The PHOSPHORUS project supported infrastrucutre

PHOSPHORUS TEAM

Artur Binczewski

artur@man.poznan.pl

GLIF meeting, 17-18.09.2007

PHOSPHORUS OVERVIEW

- Instrument: Integrated Project under FP6
- Activity: IST-2005-2.5.6 – research networking test-beds
- Project duration 30 months
- Project started 01 October 2006
- Project budget 6 868 969 euro (5 125 098 euro EC contribution)
- Project resources 814 personmonths
- http://www.ist-phosphorus.eu

PHOSPHORUS PROJECT

- **European and Global alliance of partners to develop advanced solution of application-level middleware and underlying management and control plane technologies**
- Project Vision and Mission
  - The project will address some of the key technical challenges in enabling on-demand end-to-end network services across multiple heterogenous domains
  - In the Phosphorus' implementation the underlying network will be treated as first class Grid resource
  - Phosphorus will demonstrate solutions and functionalities across a test-bed involving European NRENs, GÈANT2, Cross Border Dark Fibre and GLIF

MEMBERS OF THE CONSORTIUM

NRENs & RON:

- CESNET
- Poznan Supercomputing and Networking Center
- SURFnet
- MCNC

Manufacturers:

- ADVA Optical Networking
- Hitachi Europe Ltd.
- NORTEL

SMEs:

- Nextworks

Universities and Research Institutions:

- Communication Research Centre
- Fraunhofer-Gesellschaft
- Fundació I2CAT
- Forschungszentrum Jülich
- Interdisciplinair instituut voor BreedBand Technologie
- Research Academic Computer Technology Institute
- Research and Education Society in Information Technology
- SARA Computing and Networking Services
- University of Bonn
- University of Amsterdam
- University of Essex
- University of Leeds

PHOSPHORUS


Communications Research Centre
NORTEL Networks
MCNC
UESSEX
UWS
UvA
SARA
SURFnet
IBBT
UniBonn
ADVA
FHG
FZJ
PSNC
CESNET
Hitachi Europe SAS
Nextworks
I2CAT
RACTI
AIT

PROJECT KEY FEATURES 1/3

- **Demonstrate on demand service delivery across multi-domain/multi-vendor research network test-beds on a European and Worldwide scale. The test-bed will include:**
  - EU NRENs: SURFnet, CESNET, PIONIER as well national test-beds (VIOLA, OptiCAT, UKLight)
  - GN2, GLIF and Cross Border Dark Fibre connectivity infrastructure
  - GMPLS, UCLP, DRAC and ARGON control and management planes
  - Multi-vendor equipment environment (ADVA, HITACHI, NORTEL, Vendor's equipment in the participating NREN infrastructure)

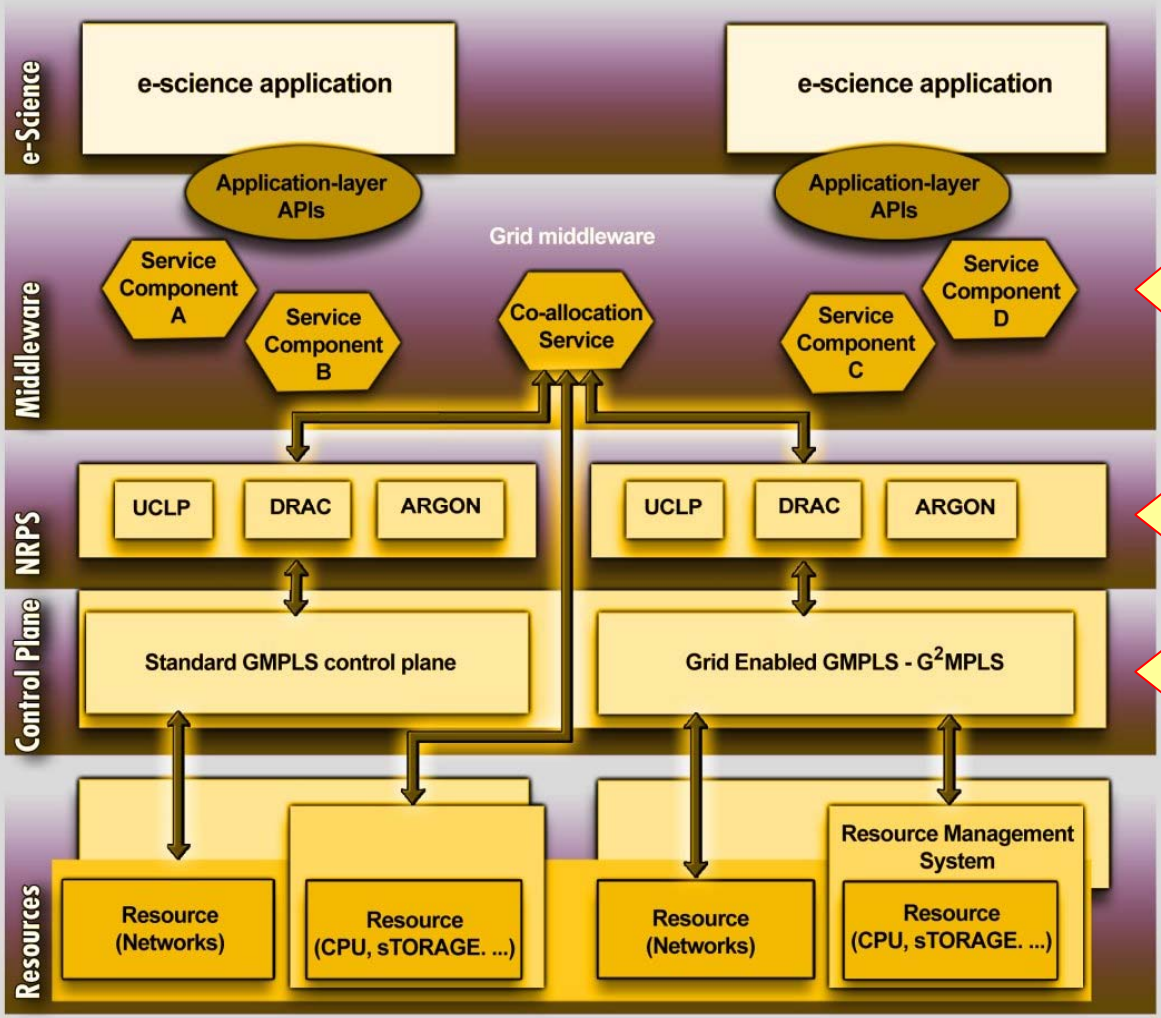
PROJECT KEY FEATURES 2/3

- Develop integration between application middleware and transport networks, based on three planes:
  - Service plane:
    - Middleware extensions and APIs to expose network and Grid resources and make reservations of those resources
    - Policy mechanisms (AAA) for networks participating in a global hybrid network infrastructure, allowing both network resource owners and applications to have a stake in the decision to allocate specific network resources
  - Network Resource Provisioning plane:
    - Adaptation of existing Network Resource Provisioning Systems (NRPS) to support the framework of the project
    - Implementation of interfaces between different NRPS to allow multi-domain interoperability with Phosphorus' resource reservation system
  - Control plane:
    - Enhancements of the GMPLS Control Plane (G^2MPLS) to provide optical network resources as first-class Grid resource
    - Interworking of GMPLS-controlled network domains with NRPS-based domains, i.e. interoperability between G^2MPLS and UCLP, DRAC and ARGON

PROJECT KEY FEATURES 3/3

- Studies to investigate and evaluate further the project outcomes :
  - Study resource management and job scheduling algorithms incorporating network-awareness, constraint based routing and advance reservation techniques
  - Develop a simulation environment, supporting the Phosphorus network scenario
- Disseminate the project experience and outcomes, toolkits and middleware to NRENs and their users, such as Supercomputing centres

