PHOSPHORUS

Lightpaths to the User using *Harmony*

A Phosphorus Approach for Multi-domain Resource Brokering

Joan Antoni García-Espín
i2CAT Foundation

Phosphorus Workshop @ TERENA Networking Conference 2009
Málaga, Spain, June 7th

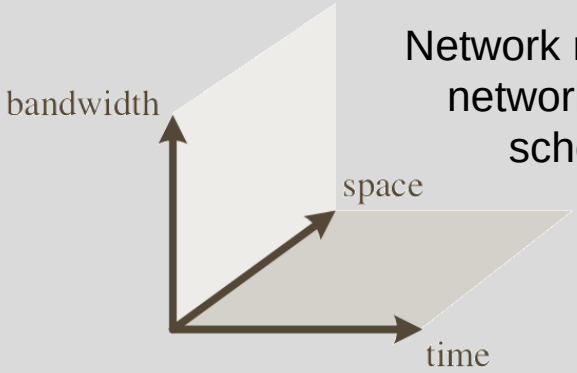
Outline

1. Lightpaths to the User, a matter for research.
2. The Harmony system
   - Challenges
   - Network Resource Provisioning Systems (NRPSs)
   - Architecture, Services and Test-bed
3. Evaluation of Harmony: performance and scalability analysis
4. Harmony and the outer world
5. Related work and Dissemination

LP to the User, a matter for research

- Why?
  - Resources and/or Services are spread over the Network
  - User is not always located where resources inhabit
  - Concretely, is the network slowing down Grid growth?
- Our motivation…

Resource co-allocation and interconnection



Network reservations, network resource scheduling

NRPS Interworking, including G(2)MPLS!

ARGON (MPLS/GMPLS)

ARGIA/ UCLP

DRAC (Nortel)

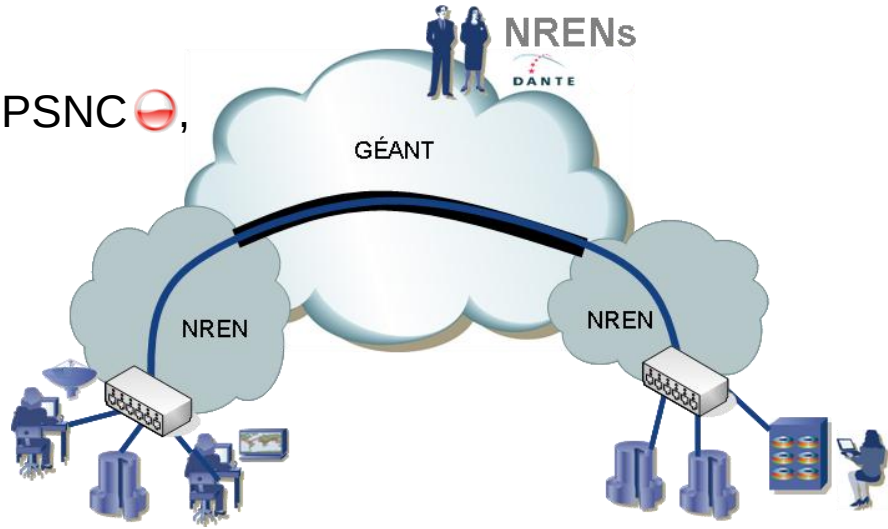
Interfacing towards the outer world: IDC, AutoBAHN, Grid applications, (big **etc.** goes here!)

Harmony system

Harmony system is…

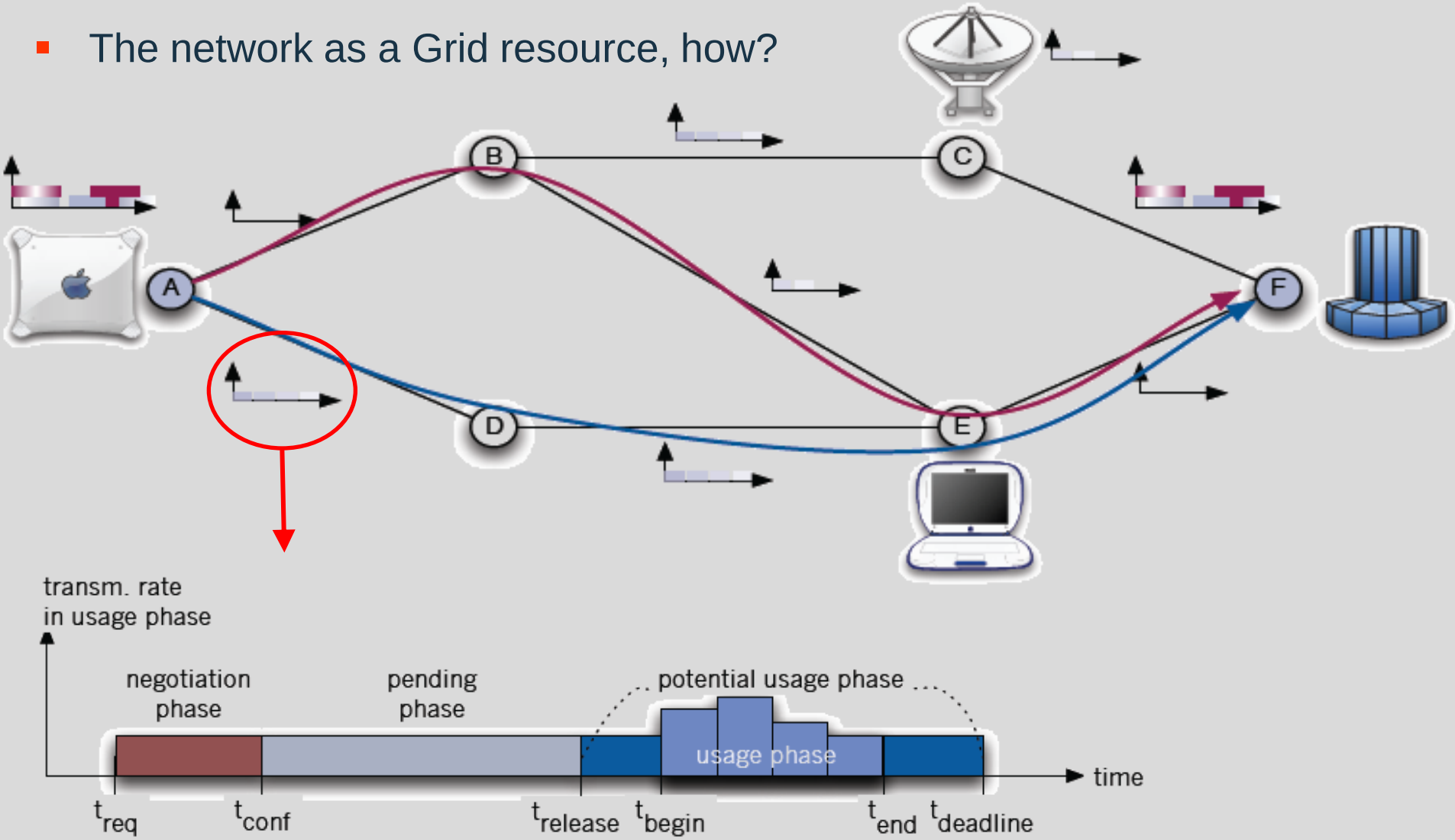
“Harmony” branding, presented in OGF23 (Barcelona, May 2008)

- A **multi-domain path provisioning system** where users and Grid applications can **book in advance** end-to-end paths and network resources with **AAI**
- Harmony allows domain interoperability by performing **inter-domain network resource brokering** over Network Resource Provisioning Systems (NRPS)
- **Harmony Service Interface** (HSI), the common language.
- NRPS considered:
  - **Argia/UCLP** (CRC, i2CAT, UEssex, PSNC, KISTI)
  - **ARGON** (VIOLA test-bed)
  - **DRAC** (SURFnet)
  - **GMPLS** Control Plane (VIOLA test-bed)



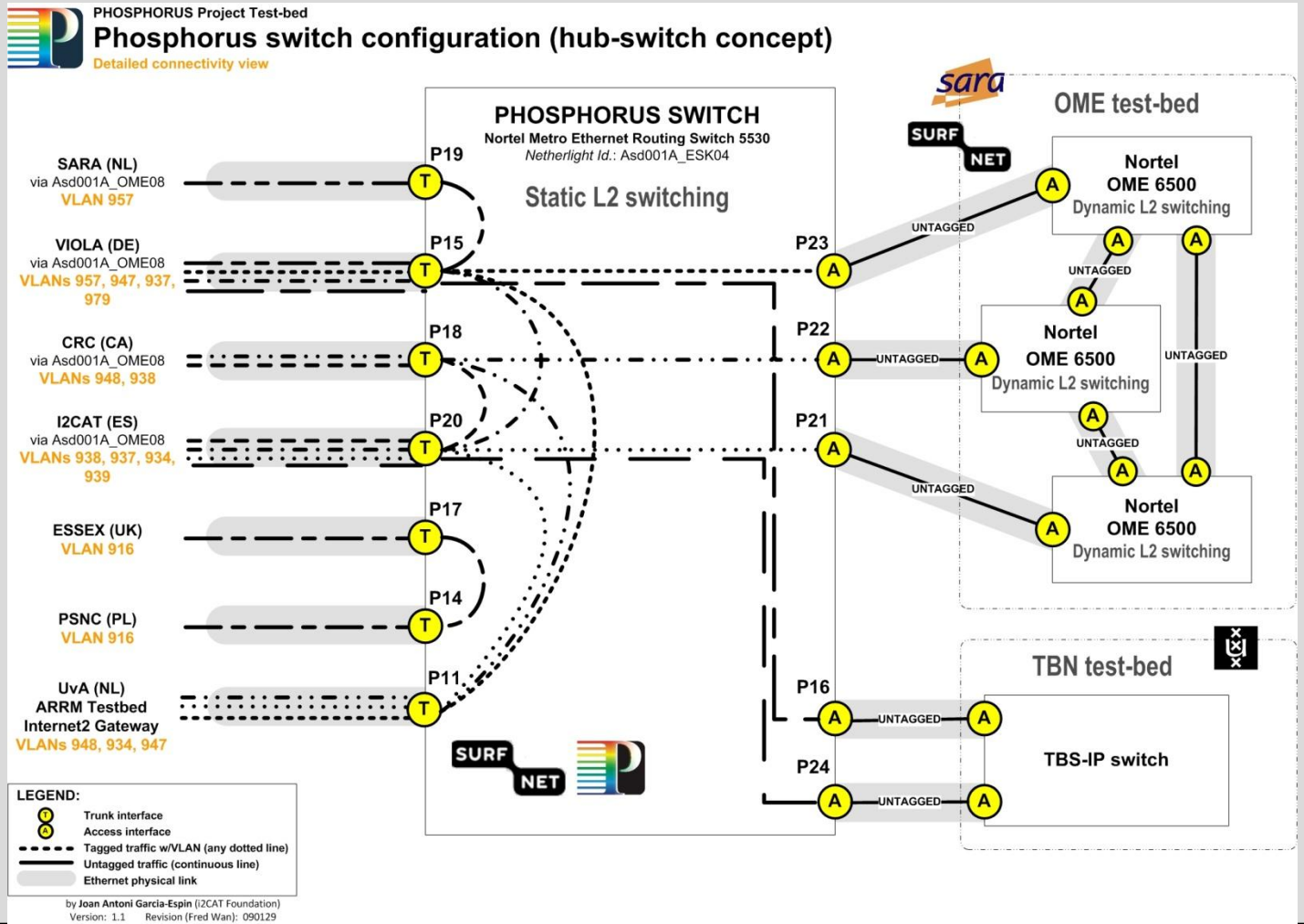
Challenges (I)

