R&D future hybrid networks

Cees de Laat

SURFnet

BSIK

EU

University of Amsterdam

SARA

TNO

NCF

Advanced School for Computing and Imaging

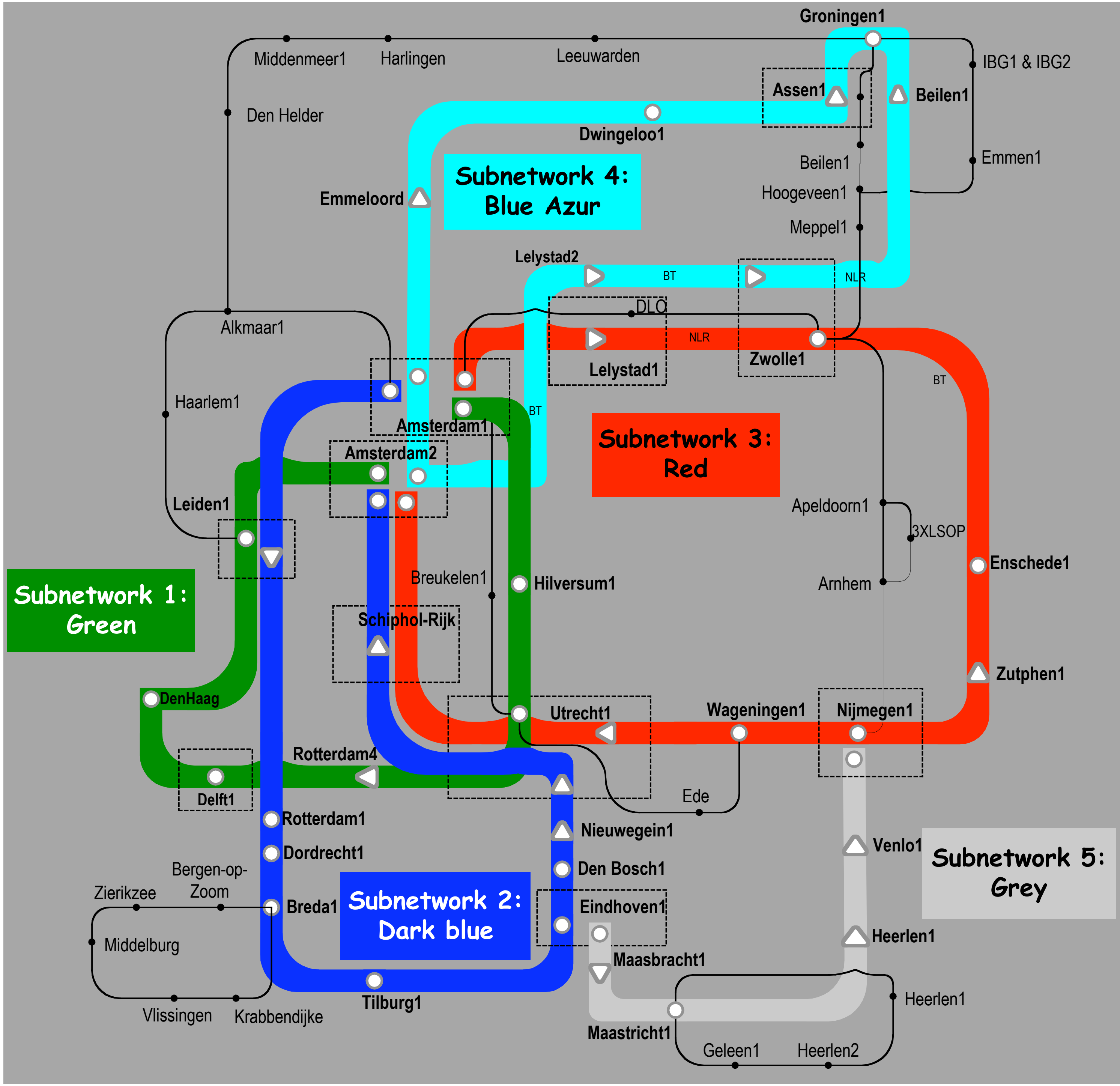
In The Netherlands SURFnet connects between 180:

- universities;
- academic hospitals;
- most polytechnics;
- research centers.

with an indirect ~750K user base

~ 6000 km

scale comparable to railway system

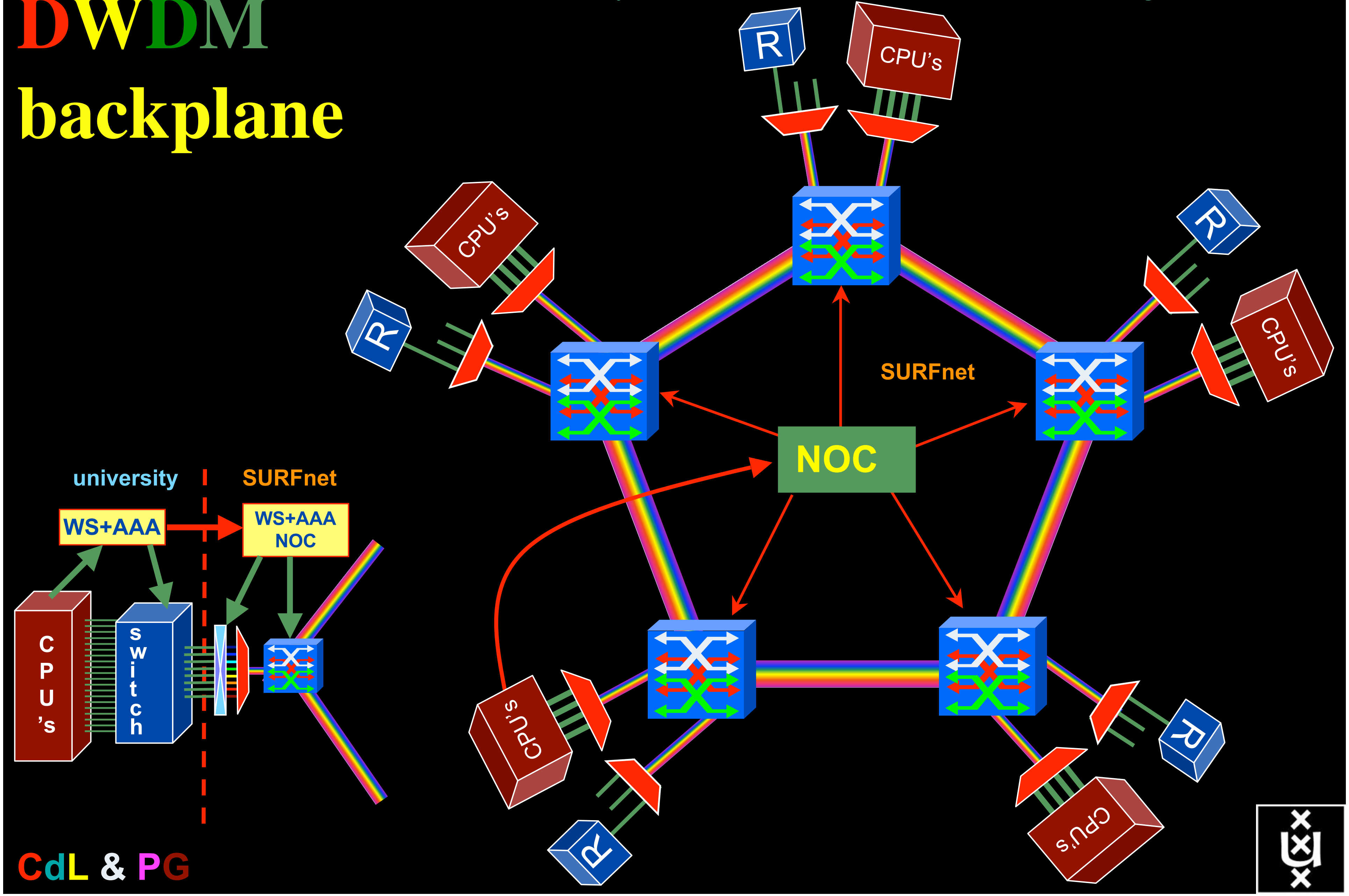


Common Photonic Layer (CPL) in SURFnet6

supports up to
72 Lambda's of
10 G each
40 G soon.

