Network monitoring in DataGRID project

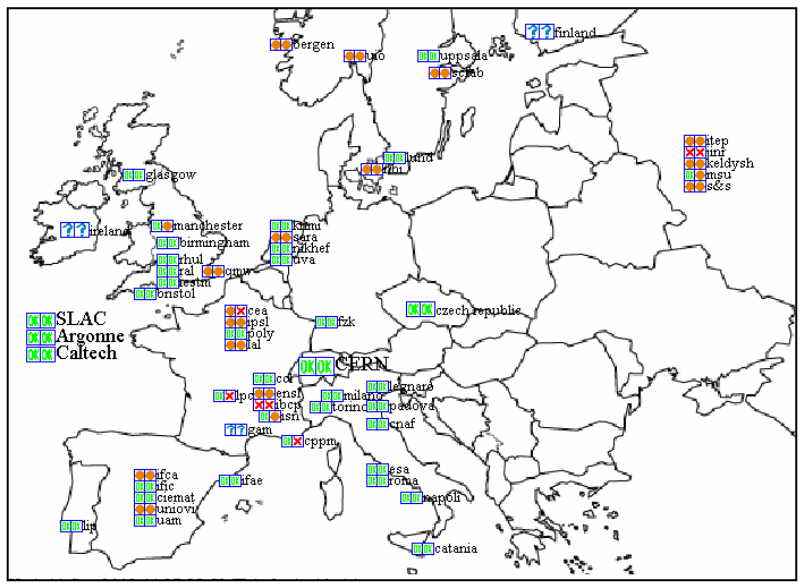
Franck Bonnassieux (CNRS)
franck.bonnassieux@ens-lyon.fr

Outline

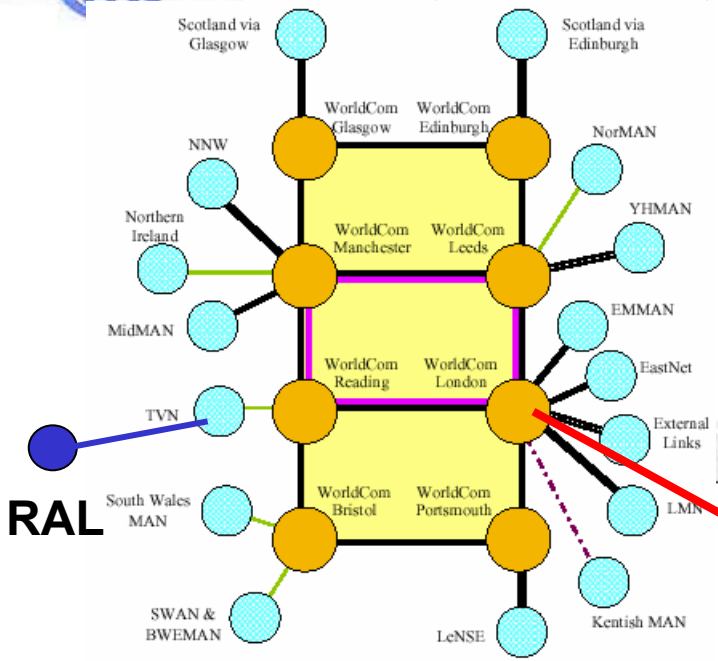
- DataGRID network
- Specificity of Grid network monitoring
- Network metrics and measures
- Architecture of DataGRID network monitoring
- Network sensors and tools
- High level tools
- Perspectives

DataGRID http://www.eu-datagrid.org/

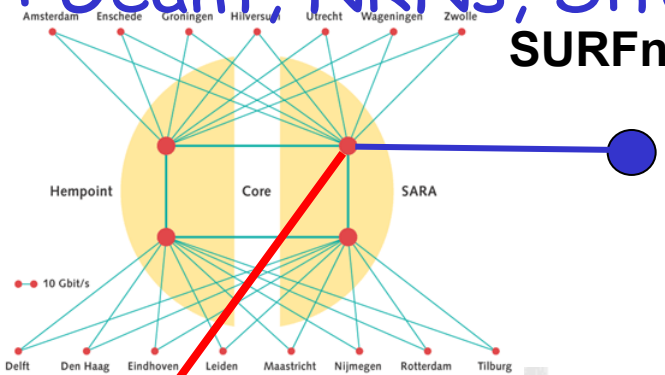
- 7 applications distributed among 6 virtual organisations
- 11 organisations over 15 countries
- 40 sites in Europe
- 12 work packages
- WP7 : network work package
  - provisioning of infrastructure
  - Network and Transport Services
  - Network and Grid traffic monitoring
  - Grid Security