- The network as a Grid resource, how?



Challenges (II)

- Dynamic switching brought to the lowest layers.
- Glif community nowadays looking for a L1/L2 switching solution to be implemented in Dynamic GOLEs.

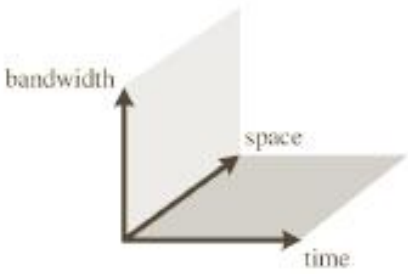


Challenges (III)

- Several derived from multi-domain scenarios (real life!)

integration

heterogeneity / topology exchange / security



malleable reservations

path computation

centralized

hierarchical

distributed

daisy chain

Harmony Achievements

- **Grid Middleware integration** One of the main challenges achieved by Harmony is the capability of “filling the gap” between the application middleware and the network control plane when dealing with several administrative domains.
- **Multi-domain path computing** The Network Service Plane (NSP) performs on-demand path computation involving resources located at several independent, heterogeneous domains.
- **Malleable multi-domain network resource allocation** Users or Grid Middleware can book two type of advance reservations: fixed or malleable. This last case provides a lot of flexibility to find a slot to serve reservations and maximizes the usage of the network resources.
- **Topology knowledge** The knowledge of the global topology is restricted, due to reasons of confidentiality, to a set of basic information based on three main elements: the endpoint (user or border), the inter-domain link, and the domain itself.
- **AAI integration** Harmony drives AAI management as another hot topic for the prototype: once the user is authenticated and authorised, he or she can use any of the services offered by the Network Service Plane.

NRPS involved in Harmony

- **ARGON (Allocation and Reservation in Grid-enabled Optic Networks)** was developed to manage resources of advanced network equipment as it is present in the German VIOLA test-bed. The advance reservation service of ARGON is able to operate both GMPLS and MPLS CPs. It guarantees a certain QoS for applications for the requested time interval.

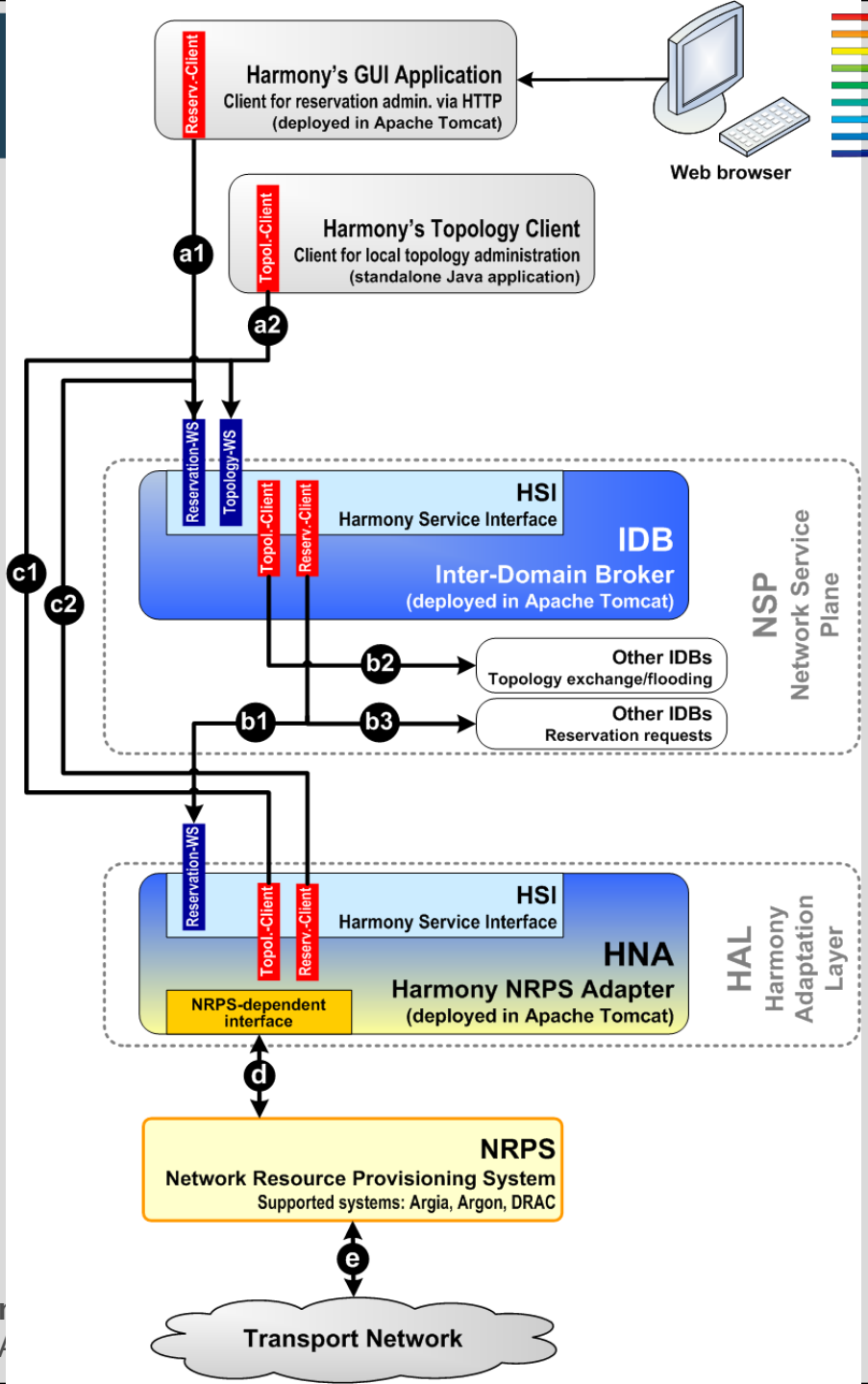
- **Nortel's DRAC (Dynamic Resource Allocation Controller)** was the world's first commercial-grade network abstraction and mediation middleware platform, acting as an agent for network clients (users, applications, compute resource managers) to negotiate and reserve appropriate network resources on their behalf. DRAC uses client's QoS requirements and pre-defined policies to negotiate end-to-end connectivity across heterogeneous in support of just-in-time or scheduled computing workflows.

- **Argia/UCLP (User Controlled Light Paths version 2)** provides a network virtualization framework upon which communities of users can build their own services or applications. Articulated Private Networks (APNs) are presented as the first services. APNs can be considered as a next generation Virtual Private Network where a user can create a complex, multi-domain network topology by binding together network resources, time slices, switching nodes and virtual or real routing services.

