



Grid vs. Cloud Computing and Why This Should Concern the Optical Networking Community

OFC/NFOEC 2009

« Virtualizing and Scheduling Network resource for Emerging IT services: the CARRIOCAS approach »

Pascale Vicat-Blanc Primet (INRIA) Dominique Verchère (Alcatel-Lucent Bell Labs)

March 23rd 2009

Optical Fiber Communication Conference and Exposition & the National Fiber Optic Engineers Conference

* Distributed/Grid Computing services delivered by Internet Optical Networks with Ultra-high transmission Capacities

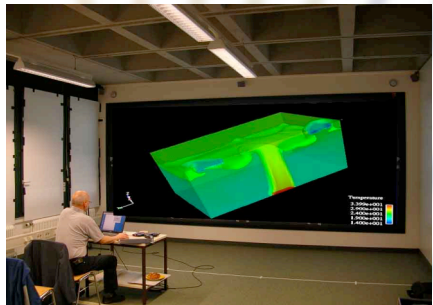


1. CARRIOCAS Challenges and Project presentation
2. Efficient Network architecture combining capacity and flexibility
3. Scheduling, Reconfiguration & Virtualization (SRV) service
4. CARRIOCAS Pilot Network and Experimentations
5. Conclusion

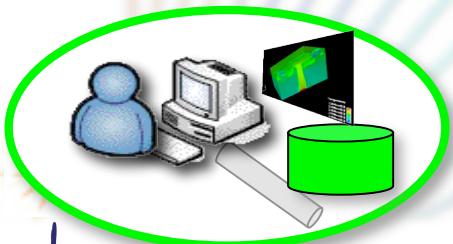
CARRIOCAS Challenges



- To research, design and implement a ultra-high bit rate network reconfigurable according to grid applications connectivity requirements
 - **Transmission capacity 40Gbit/s per wavelength***
 - Enhanced **application-network interactions** with **SRV**
 - Guaranteed **network service deliveries** in **function** of **Application QoS requirements**.



- To develop new **Services & Usages**
 - collaborative engineering on virtual prototypes
 - Computing Intensive Simulation Applications
 - interactive visualization



- To specify new business models of services integrating IT access and connectivity

The 40Gb/s transponder prototypes have been successfully deployed on Bruyères-le-Châtel→Saclay and Saclay→Clamart links, performance are at nominal level.



Optics

Networking

Computing

Application

Industrials



SME's



Oct 2006-Sept 2009



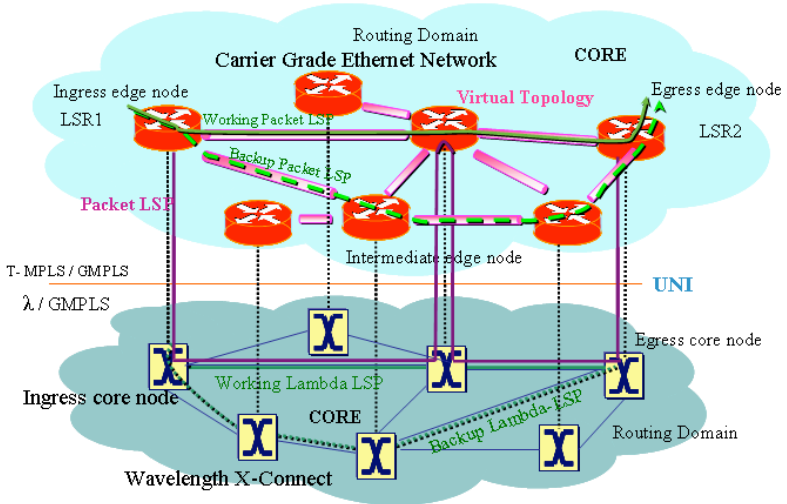
Academia



Financial partners:



- Network Traffic flow characterizations «Data Grid» generated by large organizations (companies and institutions) users (clients) and datacenters (IT resource providers)