European Topology : Geant, NRNs, Sites

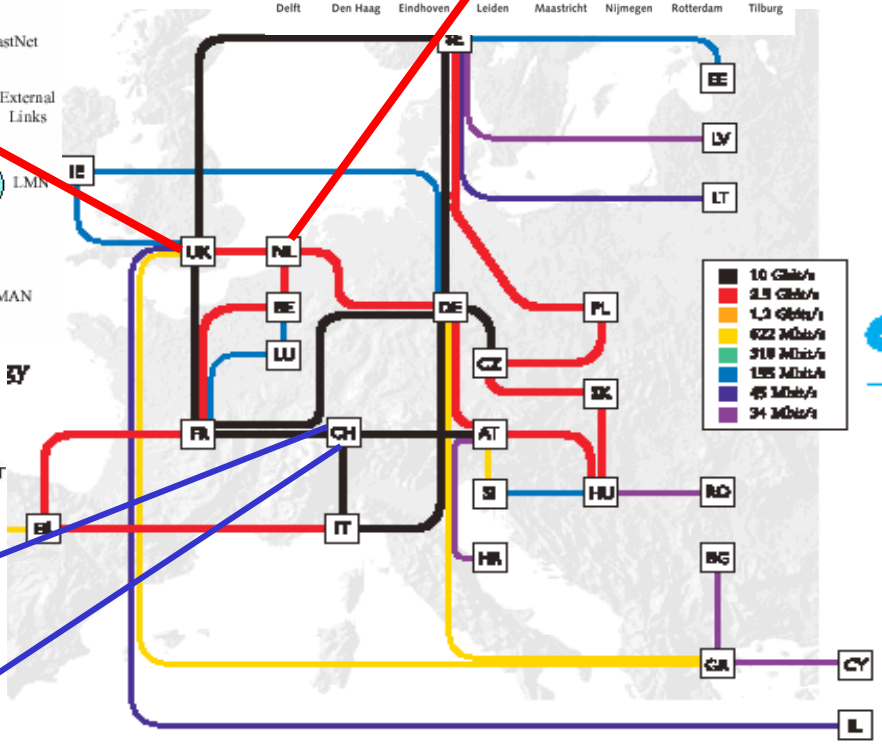


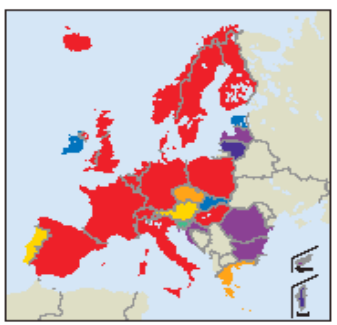
RAL

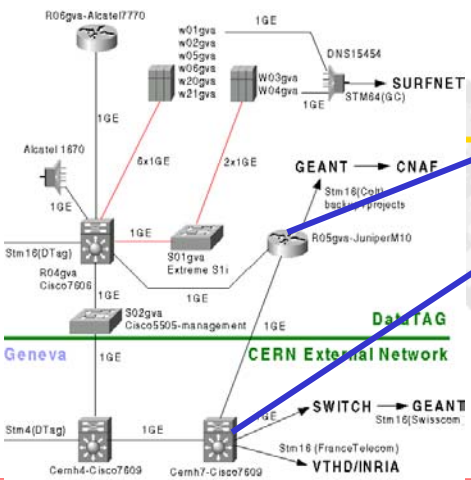


SURFnet

NIKHEF





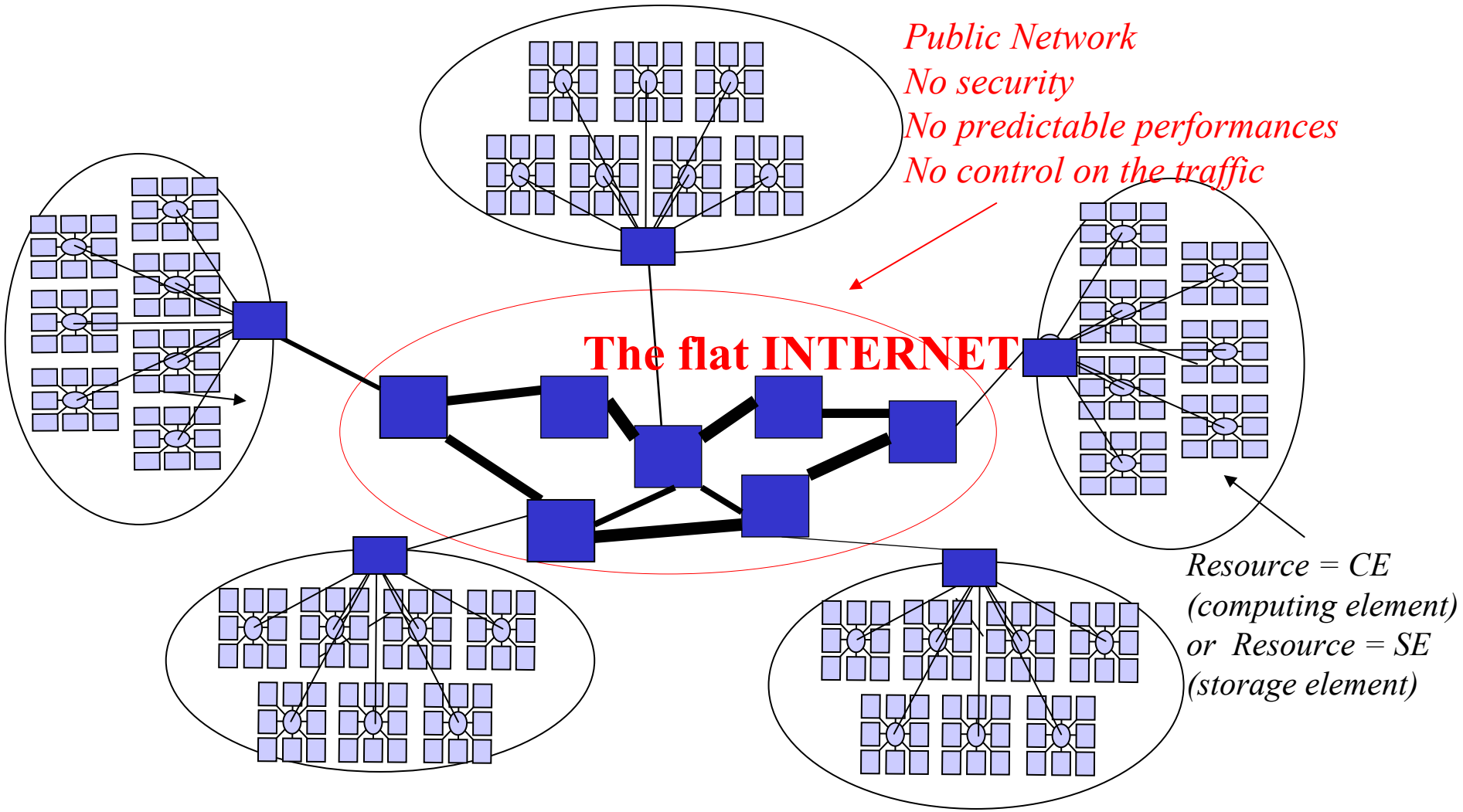


CERN

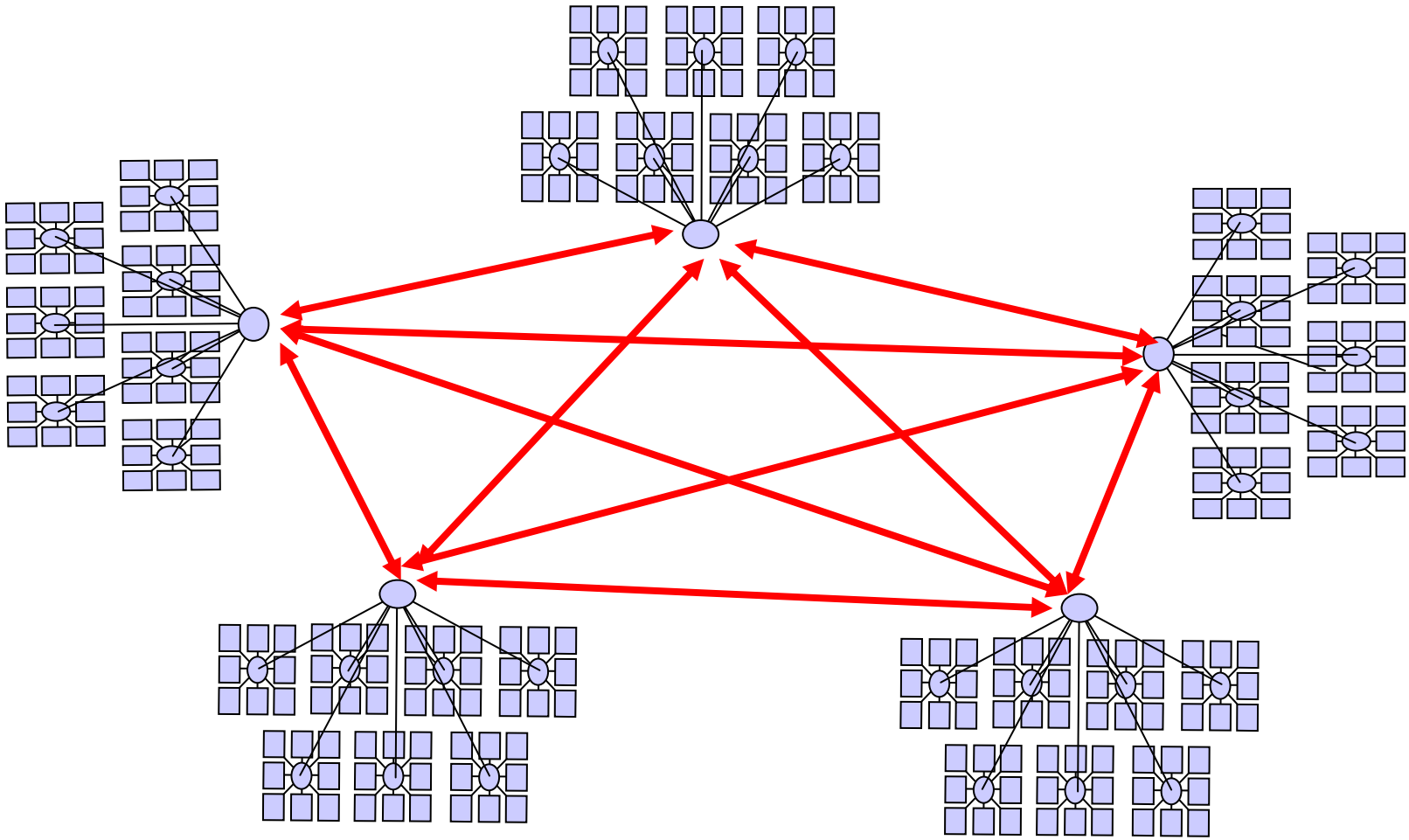
Why network performances monitoring in grid environments ?

- 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
- For Resource Brokers:
  - Network performance parameters are used for optimizing resource allocation (replication, Remote file access…)

"Physical" view of the Network



"Logical" view of the Grid Network



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 of traffic information in one point of the network : router, switch, dedicated passive host, computing element…
  - Problem : give network usage, not capacity

Network metrics and tools

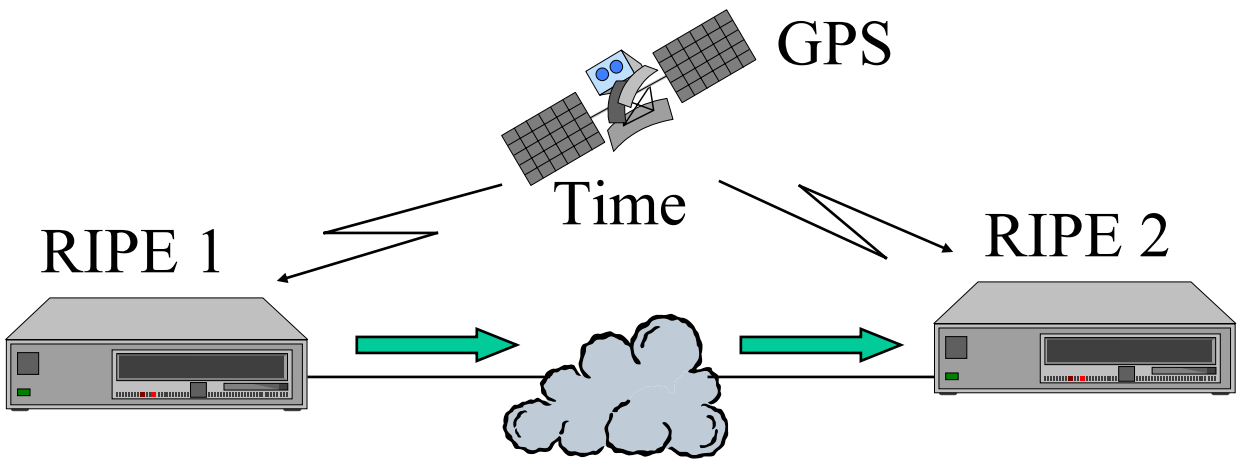
- One Way Delay => Ripe Boxes
- Round Trip Delay => PinGEr
- Packet Loss => PinGEr
- TCP throughput => IPerfEr
- UDP throughput
- Jitter => UDPMon
- Routers traffic => NetLoad Agent