Harmony Architecture (I)

Key points:

- Simple WS architecture implementing WSRF1.2 based on Apache Muse 2.2.0 for the **Network Service Plane** (NSP)
- IDB entities can adopt several roles: **parent**, **child** or **peer**. HNA only are childs.
- Therefore, several operation modes of the NSP are available: **centralised**, **hierarchical**, **distributed**, daisy chain...
- **Harmony Service Interface** is aligne with current ongoing discussions in the NSI-wg
- The distinct **IDBs** flood the information of each domain they control



Harmony Architecture (II)

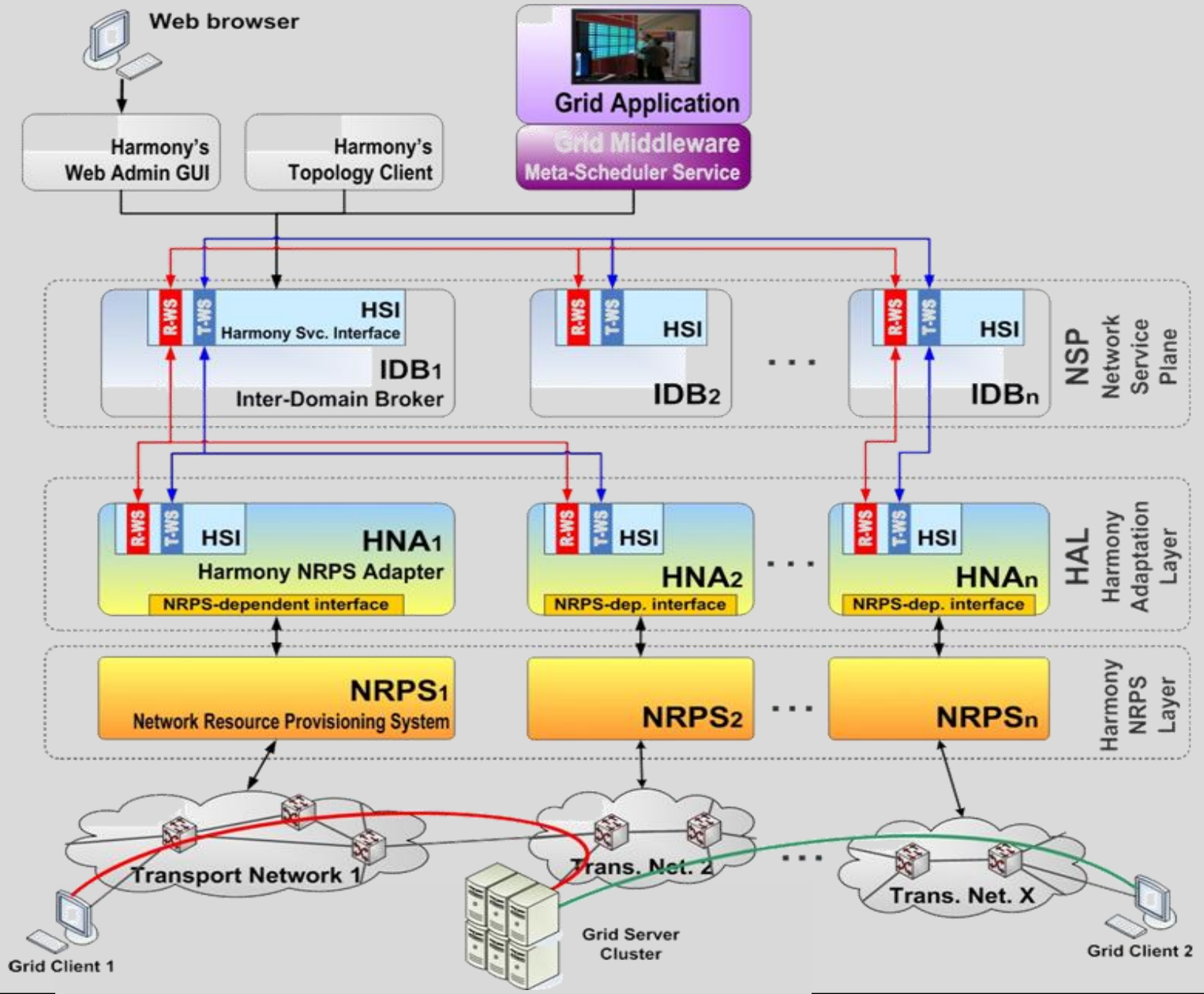
HSI
Harmony Service Interface

IDB
Inter-domain Broker

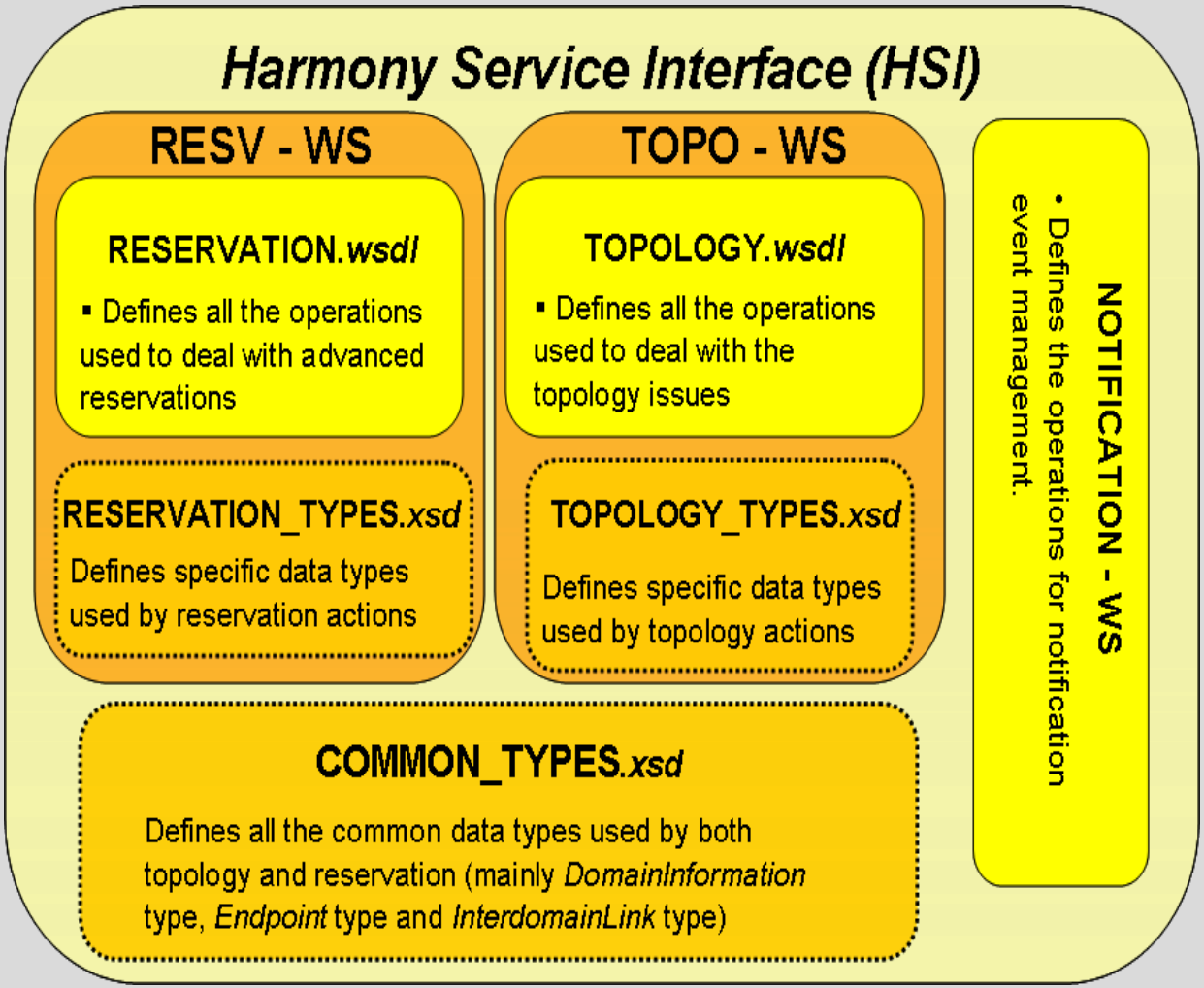
HNA
Harmony NRPS adapter

NSP
Network Service Plane

NRPS
Network Resource Provisioning System



Harmony Services: HSI

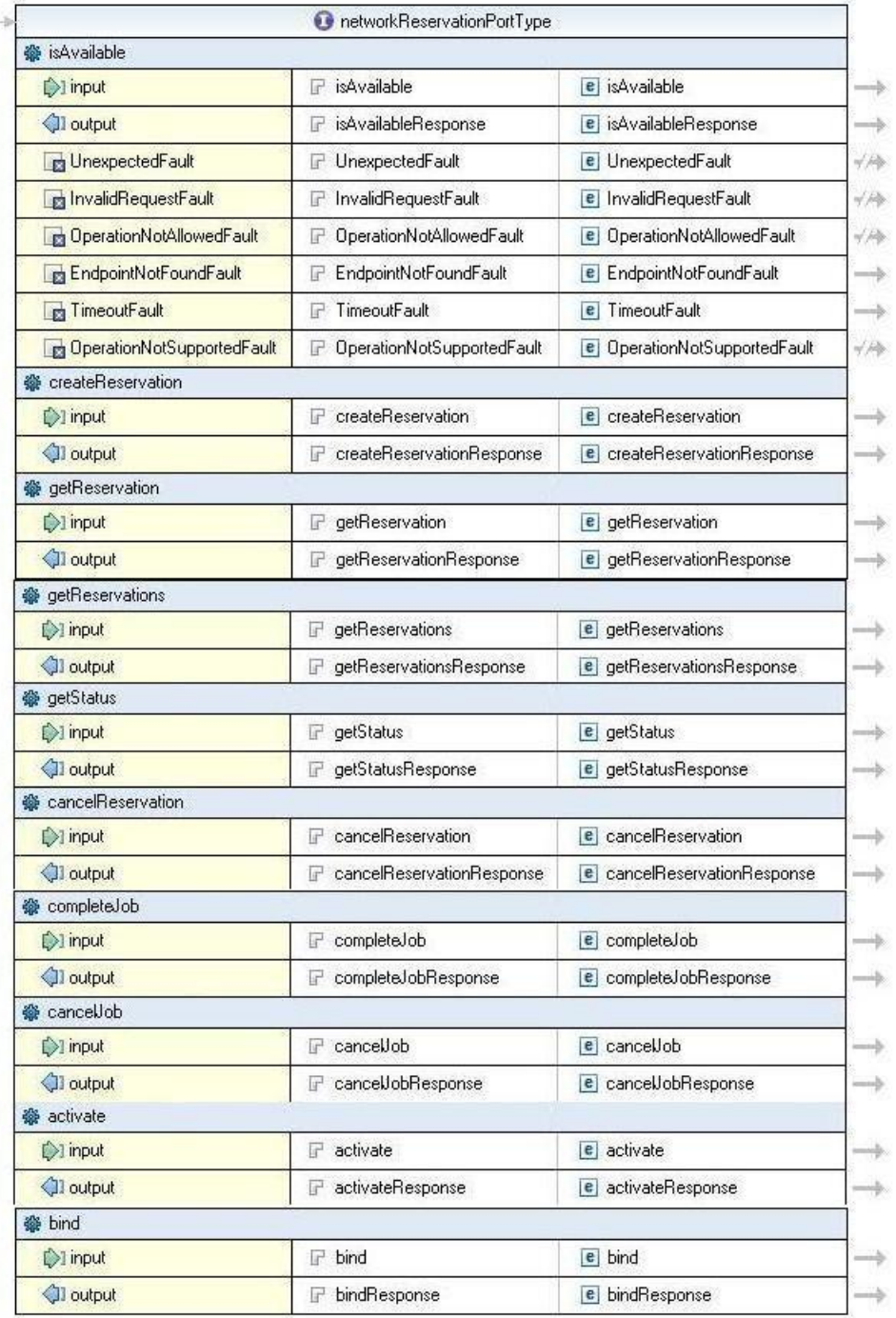


- **HSI is composed mainly by three components (Web Services):**
- ***Reservation-WS***: This WS enables the middleware to create, cancel and query advance reservations.
- ***Topology-WS***: This module enables Harmony to store, retrieve, modify and delete the resource-related information according to the topology of the controlled network domains.
- ***Notification-WS***: This WS is in charge of the event notification management. The motivation for this WS is to eliminate the need for the system to be polled periodically.
- **HSI is WSRFv1.2 compliant**