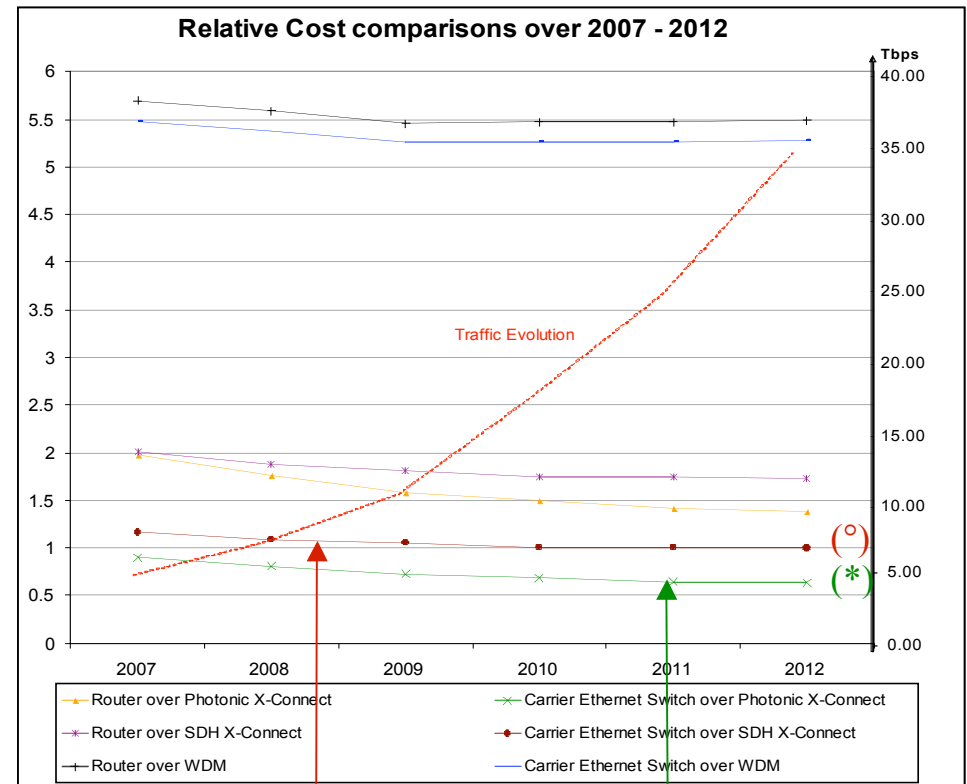


28 PoPs (Router + X-Connects), 10 Gbps / 40 Gbps interfaces

43 WDM Links, Network Connectivity $d \approx 3.1$

Input Traffic Matrix: from 5.26 Tbits on 2005 to 34.5 Tbits on 2010 with different Traffic Flow entries:

- Voice data, enterprise transactional data, datacenter grid data, internet residential data



Carrier Grade Ethernet Switch over SDH X-Connect(°)

TSS over ROADM: CARRIOCAS based architecture (*)

- Layer 2 (Carrier Grade Ethernet) and Layer 1 (WDM) reconfiguration capabilities
- Most Scalable Solution absorbing the increases of data grid traffic demands



Towards Virtualized Infrastructures and «Extended Service Providers»

High Energy Physic engineers

Bio-informatics Scientists

Car Designers – Manufacturers

Usage SLA

Service Storefront

Grid Service Provider



Provider SLA

Service Integrator & Presentation

Connectivity Services

Specific Software (AppsResources)

Storage Resources

CPU Resources

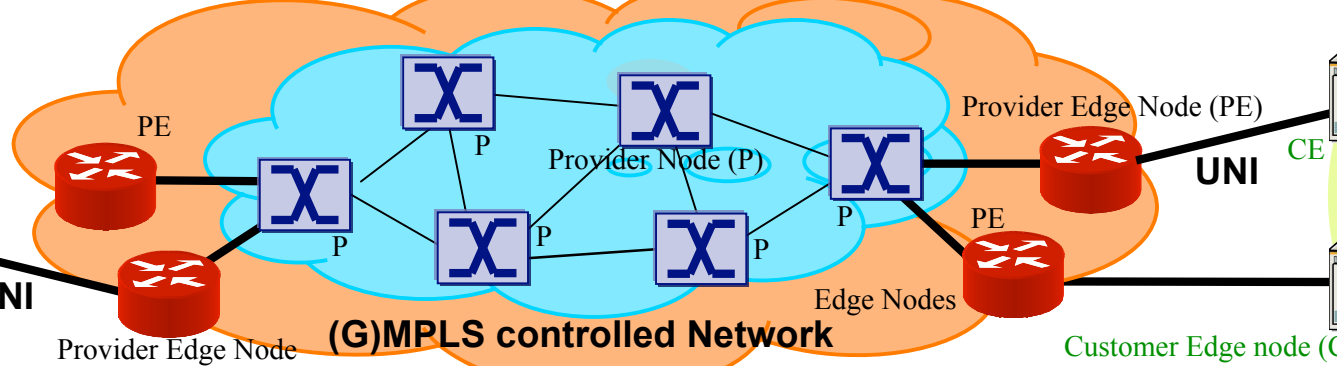
Picture Walls (Visualization Resources)

Scientific Instruments/Sensors (Acquisition Resources)

Network Service Management (S.R.V.)