Ripe boxes

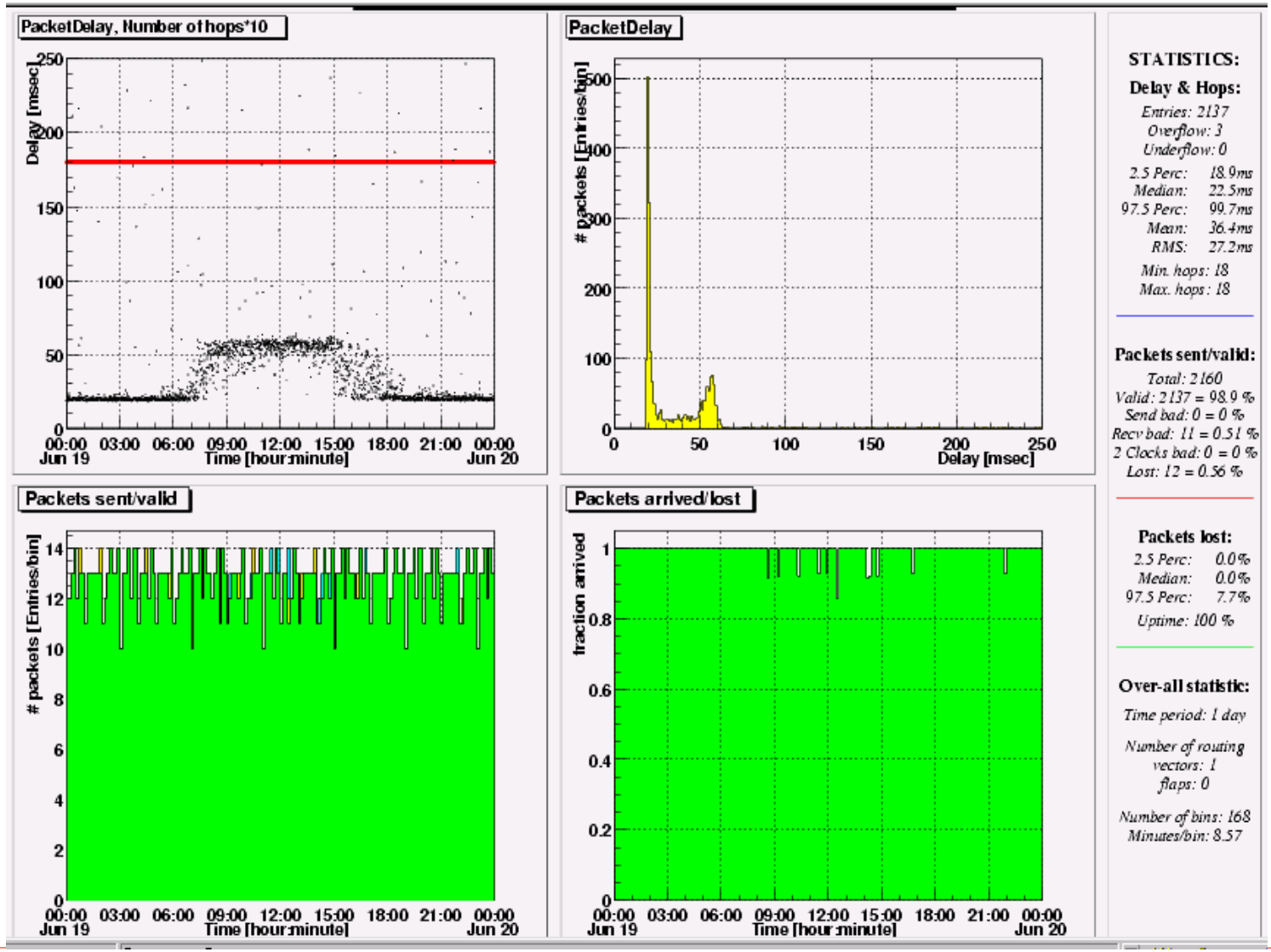
RIPE box or Surveyor

one-way delay + *loss*
UDP protocol

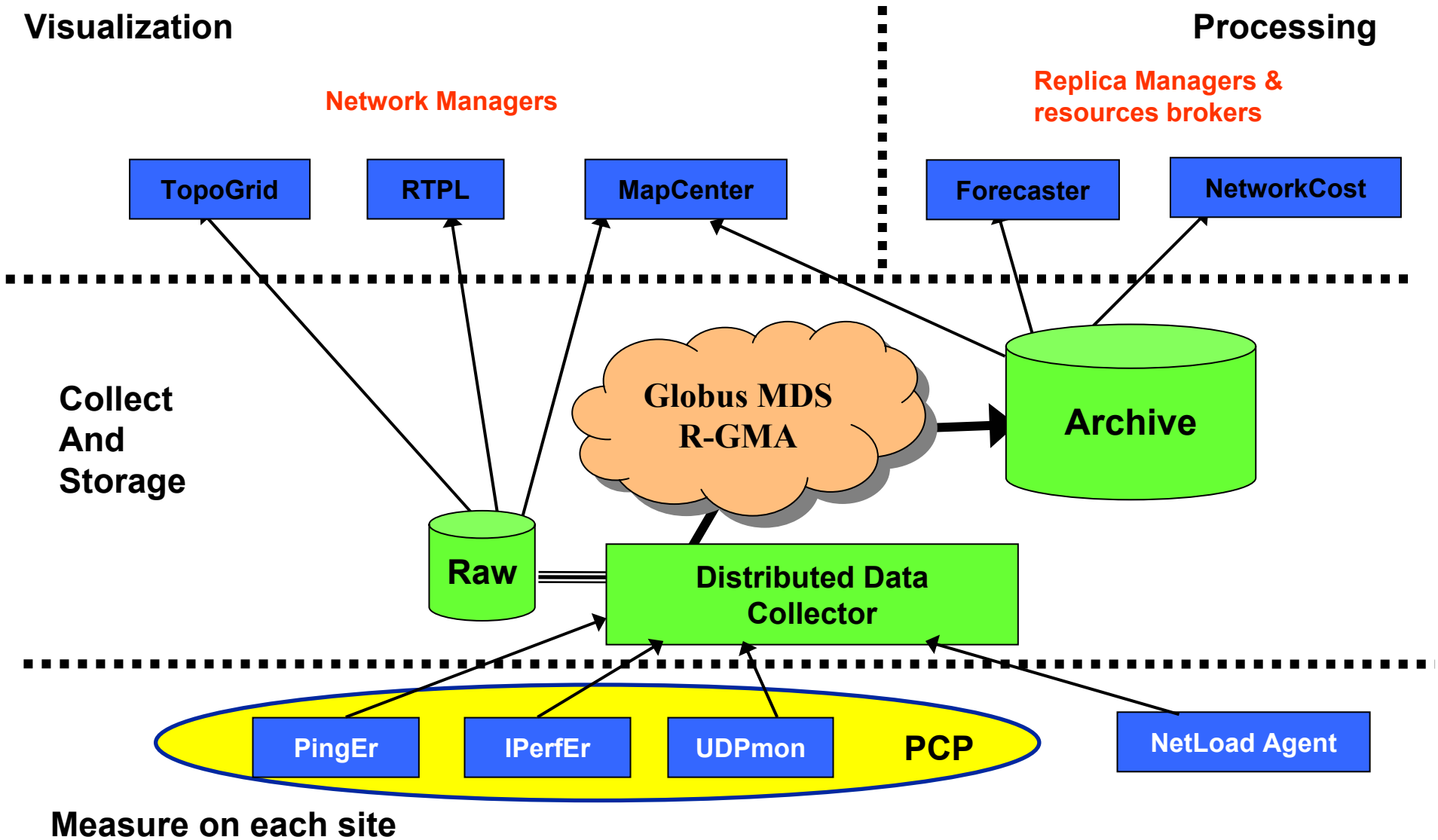
- specialized boxes
- Measurement between 2 RIPE boxes
- GPS time synchronization



