Connect. Communicate. Collaborate

Design and implementation issues of a multi-domain BoD-service for the NREN community

Afrodite Sevasti, GRNET

TNC'06, Catania, 17 May 2006

The JRA3 Activity of GN2

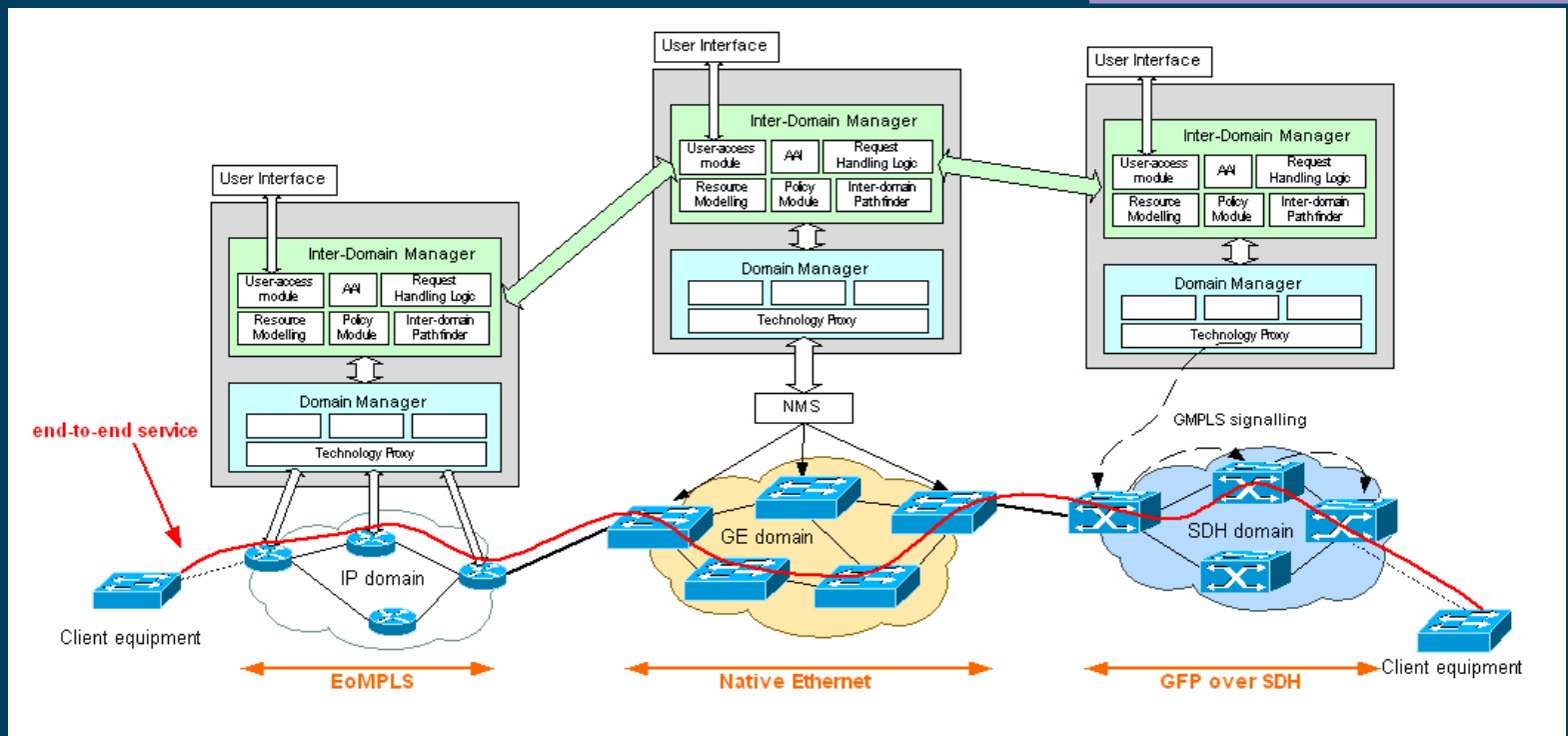
Connect. Communicate. Collaborate

- A ‘Joint Research Activity’ investigating the provision of ‘Bandwidth on Demand’ services to the NREN community
- The environment:
  - Multi-domain
  - Multiple technologies
    - GFP over SDH, L2 MPLS VPN, Native Ethernet
  - Requirements for:
    - end-to-end non-contended capacity
    - a standardized interface for service requests at end-points
    - service level indication to end-users
    - advance reservation (scheduled)