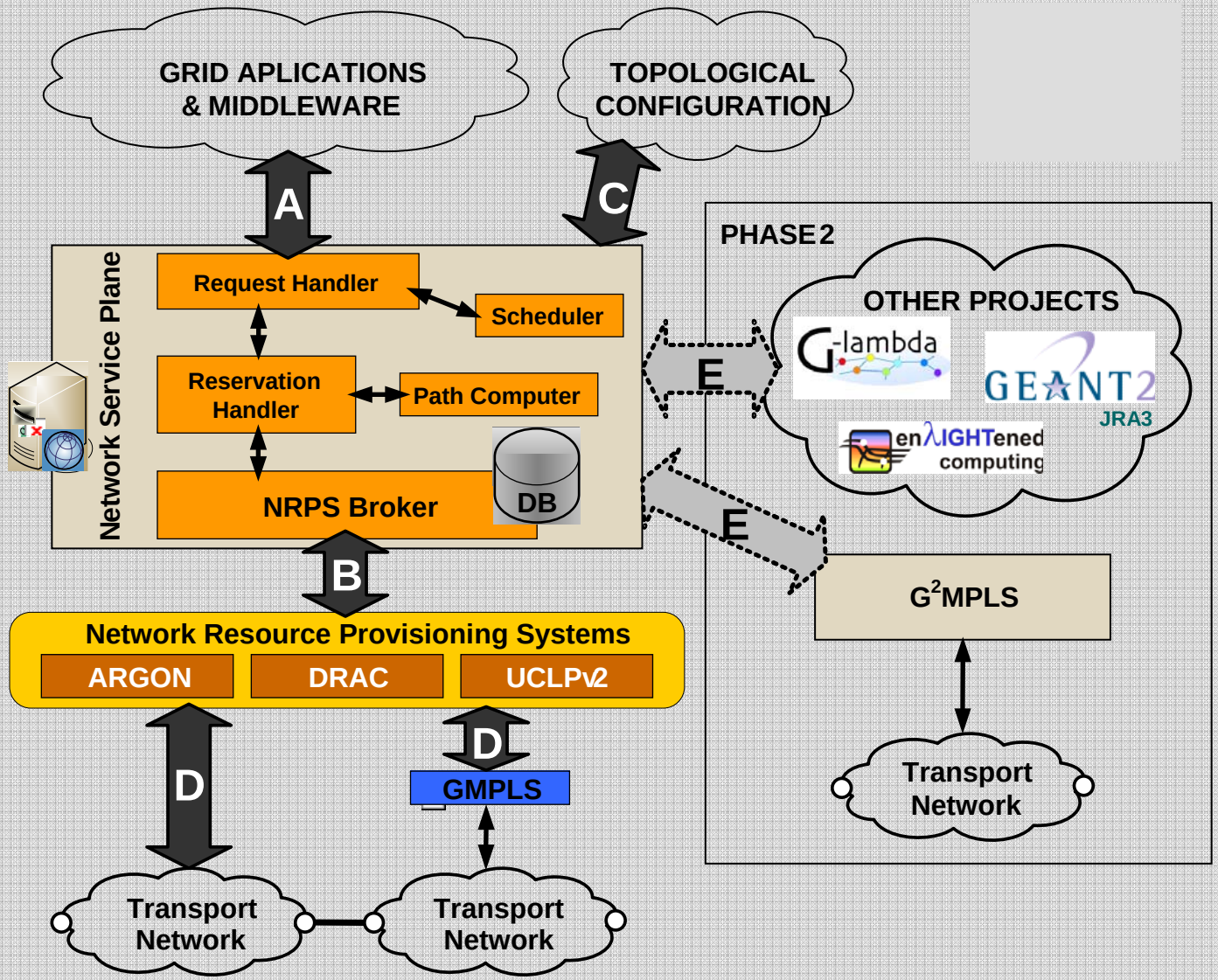
PHOSPHORUS ARCHITECTURE



The different domains of the **Phosphorus'** test-bed will have:

- Grid middleware
  - UNICORE as a reference point
  - AAA policies
- Three types of NRPS:
  - UCLP
  - DRAC
  - ARGON
- Two "flavours" of GMPLS
  - standard (*Ph. 1*)
  - Grid-enabled (*Ph. 2*)

NRPS and NSP system interfaces



A) Northbound IF:
It receives the reservation requests from the GRID Middleware.

B) East-West IF:
It is in charge of the communication between NRPSs.

C) Topological IF:
It is used to indicate to the NSP which are the resources under control (NRPSs, endpoints, links).

D) Southbound IF:
It Communicates the NRPSs and the lower layers (GMPLS or transport layer).

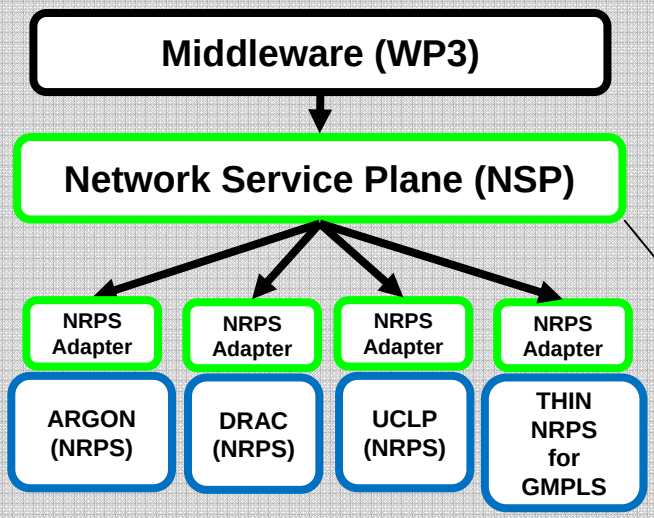
E) Phase 2 IF:
It provides interoperability between the NSP and the G²MPLS CP or other projects.

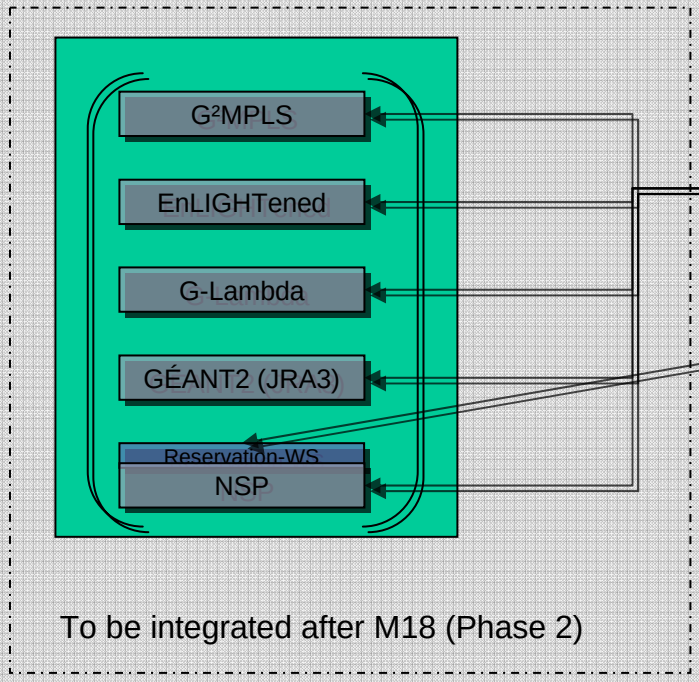
NRPS developments Phase 1 (till 2Q2008)

Provided by the partners

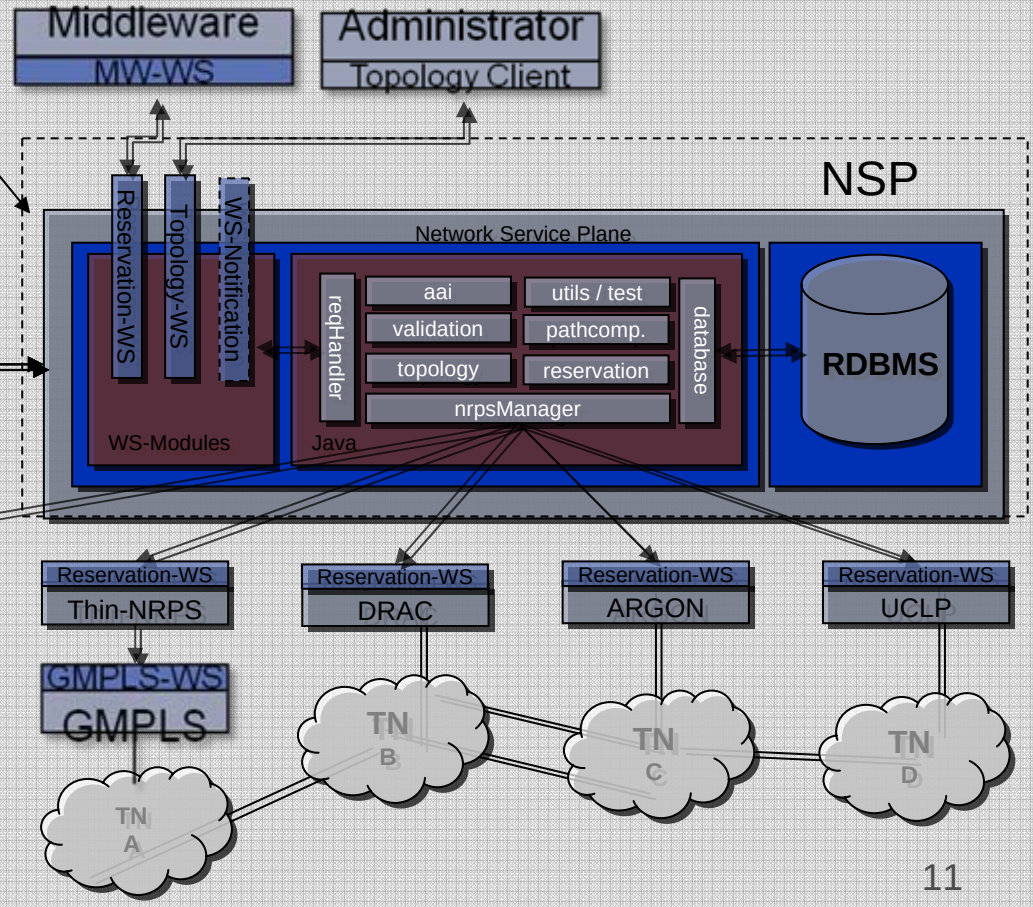
Implemented in WP1

- The NRPS Adapter is a generic open source module with a specific interface towards the NRPS
- After M18 the NSP will be enhanced to implement a distributed system architecture