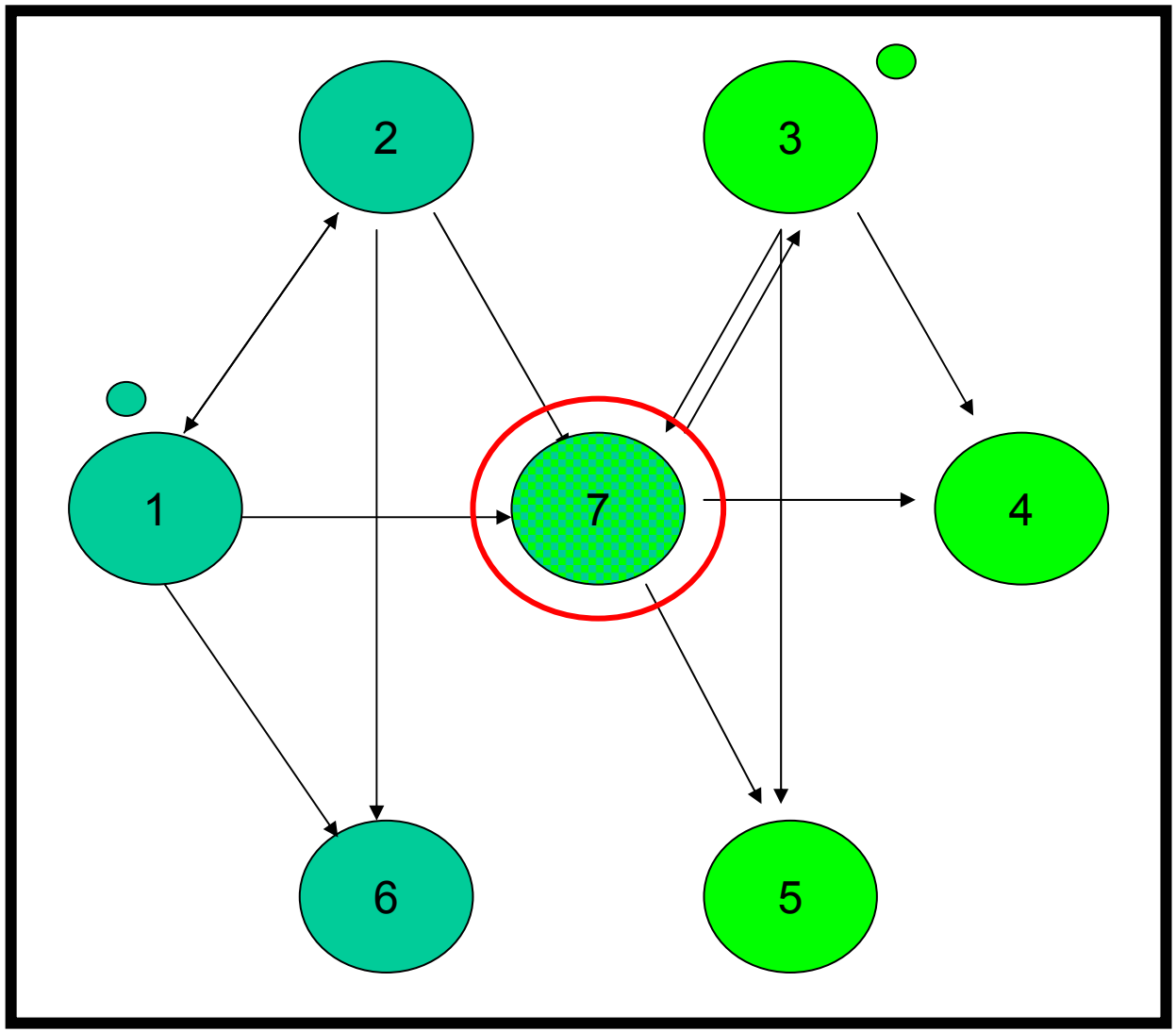
RIPE output entre UK et CERN

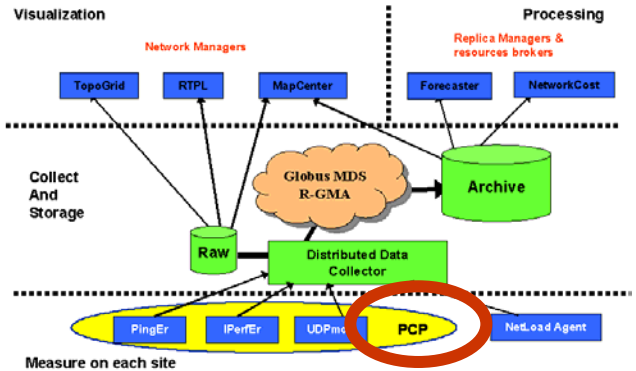


Network Monitoring Architecture



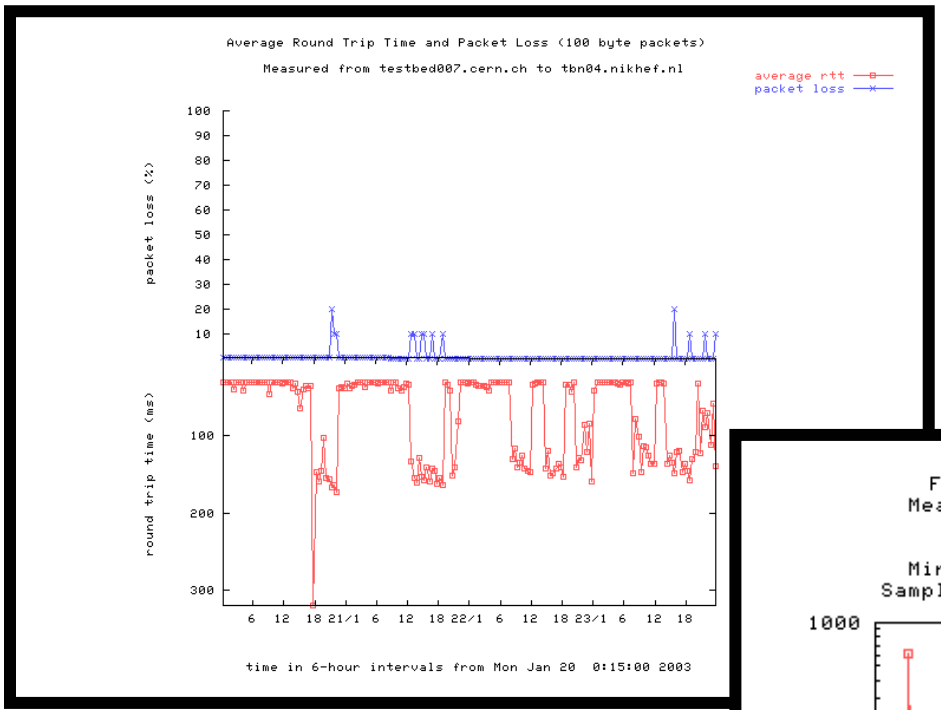
PCP (Probe Coordination Protocol)

