Grid networking in EU DataGRID

TERENA conference

Limerick - 5th of June 2002

Cees de Laat (UvA)

on behalf of

Pascale PRIMET

Manager of the workpackage "Network" of the DataGRID project

INRIA/ RESO - ENS Lyon

Pascale.Primet@ens-lyon.fr

Contents of this talk

- To save time: Wait and see

European DataGRID project

- The EDG project http://www.eu-datagrid.org/ aims to provide production quality testbeds, using real-world applications with real data:
  - High Energy Physics
    - process the huge amount of data from LHC experimentations
  - Biology and Medical Imaging
    - sharing of genomic databases for the benefit of international cooperation
    - processing of medical images for medical collaborations
  - Earth Observations
    - access and analysis of atmospheric ozone data collected by satellites as Envisat-1
- Calendar : january 2001 to december 2003
- Funded by the European Union

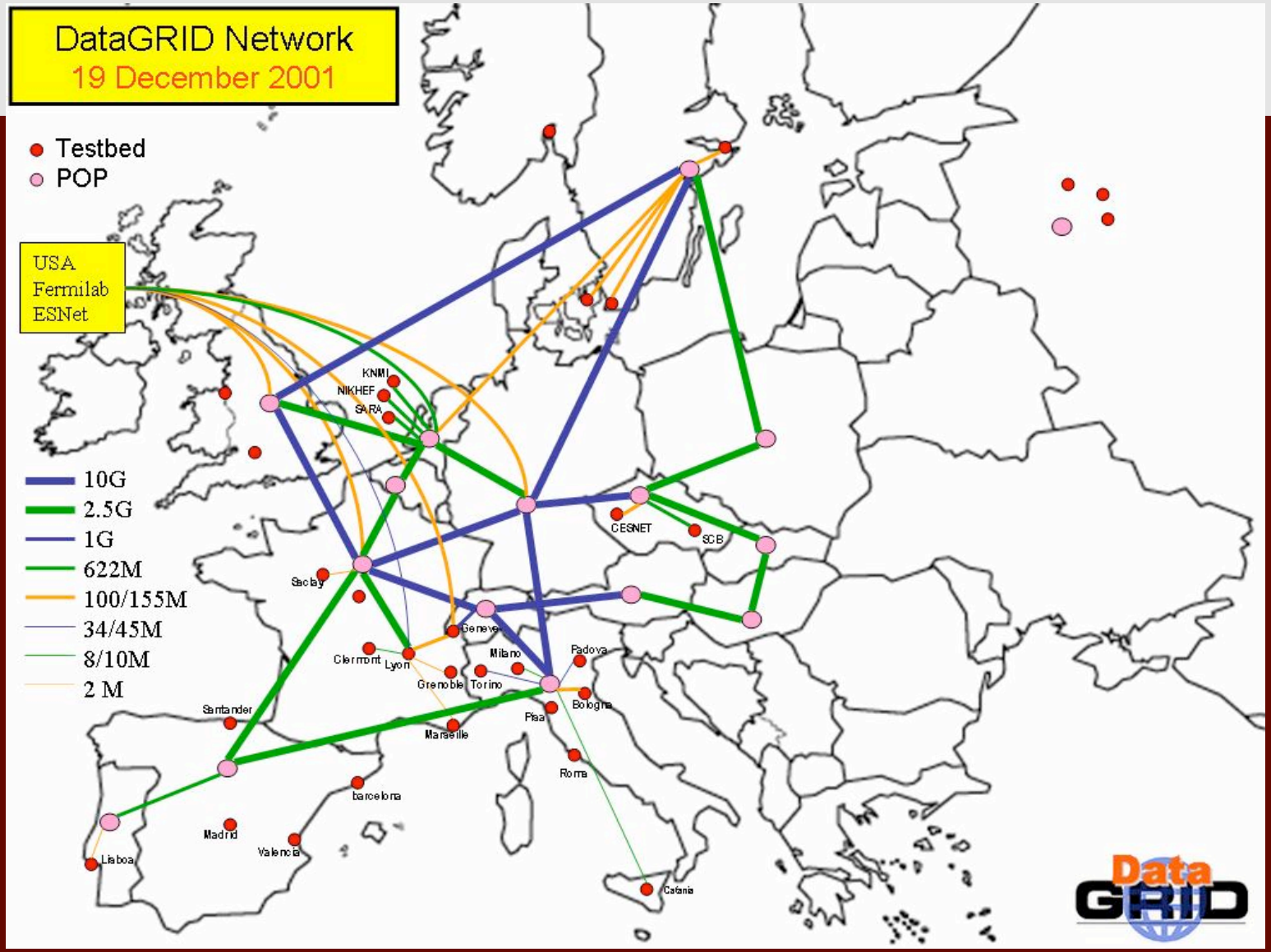
EDG - Partners

- CERN – France
- CNRS – France
  - Testbed (WP6)
  - Network (WP7)
  - Bio application (WP10)
- ESA/ESRIN – Italy
- INFN – Italy
- NIKHEF – The Netherlands
- PPARC - UK