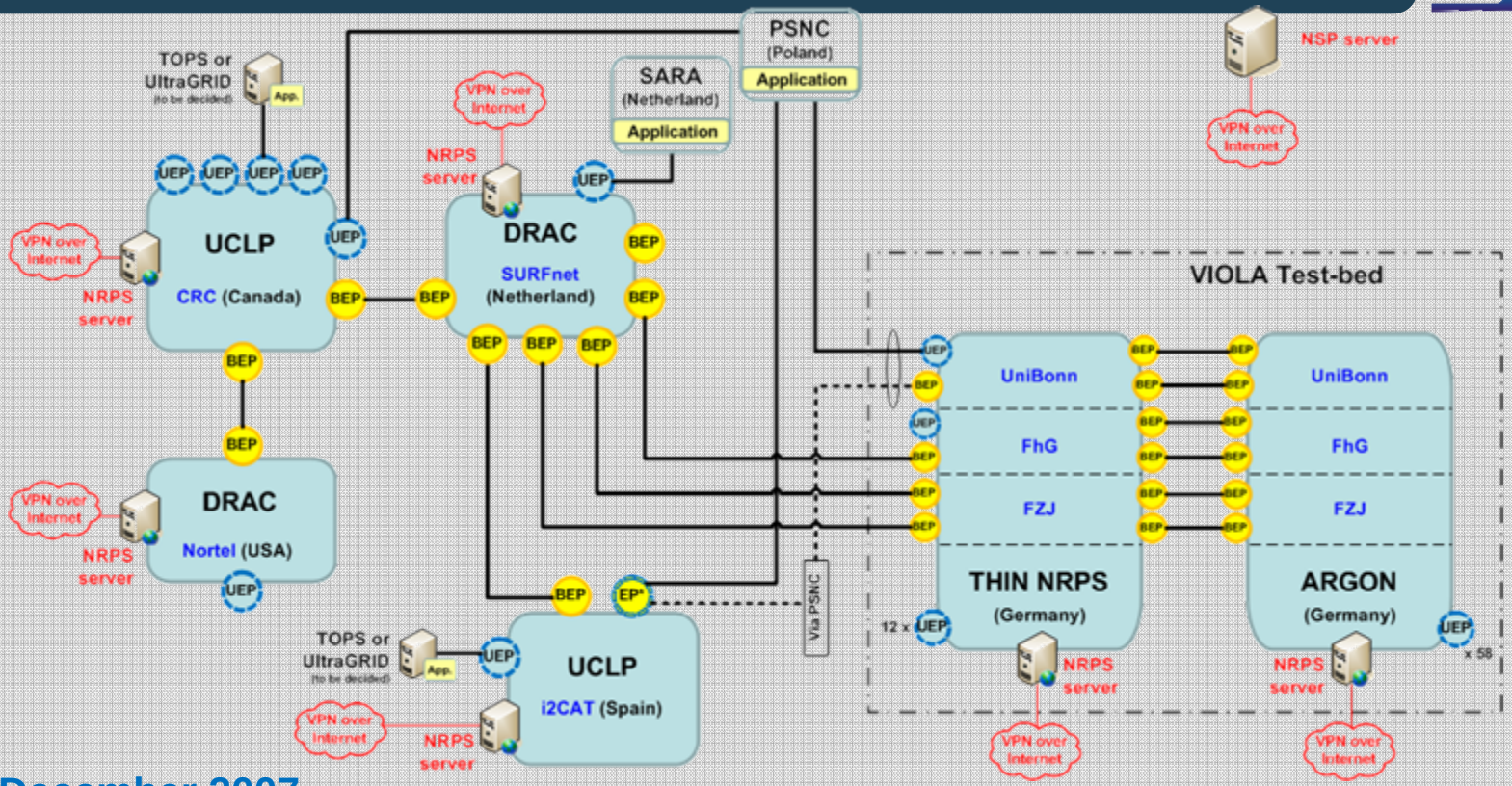




'zoom'



Management architecture. Logical view

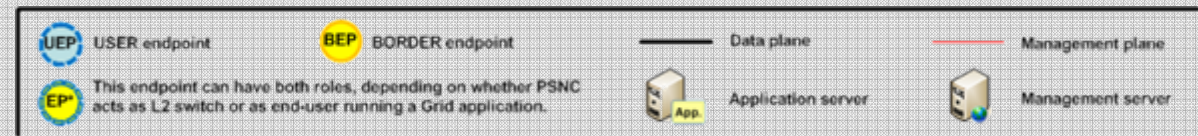


Step1 by December 2007

- Add topology information to the system (domain, link, endpoint)
- Create an immediate/advance reservation E2E across different domains/NRPS, with a selected BW and specific duration.
- 1 PHY port corresponds to 1 EndPoint
- No inter-domain BW management
- Use a THIN NRPS to control a GMPLS domain
- The NSP provides also a Path Computation algorithm

Step2 by March 2008

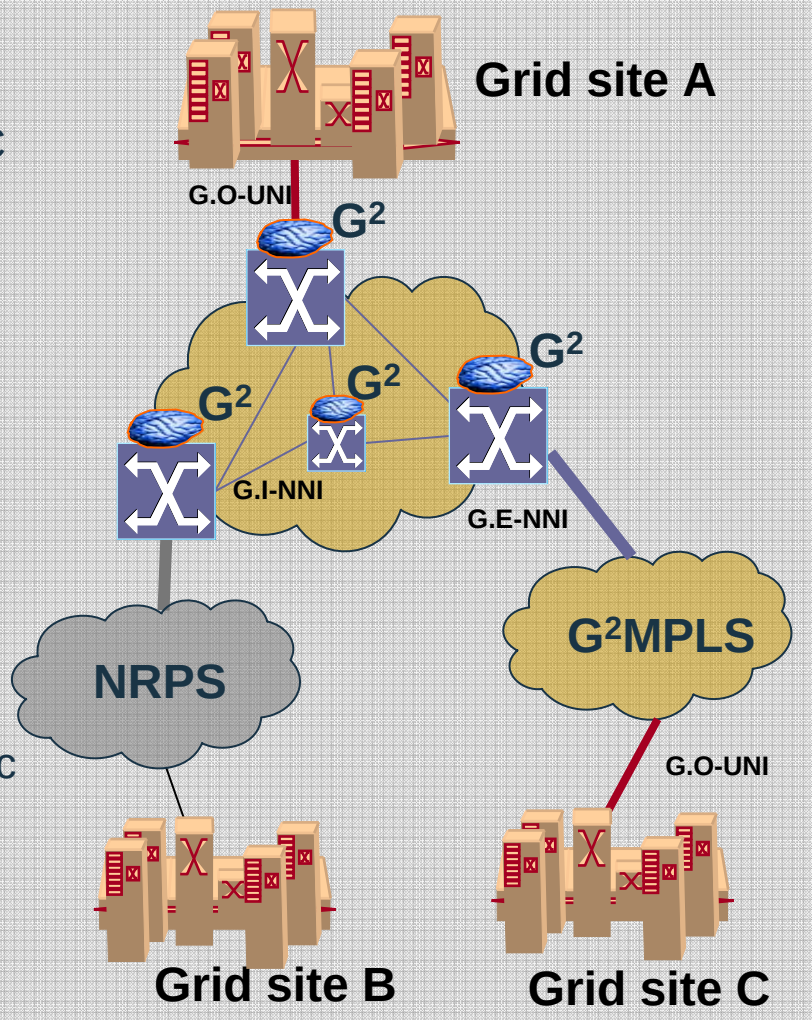
- Finish implementation of all advance reservation features to be able to create an advanced reservation end-to-end connection across domains
- Use embedded AAI in corporation with WP3
- Implement BW management that considers underlying technology
- Integrate the MW-WS



Enhancements to the GMPLS Control Plane for Grid Network Services (GNS)

Extensions to the GMPLS CP for automatic and single-step setup of Grid & network resources