Customers (Enterprises)

CE UNI



Telecom Infrastructure Operator

IT infrastructures

Datacenter Operators



■ V1 : Connectivity services

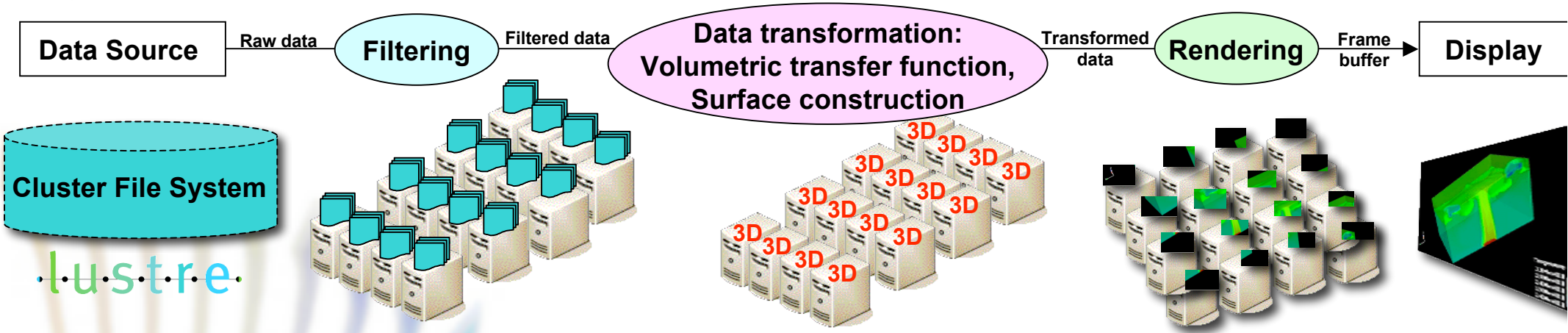
- Requirements: to guarantee Quality of Experiences for grid application users
- Adaptive connections configurations with QoS guarantee (bandwidth, latency, availability ...) negotiated according to Grid Application workflow needs
- Technology enabler: GMPLS, PCE, Timed based Network management

■ V2 : Full Virtualized services

- Requirements: To integrate grid resources management for executing the Grid Application workflows
- Selection, reservation and interconnection of Computational, Storage and Visualization Centers which may come from 3rd party resource infrastructure providers (e.g. datacenters, HPC-centers)
- Co-selection and cross-optimization of network and grid (application) resources
- Technology enablers: WSRF, and network protocols
+ VXDL & VXcore (INRIA),

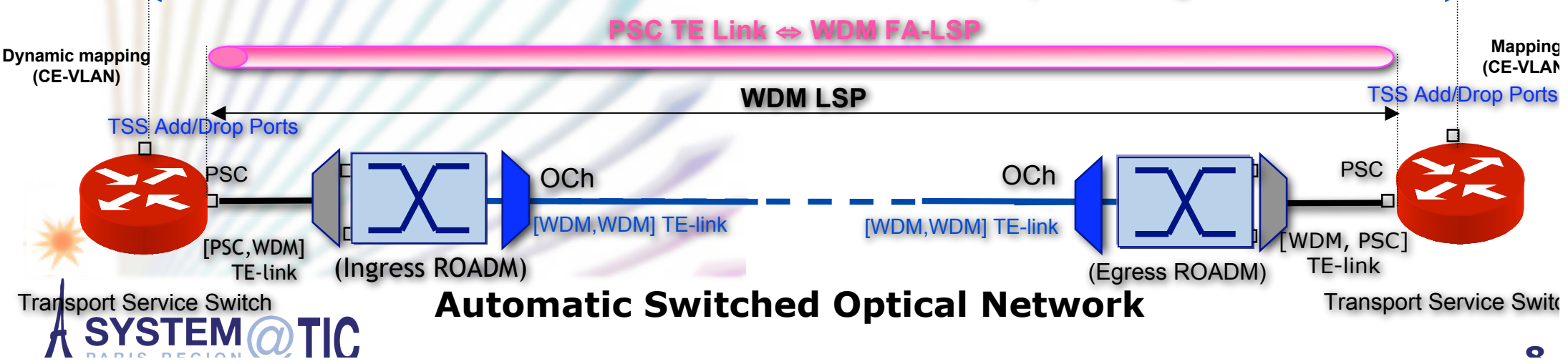
Binding of VISUPIPE flows (jobs) and Carrier Grade Networks services (resources)