HSI – Reservation Service

Key points:

- *isAvailable*
- *createReservation*: Service Info (ID, typeReservation, Connections..), jobID, Notification URL consumer
- *getReservation:* ReservationID, ServiceID
- *getReservations:* Start time, end time
- *getStatus:* ReservationID, array servicesID
- *cancelReservation:* reservationID
- *completeJob* *(currently not used)*
- *cancelJob* *(currently not used)*
- *activate*
- *bind*



HSI – Topology Service

Key points:

- *addOrEditDomain*
- *add/delete/edit/get Domain(s):*

  Identifier, Reservation EPR, Relationship, Bw, Description
- *add/delete/edit/get Endpoint(s)*

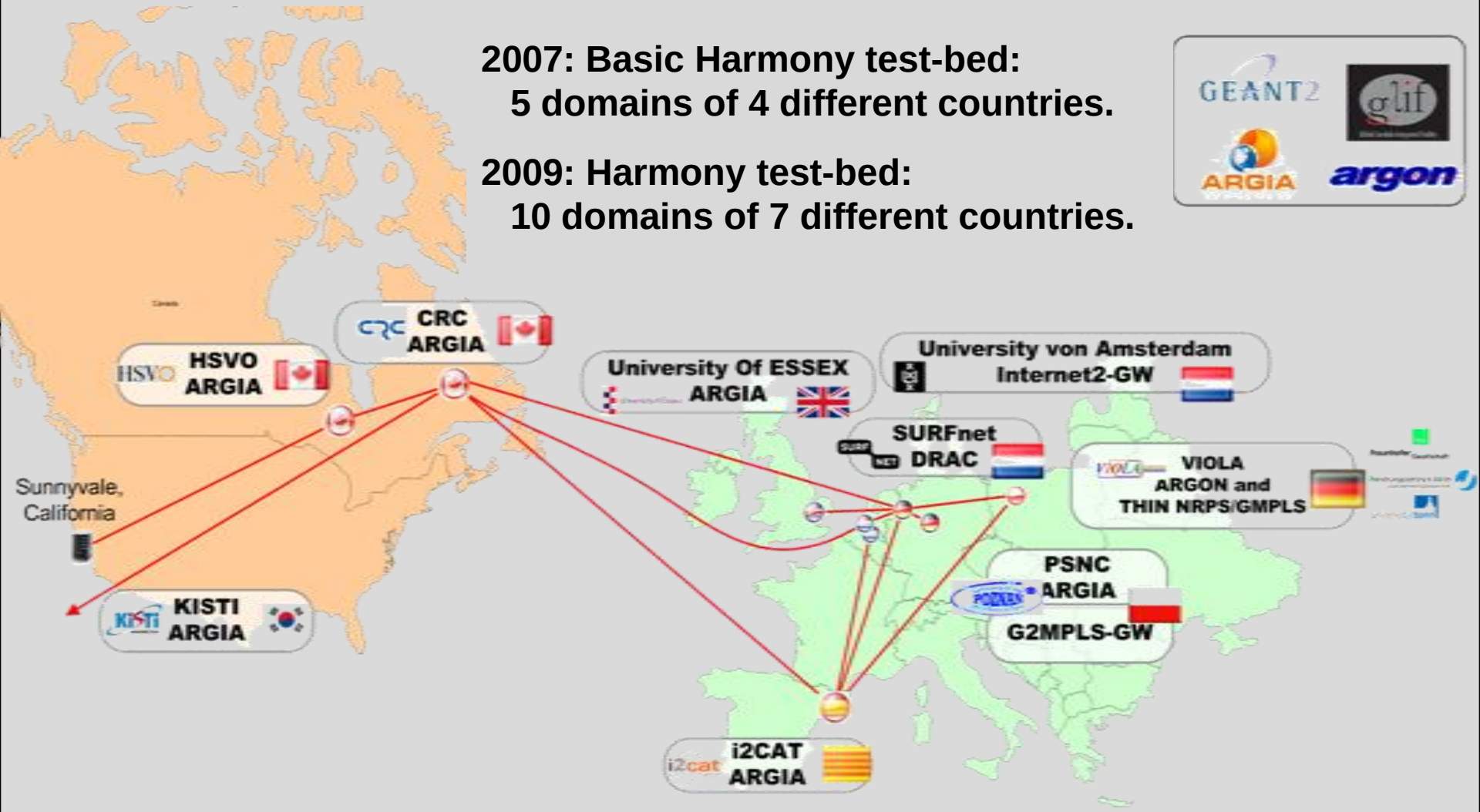
  Identifier, Name, Description, Interface, DomainId, Bw
- *add/delete/edit/get Link(s)*

  Indentifier, Source Endpoint, DomainId, Costs

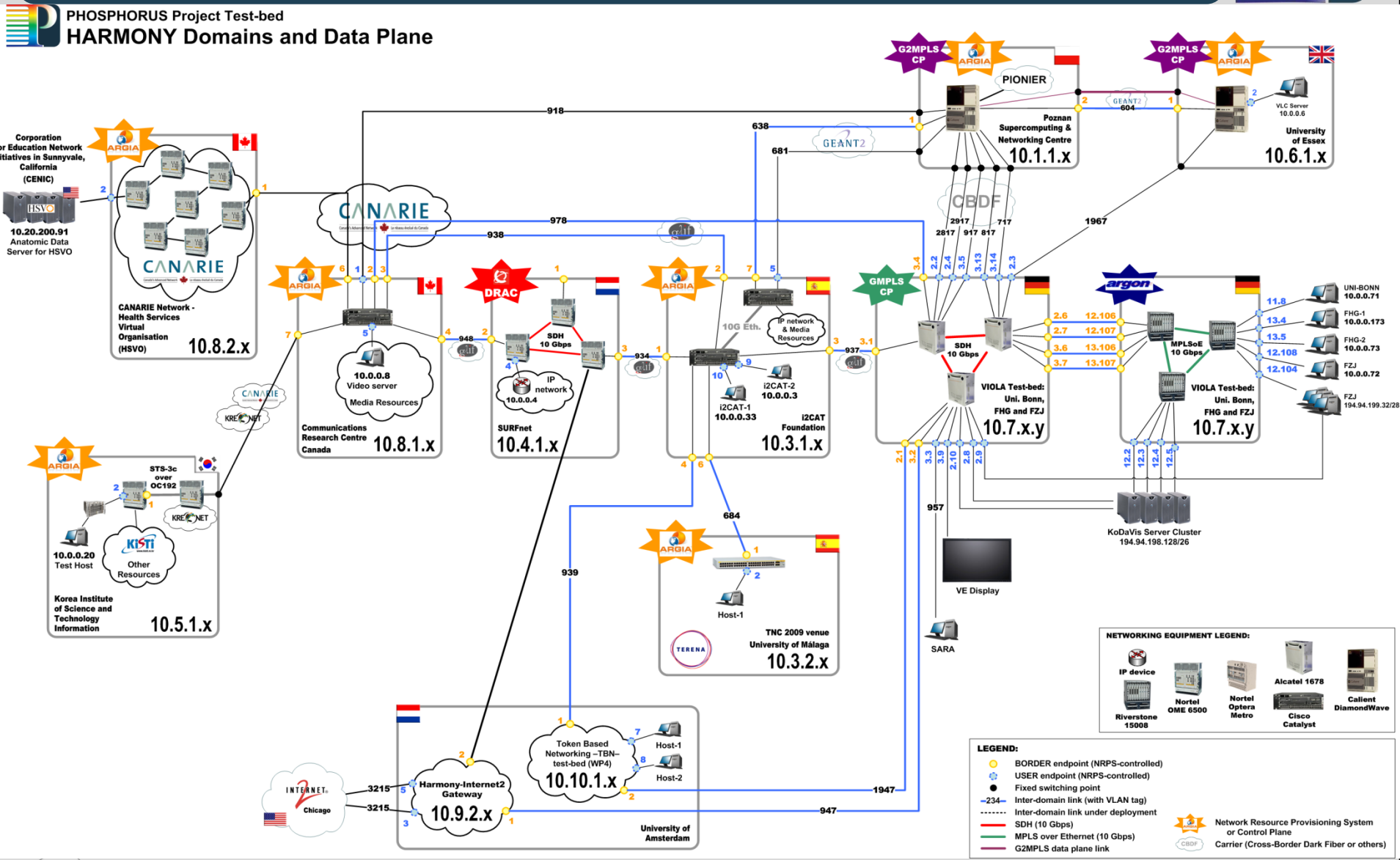
Current Harmony test-bed (I)

2007: Basic Harmony test-bed:
5 domains of 4 different countries.