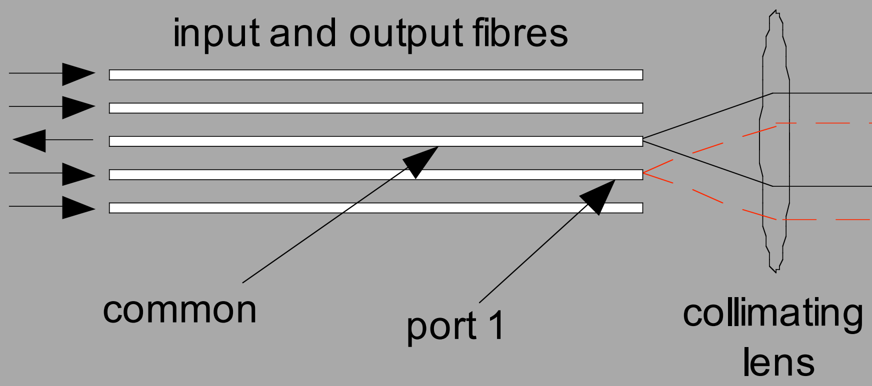
StarPlane DWDM backplane

The novelty: to give flexibility directly to the applications by allowing them to choose the logical topology in real time, ultimately with sub-second lambda switching times.