JRA3 approach

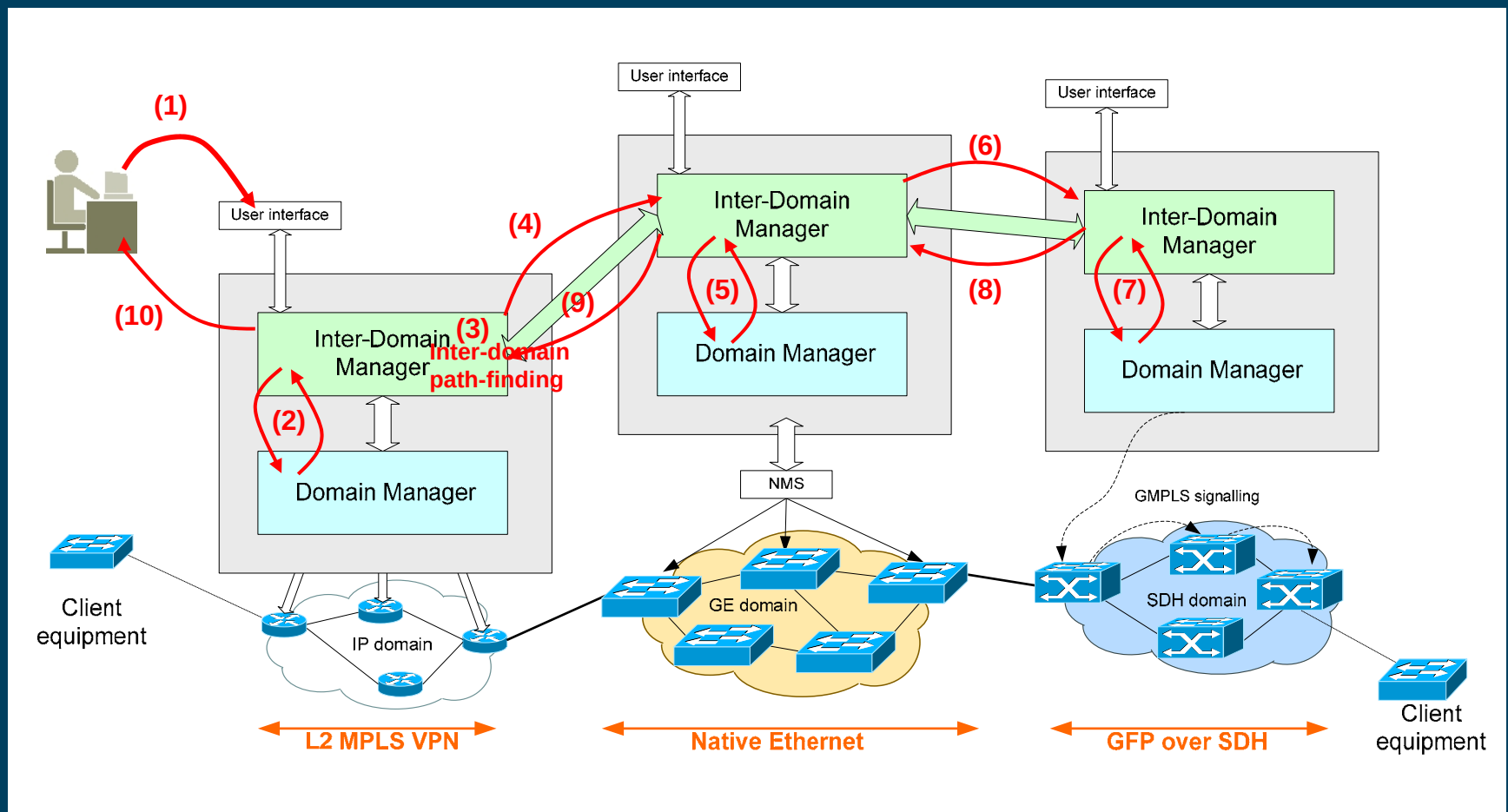
- The goal is to streamline the inter-domain setup of end-to-end paths
  - shorten the provisioning time
  - reduce the amount of human intervention
  - using existing (NREN/aggregation) networks by an overarching method
  - automate the process step-by-step; focus on inter-domain coordination process
- Service specification
  - End-to-end, connection oriented service for provisioning non-contended capacity
  - Layer 1, 2 technologies
  - AAI, policies
  - Single point of entry for users/applications
- PROTOTYPE: focus on provisioning of a deterministic non-contended bandwidth pipe between two 1Gigabit Ethernet access ports over multiple domains that employ different technologies

JRA3 architecture



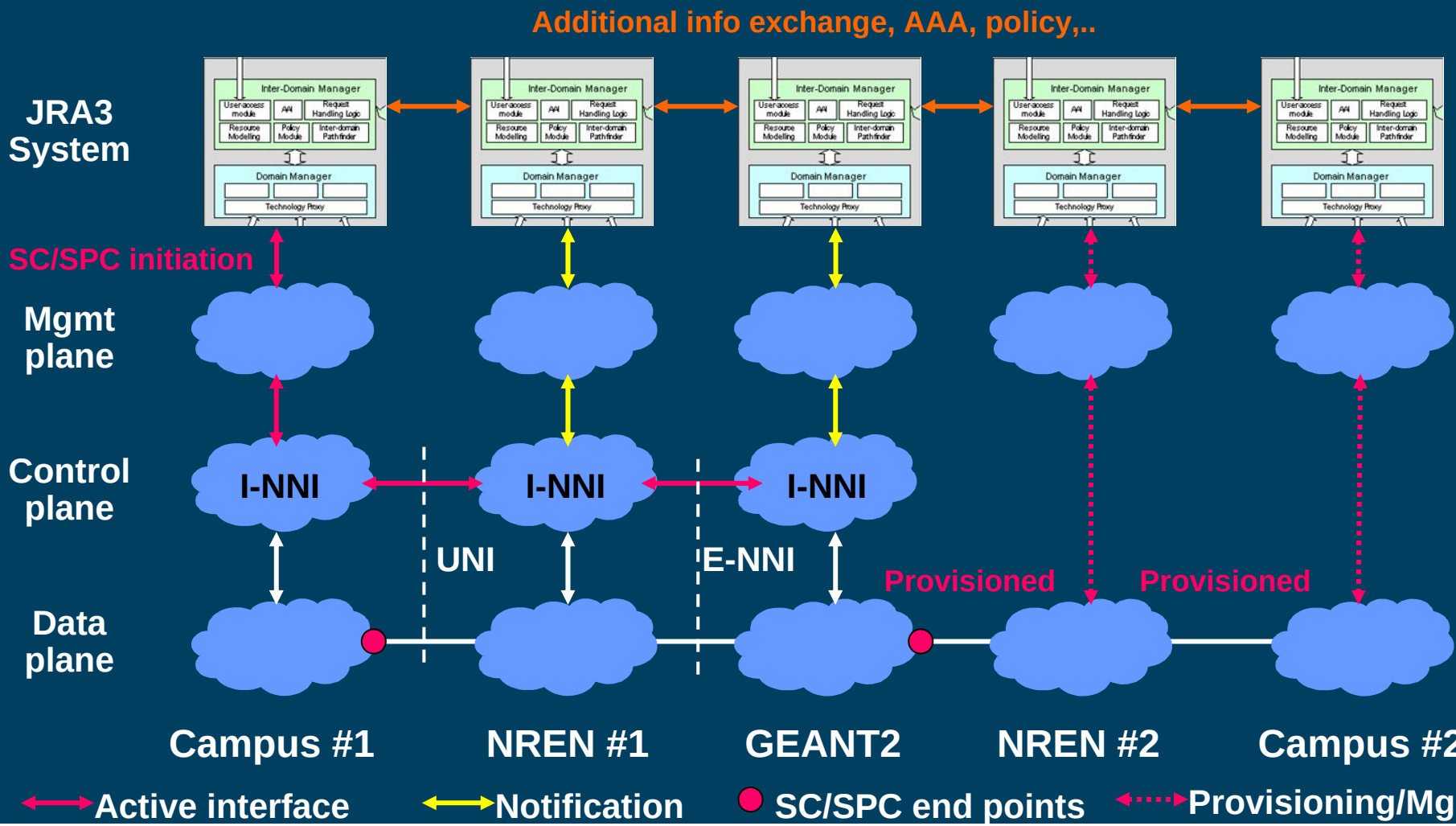
- Inter-Domain manager (IDM) - Domain manager (DM) - Standardized interfaces
- JRA3 will provide:
  - The IDM module
  - Reference implementation(s) for the DM (human NOC, Ethernet-based, etc.)
- Each domain participating in BoD service provisioning needs to operate an IDM and honor the IDM-DM and IDM-IDM interfaces