- Visualization pipeline partitioning and mapping on Grid IT Resources



- 24 Mpixels x 32 bits x 30 images/s x 2 for stereo = **46,08 Gb/s**
- Low latency requirements for interactivity

- One Data flows can be mapped on one EVC allocated on dedicated WDM
- Automatic end-to-end Ethernet Virtual Circuit provisioning



SRV position in Telecom Network



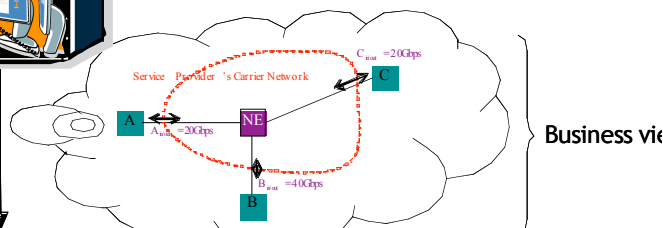
Scheduling, Reconfiguration & Virtualisation

Virtualized Services (v2)

Connectivity Services (v1)

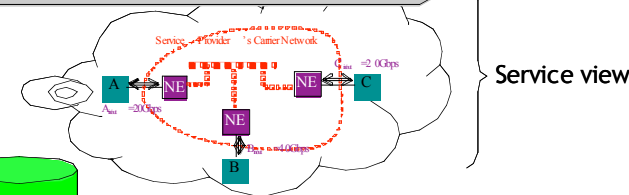
Scheduled Management Components of Network Resources

Network Management Function Layer



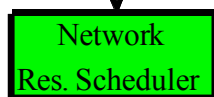
Business view

Commercial Filters (policies)

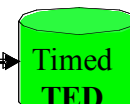


Service view

Requests / Reponses Cds



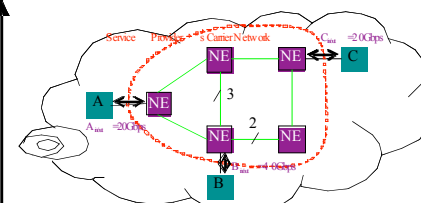
Management :
Read / Write



NetConf.



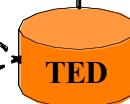
Consultation



Network view



Connection inventory



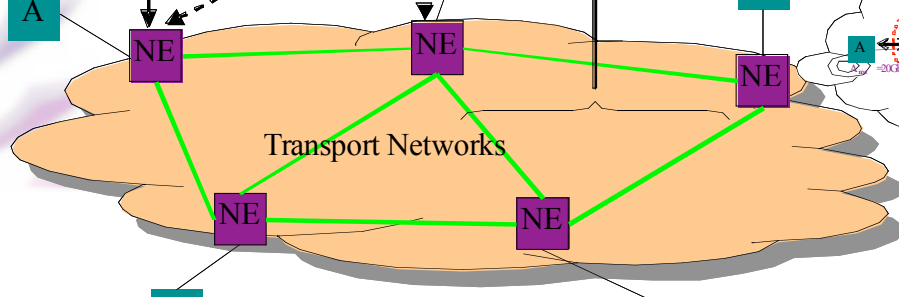
Synchronization/
Alimentation

Resource inventory

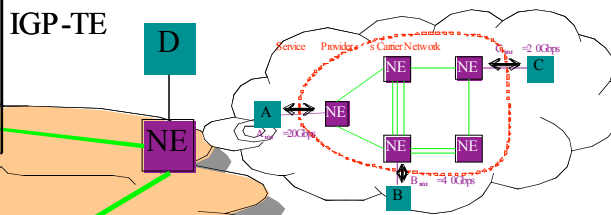
A

E

D



Transport Networks



Element view

CARRIOCAS pilot network demonstrates advanced network service functions and proving their concepts through an emulation platform

Feasibility of **automated network management and control** with **dynamic delivery of Telecom network services** to end users: e.g. company customers.

The first steps are focused on **Connectivity services** managed through a well defined Service Management System (SRV) described.