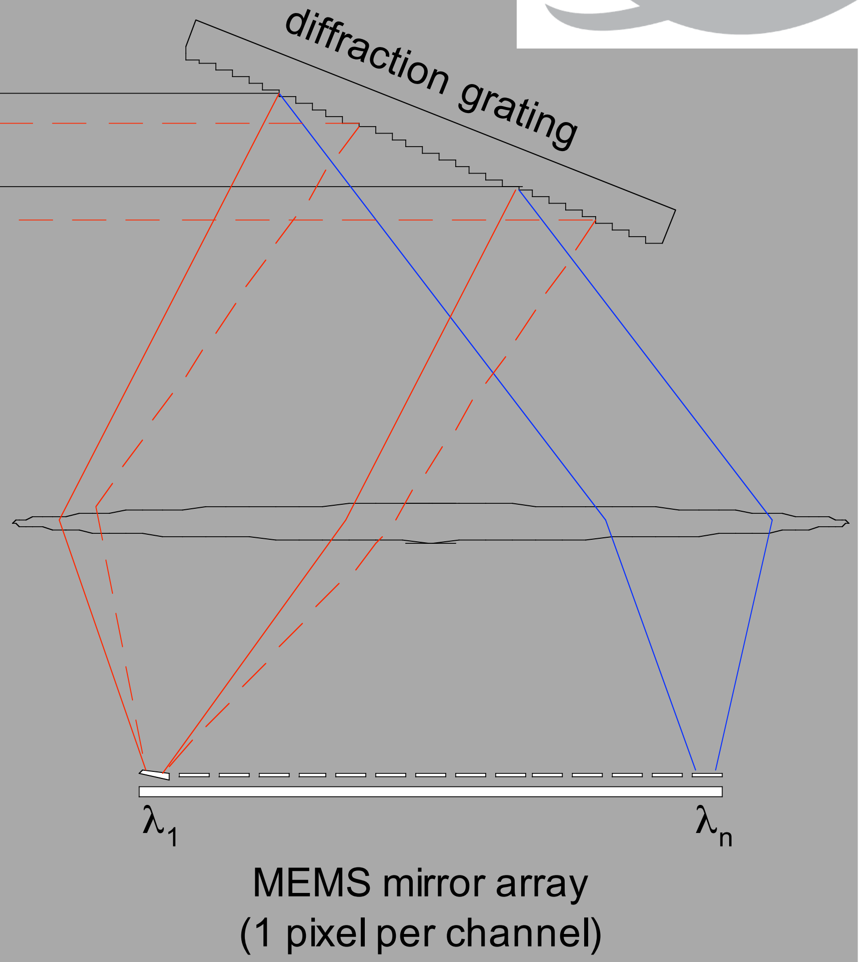
CdL & PG

Module Operation



> this schematic shows
- several input fibres and one output fibre
- light is focused and diffracted such that each channel lands on a different MEMS mirror
- the MEMS mirror is electronically controlled to tilt the reflecting surface
- the angle of tilt directs the light to the correct port

> in this example:
- channel 1 is coming in on port 1 (shown in red)
- when it hits the MEMS mirror the mirror is tilted to direct this channel from port 1 to the common
- only port 1 satisfies this angle, therefore all other ports are blocked



ref Eric Bernier, NORTEL

Dispersion compensating modem: eDCO from NORTEL

(Try to Google eDCO :-)

sender → receiver

h(t)

T(f)

transport as function of frequency T(f)

$F^{-1}(F(h(t)).T(f))$

Solution in 5 easy steps for dummy's :

1. try to figure out T(f) by trial and error
2. invert T(f) -> $T^{-1}(f)$
3. computationally multiply $T^{-1}(f)$ with Fourier transform of bit pattern to send
4. inverse Fourier transform the result from frequency to time space
5. modulate laser with resulting $h'(t) = F^{-1}(F(h(t)).T^{-1}(f))$

sender with cpu, mod → receiver

$F^{-1}(F(h(t))*T^{-1}(f))$

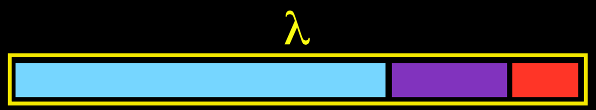
T(f)

$F^{-1}(F(F^{-1}(F(h(t)).T^{-1}(f))).T(f))$ -> h(t)

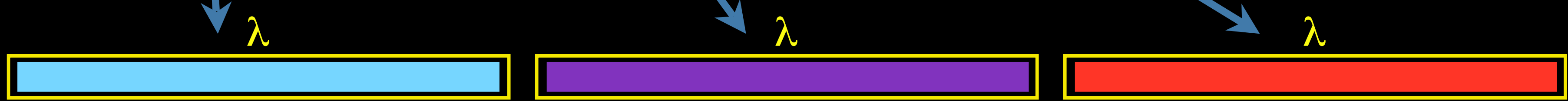
(ps. due to power ~ square E the signal to send **looks** like uncompensated received but is not)

QOS in a non destructive way!

- Destructive QOS:
  - have a link or λ
  - set part of it aside for a lucky few under higher priority
  - rest gets less service



- Constructive QOS:
  - have a λ
  - add other λ's as needed on separate colors
  - move the lucky ones over there
  - rest gets also a bit happier!



What makes StarPlane fly?

- Wavelength Selective Switches
  - for the “low cost” photonics
- Sandbox by confining StarPlane to one band
  - for experimenting on a production network
- Optimization of the controls to turn on/off a Lambda
  - direct access to part of the controls at the NOC
- electronic Dynamically Compensating Optics (eDCO)
  - to compensate for changing lengths of the path
- traffic engineering
  - to create the OPN topologies needed by the applications
- Open Source GMPLS
  - to facillitate policy enabled cross domain signalling

R&D issues

Physical layer

- filter-less networks
- tunable transmitters and receivers including dispersion compensation
- wavelength selective switches