Distributed approach



Multi-domain provisioning*

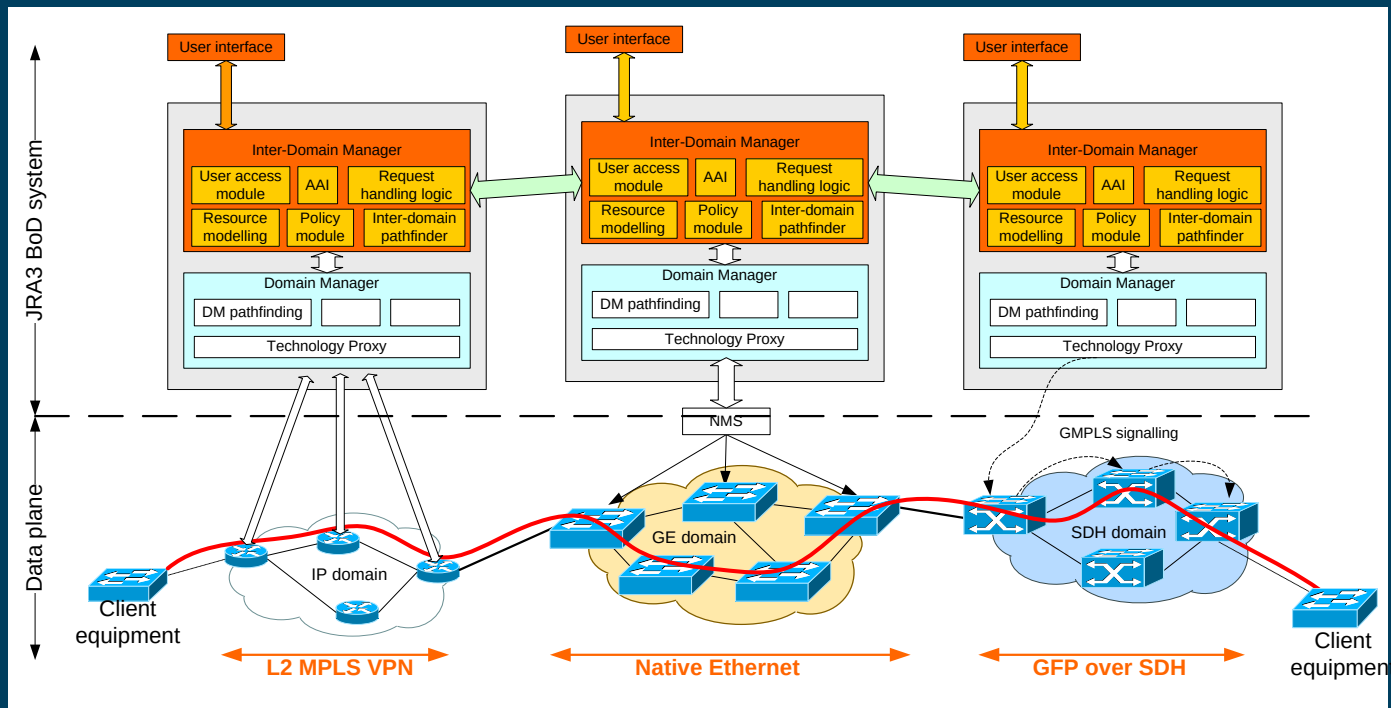
Connect. Communicate. Collaborate



*by Hans-Martin Foisel (T-Systems)

Inter Domain Manager

Connect. Communicate. Collaborate



- A set of automated procedures for the required non-technology specific inter-domain negotiations
- Can be considered as a ‘bandwidth broker’, but it is more than this

Why an Inter-Domain Manager

- The effort to provision end-to-end Bandwidth on Demand services in the European scenario requires specific developments in inter-domain collaboration
- Splitting intra-domain management functionalities from inter-domain ones in separate modules, allows multi-domain R&D to proceed autonomously and focus on this less standardized area
- At the same time, it allows to leverage existing inter-domain managers through wrappers and interfaces, exploiting a modular approach
- This effort can provide solid experience for brokering services other than Bandwidth on Demand

Domain independence

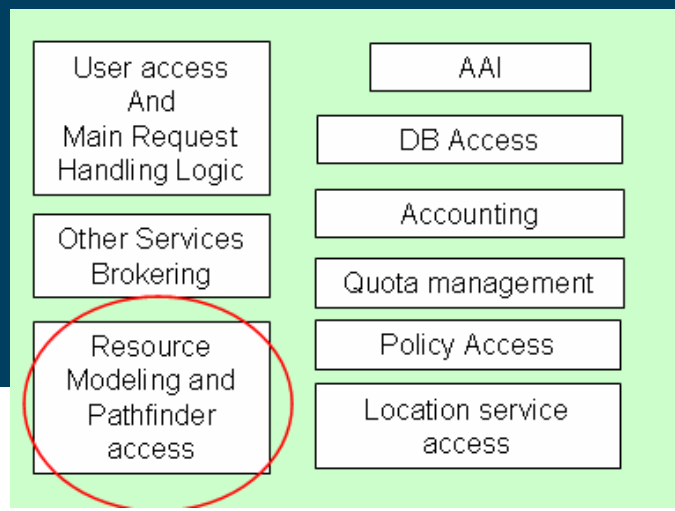
- Collaborative and distributed effort through newly defined interfaces which extend the NNI standards
- No centralised management
- Better resilience
- A common naming and addressing schema for a large amount of devices
- Possibility to hide domain internals
- Clear separation of control and data plane also at the physical level when needed