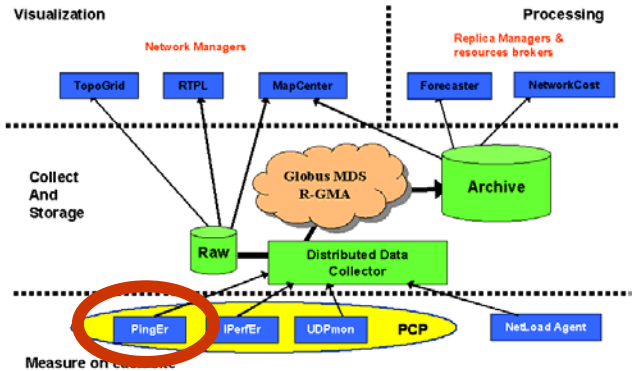


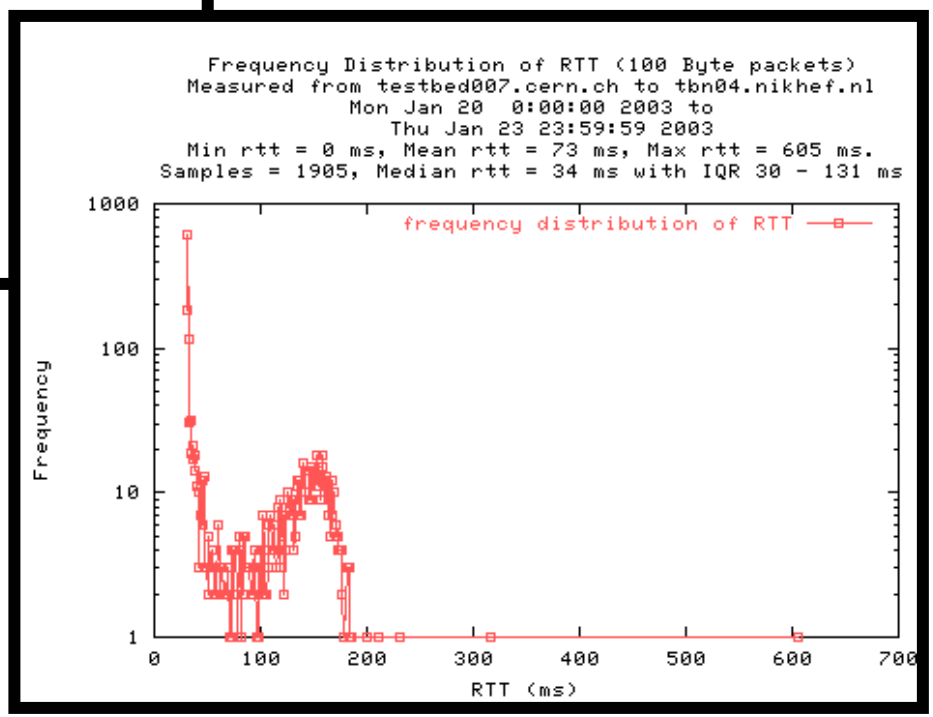


Scheduling of measures by clique

PingEr



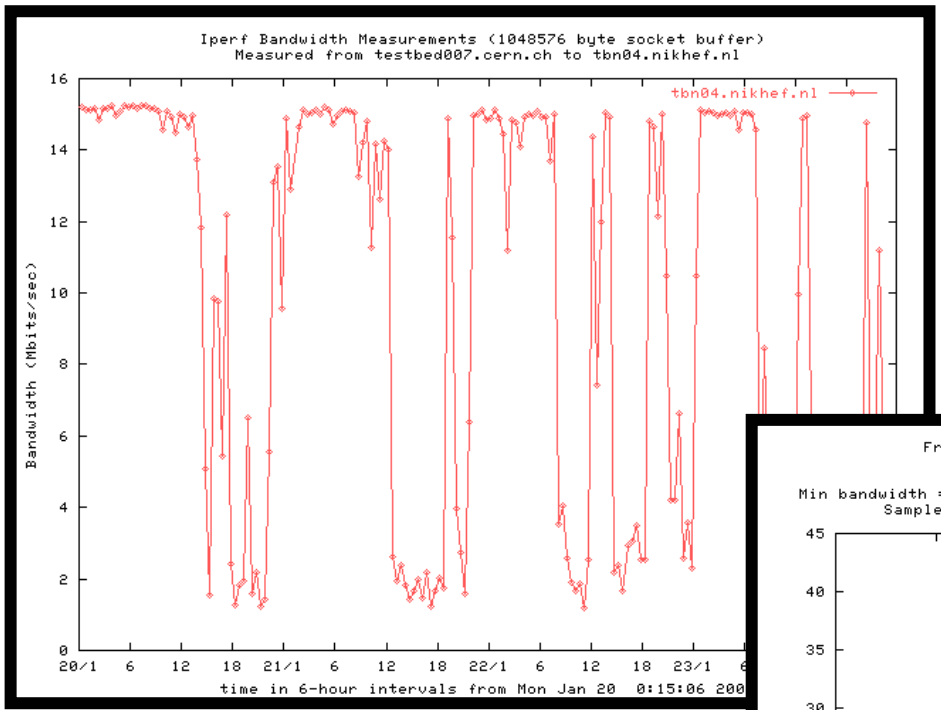


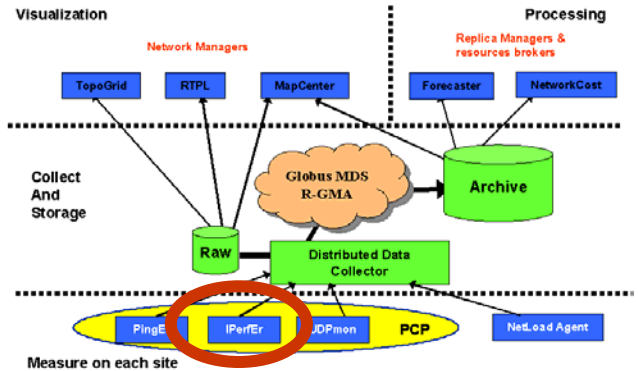


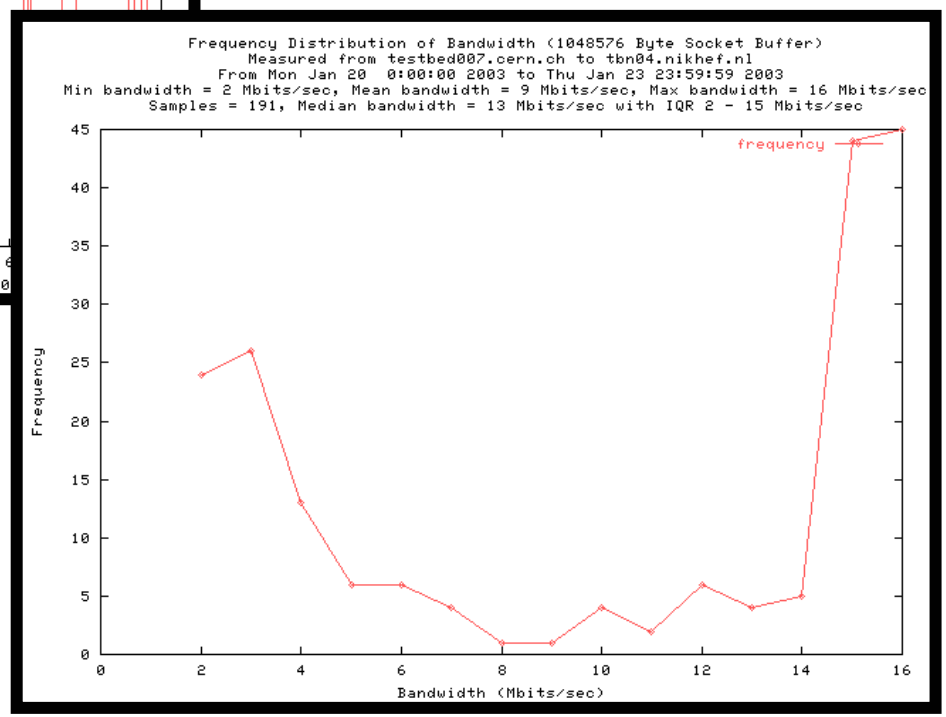
RTT

LOSS

IperfEr

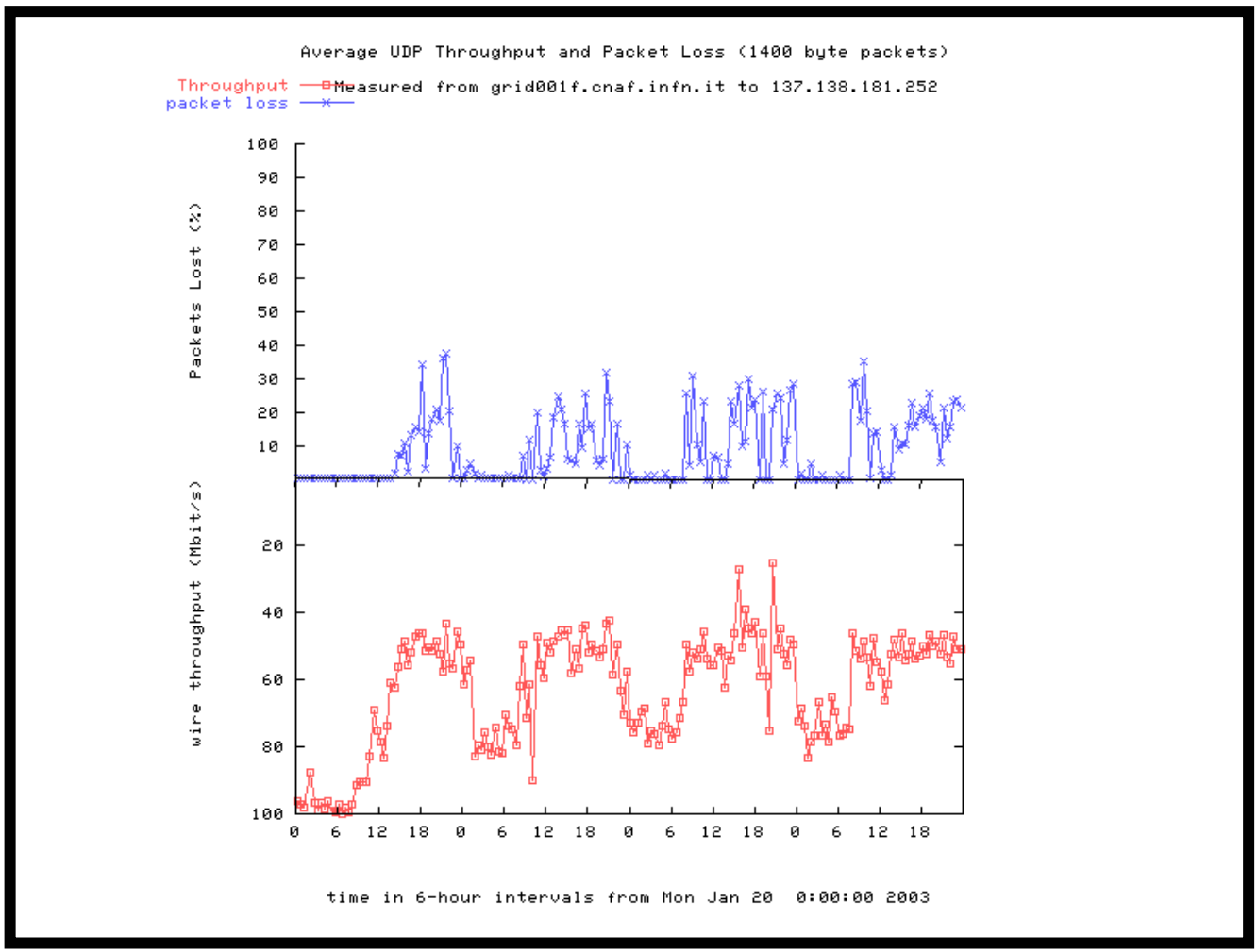


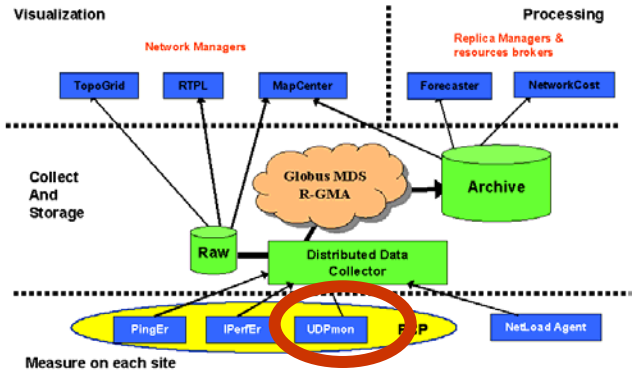




TCP Throughput

UDPmon

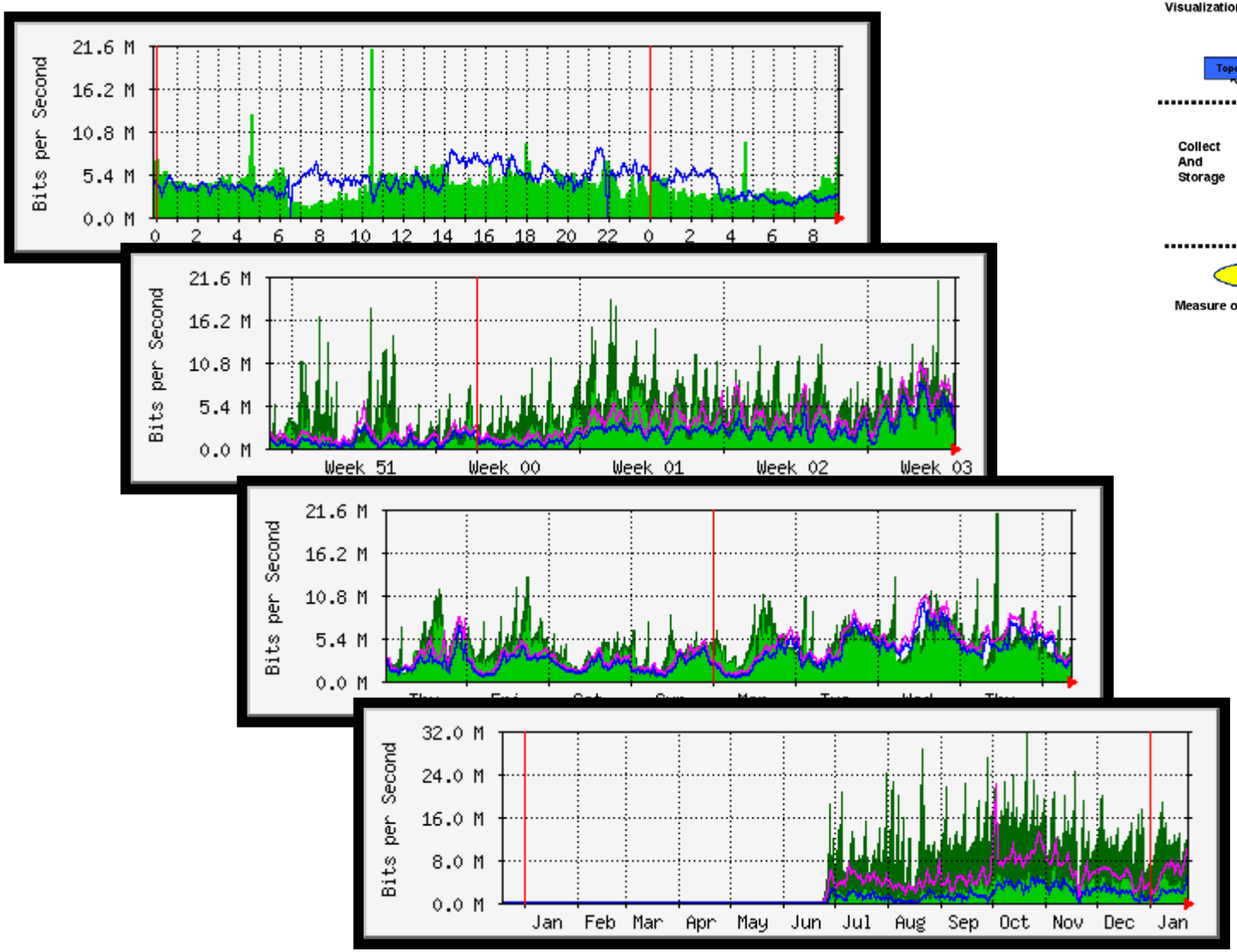


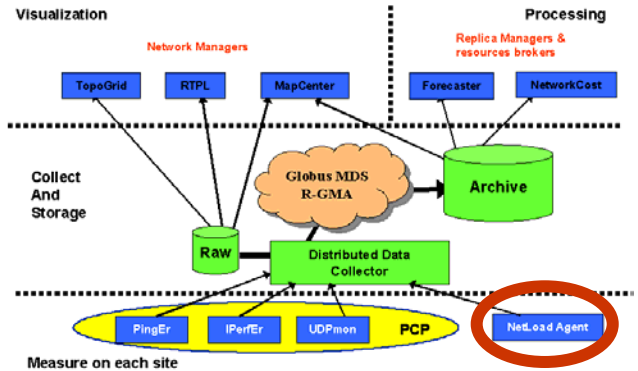


UDP Throughput

Loss

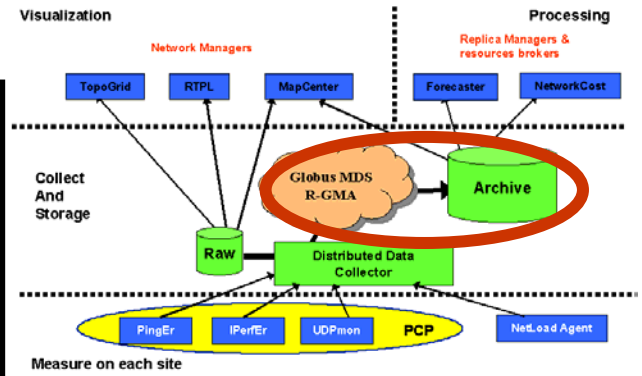
NetLoad Agent





In/Out Traffic

Metrics stored in LDAP (Globus MDS)

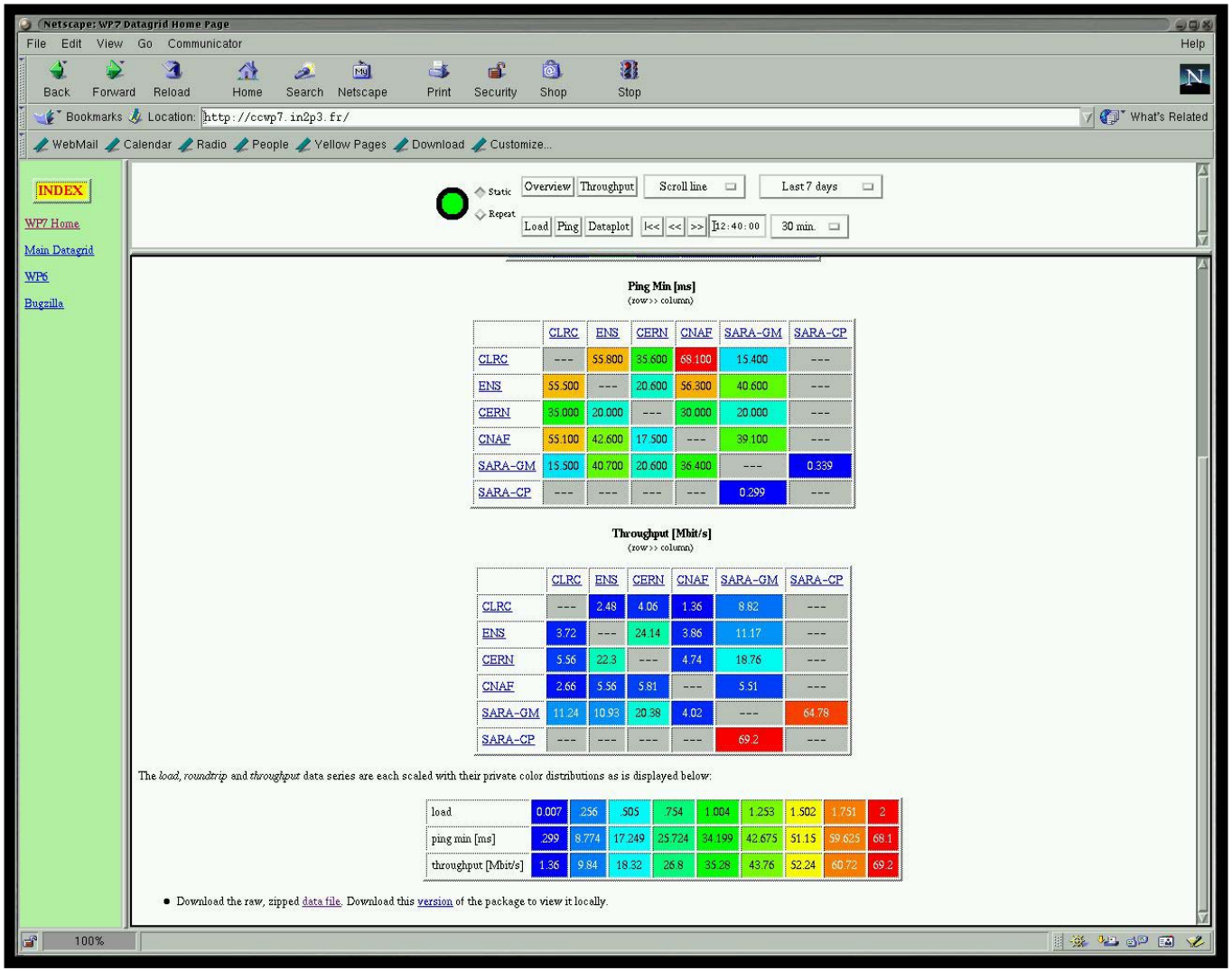


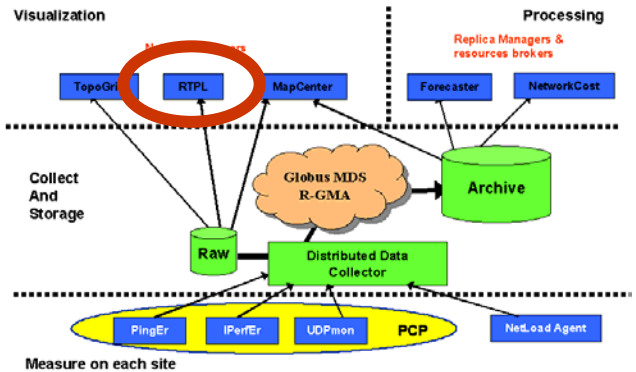
- testbed001.cnaf.infn.it
  - hn=testbed001.cnaf.infn.it,Mds-Vo-name=local,
    - Mds-Vo-name=local
    - ceId=testbed001.cnaf.infn.it:2119/jobmanage
    - ceId=testbed001.cnaf.infn.it:2119/jobmanage