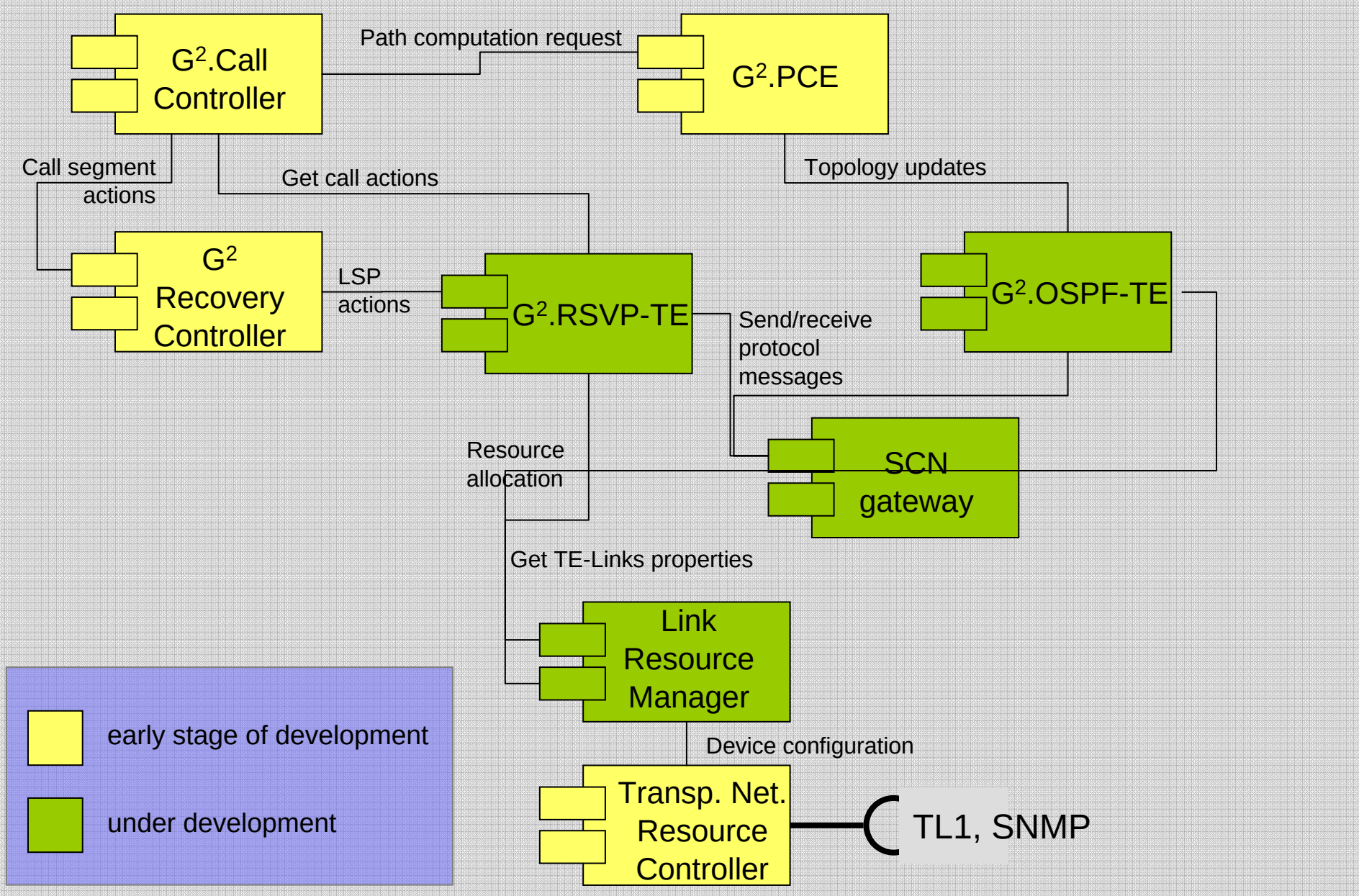
- Grid-GMPLS (G²MPLS) main research tracks:
  - seamless coexistence with NRPS (UCLP, DRAC and ARGON) & Grid MW
  - Grid-aware network reference points (G.O-UNI, G.E-NNI, G.I-NNI)
  - CBR algorithms for recovery and TE
  - Integration with AAA system
- Expected innovation in the field of co-allocation of Grid and network resources, because of
  - faster dynamics for service setup
  - adoption of well-established procedures for traffic engineering, resiliency and crankback
  - uniform interface for the Grid-user to trigger Grid & network transactions



G²MPLS roadmap

- In a short-term
  - Open source GMPLS Control Plane prototype for:
    - Optical LSP setup/teardown via UNI interface compatible with OIF UNI 2.0
    - LSP protection/restoration and crankback
    - Flexible adaptation to the management interfaces exported by the underlying Transport Plane (e.g. TL1, SNMP)
    - Multi-domain operations through OIF ENNI compliant interface
  - Basic modules for routing and signalling in single control domain going to be delivered by Q3-2007
  - Recovery strategies and inter-domain expected by Q1-2008
- In a longer-term (but within project lifetime)
  - Grid-GMPLS enhancements to the Control Plane prototype, in terms of
    - Discovery and advertisement of Grid capabilities and resources of the participating Grid sites
    - Service setup and maintenance
      - Coordination, co-allocation and configuration of grid and network resources associated with a Grid job
      - Recovery of the installed network services and possible escalation of procedures
      - Advance reservations of Grid and network resources
    - Service monitoring: retrieving of the status of a Grid job and the related network connections

G²MPLS high level software architecture



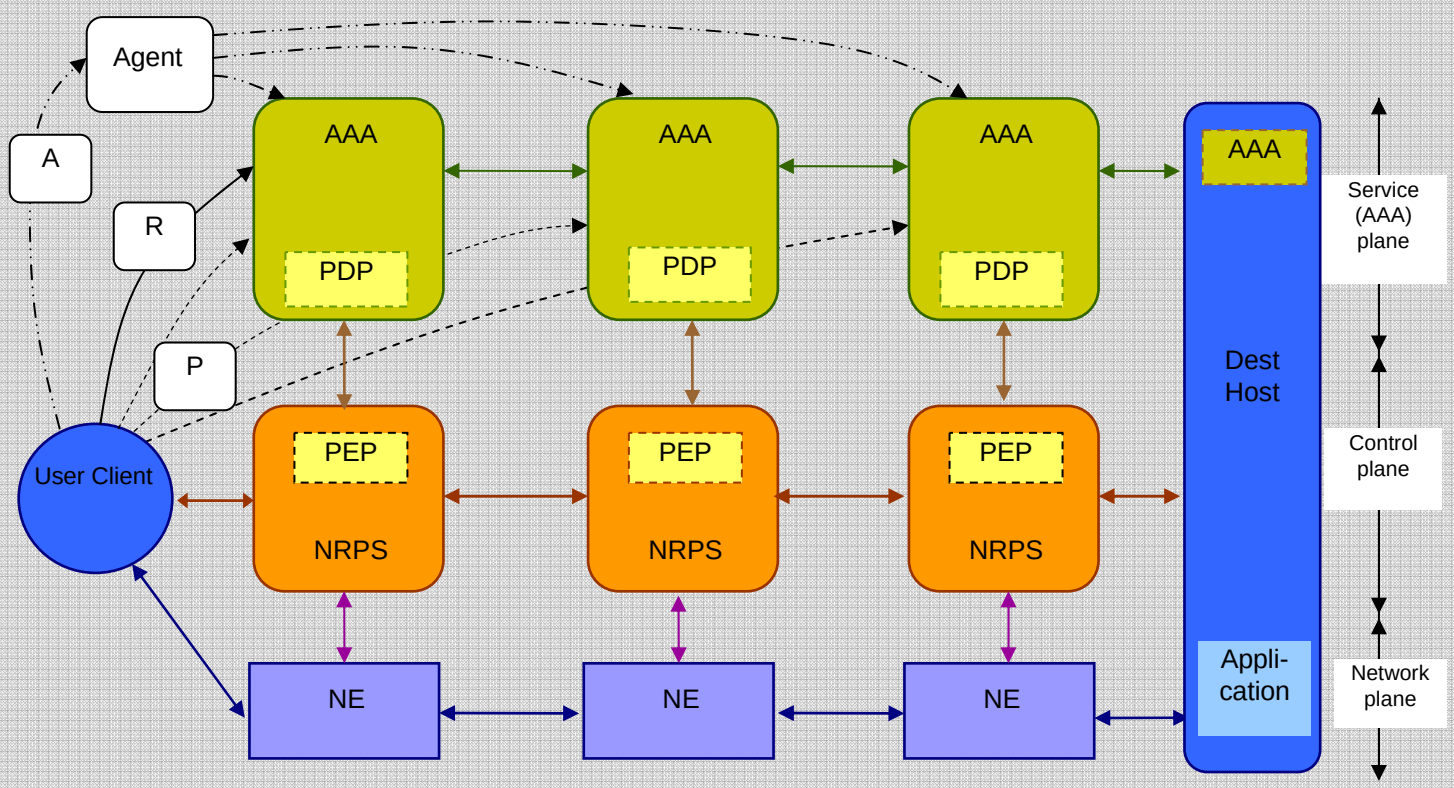
Authentication, Authorisation, Accounting