IDM multi-domain issues

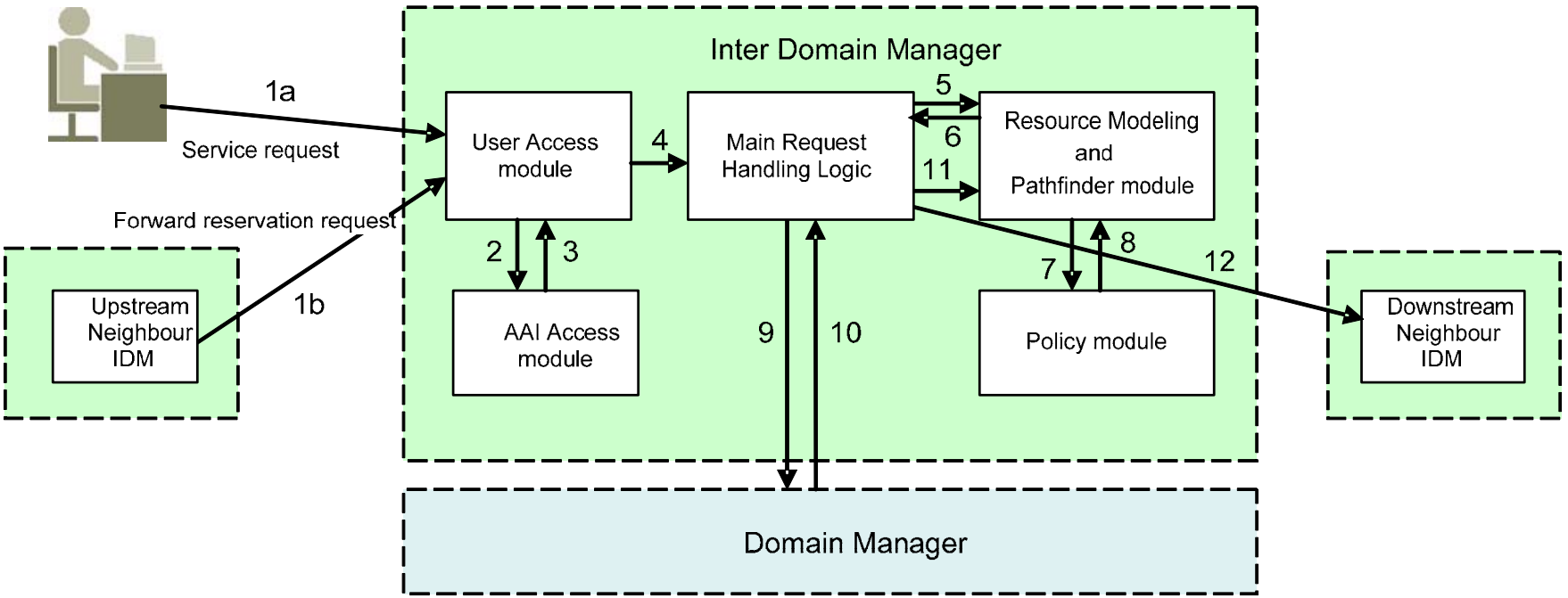
- The IDM faces a number of challenges related to its multi-domain scope:
  - domain independence for resource usage policies and technological choices
  - a service and network abstraction schema to describe implementation over very different networks
    - a schema which allows to clearly specify which type of service is requested
    - a network abstraction which allows inter-domain information exchange independently of the underlying technologies
  - advance reservation
  - multi-domain path finding procedure
  - monitoring
  - Authentication and Authorization

IDM: Overall functionality

- Ingress point to the BoD system
  - It receives and processes BoD reservation requests from users or from other IDMs of neighboring domains
- Selection of the chain of domains to be involved in each end-to-end path establishment for serving a reservation request
  - Inter-domain path-finding
  - Based on topology and traffic engineering information
- Pre-reservation and commit process between all IDMs along the end-to-end path
- Interaction with the AAI service, to authenticate the identity of BoD service requestors and authorize access to the BoD service
- Credit management sub-module for the controlled allocation of bandwidth resources
- Accounting and logging sub-module
- Policies for allocation of BoD resources and for management of the BoD service



IDM internals



IDM pathfinder (1)

- The main task of the Pathfinder sub-module in the IDM is to provide all inter-domain feasible paths to fulfill a BoD request
- A separate Routing Protocol sub-module within each IDM is used for distributing the inter-domain routing information among IDMs
  - Link-state protocol
  - Traffic Engineering extensions are used to carry information such as link capacity, resiliency, policy related information
  - Path selection is based on technology-agnostic parameters
  - Each inter-domain link is advertised as-is, for intra-domain links, each domain can adopt the level of abstraction considered appropriate*