2009: Harmony test-bed:
10 domains of 7 different countries.



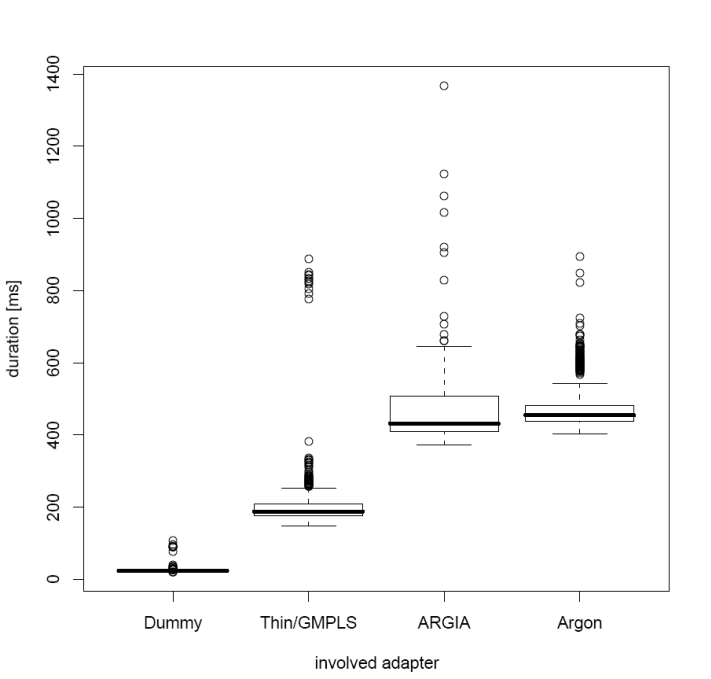
Current Harmony test-bed (II)



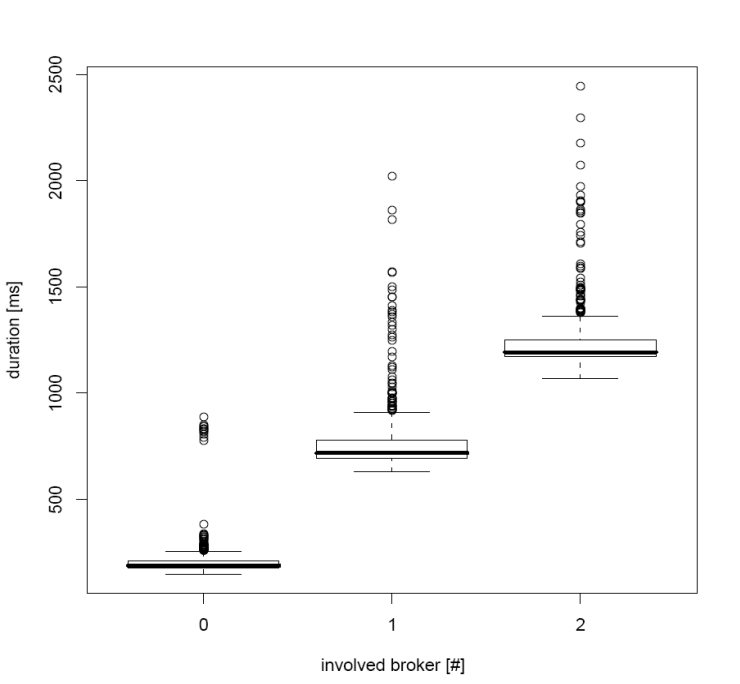
Harmony Performance and Scalability Analysis (I)

- System behaviour under different stress situations analysed.
- Focus: time response measurement of both the Network Service Plane and the entities which compose it. Request blocking ratio of the overall system.
- Several scenarios and different simulation environments considered.

• Time response of the adapters does not exceed of 500 ms (average).



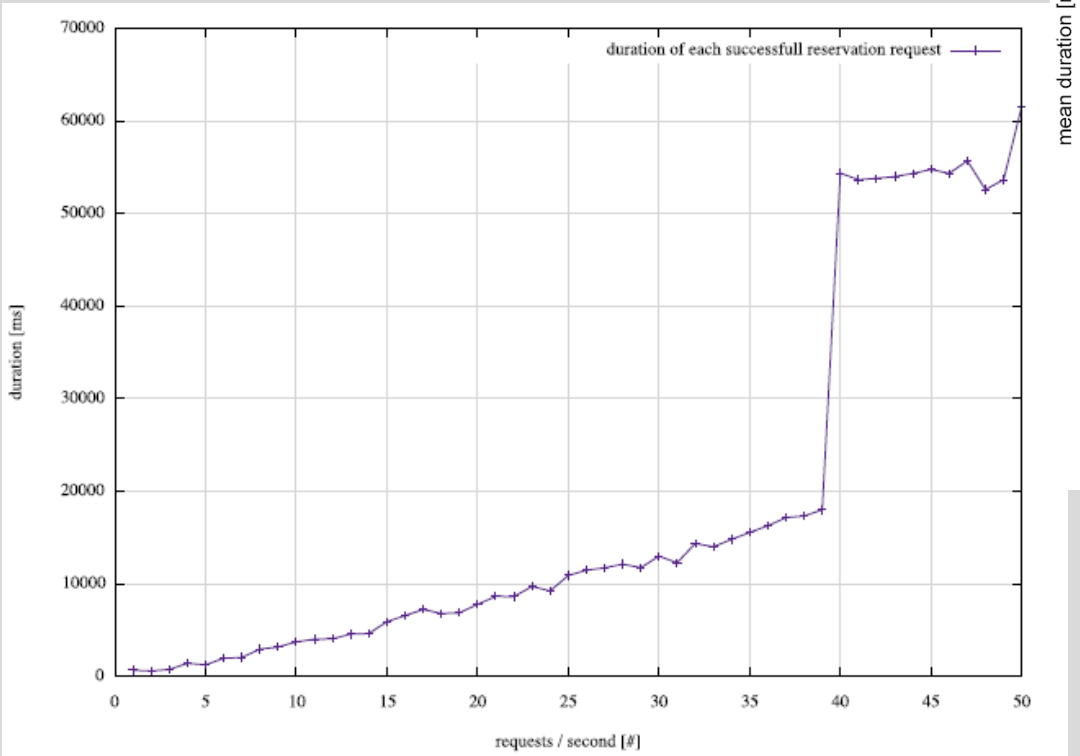
• Time response of the system increases linearly as the hierarchy level increases.

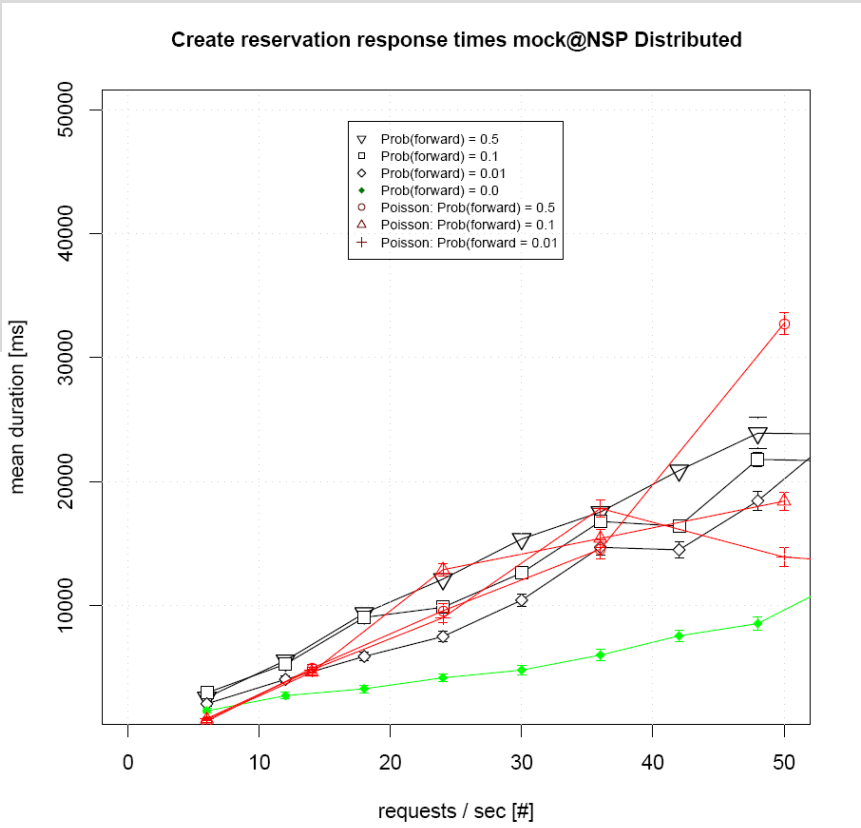


Harmony Performance and Scalability Analysis (II)

Below.