- The purpose of AAA activity is to define the architecture and subsequently design, implement and test an AAA AuthZ infrastructure to support policy based on-demand network resource provisioning and access across different administrative domains
- It will respond to practical need for policy based and customer/project centric resource provisioning
- The development is based on the experience available at UvA in developing AuthZ services for Network resources provisioning and Grid applications and experiences of other activity partners
- This activity contributes to the standardisation activity at IETF and OGF on developing standard AuthZ architecture and components for Grid based applications and token-based policy enforcement mechanisms in Network resource provisioning
- It will ensure future compatibility and interoperability of the proposed solutions by using wide use cases base and coordinating development with different Grid and network related projects such as Internet2, GN2, EGEE, Globus

AAA Authorisation Service Architecture

- Operates as integral part of the general Network Resource Provisioning Service (NRPS)
  - Can be called from NRPS domain controller, or
  - Can drive reservation/provisioning process
- Incorporates 3 basic generic AAA AuthZ sequences
  - Push, pull, and agent
- Implements 3 basic AAA AuthZ operational models for complex multidomain network resource provisioning:
  - Polling, Relaying and Agent
- Supports different and multiple policy decision mechanisms
  - Multiple policy combination and/or mapping
  - Creates and supports dynamic user and resource associations
  - Supports different AuthZ frameworks interoperation
- Supports different policy enforcement models
  - Using AuthZ tickets and tokens at Service and Control planes
  - Integration of token based enforcement mechanism inside GMPLS control plane using RFC2750 policy data object

AAA AuthZ components and basic operational models



TVS – Token Validation Service

DRAM – Dynamic Resource Allocation and Mngnt

PDP – Policy Decision Point

PEP – Policy Enforcement Point

- **Polling sequence (P)** when the user client polls all individual network domains to make a reservation.
- **Relay (R) or hop-by-hop reservation** when the user client contacts only the local network domain/provider and each consecutive domain provides path to the next domain.
- **Agent (A**) sequence when the User delegates network provisioning negotiation to the Agent that will do all necessary negotiations with all involved domains

AAA AuthZ implementation and products

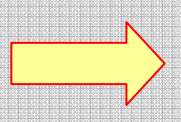
- AAA AuthZ server that can be used as a central AuthZ service to support multidomain network resource provisioning or as external AuthZ service for different NRPS's
- GAAAPI supporting all necessary interactions and interfaces (local and remote Web services based) between application, PEP and PDP
  - Pluggable component used on the server side and NRPS/client side
  - Will support standard AuthN/AuthZ protocols and credential formats such as SAML, XACML, Shibboleth, VOMS
- Token Validation Service (TVS) to support token based policy enforcement mechanism at service/control plane (using XML tokens) and at GMPLS control plane
- Common policy format that support major network resource management and provisioning models and can be mapped to different policy formats used in currently existing NRPS
  - Will be implemented as special XACML policy profile
- IETF ForCES based Token Based Switch (TBS) that allows creating overlay virtual infrastructure of the provisioned network by using token based GMPLS signalling

Middleware and Applications goals

- Main goals
  - Adapt and extend Grid middleware to present PHOSPHORUS services to applications
    - Coordinated reservation and allocation of compute, storage, and network resources
    - Integration of MetaScheduling Service (MSS) into UNICORE 6
    - Allow semantic annotation of resources to improve resource selection
    - Interface with NPRS, Control Plane and AAA activities
  - Adapt and extend network-based applications to evaluate and demonstrate the PHOSPHORUS developments in the test-bed
    - Make use of network / grid resource reservation services
    - Prepare for deployment in test-bed, support tests

Middleware and Applications

- Integration of network reservation services into existing Grid middleware
  - services for user-driven or application-driven set-up of execution environments with dedicated capabilities & performance
    - Compute nodes, storage systems, visualization devices
    - Network resources with defined QoS
- Integration of applications
  - WISDOM: Wide in silicio docking on Malaria
  - KoDaVis: collaborative, distributed visualization of huge data sets
  - TOPS: Streaming of ultra high resolution data sets over lambda networks
  - DDSS: Distributed Data Storage System
  - INCA: Intelligent Network Caching Architecture



Provide application access to PHOSPHORUS services and showcase their benefit via applications

Applications, part 1

- WISDOM - Wide In Silica Docking On Malaria:
  - large scale molecular docking on malaria to compute million of compounds with different software and parameter settings (in silico experimentation)
  - The **goal within Phosphorus** is the deployment of a CPU-intensive application generating large data flows to test the Grid infrastructure, compute and network services
- KoDaVis – Distributed visualisation
  - The main **objective in Phosphorus is to adapt KoDaVis to the Phosphorus' environment to make scheduled synchronous reservations of its resources via the UNICORE middleware**:
    - Compute capacity on the data server and the visualisation clients
    - Allocate network bandwidth and QoS between server and clients

Applications, part 2

- TOPS – Technology for Optical Pixel Streaming
  - Streaming of Ultra High Resolution Data Sets over Lambda Networks
  - Adapt the idea of video streaming: constant stream of pixels from renderer to display
  - Use lossy protocols for long distance connectivity: High performance TCP hard to achieve, UDP performance trivial
  - Light weight application – scalable bandwidth usage
- DDSS - Distributed Data Storage System
  - GridFTP - a high performance, secure, reliable data transfer protocol optimized for high-bandwidth wide-area IP networks.
  - Backup/archive copies with TSM (Tivoli Storage Manager) over the test network
  - possible scenarios of usage
    - data gathering or data distributing
    - backups of large medical data volumes (bandwidth demanding) from one or many clients
    - from one or many clients

Supporting Studies

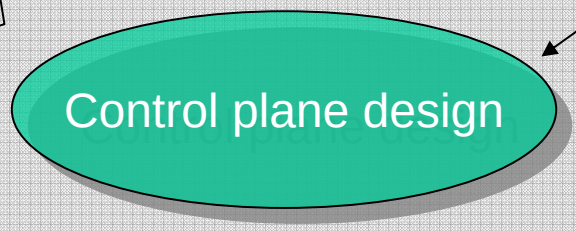
- Job demand models
- QoS resource scheduling
- Grid job routing algorithms
- Physical layer constraints
- Advance reservations

- Optical network
- Advanced control plane
- Network service plane

Middleware and Applications activity

Control Plane activity

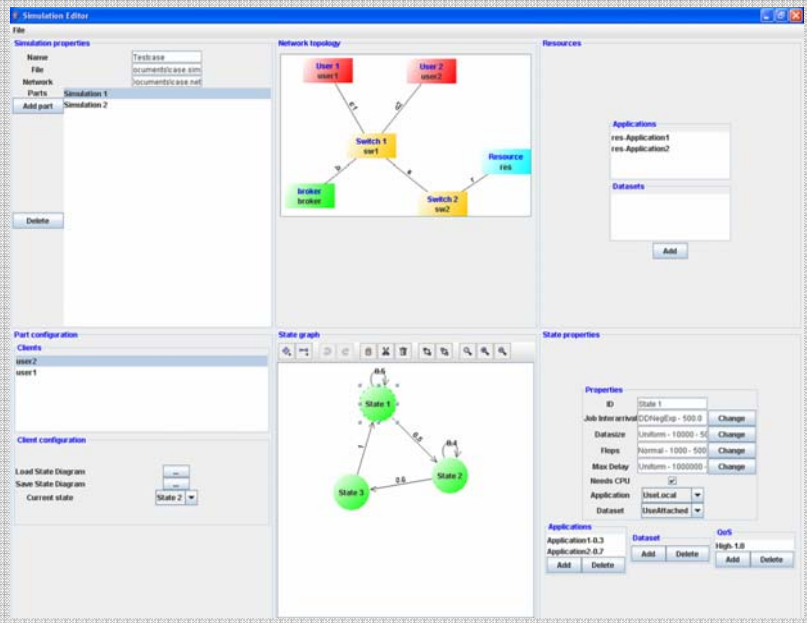
NPRS activity



- Architectural issues
- Integration strategies
- Recommendations

Supporting Studies results

- Research-oriented results
  - Job demand models
  - Grid job routing algorithms
    - Include physical impairments in routing decision
    - Multi-domain job routing
  - QoS-aware resource scheduling
  - Support for advance reservations