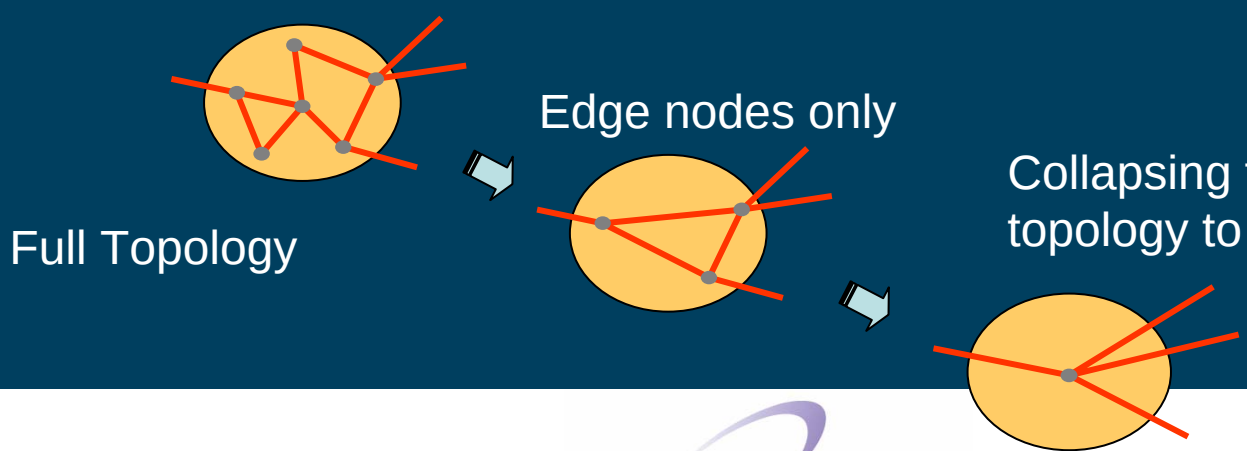
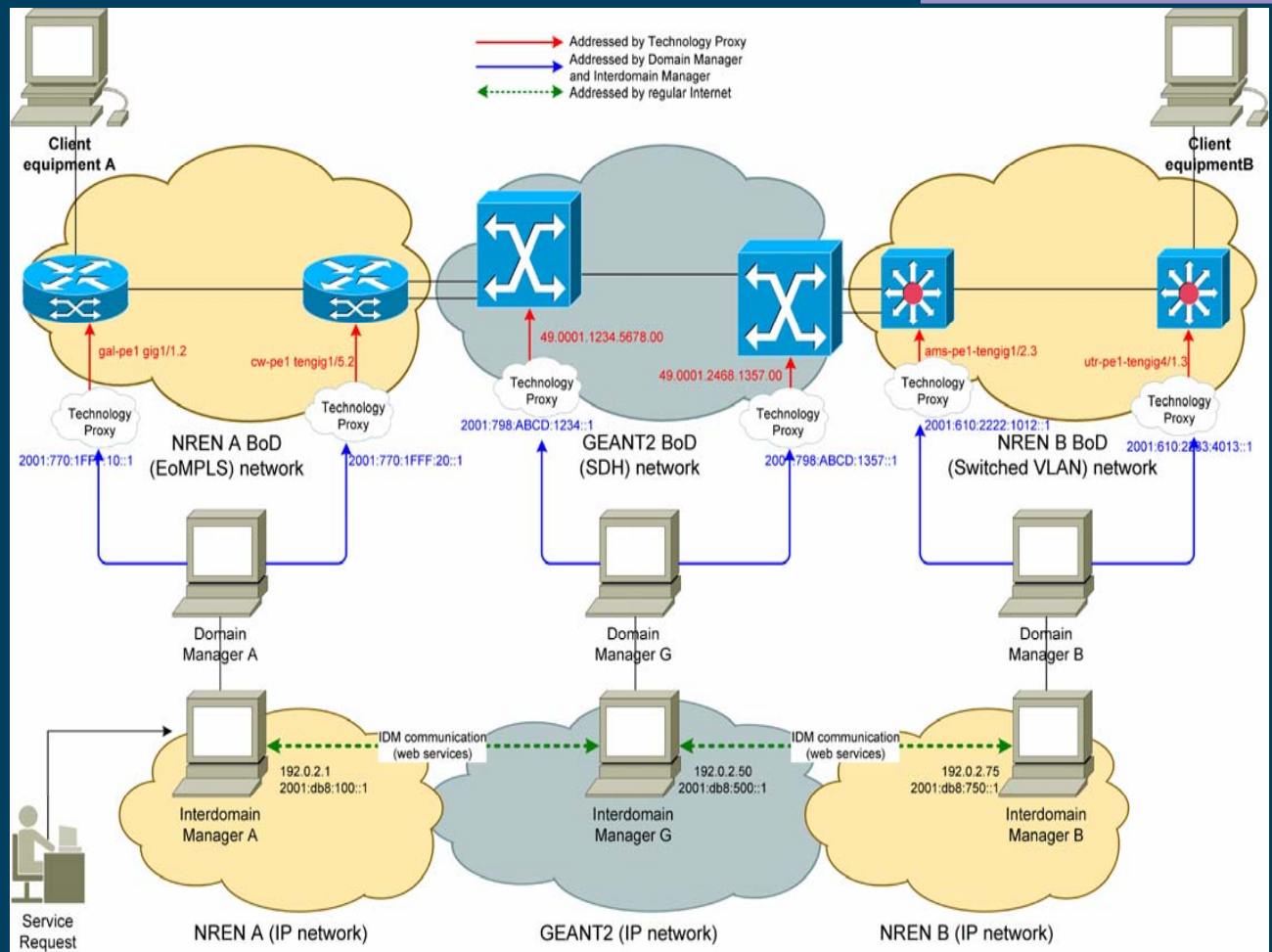


*figure adopted from DRAGON project material

IDM pathfinder (2)

- For the IDM prototype, the Quagga OSPFv2 routing daemon implementation with custom defined Opaque LSAs will be used as the Routing Protocol sub-module
  - As the Quagga OSPFv2 daemon is a SPF (shortest Path First) engine and not a constraints-based SPF engine, the Pathfinder module is required to perform additional CSPF computations
- Based on TE information for the advertised topology, the Pathfinder sub-module applies a constraint-based algorithm to create a list of paths to be handed back to the Reservation module
  - Each path in the list represents an inter-domain route over a set of interconnected domains, and includes the ingress and egress interface in each transit domain