Each connectivity service is designed with specific attributes including:

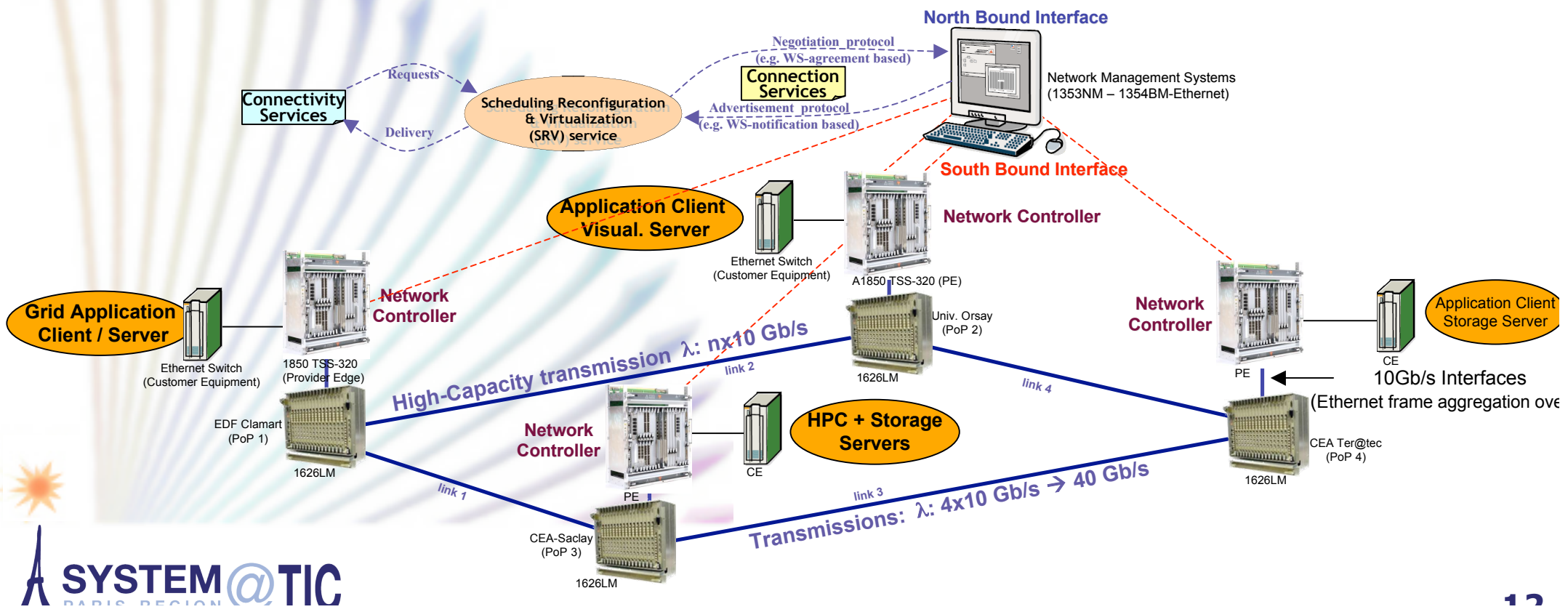
- Service Access points (UNI), throughput (x Gbps), end-to-end latency (y ms), time interval $[T_1, T_2]$

GMPLS based control protocols, Timed based PCE, Network Resource Scheduler with Carrier Switches and Wavelength Cross-connects as connection services are based on Ethernet Virtual Circuit over Wavelength connections:

- Provisioned: through NMS with automated network Operations through MTOSI 2.0 (triggered by SRV)
- Signaled: RSVP-TE **provides dynamic E2E guaranteed QoS** on two routing areas

- **Connection Service Component Activation: EVC Services**
 - SRV → NRS (phase 1): Timed EVPLINE, EVPLAN, E-TREE provisioning
 - TE management → Connection Provisioned Monitoring → NRS → SRV-DB
 - Background and periodic process (from Network States)
- **Downward information: SRV → NRS**
 - Connectivity topology is an Ethernet Service graph
 - Access points: location,
 - Bandwidth (pipes)
 - Latency on the links
 - Commands for new connectivity services
 - Feasibility
 - Reservation
 - Provision
 - Activation
- **Upward information: NRS → SRV**
 - Connection status:
 - Reserved / Available
 - Provisioned / Unprovisioned
 - Activated (i.e. used) / De-activated
- **Shared Info/Data Model based on MTOSI 2.0 / TMF**
 - SCAI: Resource facing Interfaces (Connection data model)

- Ethernet → MPLS-TP over WDM Network architecture referencing :
 - Reconfigurable-OADM based WDM transmission: 1626LM (with 40 Gbps)
 - Carrier Grade Ethernet based Transport Networks: 1850TSS
 - Network Configurations based on Management: 1350-OMS service functions
- Large Scale Distributed Applications (Grid Applications):
 - Computing and Storage Service on Demand (Ter@tec): LUSTRE CFS Application
 - Instruments: Industrial (Remote Visualization): EDF Picture Walls



CARRIOCAS Network Pilot and test bed configuration



1354BM

Network Management

1353NM

Issy les Moulineaux

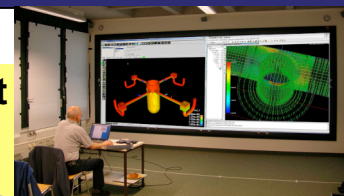


Local monitoring



4 x 10GEth.