    - ceId=testbed001.cnaf.infn.it:2119/jobmanage
    - ceId=testbed001.cnaf.infn.it:2119/jobmanage
    - nmId=grid001f.cnaf.infn.it/iperfer
      - Mds-Vo-name=local
      - SourceHost=grid001f.cnaf.infn.it
        - DestHost=ccwp7.in2p3.fr
          - Mds-Vo-name=local
          - NMMeasureId=tool.iperfer/buffersize:
          - NMMeasureId=tool.iperfer/buffersize:
        - DestHost=gppnm.gridpp.rl.ac.uk
        - DestHost=tbn04.nikhef.nl
        - DestHost=testbed007.cern.ch

Distinguished Name =	NMMeasureId=tool.iperfer/buffersi
objectClass =	NetworkMeasurement
SourceHost =	grid001f.cnaf.infn.it
DestHost =	ccwp7.in2p3.fr
DestSite =	IN2P3
DestNE =	in2p3.fr
MonitorTool =	iperfer
NMMeasureId =	tool.iperfer/buffersize:1048576/strea
MetricName =	tcpthroughput
MetricValue =	1498
MetricUnit =	bit/s
MetricTime =	200209111124019Z
Parameter =	buffersize:1048576
Parameter =	streams:1
Parameter =	duration:549.1
Parameter =	protocol:tcp