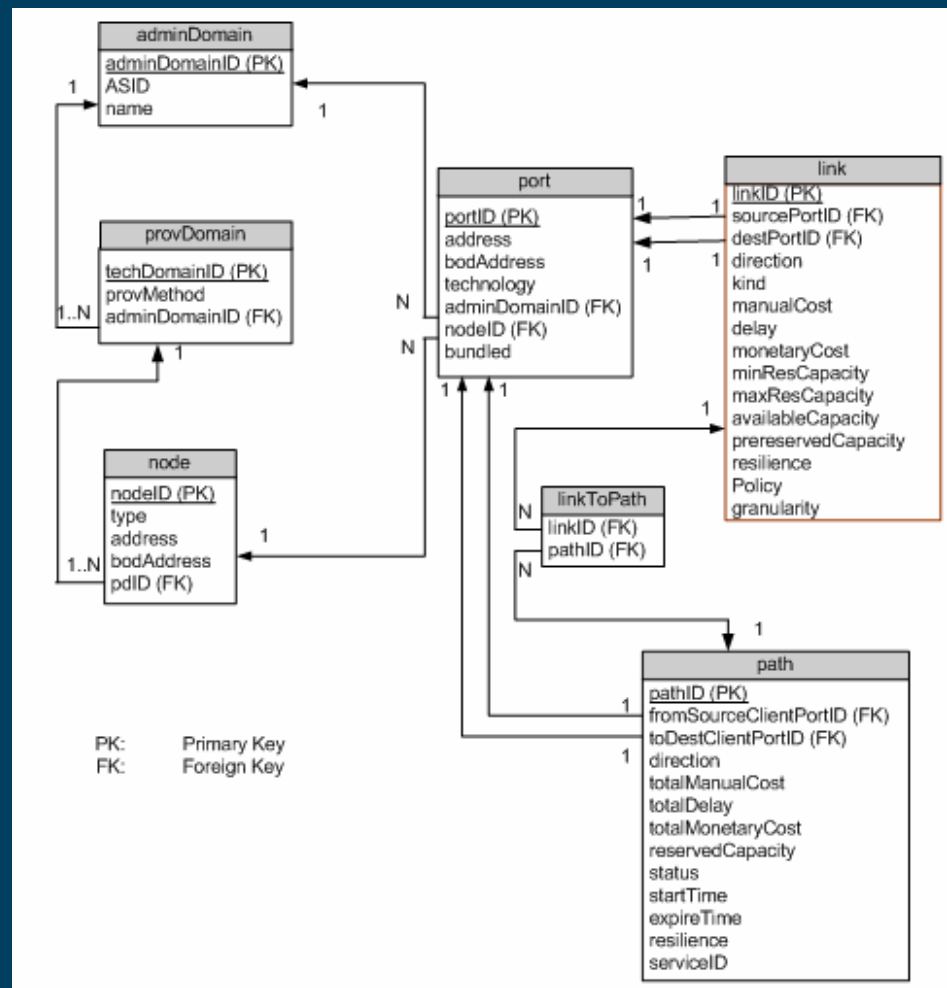
IDM addressing



Use of IPv6 addresses to identify data plane entities at the BoD system level

IDM abstract representation

- A network abstraction which allows each module and each domain to exchange information independently of the underlying technology
  - the IDM assumes the DM responsibility for keeping the abstract representation up-to-date
  - start with an Entity-Relationship schema, implement using XML
  - need to define both local and global entities



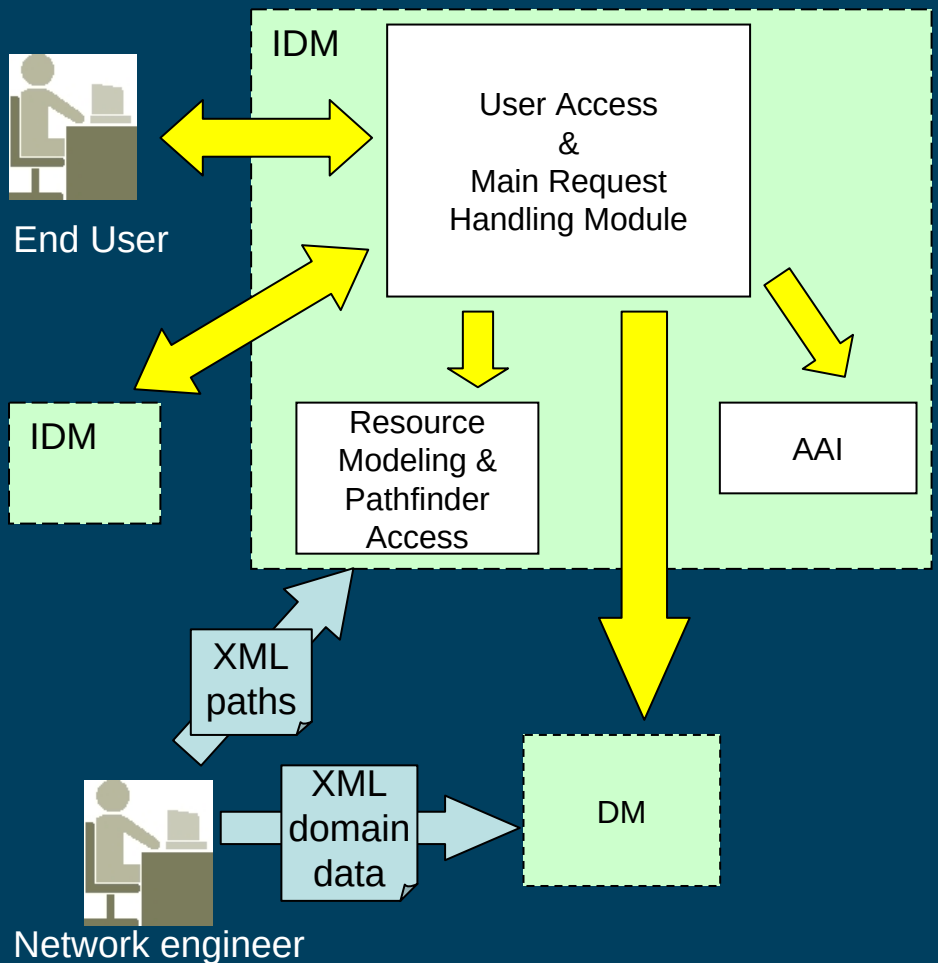
IDM Prototype implementation

- Objectives
  - to validate design and architectural assumptions
  - to define potential risk points and bottlenecks
  - to test IDM reservation procedures and communication schemas
- Modular implementation
- Web-services' technology

IDM Prototype

- Minimal functionality
  - IDM prototype is designed to provide minimal required functionality for test purposes. Therefore, the following limitations are introduced:
    - Simplified network abstraction schema
    - Pathfinder and DM processing are pre-defined manually
    - Authentication is based on X.509 certificates for SSL connections and authorization is based on „always allow” policy
    - Data life-time is limited to application run-time

IDM Prototype



- **IDM prototype features**
  - accepts UNI service request (request, cancel, status)
  - NNI communication is implemented, so domains can agree on reservation parameters and schedule resources' booking
  - performs reservation process at inter-domain level (inter-domain link capacity check, VLAN numbers, path costs validation)
  - the pathfinder supports IDM with manually pre-defined inter-domain paths
  - DM supports IDM with manually pre-defined information about domain topology

IDM Prototype

- Future development after prototype tests:
  - design and implementation of DM functionality (may include manual provisioning)
  - design of network resources' representation at the IDM and DM level
  - extensions to the current transaction mechanism (data life-time will exceed application run-time)
  - full implementation of pathfinder functionality
  - AAI extensions, incorporating the federated model of JRA5 activity in GN2 project