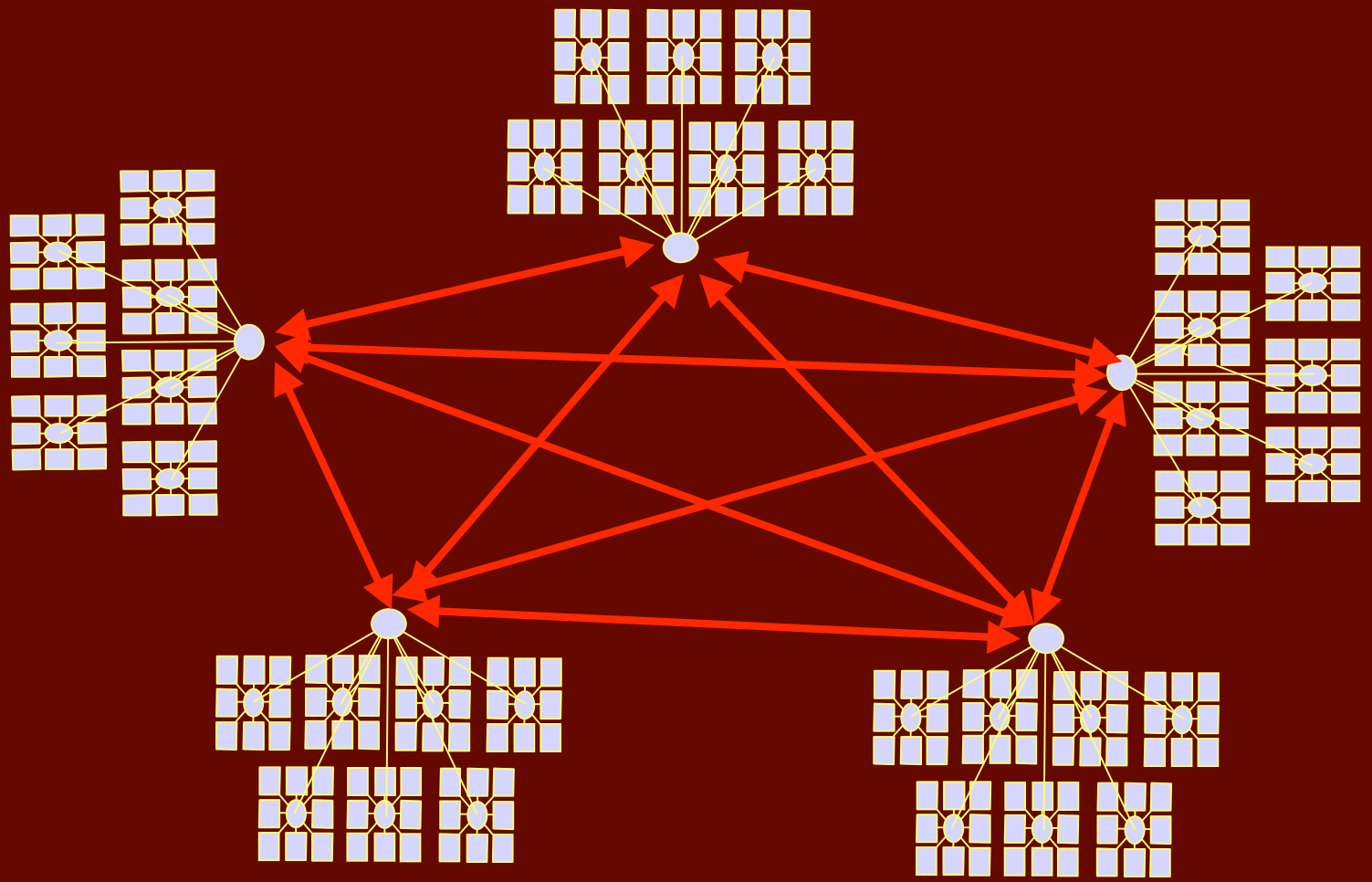
European DataGRID project

- 7 applications distributed among 6 virtual organisations
- 11 organisations over 15 countries
- 40 sites in Europe
- Based on the European GEANT backbone and National NREN's

http://ccwp7.in2p3.fr


DataGRID Network
19 December 2001
Testbed
POP
USA
Fermilab
ESNet
10G
2.5G
1G
622M
100/155M
34/45M
8/10M
2 M
KNMI
NIKHEF
SARA
CESNET
SCB
Saclay
Geneve
Clermont
Lyon
Grenoble
Torino
Milano
Padova
Bologna
Pisa
Roma
Marseille
Santander
barcelona
Madrid
Valencia
Lisboa
Catania