Data link layer

- PBT, PLSB, addressing

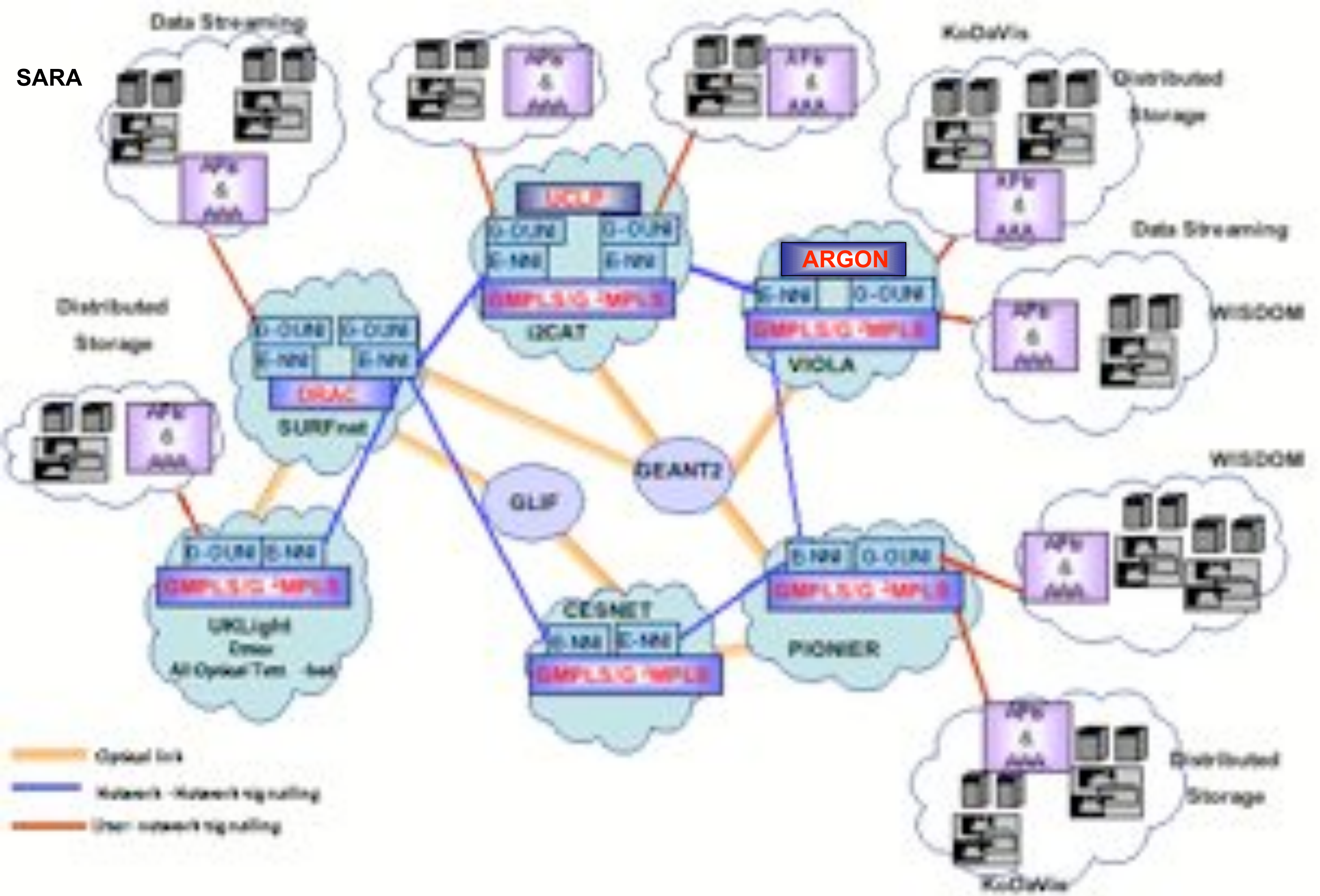
Network transport Layer

- addressing, routing updates, etc.

