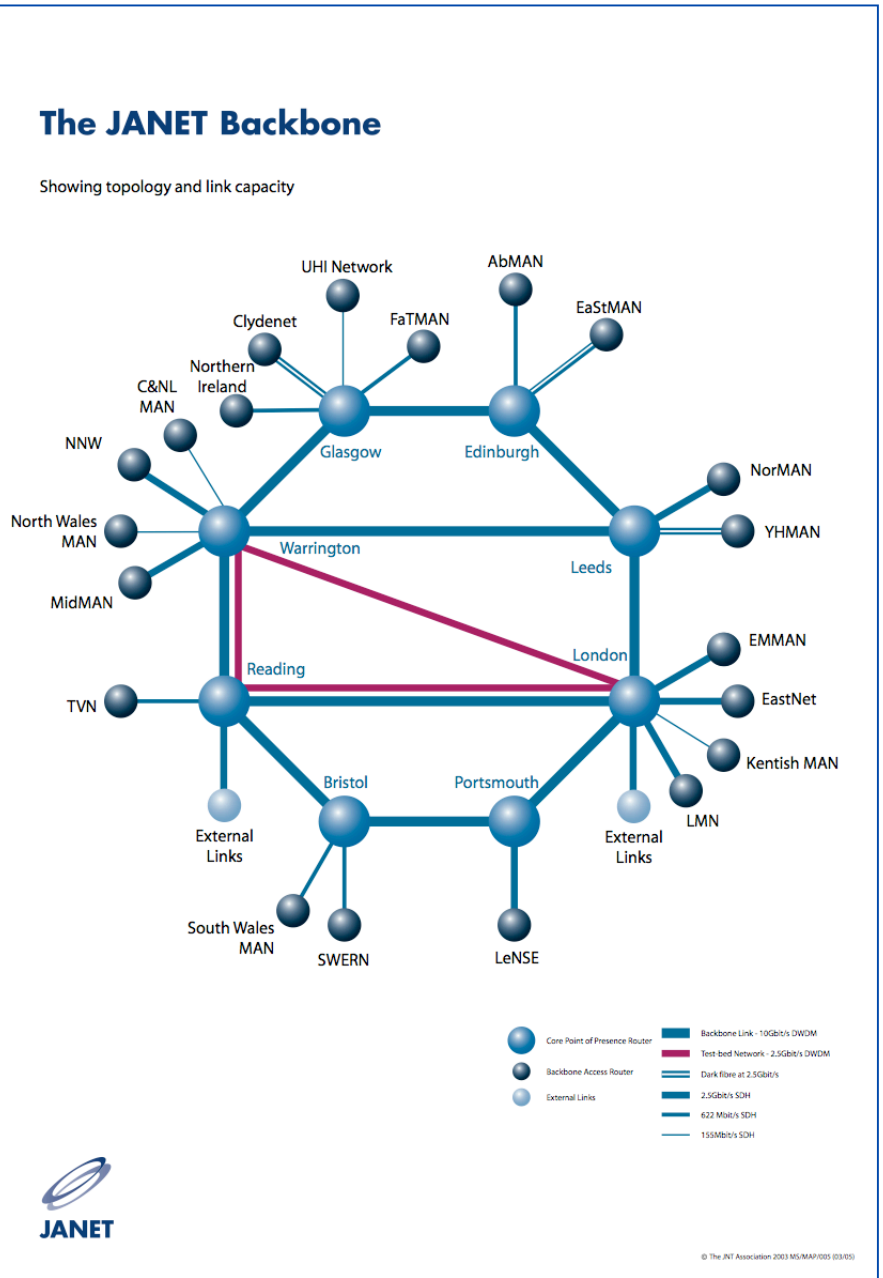
MB-NG http://www.mb-ng.net/
Managed Bandwidth Next Generation
~£900K