Logical view of the Grid Network



EDG WP7 activities

T7.1 : Technical Collaboration with Dante/NRENs

- Pilot services test (QoS, multicast)
- Dedicated machines in GEANT PoPs

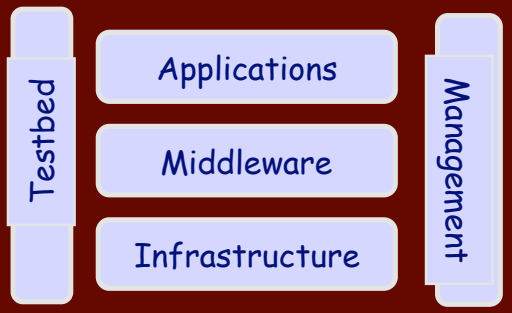
T7.2 : QoS and advanced services

- QoS services test with biological/medical applications
- Reliable Multicast Protocol test and deployement
- High performance transport protocol (TCP/nonTCP)

T7.3 : Network Monitoring Architecture

- Design and deploy a Network Monitoring Infrastructure
- Visualize and analyze monitoring data

T7.4 : Security => EDG Security team



Collaboration with GEANT

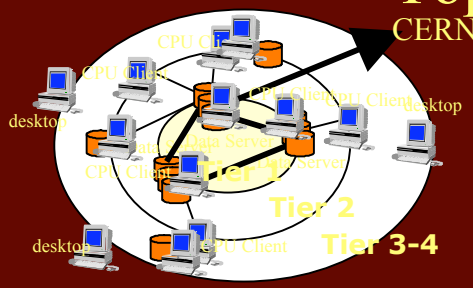
- E2E : Close participation to pilot services
  - Test of IP Premium service/WP10
- In Backbone : *(our proposal)*
  - Use of dedicated machines in GEANT POPs
    - Amsterdam, Geneva, London
  - Tests of high throughput transfers
  - Test of IP multicast for Reliable Multicast
  - Sharing WP7 monitoring and DANTE monitoring data

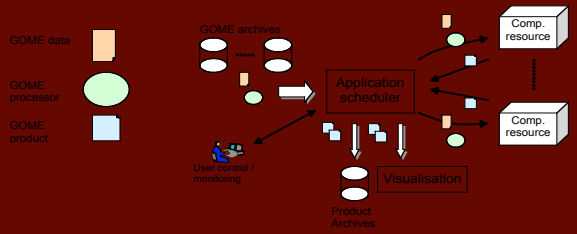
Network provisioning

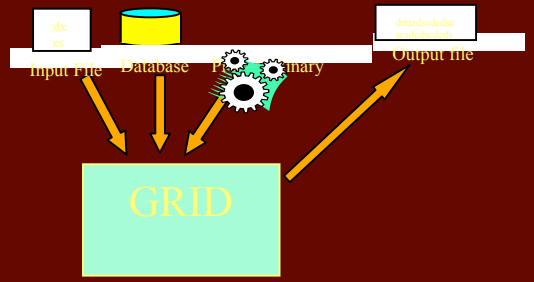
- Network Requirements studies
  - Application Requirements (WP8, WP9, WP10)
  - Middleware Requirements
- Physical Networks
  1. GEANT : 2.5 Gbps to 10 Gbps
  2. NRENs : from 155Mbps (or less) to 2.5Gbps
  3. Regional networks: from 2Mbps to 155Mbps
  4. Local Area Networks : from 10Mbps to 1Gbps)
- Is a « Virtual Private Network » required for the DataGRID ?
  - concept definition / VPN technologies review
- See our D7.1 document on WP7 EDG site

Application requirement studies

Top down stream identification







WP8 WP9 WP10

Flows list

N°	Name	Application	WP	Type	Transfert volume (Mbytes)	Frequency (in days)	Average bitrate (Kbit/s)	constraints	Observations
1	Monté Carlo Data Réplication	LHCb	8	gridftp	30 000	24	66 667		
2	ENVISAT Data from ground station to storage centre	METEO	9	tcp	5 000 000	0	66 138		
3					10	86 400	80 000		