Duration of setup phase per each successful reservation requested under different request arrival rates

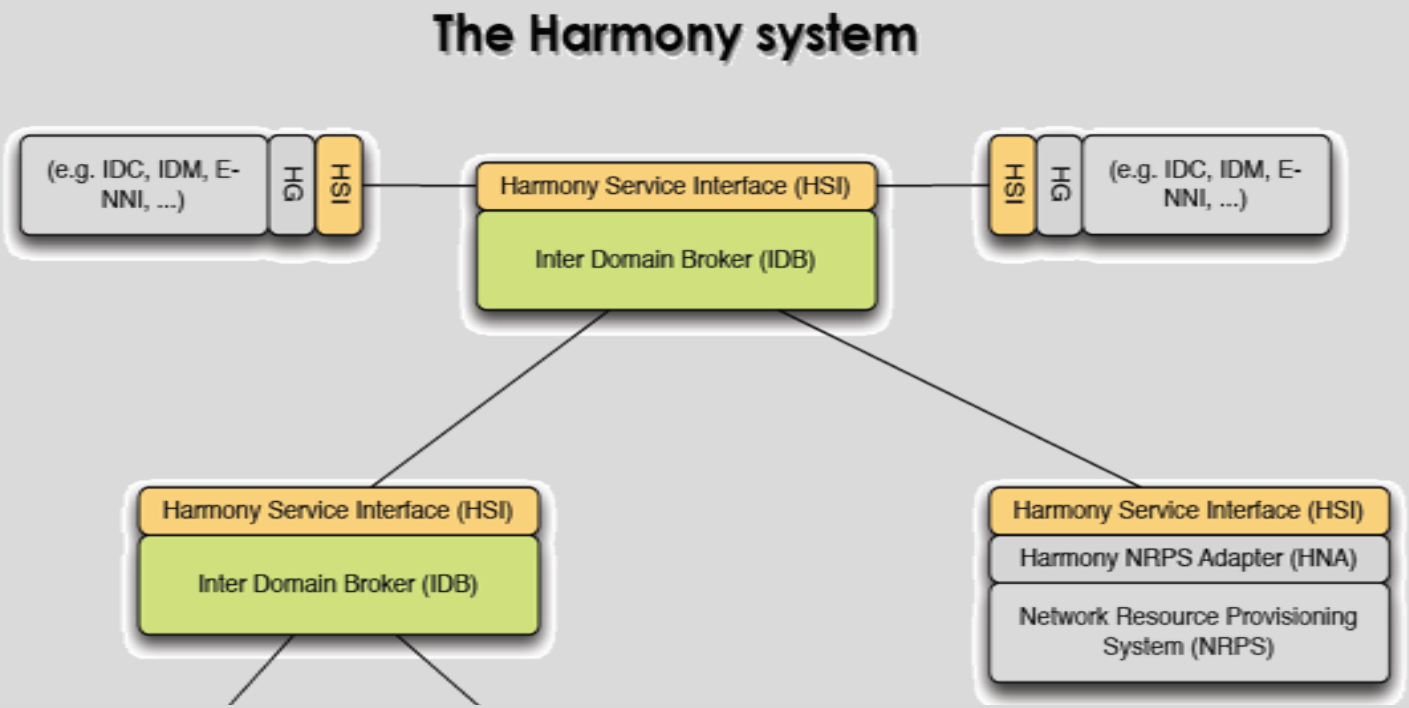




Above.

Create reservation response time under different request inter-arrival distribution (deterministic and Poisson)

Harmony and the outer world



Current implement.:

- **Harmony-IDC gateway** (demonstrated on SC08, booth #2603)
- **Harmony-G²MPLS** (signaling, working on Topology)
- **Harmony-AutoBHAN** (designed, using IDC)

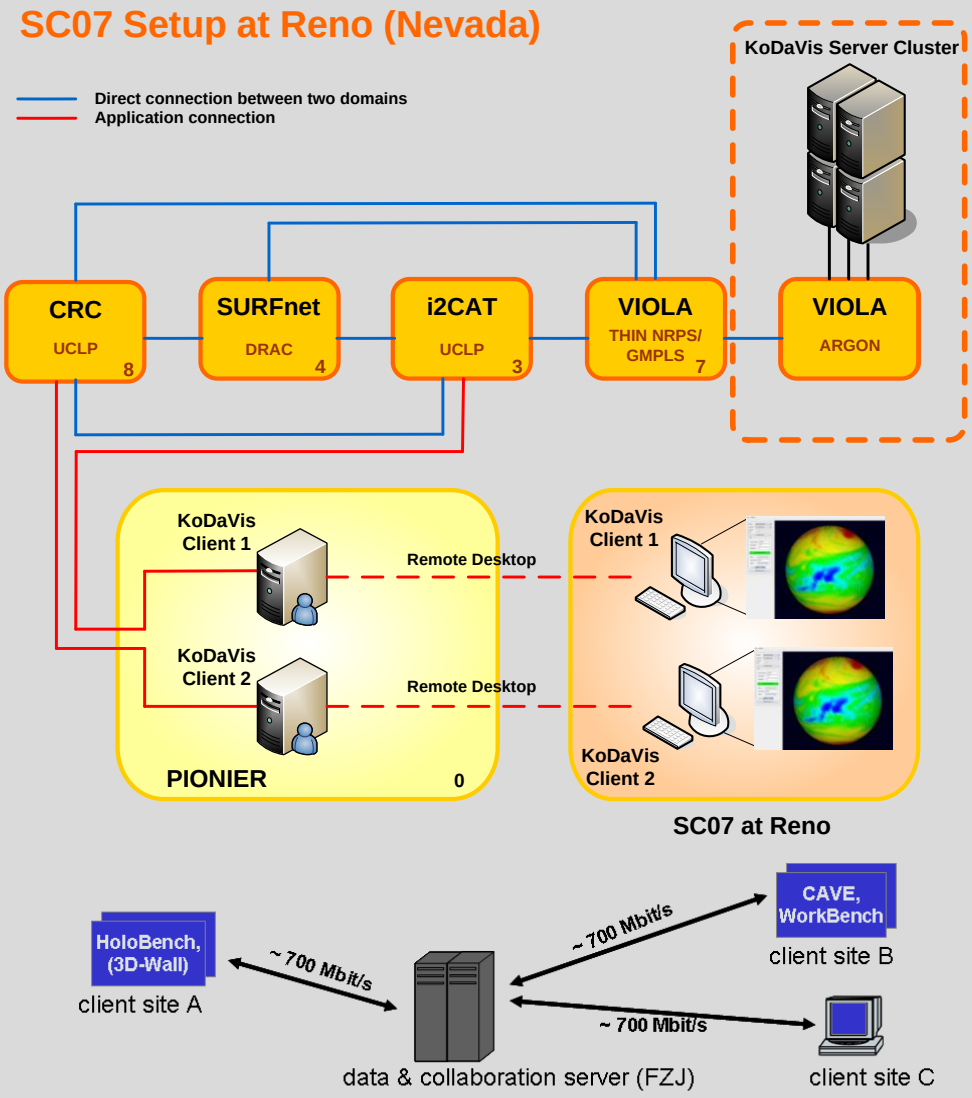
Key points:

- For any integration it is necessary to build a **Harmony Gateway (HG).** Both **HSI** and the **interface of the other system** are implemented, with the correspondent adaptation logics.
- Modularity in the internals of the HG allows stateful or stateless interoperability, depending on HG implementation and other system's requirements.

Unicore 6 Middleware and KoDaVis Grid application

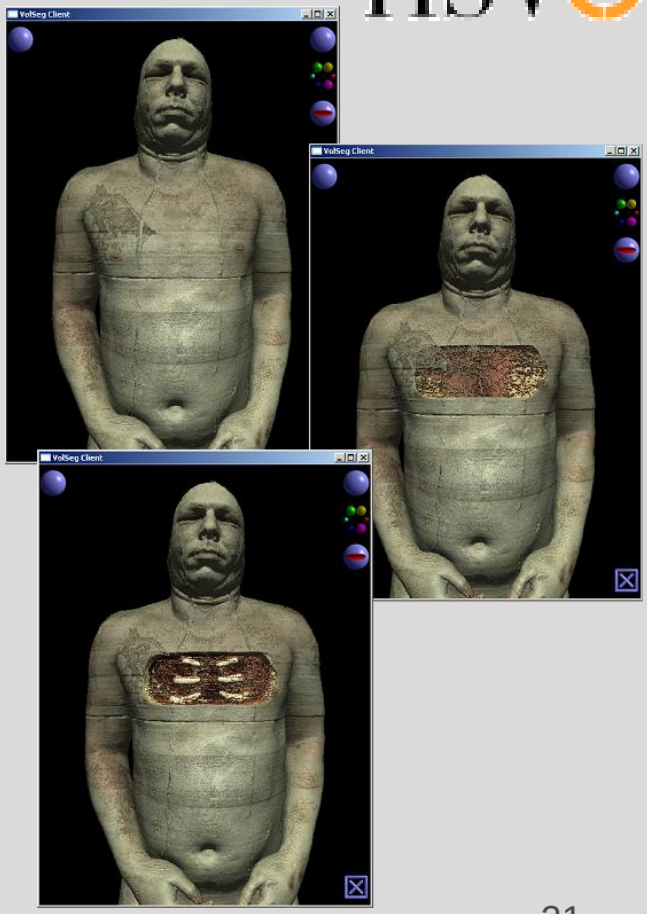
KoDaVis
Making Atmospheric Processes visible

- Collaborative visualization of data provides insight into processes
- Simulations of physical and chemical processes in the atmosphere help to understand the effect of human activities on the climate
- Data sets of ~ **1000 GigaByte** stored on 1 or more central servers
- Each client typically receives a throughput of 700 to 800 Mbps and at least two of them access serve at the same time