GRS http://www.cs.ucl.ac.uk/research/nrs/
Grid Resource Scheduling
~£138K

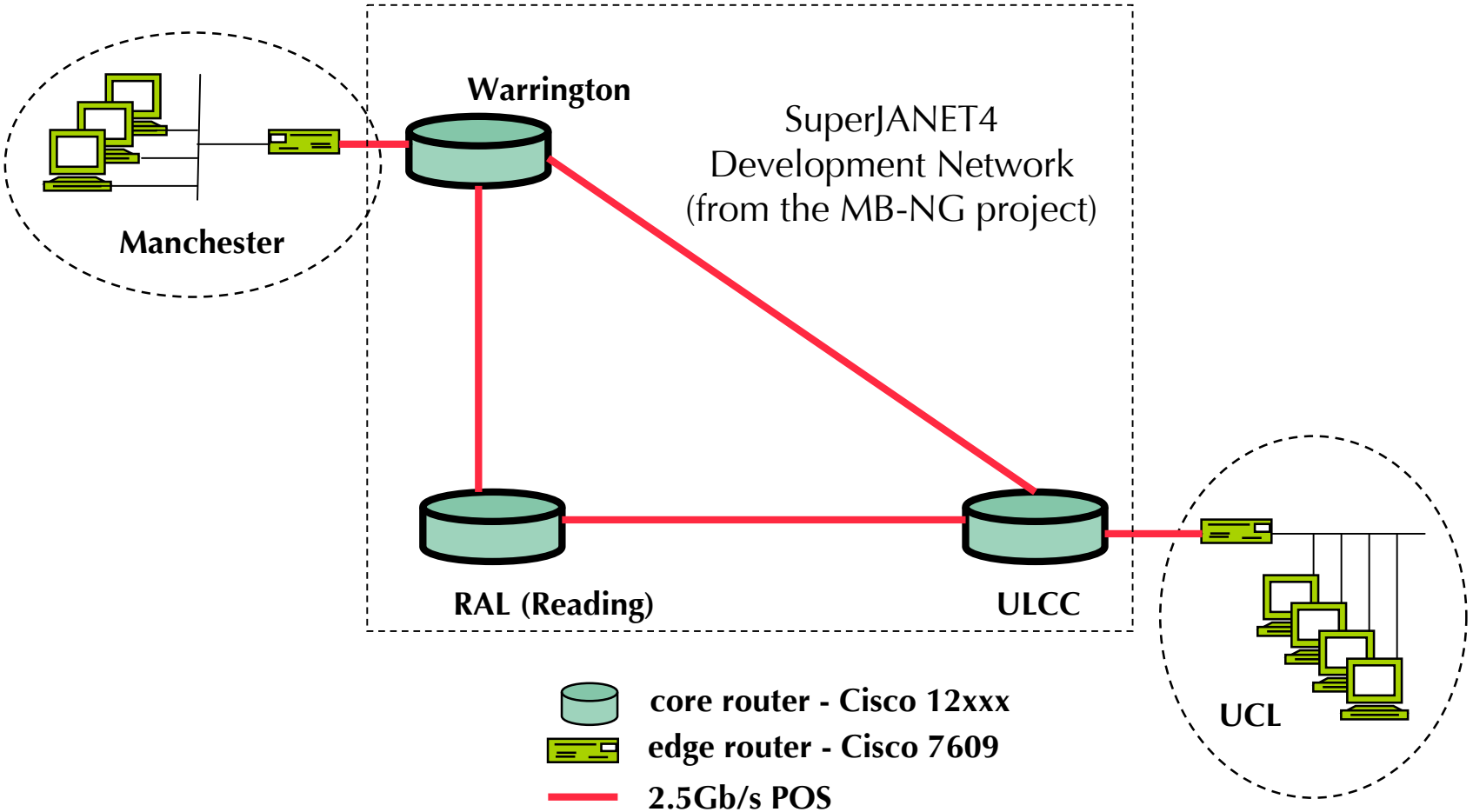
NRS and MB-NG

- MB-NG:
  - SuperJANET4 **development network**
  - DIFFSERV across a **network core**
  - high-speed TCP experiments
  - (dynamic QoS-control using NRS)
- NRS:
  - **site-to-site *per-flow* QoS reservations**
  - decentralised system
  - application-level APIs
  - (use MB-NG network as test-bed)