Clamart



Picture wall

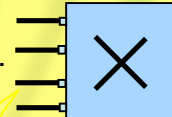
Transport Service Switch
Carrier Grade Ethernet with T-MPLS capabilities

1850 TSS

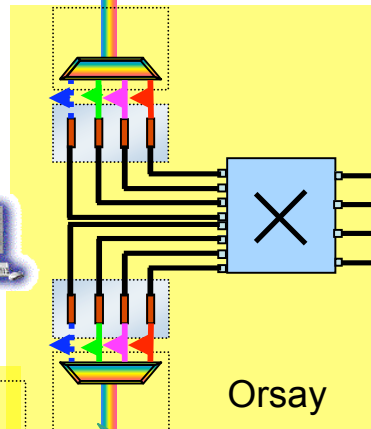


4 x 10GEth.

Saclay



Local monitoring



Orsay



4 x 10GEth.

storage

storage

4x10 GEth. Access

40 Gbps Transmission

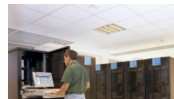
Nx10 Gbps Transmission

1626 LM

Super-computers

4 x 10GEth.

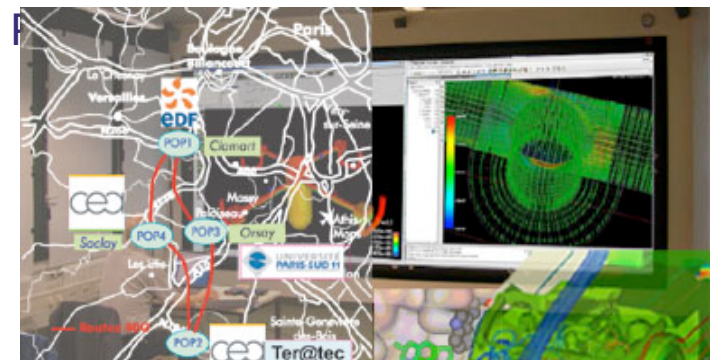
Bruyères Le Châtel



Deployment phase 1 (10Gb/s): June 2007
Deployment phase 2 (40Gb/s): June 2008



- Respond to strong demands by Organizations on accessing High-Performance IT services remotely through WAN with strict QoS:
 - Grid → Cloud Computing towards Utility Computing
- Service optimized Transport Networks & Data Grid Facilities will be built on different requirements to offer e-science services
 - Depending if Public or Private owned infrastructures
 - Needs to coordinate our efforts for standardizations
- SRV v1 software, developed by INRIA, automates:
 - Support for different network services offered by multilayer networks (Lx-VPN)
 - Mapping of applications/users requirements over connectivity services
 - Fine grain Dynamic Bandwidth Requests scheduling
 - Connection service Reconfiguration → Optimization → X-optimization
- Standardization efforts at ETSI (Grid TC) led by BT and Alcatel-Lucent with contributions from BT, FT, Telefónica, Nokia-Siemens,
 - close follow-up at Open Grid Forum (through I



Many Thanks to

CARRIOCAS-SP2 Colleagues for their
contributions

French Ministry of Industry, Essonnes,
Haut-de-Seines and Paris General Council
for their supports on CARRIOCAS project