Some numbers

- HEP applications:
  - Bulk Data transfer : from 100Mb/s (TB1) to 1Gb/s cont. (TB3)
- Medical applications:
  - Interactive Traffic with burst of more than 1Gbyte
  - Real Time High Performance Vizualisation/Simulations

Network performances measurement (1)

For Provisioning:

- To be available, via visualization to human observer (user, network/system administrators)
- To provide tools for network performances measurement, problems identification and resolution (bottlenecks, point of unreliability, quality of service needs, topology…)
- To achieve network performance forecast and optimization – Capacity planning

Network performances measurement (2)

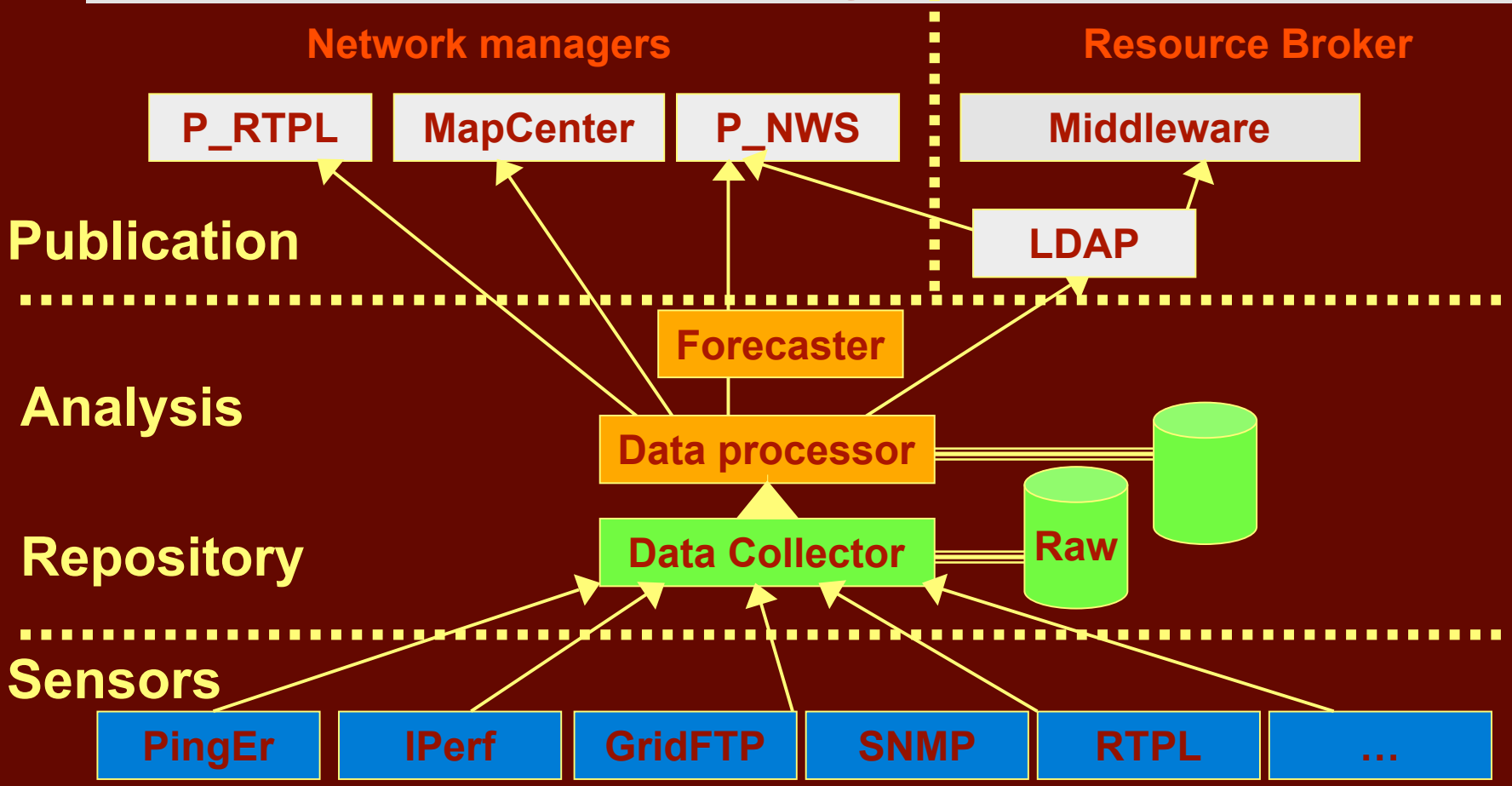
For Resource Brokers:

- Network performance parameters are used for optimizing resource allocation (replication, MPI, Remote file access…)
- Network performance metrics must
  - be published to the Grid Information System
  - Be accessible through aggregated functions called by Grid resource broker services (computing and data storage).