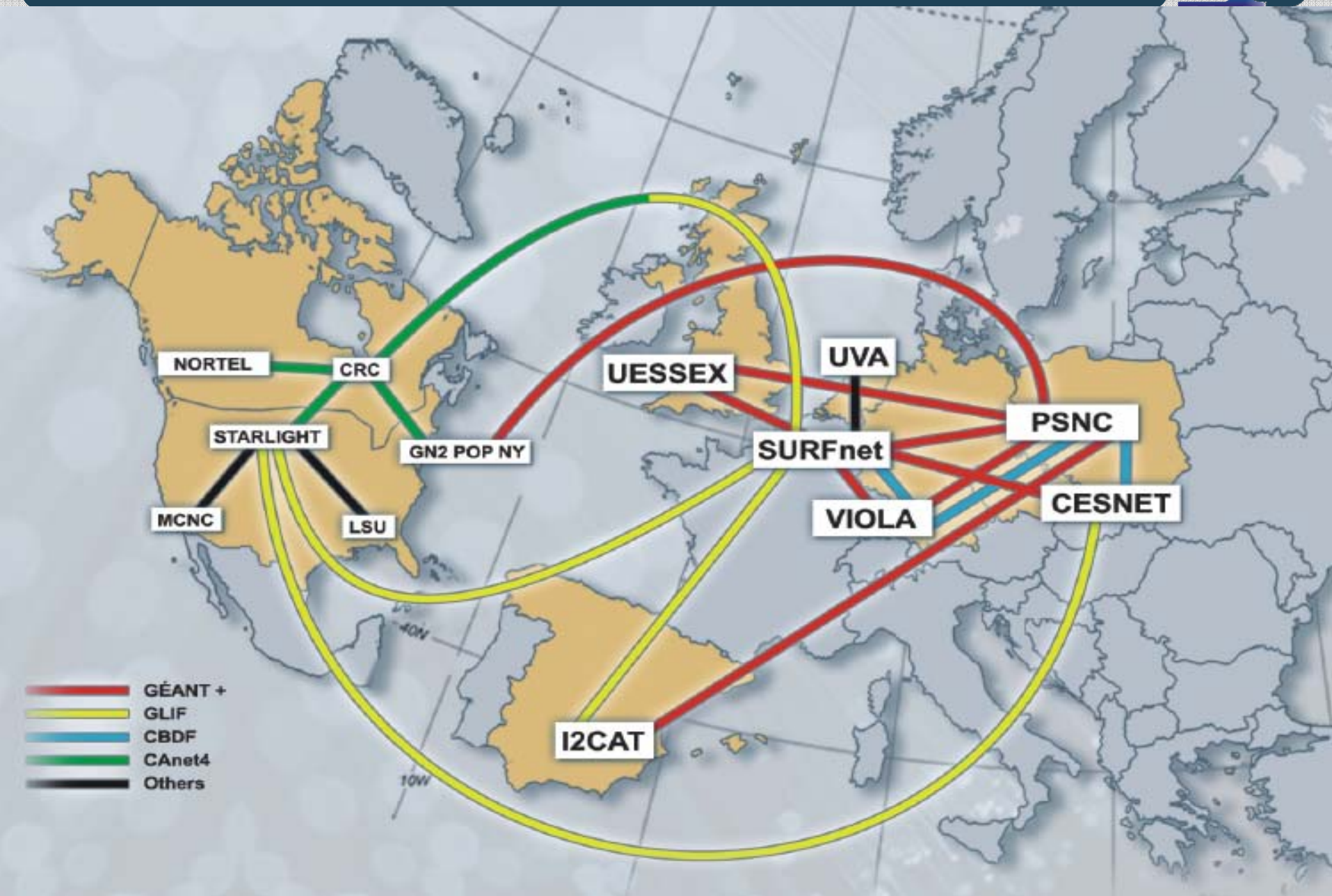


- Simulation environment
  - Java, no dependencies, discrete event
  - Modeling network and Grid resources
  - Dynamic OCS path set-up and tear-down
  - Flexible job models (based on Markov states)
  - GUI to define network topology and traffic models
  - Code available on svn: http://phosphorus.atlantis.ugent.be/phosphorus
    - Mail marc.deleenheer@intec.ugent.be for login details

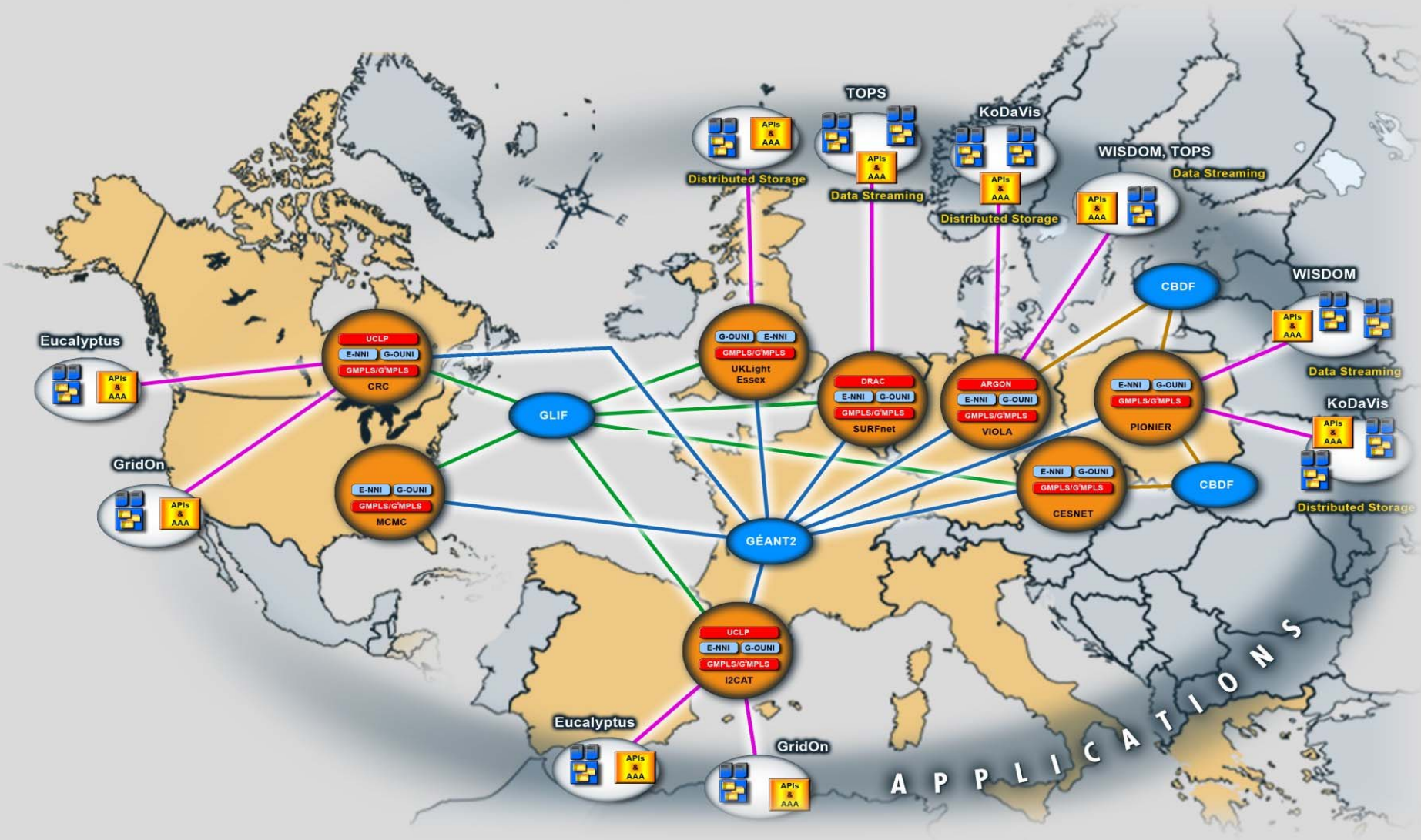
Test-bed & Demonstration activities

- Objectives:
  - Requirements analysis and design of the test-bed
  - Construction of the test-bed and configuration of all related software components, middleware and applications
  - Tests of project's developments
  - Demonstration of project's results
  - Real scientific application and real life scenarios used to verify project developments
- Architecture:
  - Several local test-bed supporting different technologies (optical, Ethernet, SDH switching DWDM, etc.), GRID & HPC resources and applications
  - Lightpaths to interconnect local test-beds
- Current status:
  - Most resources in local test-beds available for applications, almost all interconnection in place and NPRSes installed
  - Verification of the test-bed will be finished till end of September 2007
  - Installation of middleware and applications in progress