<http://www.carriocas.org>



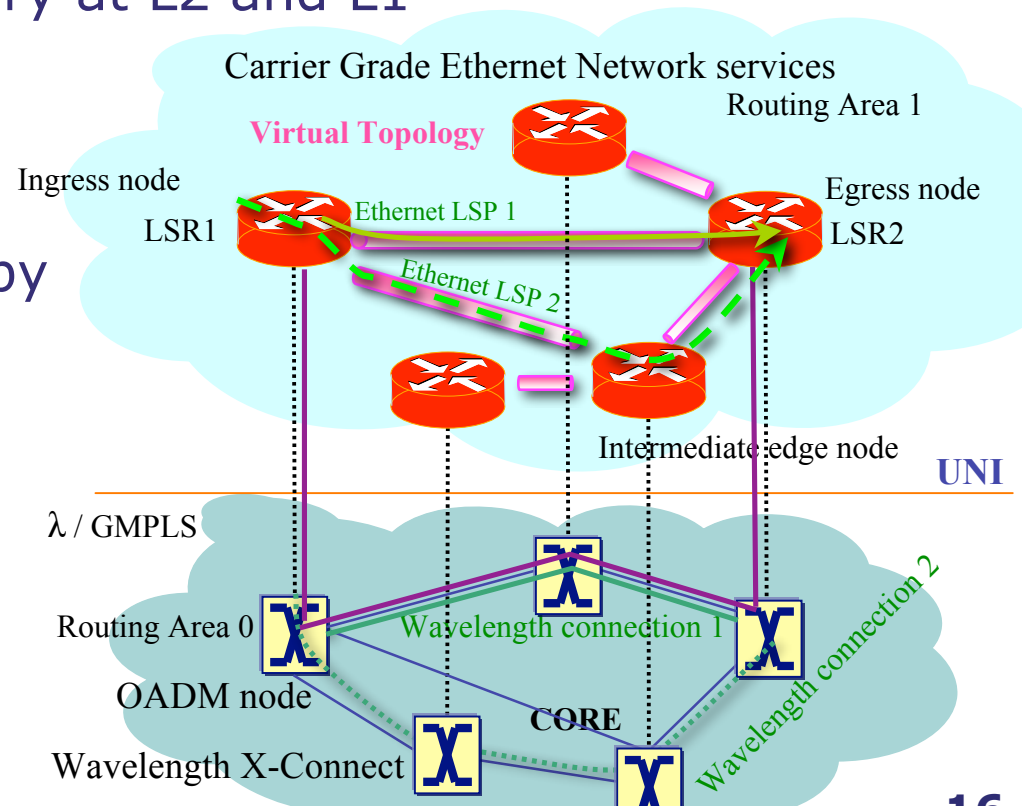
CARRIOCAS

Special thanks to PHOSPHORUS and G-Lambda
projects for their stimulating and valuable
interactions

■ Questions ?

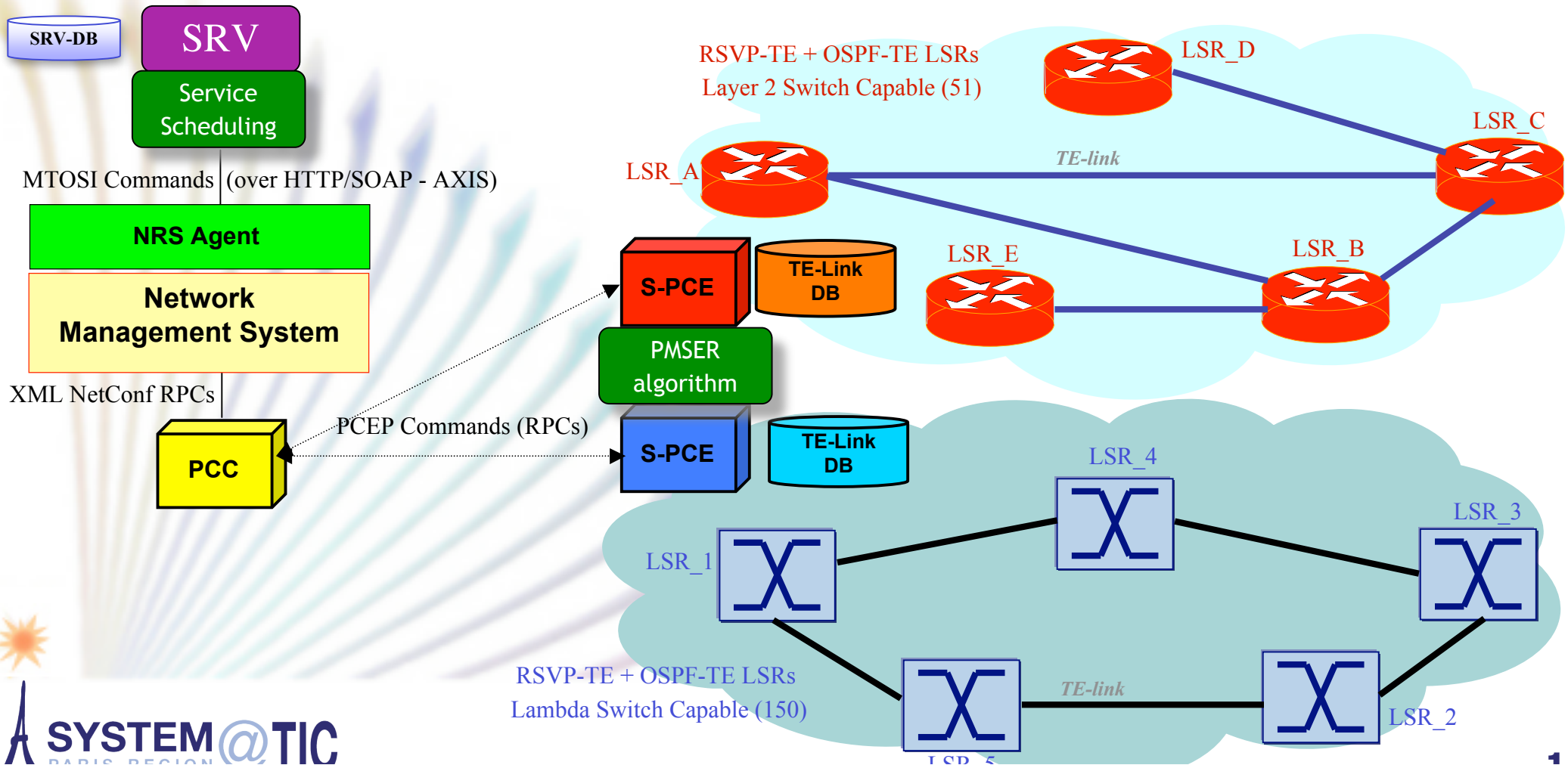
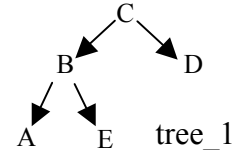
- Emulated Platform is based on CARRIOCAS network pilot architecture*
 - Two Routing Areas (1 per switching layer): (i) CGEthernet (ii) Wavelength
 - 5 PoP emulated, each PoP is
 - Carrier Switch (Ethernet-switching capable) + PXC (Lambda-switching capable)
 - 10 Node controllers (5+5): RSVP-TE + OSPF-TE
 - 2 Path Computing Engines (1 per layer): PCEP, OSPF-TE
 - 1 NMS: performs resource inventory at L2 and L1
- Dynamic Ethernet LSP establishments
 - 5 EVPLINE services
 - 1 EVPLAN service (4 locations)
- Static Wavelength LSP characterized by
 - Capacity: 3x10Gbps, 1x30Gbps, 1x40Gbps
 - Latency: $1\text{ms} \leq t \leq 10\text{ms}$