Architectural design

- four functional units :
  - monitoring tools or sensors
  - a repository for collected data;
  - the means for data analysis to generate network metrics;
  - the means to access and to use the derived metrics.
- See our D7.2 document on WP7 EDG site

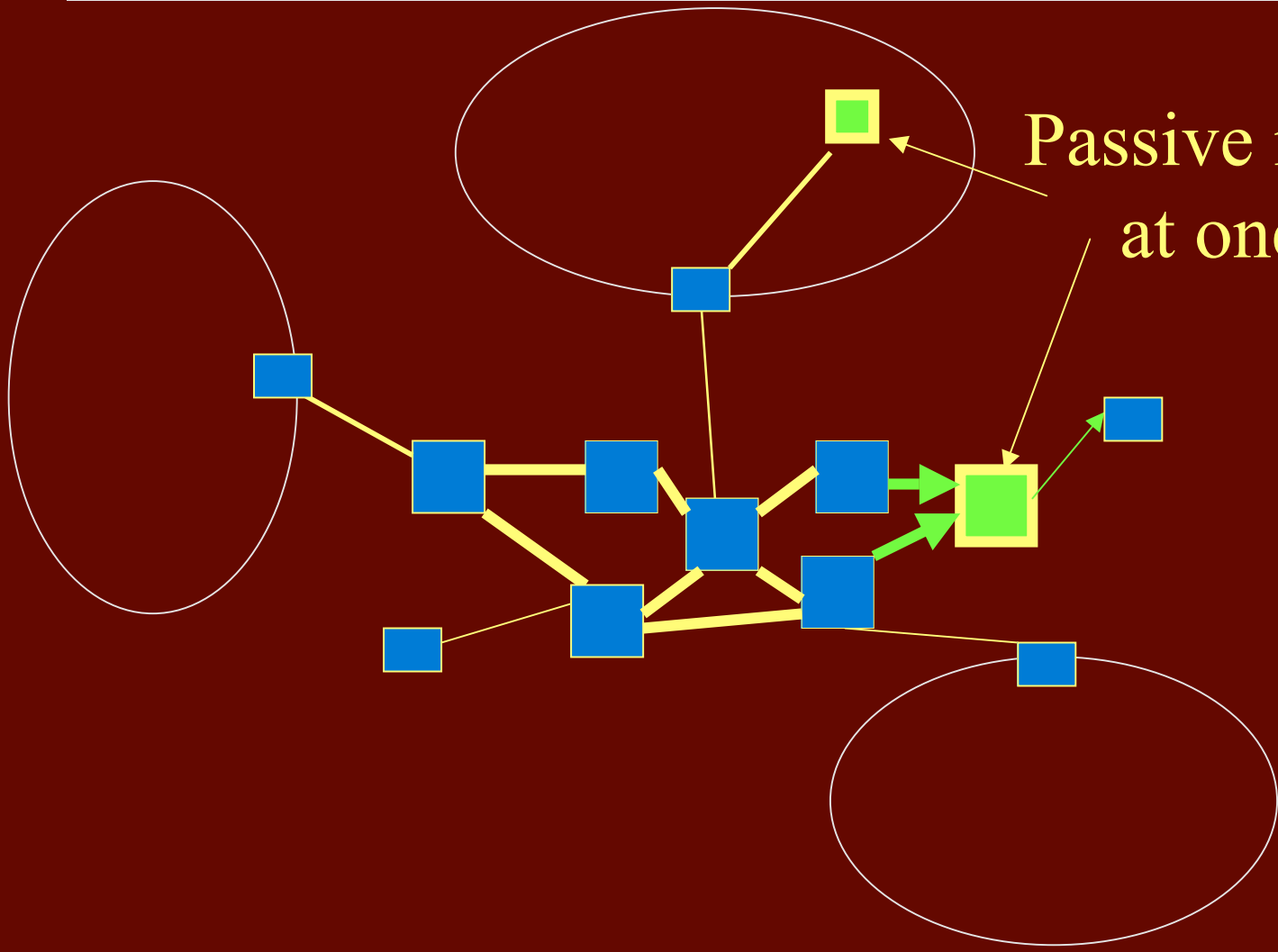
Network Monitoring Architecture



Measurement methods

- Active methods
  - Injection of traffic inside the network for testing performances between two points
  - problem: may be intrusive (TCP/UDP throughput)
- Passive methods
  - Collect traffic informations in one point of the network : router, switch, dedicated passive host, computing element or storage element (GRIDftp logs)…
  - Problem : give network usage, not capacity