- R-GMA and Globus MDS Producer for storage of network metrics
- Storage and aggregation of all historical data in archives

rTPL

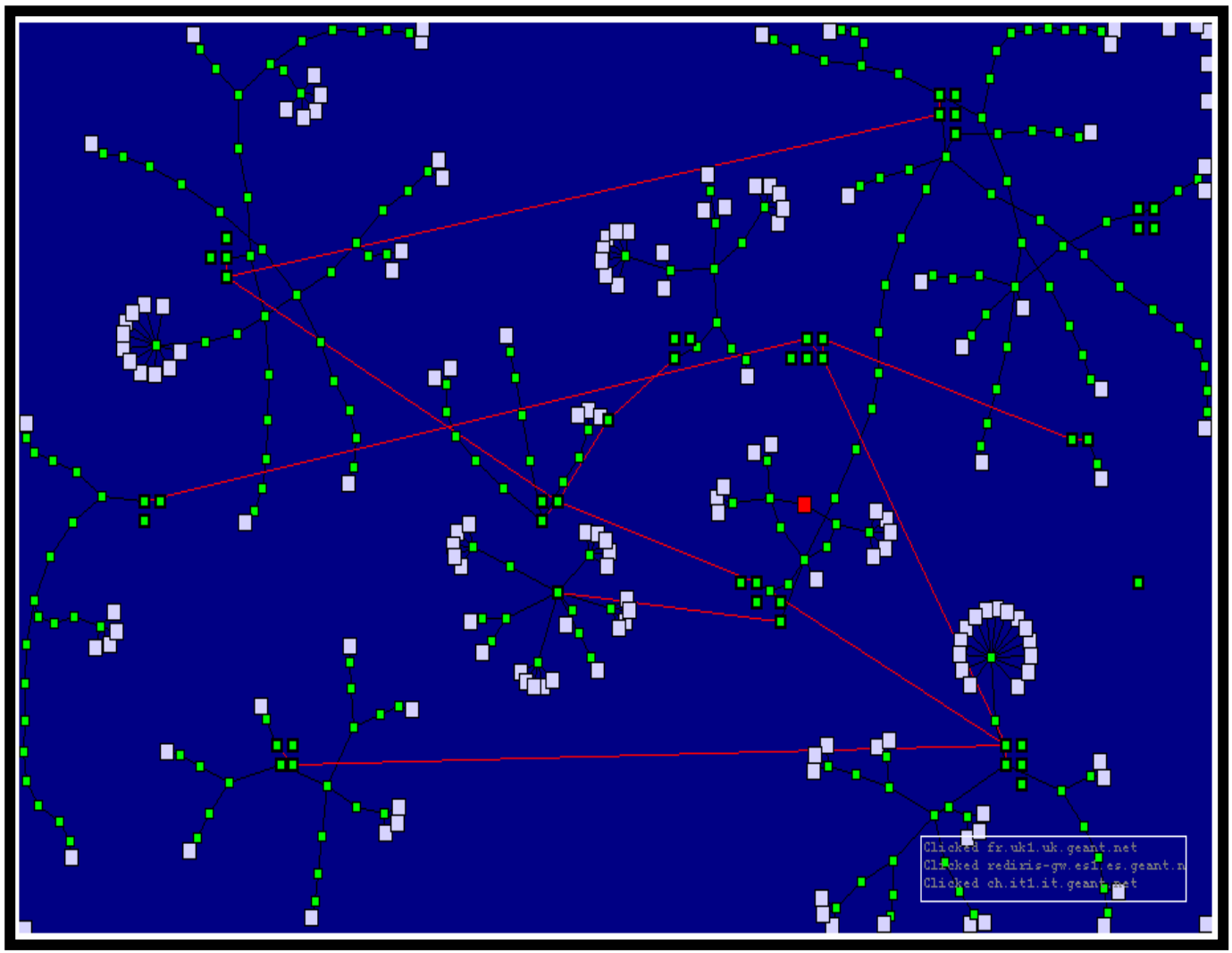


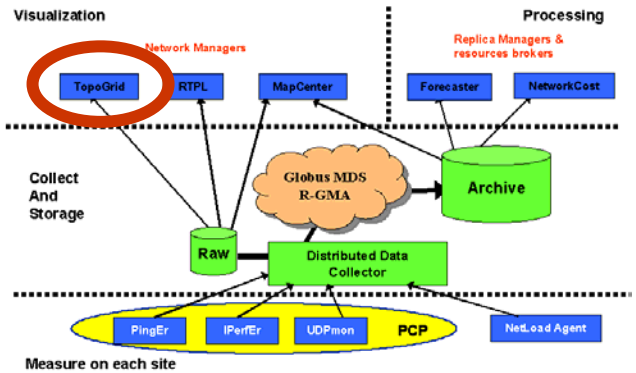


- High-level performance presentation interface
- Built-on-demand network metrics matrixes
- Customizable views for network administrators of all testbed sites

TopoGRID

http://ccwp7.in2p3.fr/topogrid

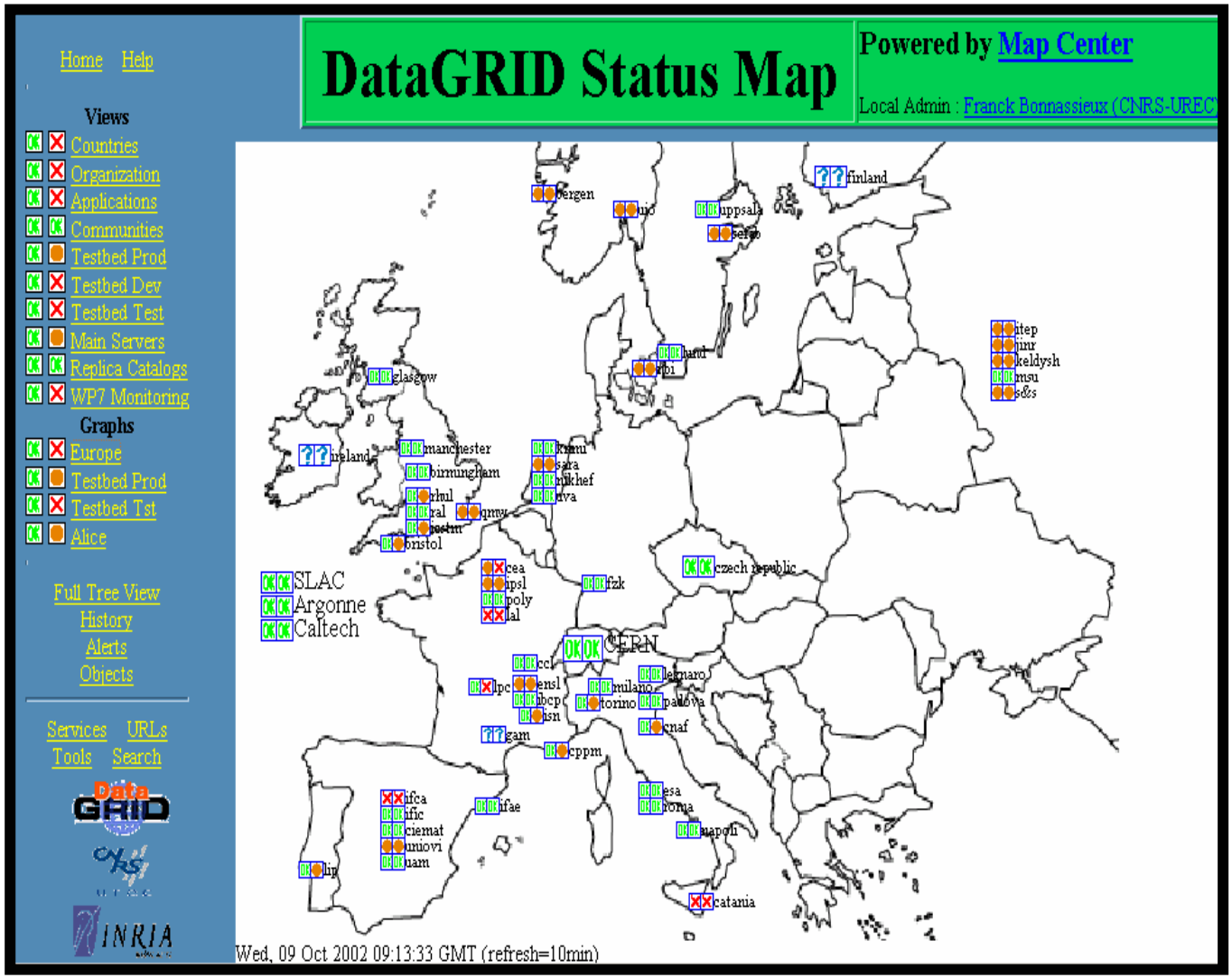


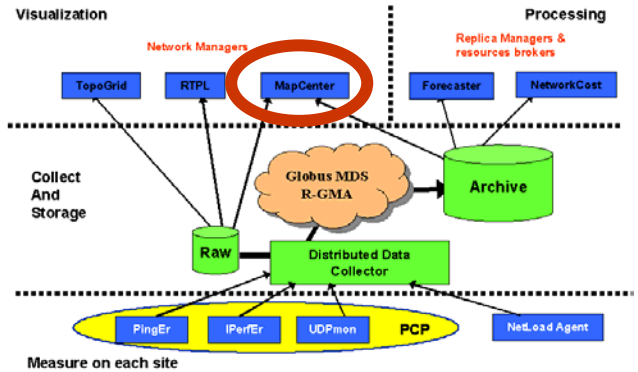


- Automatic discovery of routers and end nodes
- Identification of routes and bottlenecks
- Clickable Java Applet for in deep analysis of network paths