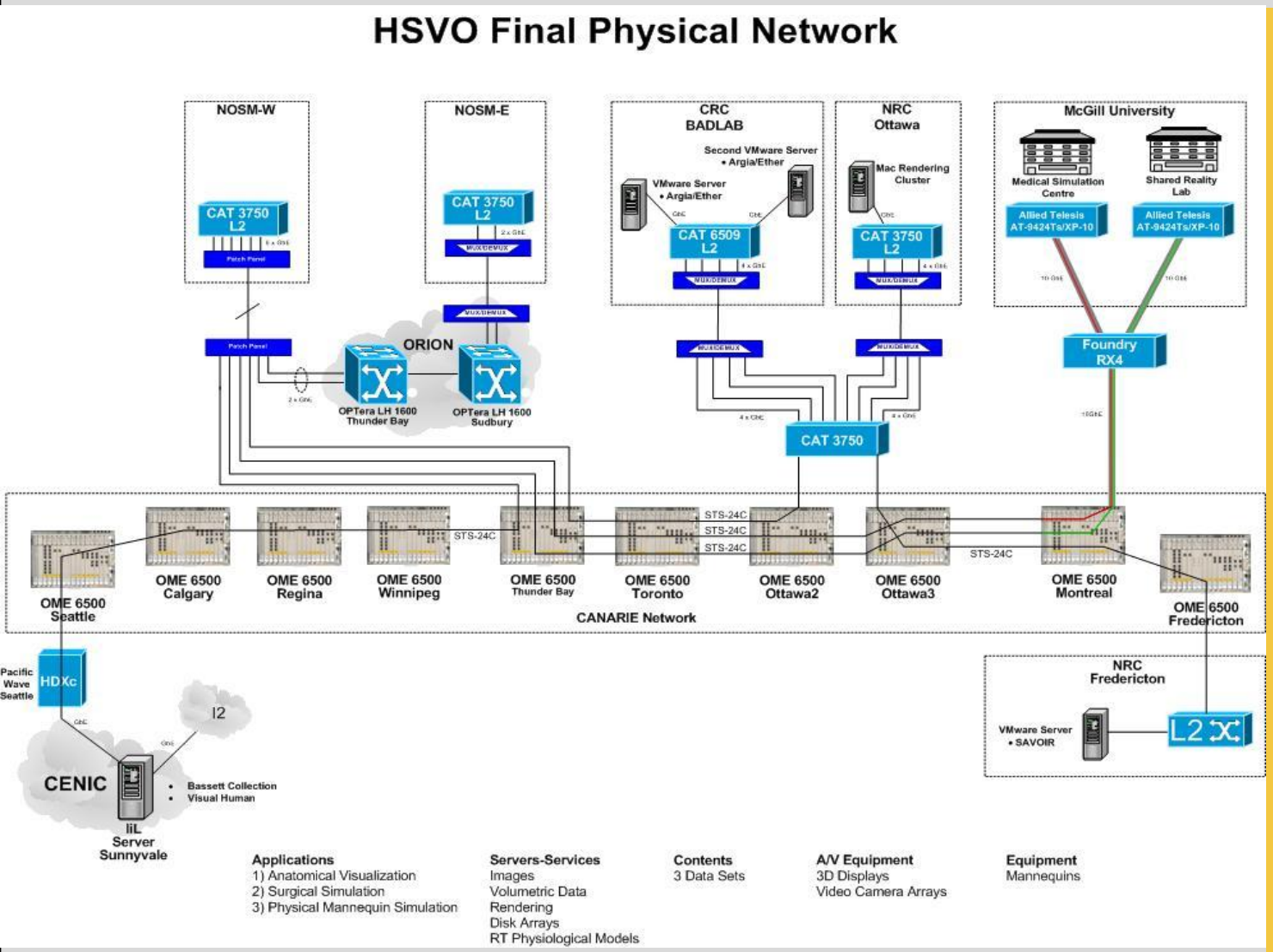


Harmony and HSVO (I)

- Health Services Virtual Organisation (HSVO) is an application of the SAVOIR framework developed in a Canadian initiative from CRC and Canarie.
- HSVO is now an operative domain in Harmony using Argia NRPS, crossing 7+ OXCs in the Canarie network from Ottawato Seattle.
- HSVO Use cases:
  - Cadaver Dissection
  - Virtual Patient Simulation
  - Anatomical Visualization

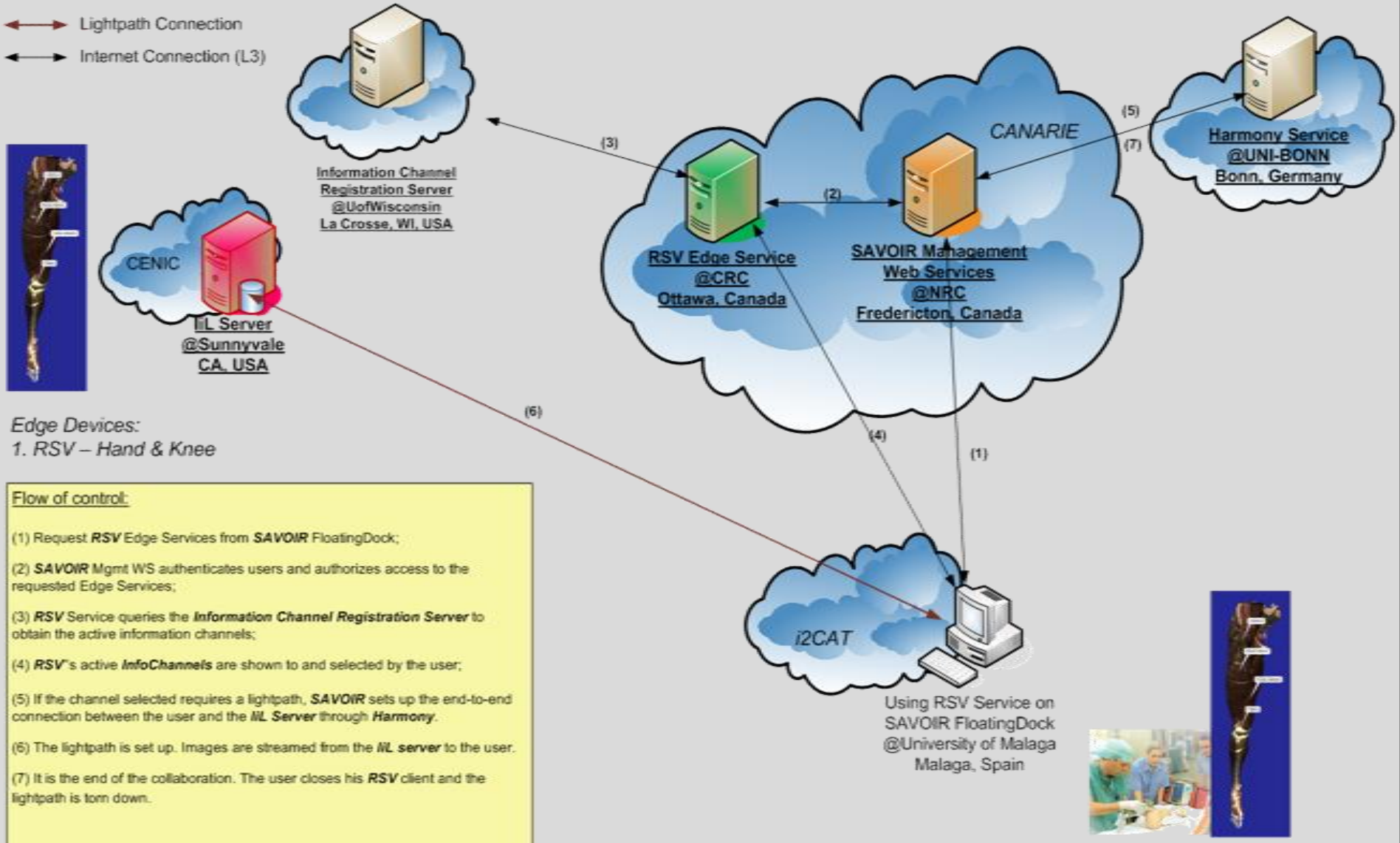


Harmony and HSVO (II): the Network



- HSVO Network is composed of 7 Nortel OME 6500 crossing Canada and part of US
- HSVO Network is controlled by Argia, which is responsible of creating connections when Harmony requests any path involving this domain.
- RSV Viewer is the module of the HSVO/Savoir application responsible of the communication with Harmony

Harmony and HSVO (III): the Workflow



Edge Devices:
1. RSV – Hand & Knee

Flow of control:

(1) Request ***RSV*** Edge Services from ***SAVOIR*** FloatingDock;

(2) ***SAVOIR*** Mgmt WS authenticates users and authorizes access to the requested Edge Services;

(3) ***RSV*** Service queries the ***Information Channel Registration Server*** to obtain the active information channels;

(4) ***RSV***'s active ***InfoChannels*** are shown to and selected by the user;

(5) If the channel selected requires a lightpath, ***SAVOIR*** sets up the end-to-end connection between the user and the ***IIL Server*** through ***Harmony***.

(6) The lightpath is set up. Images are streamed from the ***IIL server*** to the user.

(7) It is the end of the collaboration. The user closes his ***RSV*** client and the lightpath is torn down.

Dissemination during last year

- A. Willner, C. Barz, J. A. García-Espín, J. Ferrer, S. Figuerola, and P. Martini: ***Harmony: Advance Reservations in Heterogeneous Multi-domain Environments***. Proceedings of IFIP/TC6 09 Networking Conference. Accepted paper.
- S. Figuerola, J. A. García-Espín, J. Ferrer, and A. Willner: ***Performance Analysis of Harmony, an Optical, Multi-domain Network Resource Broker***. Invited Paper to ICTON 09 Conference.
- S. Figuerola, J. A. García-Espín, J. Ferrer, and A. Willner: ***Harmony, a Multi-domain, Multi-vendor Network Resource Broker.*** Submitted to IEEE Communications Magazine. Acceptance pending***.***
- S. Figuerola, J. A. García-Espín, J. Ferrer, and A. Willner: ***Scalability Analysis and Evaluation of the Multi-domain, Optical Network Service Plane in Harmony*****.** 35th European Conference on Optical Communication (ECOC 09). Accepted paper.
- Conferences and workshops participated/demo done during last year:
  - April 2008: OpenNet workshop
  - May 2008: TNC 2008 (May 08)
  - June 2008: ICTON 08, TAC 08, OGF 23 (NSI bof)
  - Sept. 2008: Broadnets 2008, TERENA & NREN Grid workshop, IBT 2008
  - Nov. 2008: SuperComputing 08, ICT 08
  - Dec. 2008: 1st TERENA e2e Workshop
  - Feb. 2008: Spanish Network of Excellence
  - March 2009: OGF 25 (NSI wg)

Special thanks to:

Jordi Ferrer
Sergi Figuerola
Michel Savoie
Alex Willner
and WP1 members

Question round

Thank you!