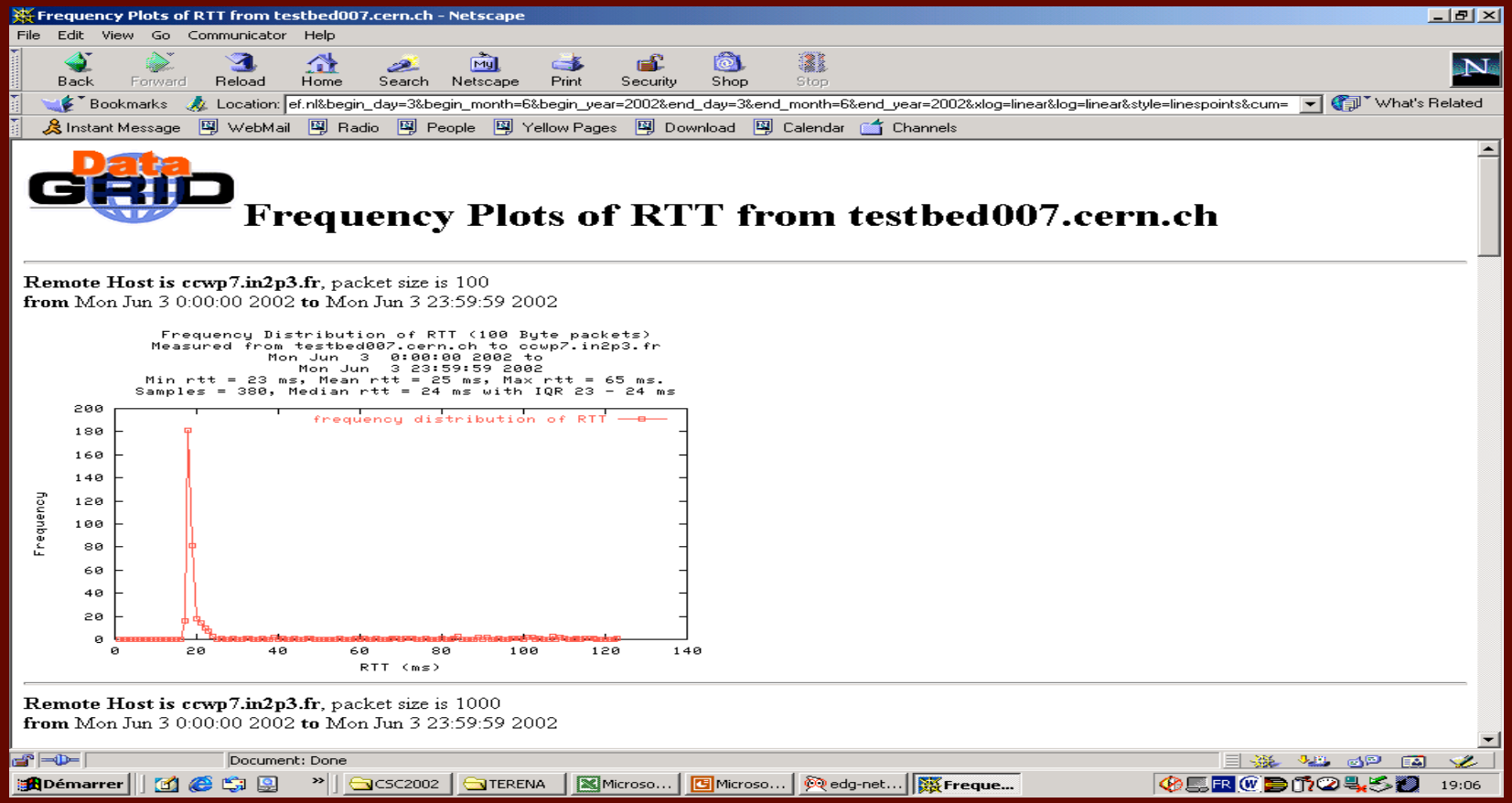
Passive measurement



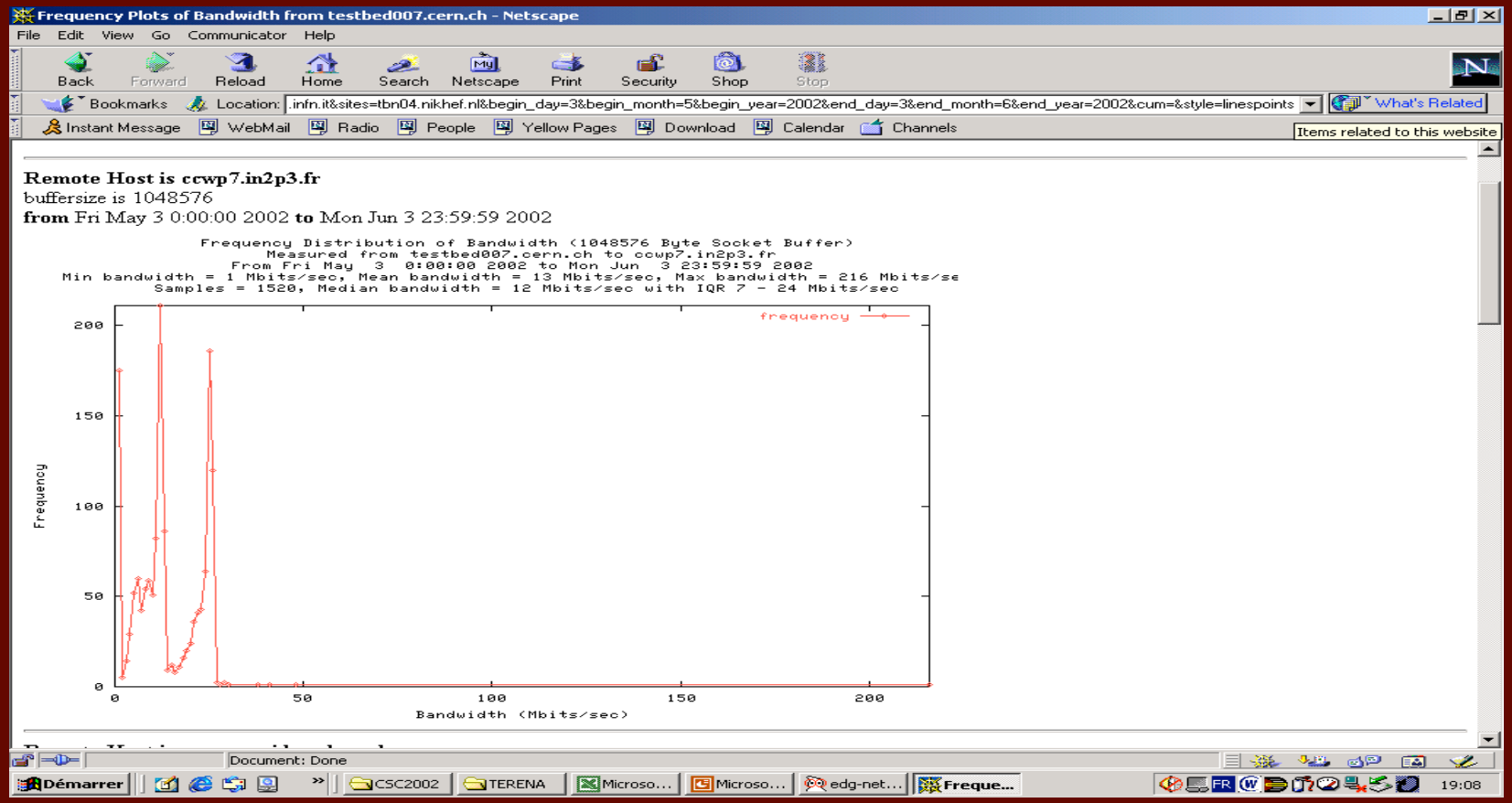
Metrics and tools

- Round Trip Delay => PinGER (Lyon->nikhef)
- Packet Loss => PinGER (Lyon->nikhef)
- TCP throughput => IPerfER (nikhef -> Ral)
- UDP throughput => UDPMon (CZ->Cern)
- site connectivity => MapCenter
- service availability => MapCenter
- OneWay metrics => RIPEncc test boxes
- End user performance => RTPL (UvA)

PingER results



IPerfER Results



Schema and LDAP backend

- Grid applications/mw are able to access network monitoring metrics via LDAP services according to a defined LDAP schema.
- LDAP back end to make measurements visible through the Globus GIIS/GRIS system has been developed.
  - that fetch, or have pushed, the current metric information from the local network monitoring data store.
- R-GMA is tested as an alternative solution to Globus MDS

http://ccwp7.in2p3.fr/mapcenter

Network Cost functions

Network metrics published in LDAP repositories are used by resource brokers and replica managers through network cost functions :