Current test-bed interconnection data links



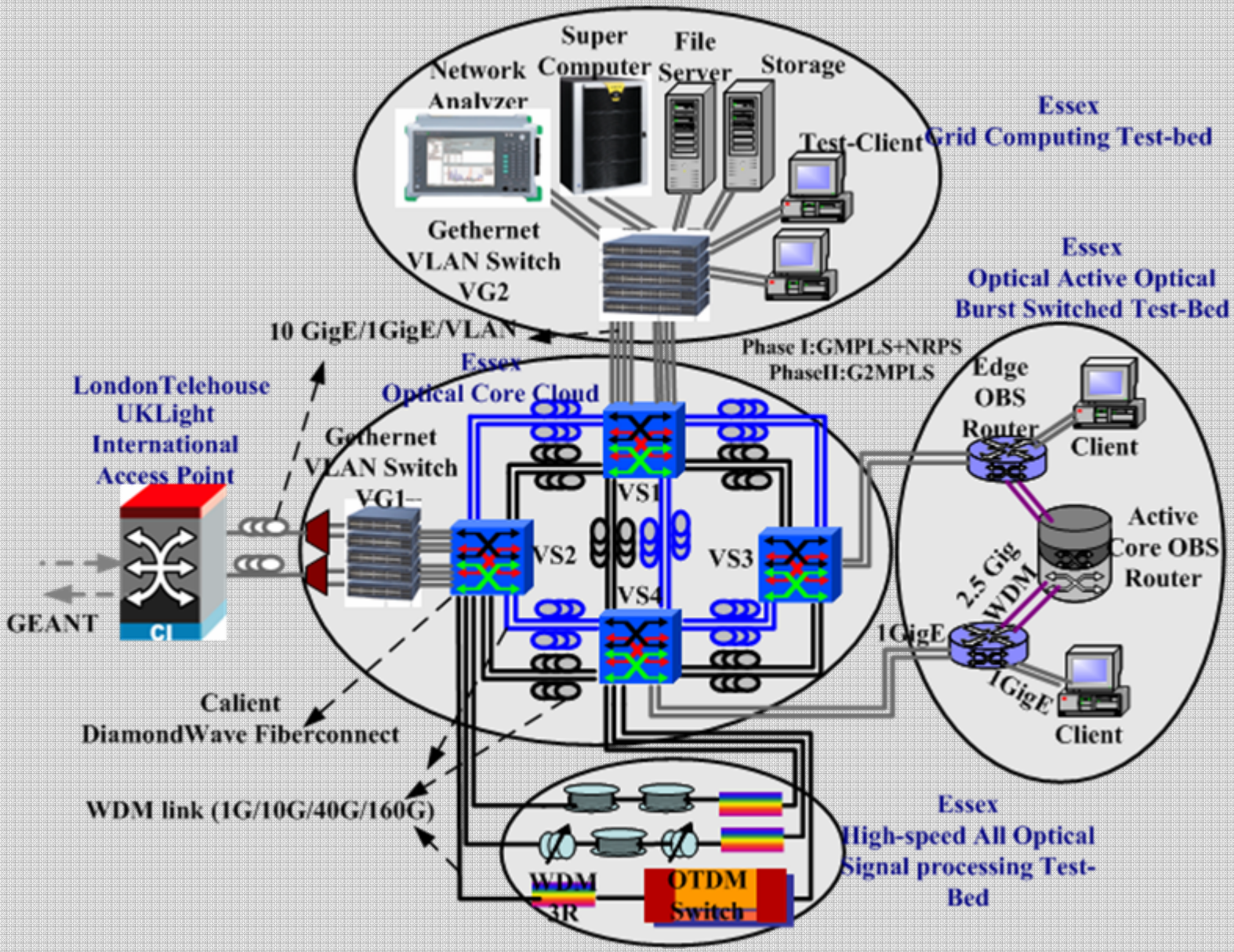
MULTI-DOMAIN TESTBED



UESSEX local test-bed

- Data Plane
  - Fiber/wavelength switched test-bed
    - Calient Diamond Wave Fiberconnect switch
  - Sub-wavelength Switched test-bed
    - OBS Test-bed
    - OTDM test-bed
  - VLAN Capable layer2/layer3 managed switch
- Control plane, NRPS & middle
  - UCLP, phosphorus GMPLS ,phosphorus G2MPLS, Globus
- Connectivity
  - 10GE & 1GE to Geant, 1GE to PSNC , 1GE to Surfnet
- GRID and testing resources:
  - 4 test-clients, 4TB storage, 1 SGI Atlix 420 super computer
- Applications: DDSS, KoDaVis

UESSEX local test-bed (cont.)



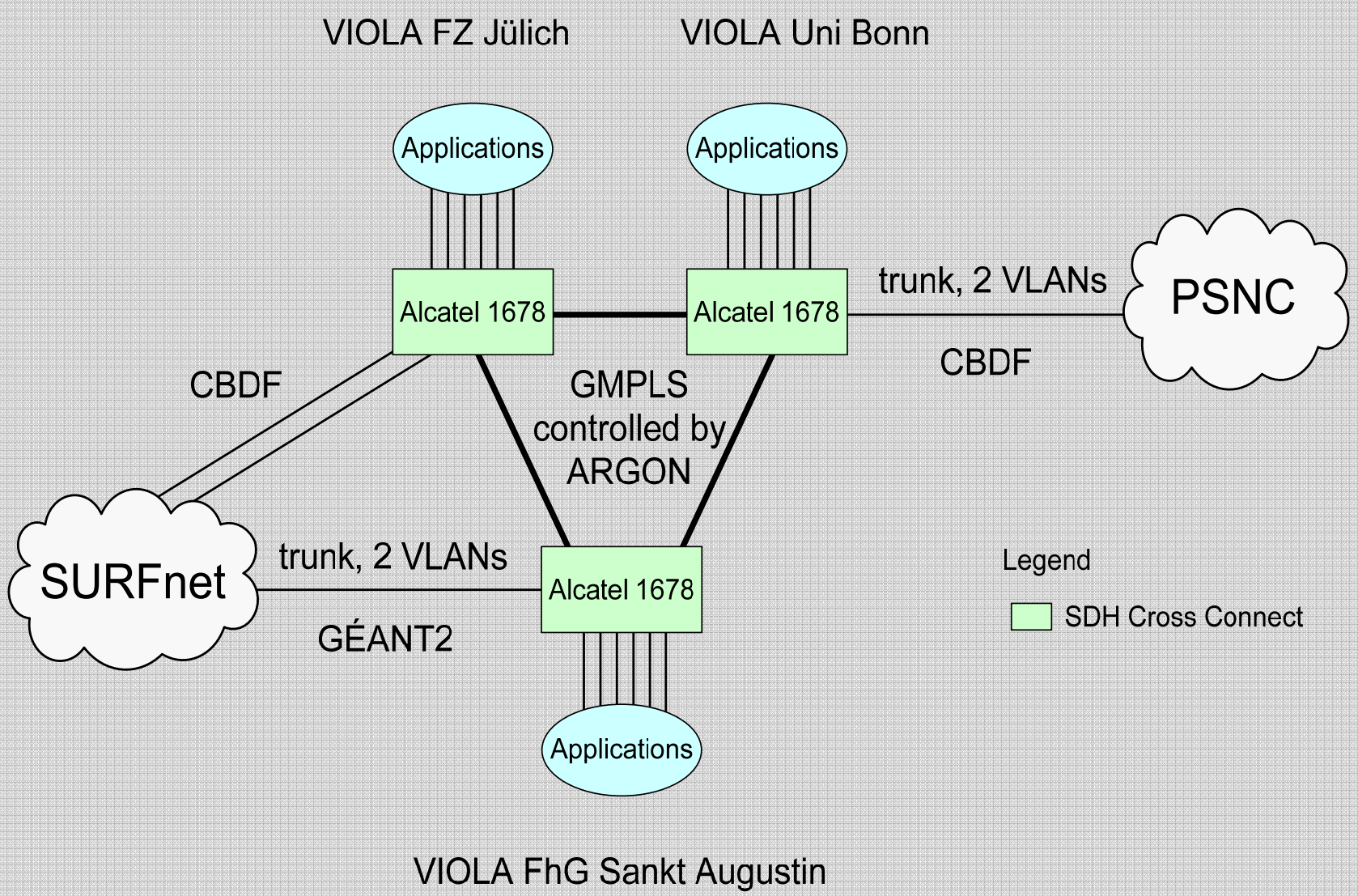
VIOLA local test-bed

- VIOLA test-bed is provided by: FHG, FZJ, UniBonn
- Switching equipment:
  - 3 SDH cross connects (Alcatel 1678 MCC)
  - 3 Riverstone 15008 Gigabit Ethernet switches
  - 3 Gigabit Ethernet switches
- Management plane – ARGON (for SDH cross connects and GE switches)
- Control plane - GMPLS (for SDH cross connects)
- GRID resources:
  - CRAY X-D1 (14 dual Opteron computing nodes)
  - 17-node dual Xeon cluster
  - access to TSM server and type library
  - i-CONE display system
- Applications: DDSS, TOPS, KoDaVis, WISDOM