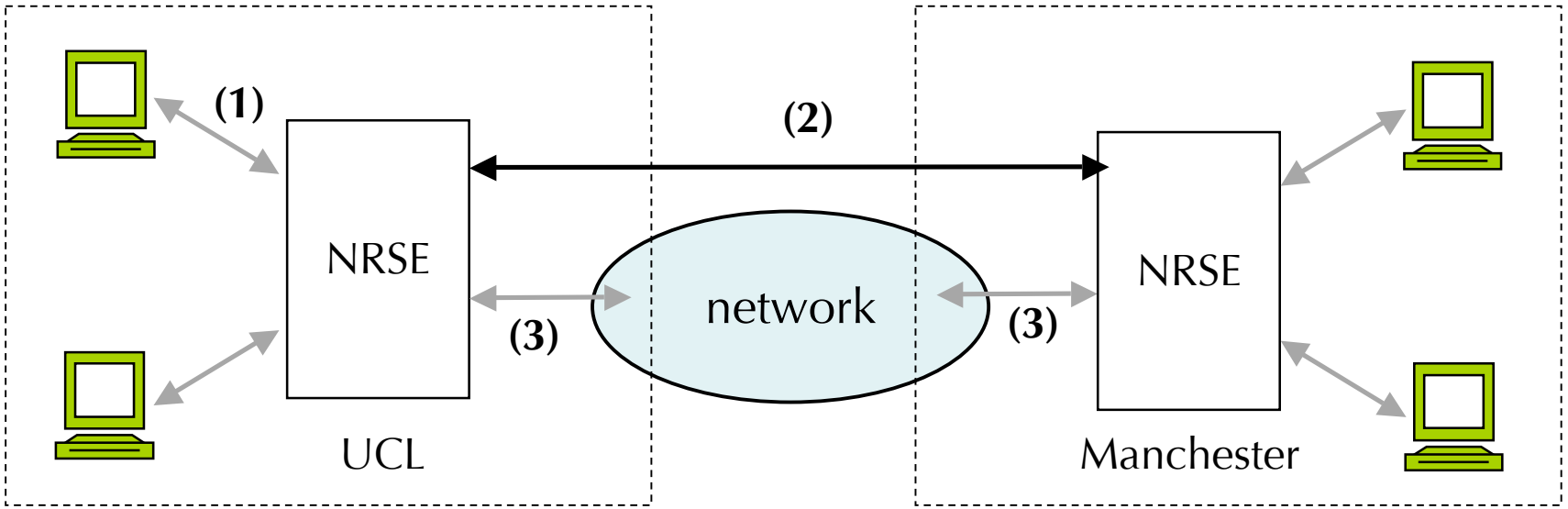

The JANET Backbone
Key
AbMAN Aberdeen MAN
C&NL MAN Cumbria and North Lancashire Area Network
Clydenet Clyde Area Network
EaStMAN Edinburgh and Stirling MAN
EastNet East of England Regional Network
EMMAN East Midlands MAN
FaTMAN Fife and Tayside MAN
Kentish MAN Kentish MAN
LeNSE Learning Network South East
LMN London Metropolitan Network
MAN Metropolitan Area Network
MidMAN Midlands MAN
NNW Net North West
NorMAN North East MAN
North Wales MAN North Wales MAN
South Wales MAN South Wales MAN
SWERN South West England Regional Network
TVN Thames Valley Network
UHIMI Network University of the Highlands and Islands Millennium Institute Network
YHMAN Yorkshire and Humberside MAN
Core Point of Presence
Regional network entry point
Backbone link
Access link
UHIMI Network
AbMAN
FaTMAN
Glasgow
Edinburgh
Clydenet
EaStMAN
NorMAN
Northern Ireland
C&NL MAN
YHMAN
Warrington
Leeds
NNW
North Wales MAN
EMMAN
EastNet
MidMAN
South Wales MAN
TVN
Bristol
LMN
Reading
London
SWERN
LeNSE
Kentish MAN
Portsmouth
© The JNT Association 2003 MS/MAP/004 (03/02)