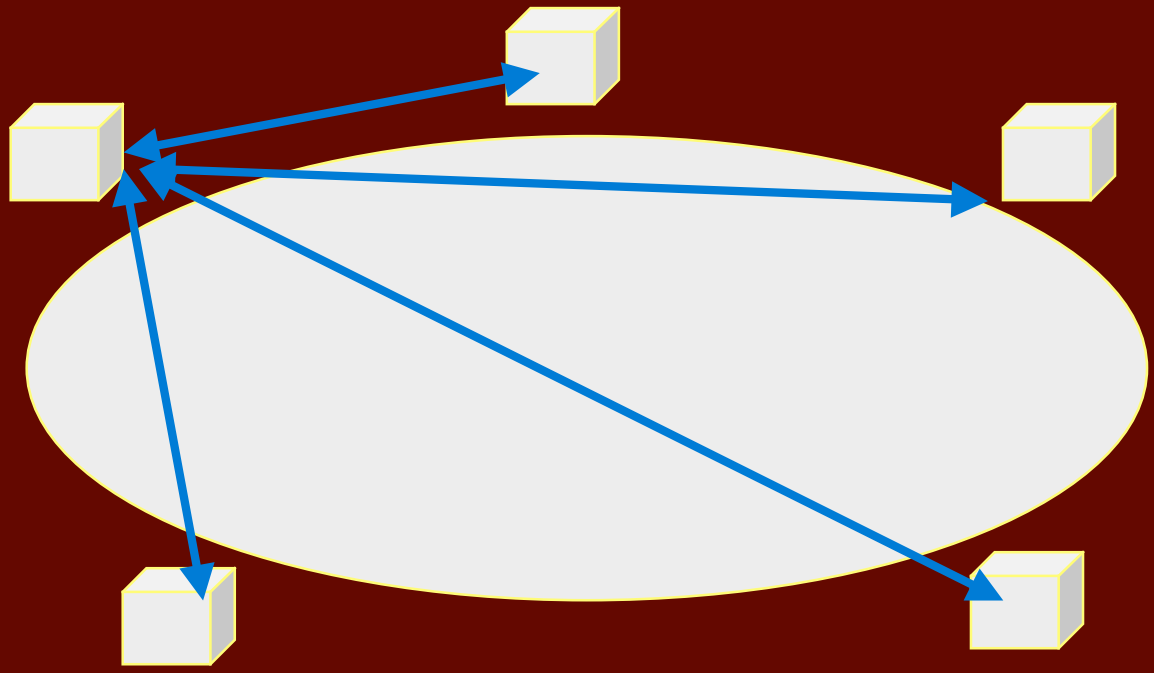
Time = networkCost (SE1, SE2, filesize)

Computed from

1. GridFTP logs
2. ***TCP throughput measurements (aggregated)***
3. RTT Measurements (aggregated)

EDG Network Cost Function

Network Element => Network COST function



EDG MapCenter Tool

- Connectivity of sites
- Availability of services running over all sites involved
- Efficient and flexible model to logically and graphically represent all communities, organization, applications running over grids.
- MapCenter enables representation of any level of abstraction (national and international organizations, virtual organizations, application etc) needed by grid environments.
- http://ccwp7.in2p3.fr/mapcenter

Network and Transport Services

- QoS:
  - Demonstrate and build experience in use of E2E diffserv services in Grid context
  - Feedback experiences to GEANT/DANTE, NRENs and LANs
- Transport
  - High performance transport protocols
  - Reliable multicast protocols tests

High Performance Transport

TCP mechanisms optimization
  – Tests of applicability of new mechanisms
- Use of QoS solutions
  – diminution of Packet Loss
  – Active queue management (WRED, ECN)
  – TCP over DiffServ (AF, EF, PDS, EDS…)
- Reliable Multicast Protocol
  – Test and deployement of JRMS and TRAM

WP7 and other collaborations

- WP7 and EU DataTAG collaboration
  - QoS service study and experiment
  - High Throughput study and experiment
  - Network monitoring and measurement
- GGF
  - GHPN research group
- Other European Grid projects (FR e-toile, UK e-science, INFN grid…)

Conclusion

- In testbed0 and testbed1 the networking functionality was here
  - IP technology: Best effort
  - GEANT has been deployed
  - A Performance Measurement Architecture developed
- In testbed 2 and testbed 3
  - Grid application performance optimization
  - End to end performance analysis
  - Test and provide enhanced network and transport services : Premium, Scavenger, Multicast

For more information

- Consult our sites:
  - http://ccwp7.in2p3.fr
  - http://eu-datagrid.web.cern.ch