* Mais contient des fonctions supplémentaires



6 Network configuration implementation (1/2)

- 5 point-to-point Ethernet LSP (E_LSP): $E_LSP(ERO, BW(\text{Gbps}), \text{delay}(\text{ms}))$
 - $E_LSP_1 (\{A \rightarrow B\}, 5, 1)$; $E_LSP_2 (\{A \rightarrow B \rightarrow C\}, 10, 5)$; $E_LSP_3 (\{A \leftrightarrow C\}, 40, 1)$; $E_LSP_4 (\{A \rightarrow \rightarrow D\}, 20, 10)$; $E_LSP_5 (\{A \leftrightarrow B \leftrightarrow E\}, 15, 5)$;
- 1 point-to-multipoint Ethernet LSP:
 - $E_LSP_6 (\{tree_1\}, 10, 10)$;



- 5 wavelength LSPs (L_LSP): L_LSP(ERO, BW(Gbps))
 - L_LSP_1 ({1 ↔ 4},10); L_LSP_2 ({1 ↔(4)↔ 3},30); L_LSP_3 ({1↔(5)↔ 2},40); L_LSP_4 ({↔ 3},10); L_LSP_5 ({1 ↔ 5},15);
- Scenarios – Connectivity services managed by SRV
 - $T_0=1$ (January 1st) , $T_1=55$ (February 24th), duration=nb of slots (1 time_slot = ½ day), e.g. duration = 60 (30 days)

End-points	T_1 /duration	Max. Throughput/ Guaranteed BW.	e2e delay (ms)	Service availability	Application_id (information)
A ↔ B	23/10	38.4 / 30 Gbps	5 ms	99.999%	Ultra-high perf. Visu Coll.
A → B	26/60	10 / 5	100ms	99.999%	Back-up Storage
A ↔ D	24/20	20 / 10	5 ms	99.999%	Ultra-high perf. Visu
E → D	38 / 6	5 / 1	1 ms	99.999%	Video-Streaming
D → C	35 / 12	30 / 10	10 ms	99.9 %	Data Storage
C → {A,E,D}	30 / 16	15 / 10	5 ms	99.99%	Tele-Call-Conf