The JANET Backbone
Showing topology and link capacity
UHI Network
AbMAN
EaStMAN
Clydenet
FaTMAN
Northern Ireland
C&NL MAN
NNW
Glasgow
Edinburgh
NorMAN
North Wales MAN
YHMAN
Warrington
Leeds
MidMAN
EMMAN
Reading
London
TVN
EastNet
Kentish MAN
Bristol
Portsmouth
LMN
External Links
External Links
South Wales MAN
SWERN
LeNSE
Core Point of Presence Router
Backbone Access Router
External Links
Backbone Link - 10Gbit/s DWDM
Test-bed Network - 2.5Gbit/s DWDM
Dark fibre at 2.5Gbit/s
2.5Gbit/s SDH
622 Mbit/s SDH
155Mbit/s SDH
JANET
© The JNT Association 2003 MS/MAP/005 (03/05)

Experiments: SuperJANET4



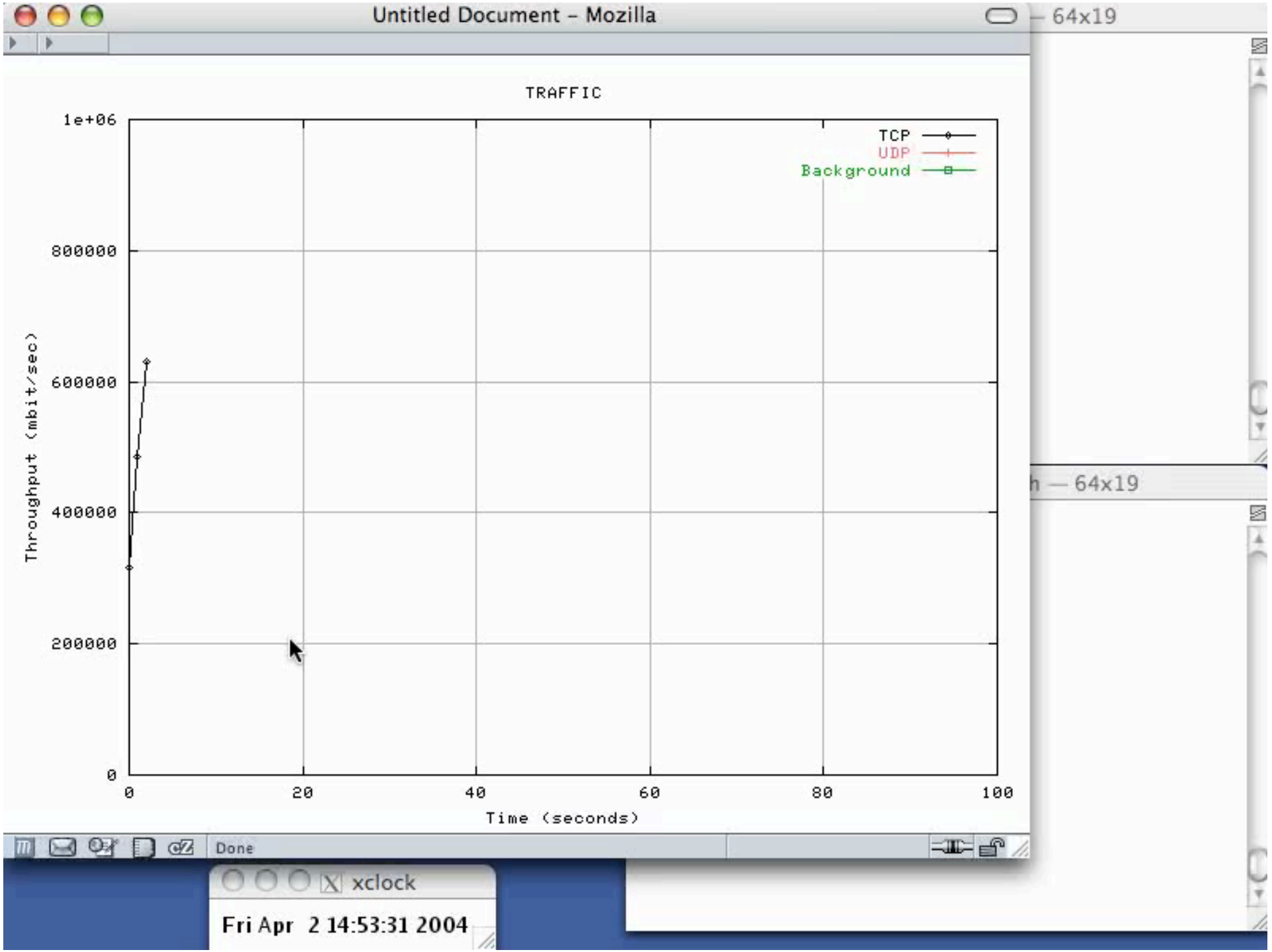
Software for multi-site QoS control



(1) Client contacts local NRSE (Network Resource Scheduling Entity) with a service-level agreement (SLA): local interaction is using a **client-server** model

(2) Local NRSE "negotiates" with remote NRSE at other site: site-to-site interaction is **peer-to-peer**

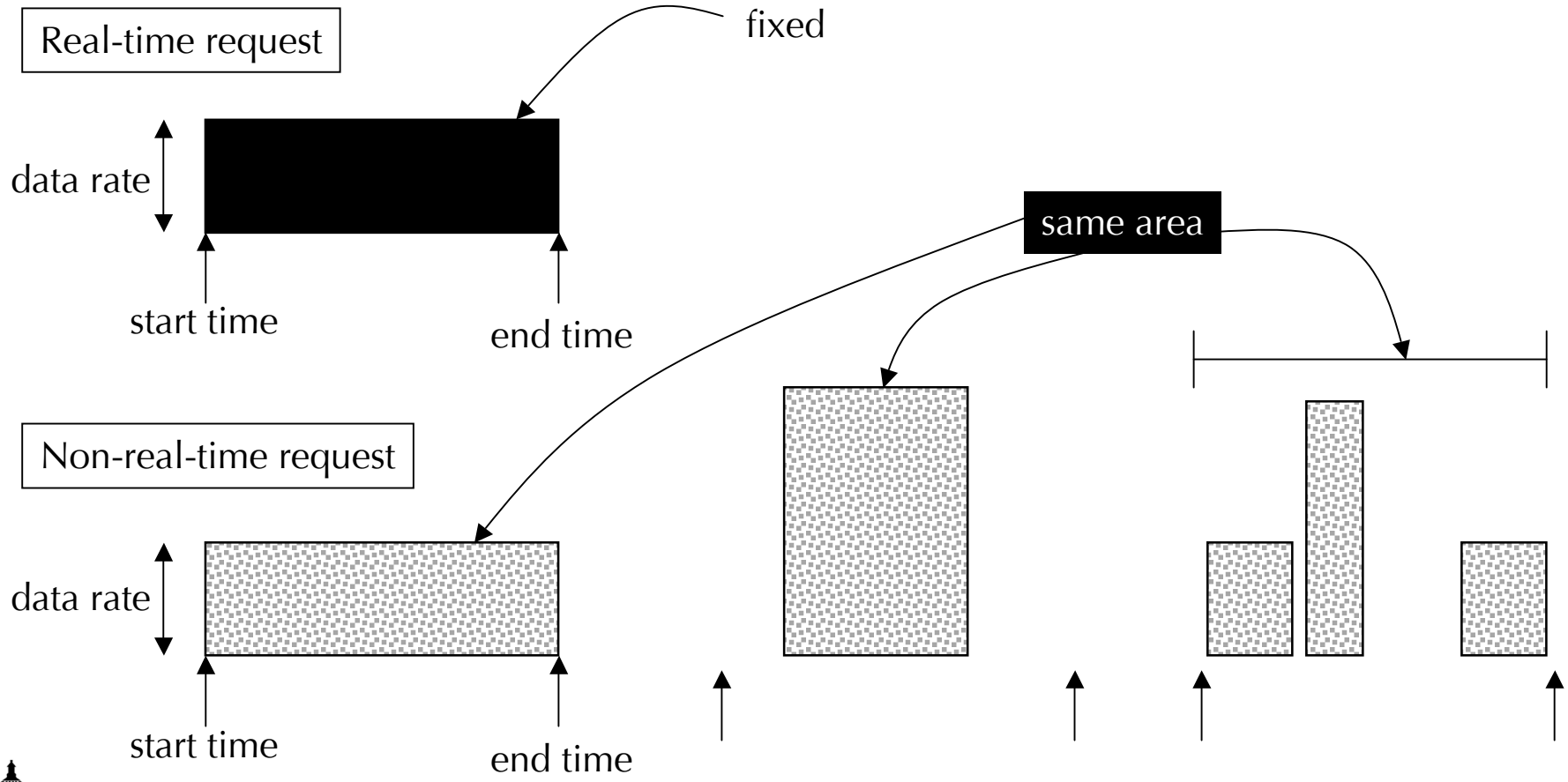
(3) At the correct times, the NRSEs send instructions to the network to implement the SLA