Intra-domain provisioning

- Manual intra-domain configurations and provisioning for the establishment of the intra-domain segments of the end-to-end path
- Intra-domain provisioning design to accommodate
  - Domains that have a G.ASON/GMPLS CP “out of the box” e.g. Generic MPLS Routing Engine (distributed control plane in their Alcatel 1678 MCC OXC)
  - Domains operated via NMS
  - Domains that may decide to adopt proprietary Bandwidth Brokers

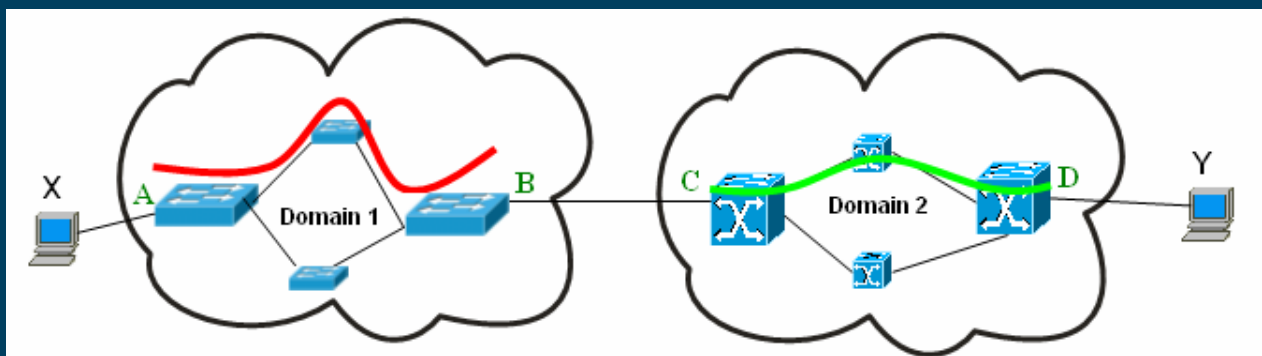
Domain Manager

- Intra-domain modules, implemented in later phases, will comprise the so-called BoD service Domain Manager (DM)
  - Processes intra-domain provisioning requests from the IDM wrt technology-specific issues
  - Provides to the IDM intra-domain topology updates
  - Includes one or more technology proxy sub-modules for the configuration of the network elements/interaction with the local NMS/interaction with the local control plane

Technology Stitching
Why is it needed?

Connect. Communicate. Collaborate

- Different network technologies exist across NRENs and this is not expected to change in the near future
- Need to provide a homogenous method to interconnect domains
- The technology stitching sub activity starts with determining/collecting (manual) procedures how to stitch technologies between two domains
- Automated Technology Stitching is the aim



Technology Stitching
Network Technology Types

Connect. Communicate. Collaborate

- Based on existing NREN technologies
  - SONET/SDH
  - Ethernet based:
    - Native Ethernet
    - L2 MPLS VPN
  - DiffServ technologies
    - PIP
    - IP MPLS QoS
- 14 different interconnection scenarios in total identified

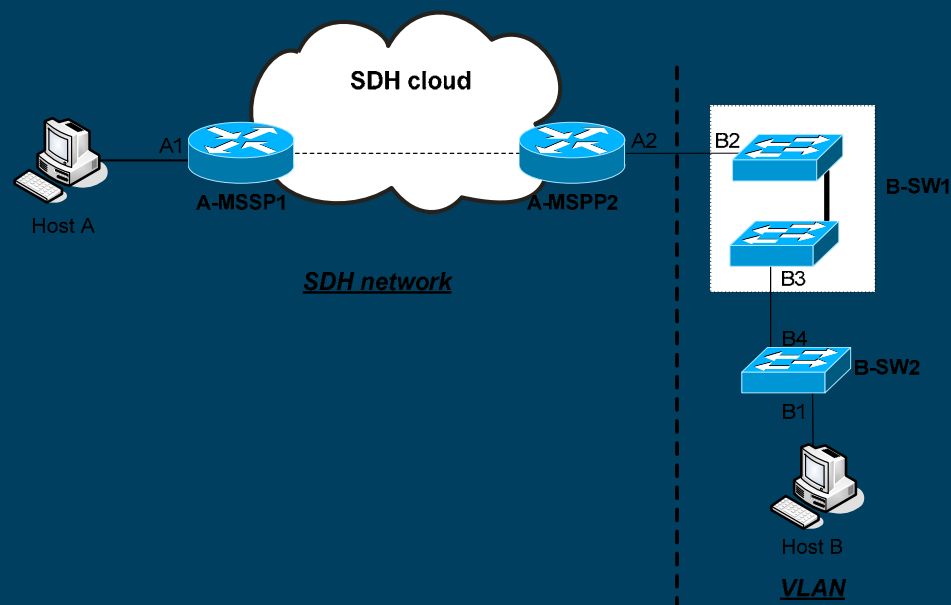
	SDH/SONET	Layer 2 MPLS VPN	PIP	IP MPLS QoS	Ethernet
SDH/SONET	x				
Layer 2 MPLS VPN	x	x			
PIP	x	x	x		
IP MPLS QoS	x		x	x	
Ethernet	x	x	x	x	x

Connect. Communicate. Collaborate

Technology Stitching
The testing process

- Simple manual stitching testing between one technological domain and another
- 1 GE circuit implemented over an SDH network using GFP and VCAT, and across an Ethernet domain using trunk connections implemented as a VLAN
- Pre-Conditions
  - Sufficient BW available on the SDH network to create the full-rate connection
  - Host A and B have an IP address in the same (sub) network range
  - Link connectivity is present
- Actions
  - Configure the circuit on the SDH network as a VC4-7v or a VC4-8c.
  - Configure the trunk circuit on the Ethernet network
  - Testing connectivity: ping hosts
  - Testing BW achieved: use a test set and loopbacks on either end of the circuit