+ info:

joan.antoni.garcia@i2cat.net / +34 93 552 2518

or http://www.ist-phosphorus.eu

Advance Reservations (detailed)

reservation$_{mar}$ = ($t_{release}$, $t_{deadline}$, s, d, S, C), where

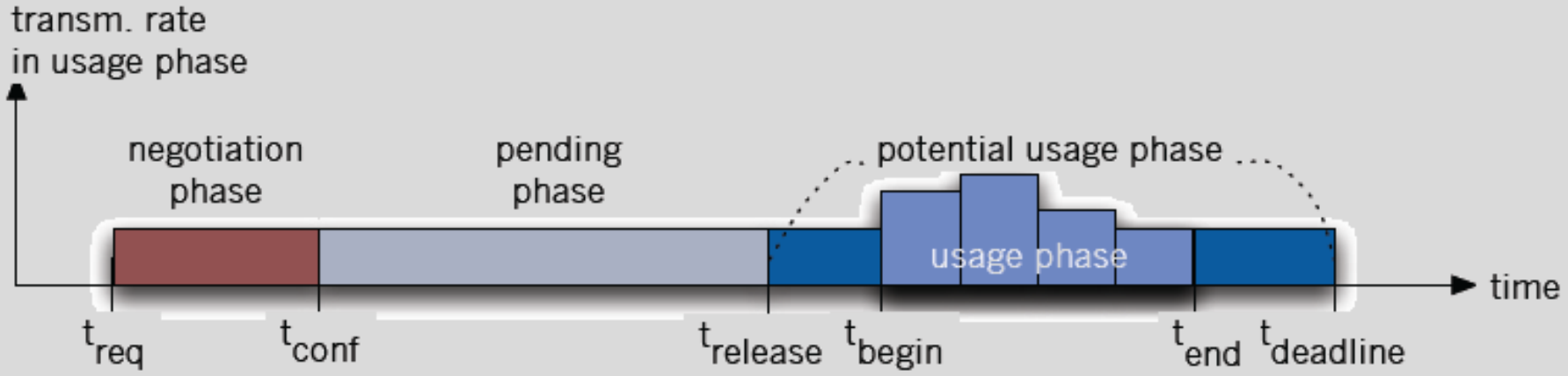
$t_{release} < t_{deadline}$

s = start endpoint

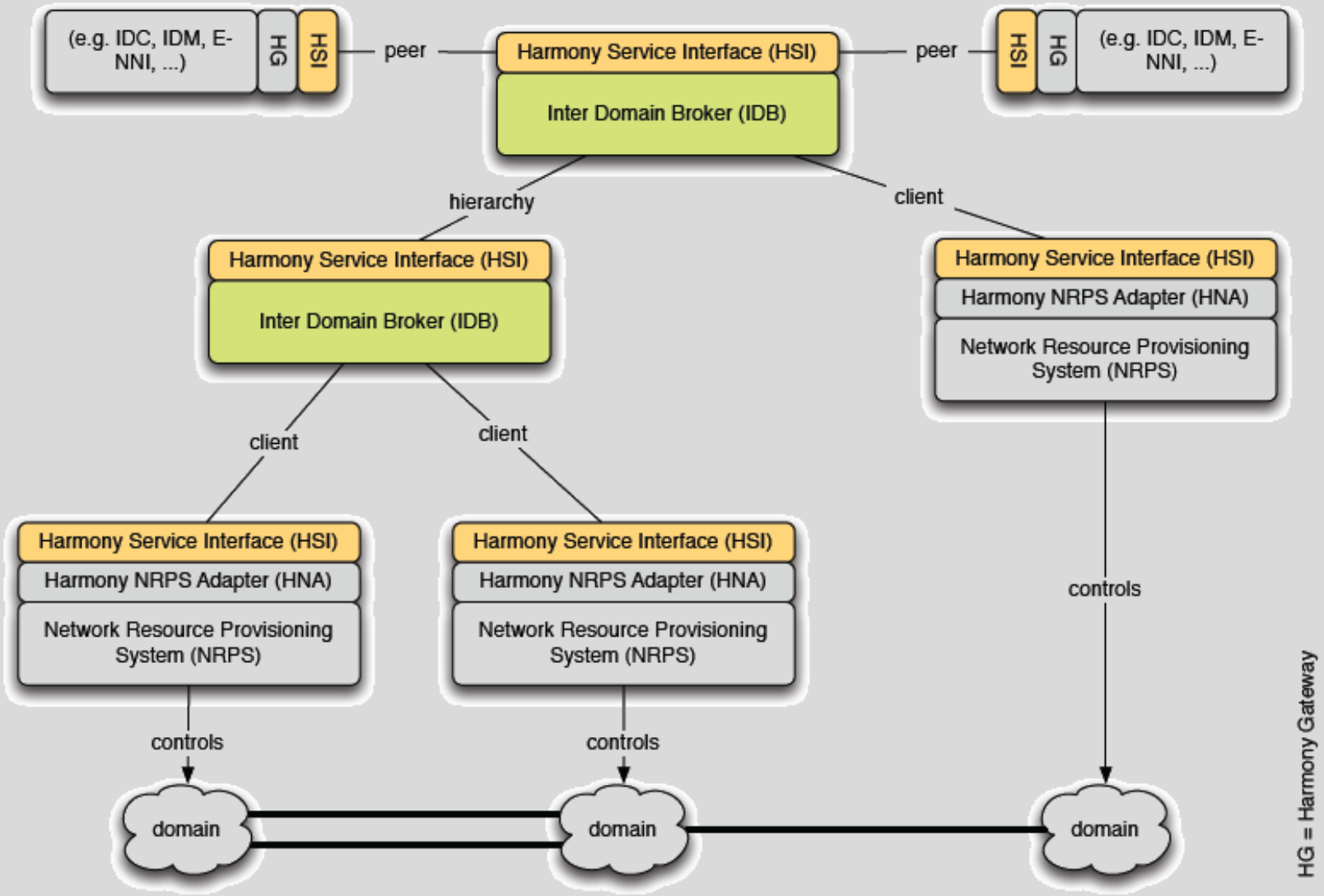
d = destination endpoint

S = data size (transmission rate and time product)

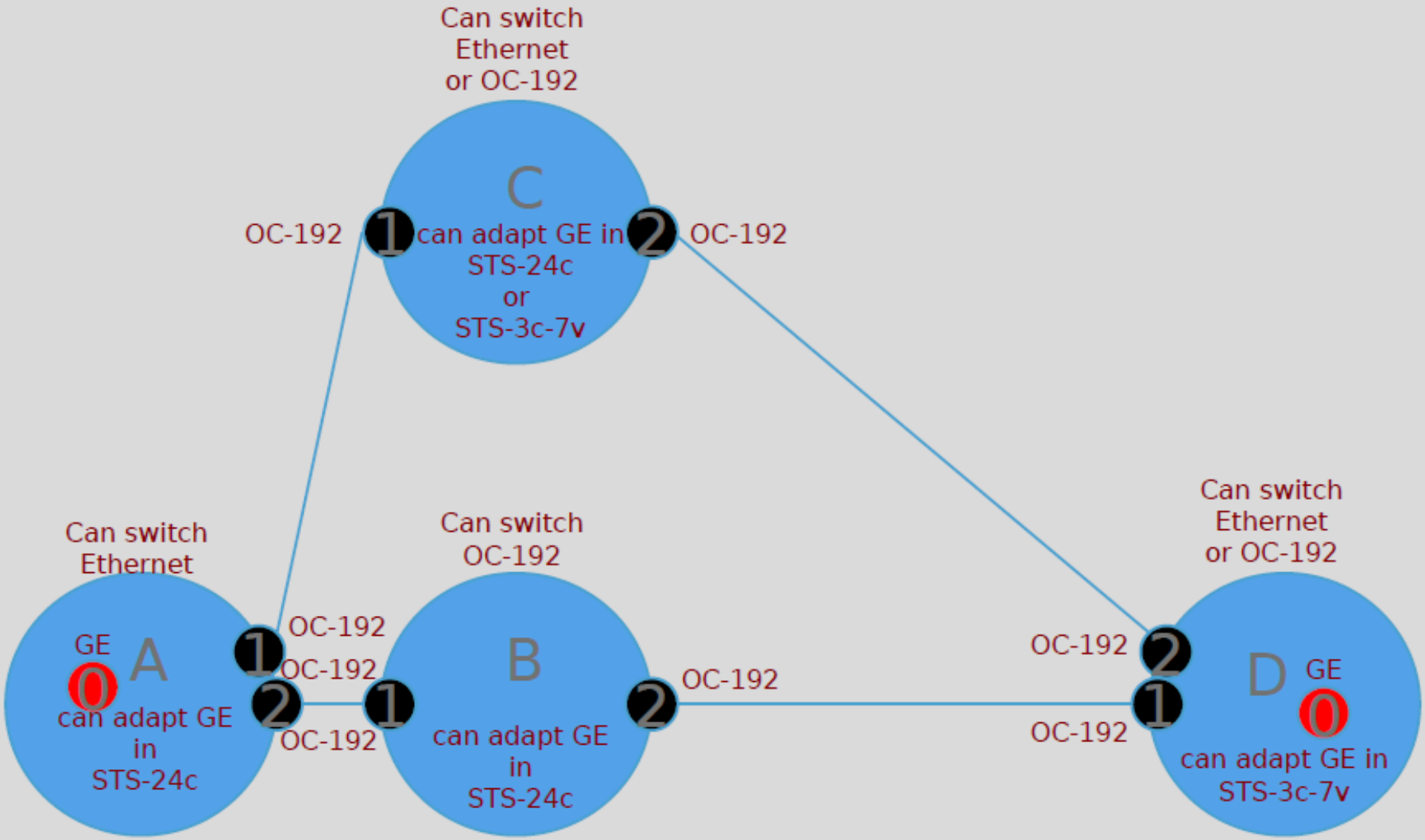
C = resource constraints (lower and upper bound for transmission rate)



Harmony Architecture (example)



Multi-technology Sample Scenario



Inter-domain Path Computer (details)