Untitled Document - Mozilla
TRAFFIC
TCP
UDP
Background
1e+06
800000
600000
400000
200000
0
Throughput (mbit/sec)
0
20
40
60
80
100
Time (seconds)
Done
xclock
Fri Apr 2 14:53:31 2004
64x19
64x19

Activities

- QoS state management between sites:
  - DIFFSERV EF
- QoS control software (NRS):
  - dynamic, on-demand, site-to-site QoS control
  - peer-to-peer trust model
  - 'single sign-on' capability
  - modular system - other sub-net technologies
  - released in 20 May 2004 (open source)
    http://www.cs.ucl.ac.uk/research/nrs/
    latest release v2.1, 9 February 2005

Flexible reservations (in progress)



The future - questions raised

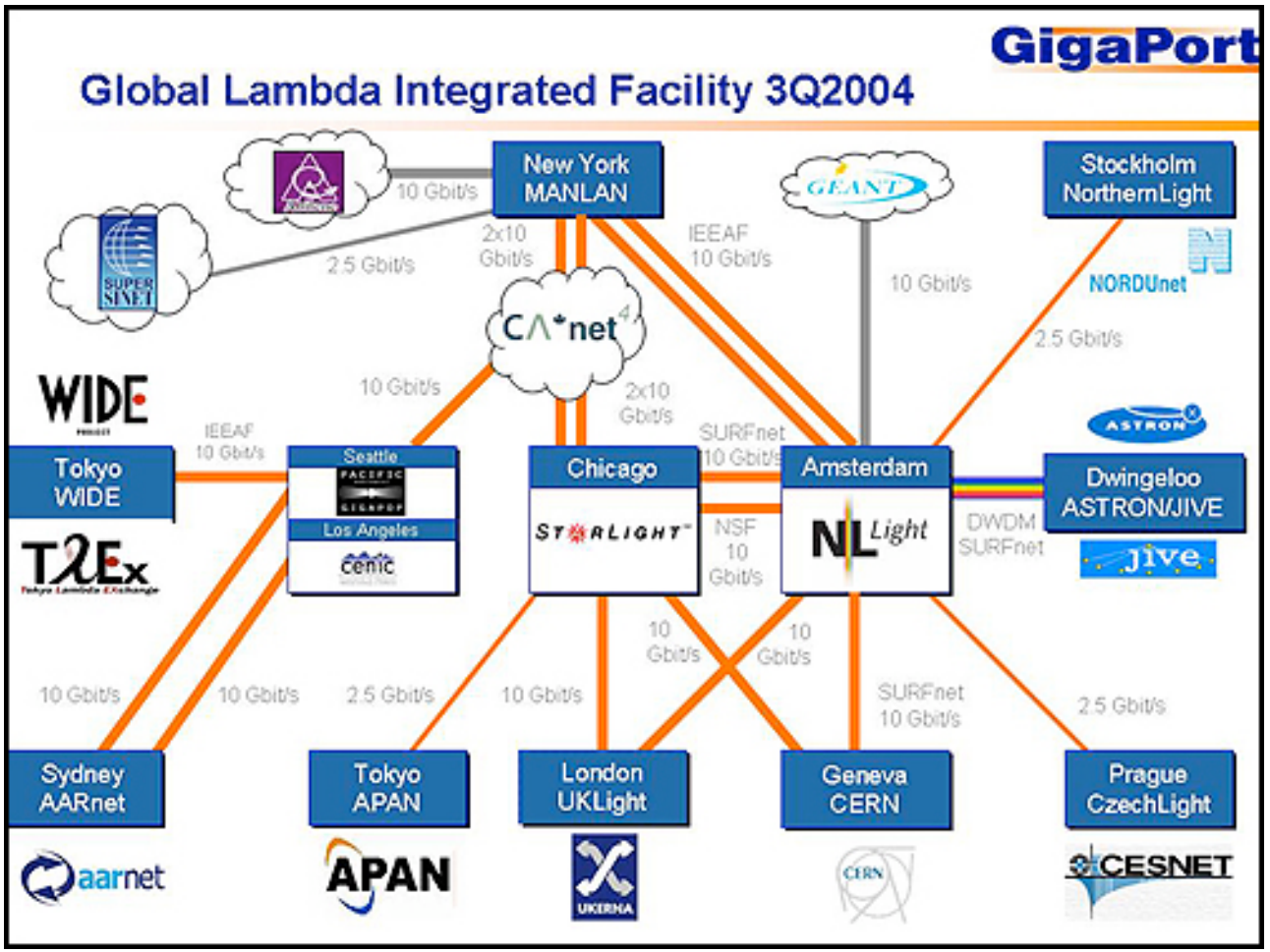
- What happens with a more heterogeneous traffic mix?
- How does the system scale?
  - number of flows?
  - amount of traffic?
  - size of network?
  - heterogeneity of traffic?
- What is a sensible profile of mechanisms to use?
  - congestion control? (HS-TCP? Scalable-TCP? ECN? XCP? etc.)
  - queue management?
  - QoS control?
- Is TCP (+ mods) actually useful for mulit-Gb/s speeds?
  - some other, **new** transport protocol perhaps? e.g. DCCP?
- How should the application get involved?