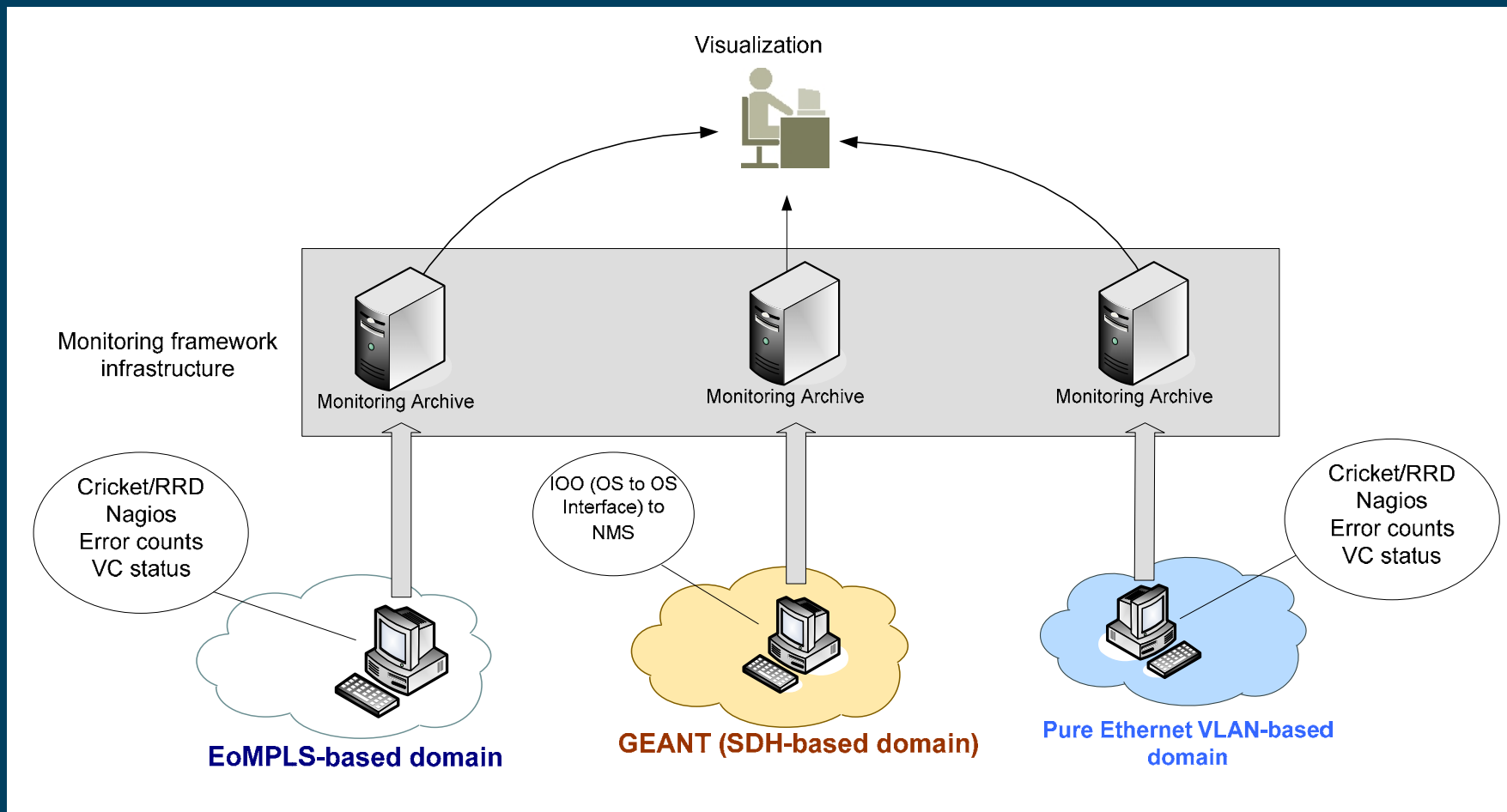
SAMPLE CASE



JRA3 BoD Monitoring

- JRA3 activity aims to use existing NRENs' network infrastructure to provide a BoD service, under a single interface
- GN2 JRA1 activity aims to use provide ubiquitous access to monitoring information for groups of uses
  - Definition of a framework
  - Easy to install, easy to configure, covering the different needs of the NRENs, easy to modify
  - Integrate the measurement tools within the framework as reference implementations
- JRA3 should build the technology-specific measurement tools for end-to-end L1-L2 services and feed them to the JRA1 framework for storage, processing, concatenation and visualization purposes

Overview


Visualization
Monitoring framework infrastructure
Monitoring Archive
Monitoring Archive
Monitoring Archive
Cricket/RRD
Nagios
Error counts
VC status
IOO (OS to OS Interface) to NMS
Cricket/RRD
Nagios
Error counts
VC status
EoMPLS-based domain
GEANT (SDH-based domain)
Pure Ethernet VLAN-based domain

Monitoring priorities

- Technologies: BoD Ethernet circuits over
  - One EoMPLS/switched Ethernet network
  - One SDH-based network

- Metrics to be monitored, in order of priority
  - Up/down
  - Degraded/not degraded
  - Level of usage (where possible)

Progress

- EoMPLS and Geant2 IOO monitoring now being implemented
- XML schema towards existing JRA1 monitoring framework
- First implementation (up/down status) across two domains (one EoMPLS, one SDH)
- Next, work on concatenating more complex metric across multiple technologies

JRA3 thinks it is also fundamental to work on …..

- Looking into developments in standardization bodies (OIF, IETF)
- Collaboration with Internet2, CANARIE and ESnet (already on-going for 2 years)
- Liaison with projects: MUPBED, NOBEL, VIOLA, DRAGON, HOPI, UCLPv2, ...
- Specifying requirements for a pan-European scale test-bed to test JRA3 prototypes and modules
- General information at: http://www.geant2.net/server/show/nav.756 (to be updated)
- Collecting user/application requirements on BoD service
  - Please send your feedback to: sevasti@grnet.gr

JRA3 team

- The JRA3 work is a joint effort of the following NRENs & DANTE
  - CARNET
  - CESNET
  - DANTE
  - FCCN
  - GARR
  - GRNET
  - HEANET
  - HUNGARNET
  - PSNC
  - REDIRIS
  - RENATER
  - SURFNET
- This presentation was co-authored by
  - Mauro Campanella (GARR)
  - Radek Krzywania (PSNC)
  - Noel McKenna (HEAnet)
  - Victor Reijs (HEAnet)
  - Dave Wilson (HEAnet)