VIOLA local test-bed

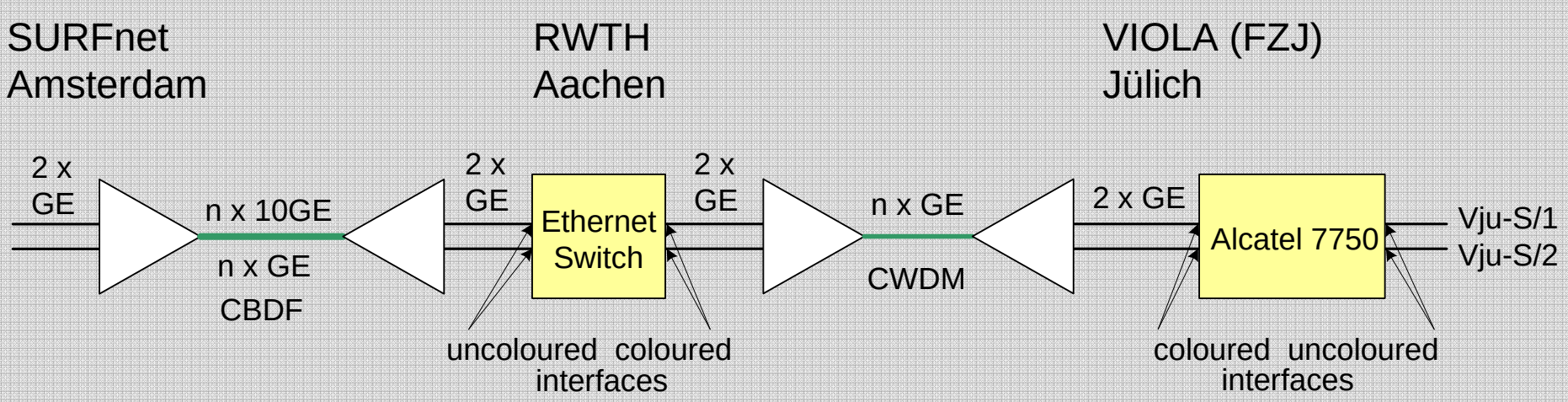
- Each of the VIOLA sites will be connected to the Phosphorus testbed by two connections
- The three VIOLA sites are connected to each other by a SDH network controlled by GMPLS
- Switching of connections
  - Via GMPLS the connections to the Phosphorus testbed can be switched to applications at all three VIOLA sites
  - Via GMPLS the connections to the Phosphorus testbed can be switched to each other, allowing Phosphorus connections to use the VIOLA testbed as a transit network
- Support of JUMBO Frames

VIOLA local test-bed (cont.)



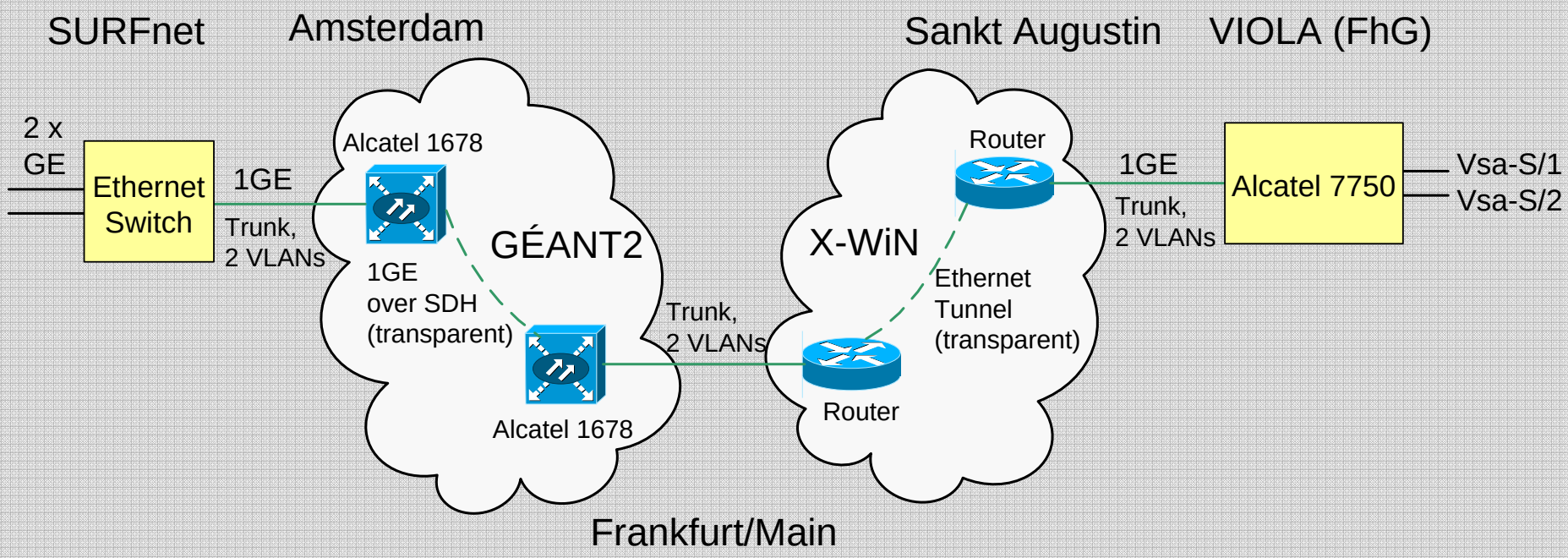
VIOLA local test-bed (cont.)

VIOLA Phosphorus testbed connections
① SURFnet – VIOLA (FZJ)



VIOLA local test-bed (cont.)

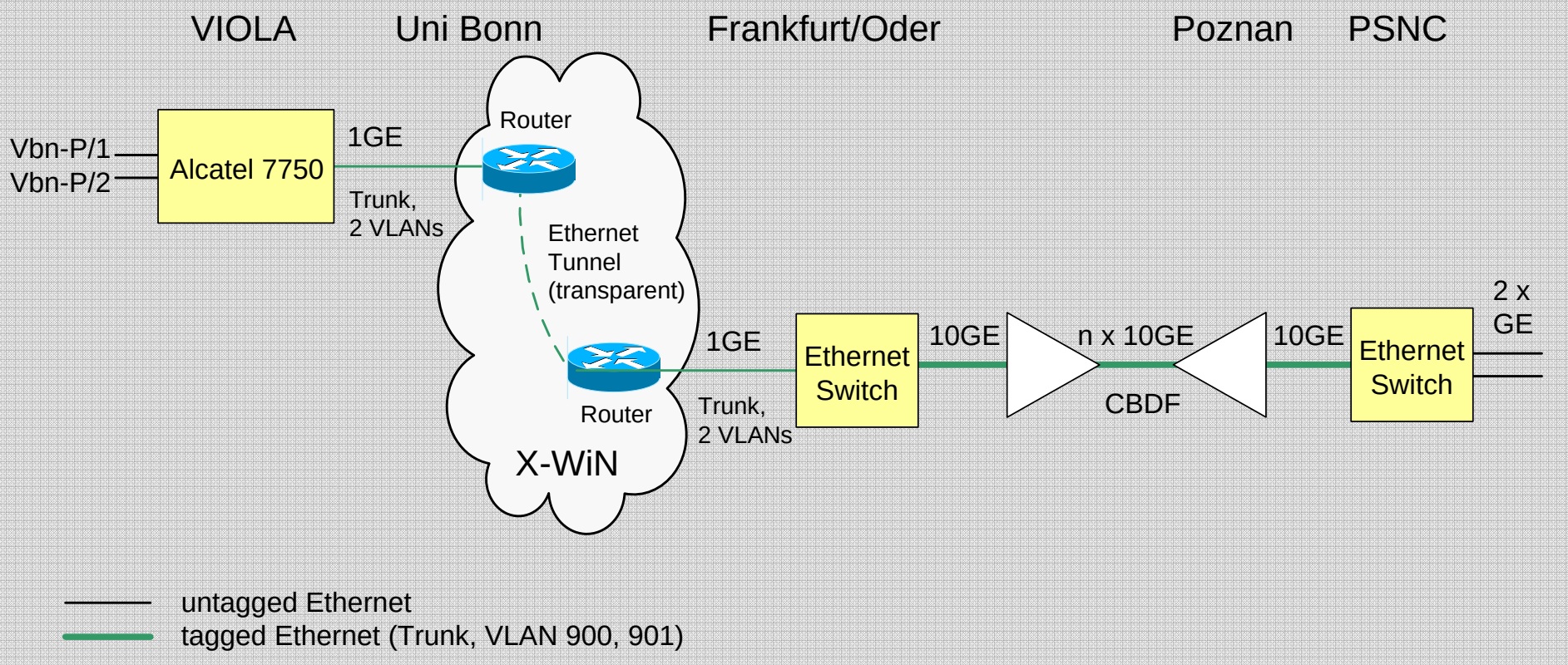
VIOLA Phosphorus testbed connections
② SURFnet – VIOLA (FhG)



untagged Ethernet
tagged Ethernet (Trunk, VLAN 900, 901)

VIOLA local test-bed (cont.)

VIOLA Phosphorus testbed connections
③ PSNC - VIOLA (UniBonn)

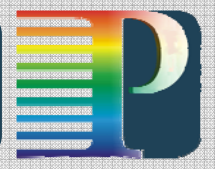


Dissemination, Cotribution to standards, Liaisons

- Disseminate information concerning the technical developments to
  - NRENs
  - Related projects:
    - GN2, NOBEL, EGEE, UCLP, Enlightened Computing, G-Lambda…
- Coordinate direct contributions to standards
- Build a collaborative framework for participation to test-bed activities from within and external to EU

The Three Continent Collaboration-MCNC Raleigh Jan.'06

PHOSPHORUS



Thank you