UKLIGHT

- The need for a high-speed network facility **for research**:
  - transmission/photonics and switching
  - networking
  - applications
- Link to other existing facilities world-wide:
  - Global Lambda Integrated Facility (GLIF)
  - http://www.glif.is/

UKLIGHT http://www.uklight.ac.uk/
~£6.5M
many contributors - http://www.cs.ucl.ac.uk/staff/S.Bhatti/uklight-proposal/

UKLIGHT - GLIF



from http://www.glif.is/

Current work - just starting

46PaQ http://www.cs.ucl.ac.uk/research/46paq/
~£600K, IPv4 & IPv6 Performance and QoS

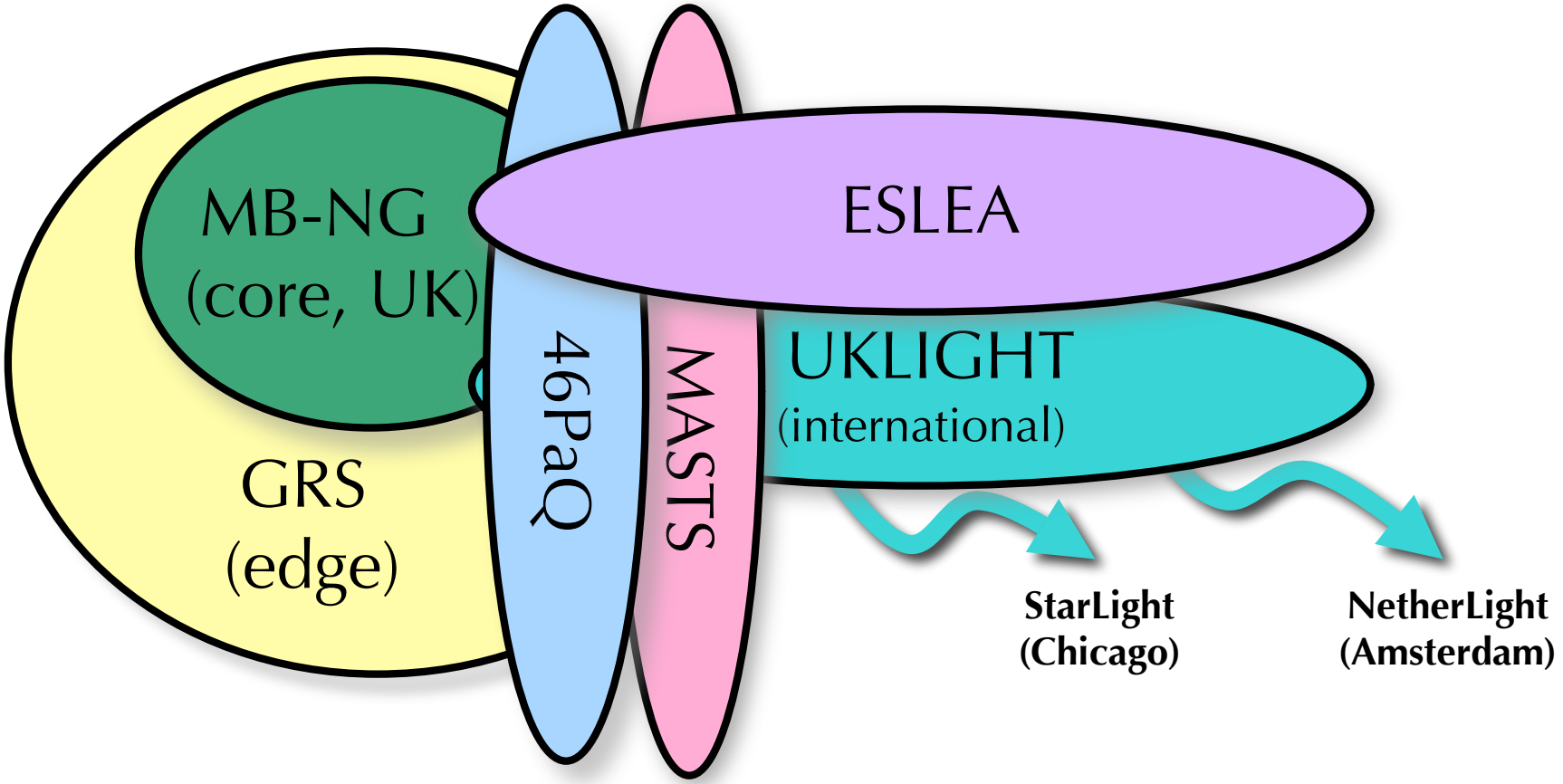
MASTS http://www.ee.ucl.ac.uk/~lsacks/acse/masts/
~£1.1M, Measurements at All Scales in Time and Space

ESLEA
Exploiting Switched Lightpaths for E-science Applications
~£1.1M

MRC



Project links



GRS - http://www.cs.ucl.ac.uk/research/nrs/
MB-NG - http://www.mb-ng.net/
UKLIGHT - http://www.uklight.ac.uk/

46PaQ - http://www.cs.ucl.ac.uk/research/46paq/
MASTS - http://www.ee.ucl.ac.uk/~lsacks/acse/masts/
ESLEA - URL TBA

Potential output

- High-speed experiments & performance analysis:
  - methods, metrics, tools
- Better (performing) implementations:
  - end-systems design and engineering
- Feedback to applications:
  - performance-based adaptation for applications
- Traffic models:
  - network management, planning and design
- Network engineering insights

Other questions forming …

- How do we design and engineer high-speed protocols and networked systems that:
  - make efficient use of QoS and congestion control?
  - have appropriate instrumentation in the network?
  - can provide feedback to applications?
- How can applications react in a timely manner to feedback from the network?
  - change the design of applications?
  - new transport protocols and/or systems APIs?
  - instrumentation and engineering of the end-systems?
- Watch this space …. ☺

Thank you!
Questions