MapCenter

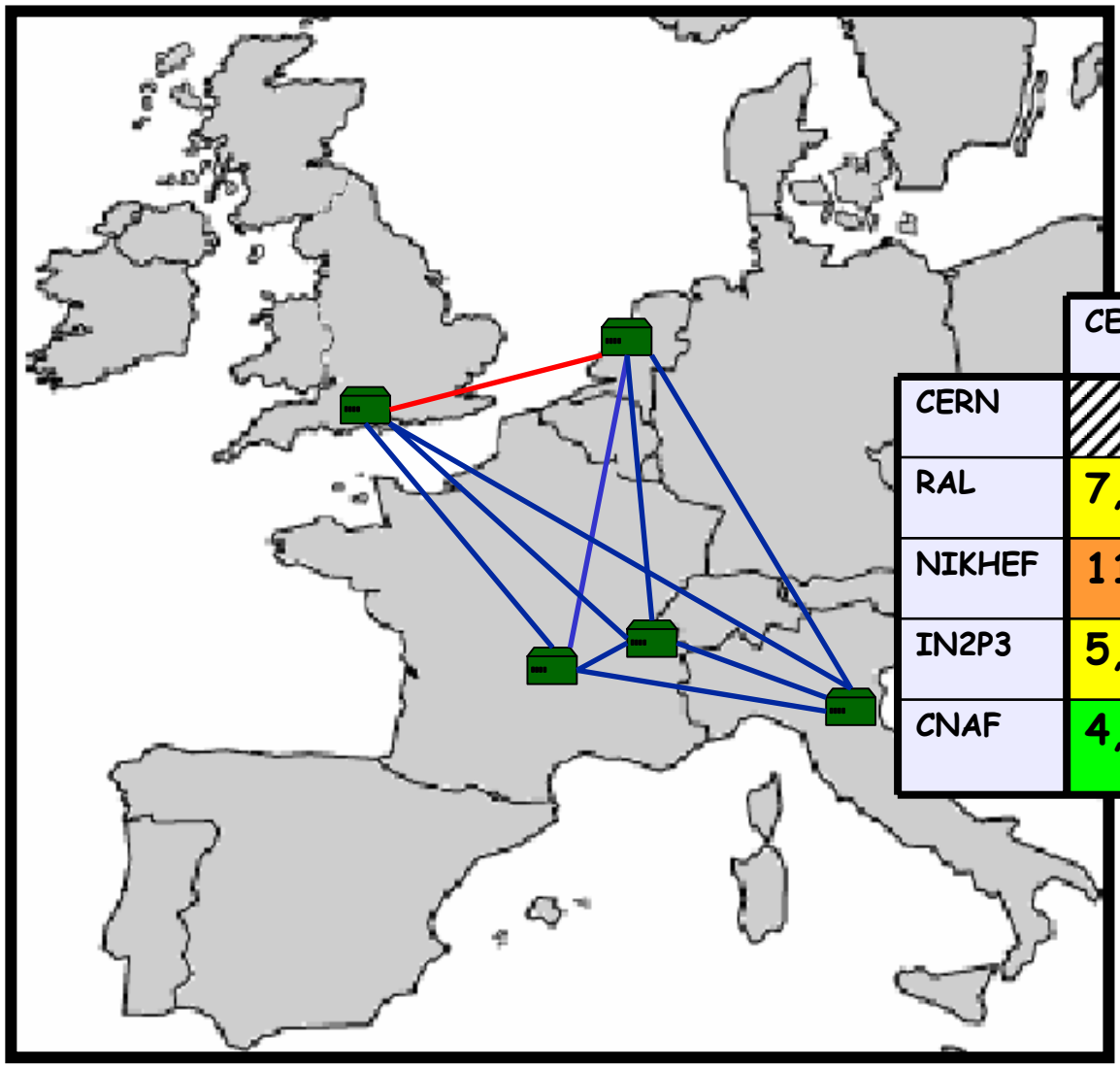
http://ccwp7.in2p3.fr/mapcenter

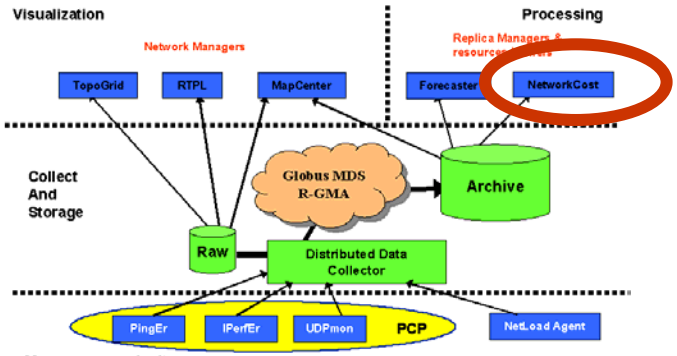




- Open and Flexible tool to visualize in real-time Grid status and Network performances
- Automatic discovery and positioning of resources in graphical maps
- Advanced stealth monitoring techniques (ICMP, TCP, UDP, HTTP, LDAP)

NetworkCost functions





	CERN	RAL	NIKHEF	IN2P3	CNAF
CERN		46,75	77,78	44,87	35,44
RAL	7,46		2,44	7,12	4,35
NIKHEF	11,13	3,25		11,86	2,66
IN2P3	5,03	10,38	6,24		7,08
CNAF	4,5	6,53	4,04	13,08	

getNetworkCost

CERN
RAL
NIKHEF
IN2P3
CNAF

CERN
RAL
NIKHEF
IN2P3
CNAF

FileSize = 10 MB

Results = time to transfer (sec.)

NetworkCost model

- The current cost model is designed for data intensive computing and especially large files transfers
  - The most relevant metric for that cost model is available throughput
- Implementation
  - Iperf Measurements (current)
  - GridFTP Logs (future)
  - Other metrics (future) : UDP, RTT, Jitter, ...
  - Synchronisation (PCP)

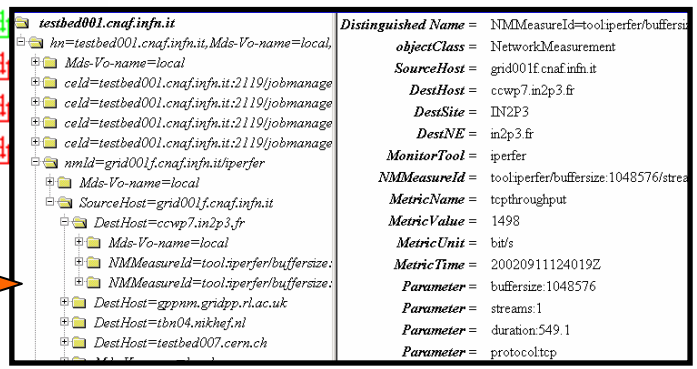
MapCenter – Network Monitoring Page

- WP7 Production Testbed
  - CERN testbed007.cern.ch (icmp | http iperfd udpmon | pinger iperf udpmon)
  - INFN-CNAF grid001f.cnaf.infn.it (icmp | http iperfd udpmon | pinger iperf udpmon)
  - NIKHEF tbn04.nikhef.nl (icmp | http iperfd udpmon | pinger iperf udpmon)
  - RAL gppnm.gridpp.rl.ac.uk (icmp | http iperfd udpmon | pinger iperf udpmon)
  - CC-Lyon ccwp7.in2p3.fr (icmp | http iperfd udpmon | pinger iperf udpmon)
- WP7 Test Testbed
  - Daresbury (UK) icfamon.dl.ac.uk (icmp | http iperfd udpmon | edg_ukwp7)
  - Rutherford (UK) icfamon.rl.ac.uk (icmp | http iperfd udpmon | edg_ukwp7)
  - SARA (NL) gridmon.sara.nl (icmp | http iperfd udpmon |)
  - INFN-CNAF (IT) grid001f.cnaf.infn.it (icmp | http iperfd udpmon | root)
  - CC-Lyon (FR) ccwp7.in2p3.fr (icmp | http iperfd udpmon | edg_frwp7)
  - CZ atmtest-eth-bm.muni.cz (icmp | http iperfd udpmon | pinger iperf udpmon)
  - IPSL (FR) amundsen.datagrid.jussieu.fr (icmp | http iperfd udpmon | pinger iperf udpmon)
  - ESA-ESRIN (IT) grid0005.esrin.esa.it (icmp | http iperfd udpmon | pinger iperf ud
  - Oslo grid.uio.no (icmp | http iperfd udpmon | pinger iperf ud
  - Uppsala grid.tsl.uu.se (icmp | http iperfd udpmon | pinger iperf ud
  - ES netmon.ifae.es (icmp | http iperfd udpmon | pinger iperf ud
- MDS metrics
  - Network Information Index ccwp7.in2p3.fr (icmp | nii)
  - CERN testbed007.cern.ch (icmp | mds)
  - INFN-CNAF testbed001.cnaf.infn.it (icmp | mds)
  - NIKHEF ce01.nikhef.nl (icmp | mds)
  - RAL gppce06.gridpp.rl.ac.uk (icmp | mds)
  - CC-Lyon UREC-glob.ens-lyon.fr (icmp | mds)
- Network Load
  - Net Load Server ccwp7.in2p3.fr (icmp | http | netlsrv)
  - CERN RB testbed007.cern.ch (icmp | netlagt)

Network Monitoring for WP7

IPerf pages

- Latest Offsite Connectivity
- Plot Offsite Iperf Data
- Plot frequency histograms
- Main DataGrid WP7 Network Monitoring site



Site Name	Agent	Out				In				Bw	Infos	
		5 m	20 m	1 h	Max	5 m	20 m	1 h	Max			
CERN RB	testbed007.cern.ch	0.035	0.038	0.041	5.799	0.025	0.025	0.028	9.787	100	Conf	Link
LPC (France)	biolpc03.in2p3.fr	0.001	0.002	0.002	0.425	0.010	0.015	0.011	3.407	10	Conf	Link
ENS Lyon (France)	UREC-glob.ens-lyon.fr	2.067	2.357	2.493	6.000	3.214	4.287	4.522	22.031	100	Conf	Link
IPSL	amundsen.datagrid.jussieu.fr	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	10		

Perspectives

- GRID -> SCAMPI
  - Testbed infrastructure
  - Huge traffic generation on gigabits links
  - End-to-end metrics
  - …
- SCAMPI -> GRID
  - Traffic analysis (protocol distribution)
    - GridFTP, Remote file access, …
  - Grid traffic measurement (per protocol or/and per address domains)
  - …