Power is a big issue

- UvA cluster uses (max) 30 kWh
- 1 kWh ~ 0.1 €
- per year -> 26 k€/y
- add cooling 50% -> 39 k€/y
- Emergency power system -> 50 k€/y
- per rack 10 kWh is now normal
- Cost of this cluster -> 500 k€/y
- **YOU BURN ABOUT ONE THIRD TO HALF THE CLUSTER OVER ITS LIFETIME!**

European Multi-Domain Test-Bed Including Phosphorus Planned Developments



IRTF - AAAARCH - RG

Authentication Authorisation Accounting ARCHitecture RG

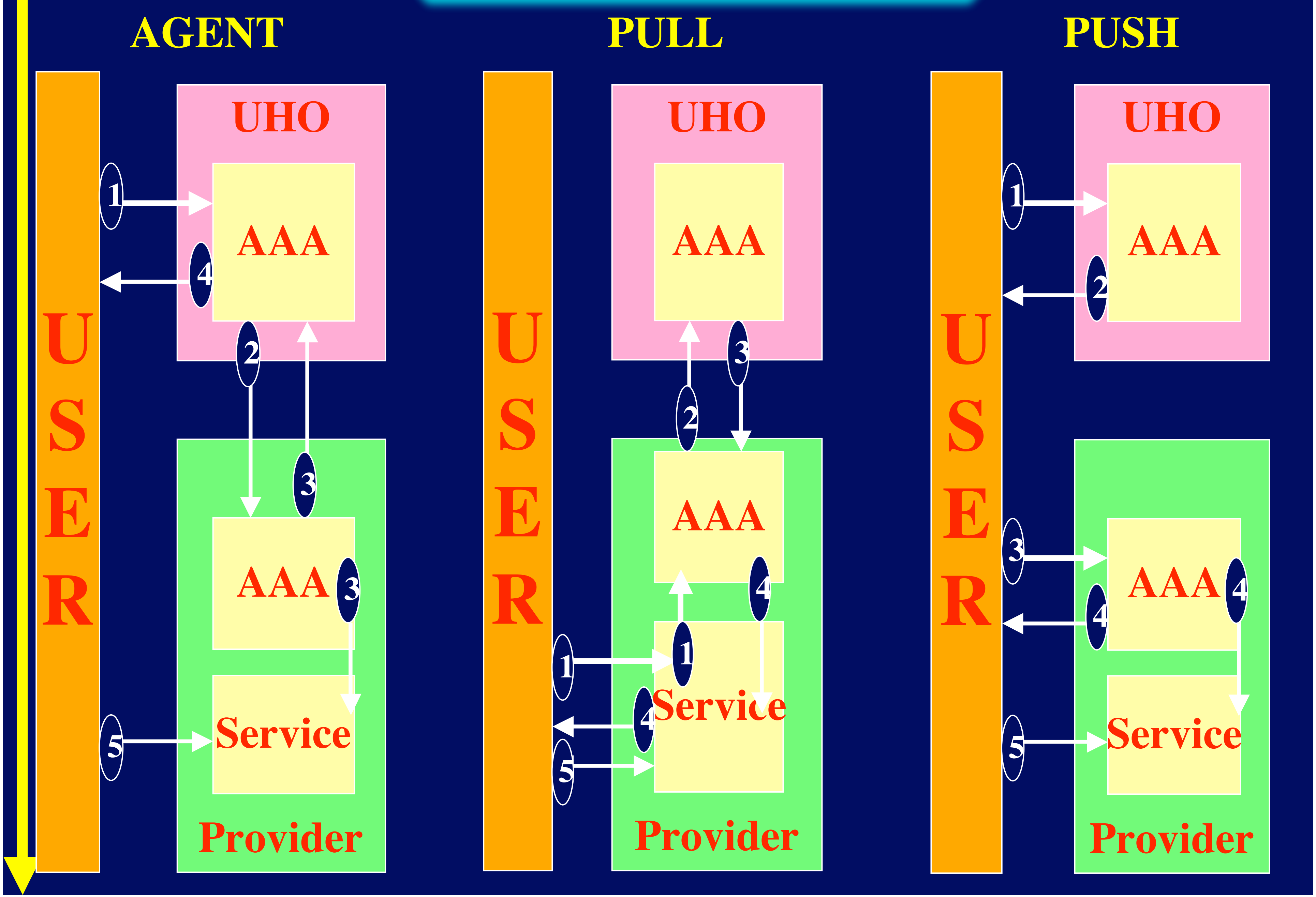
chairs:

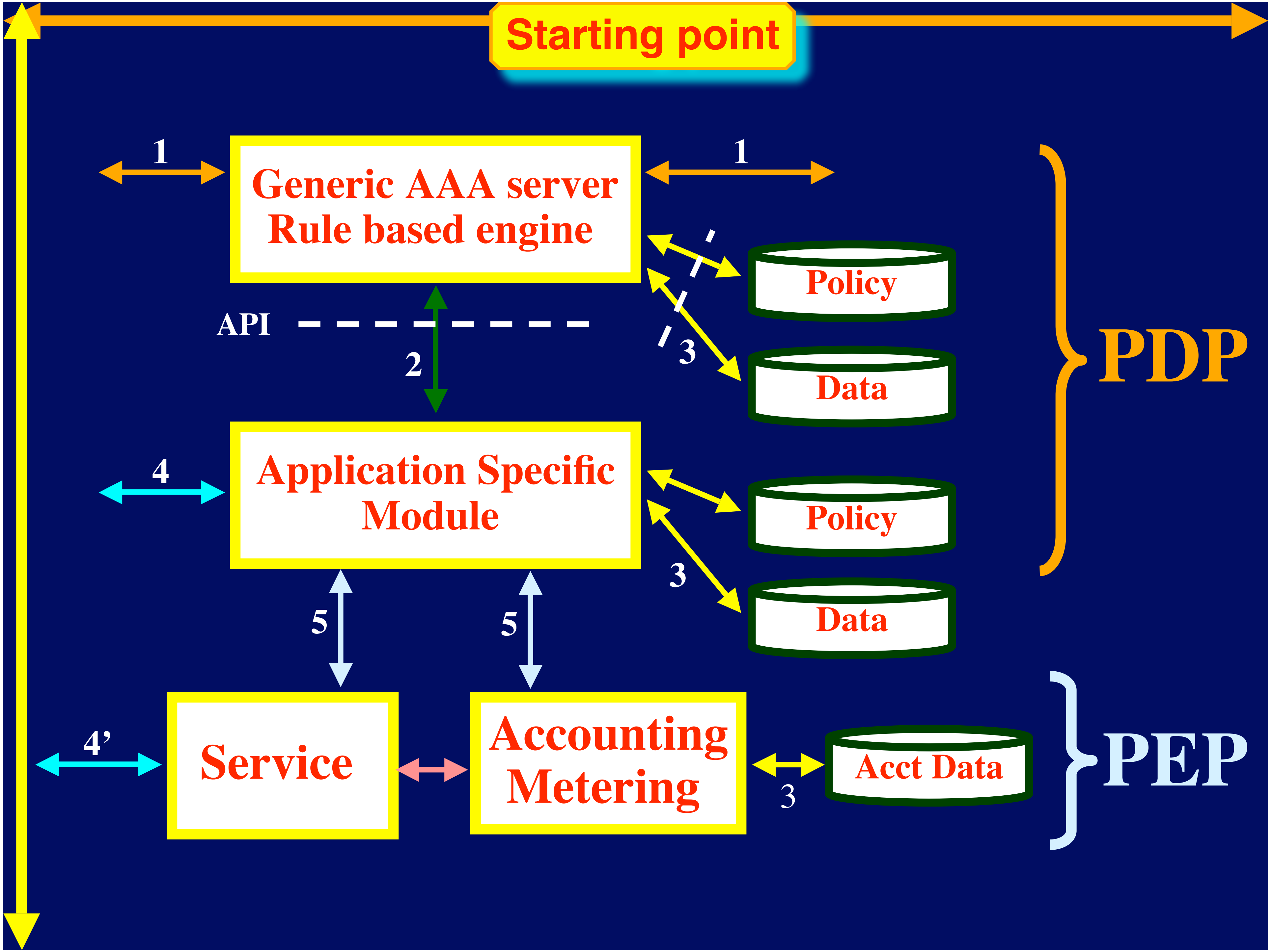
C. de Laat and J. Vollbrecht

www.aaaarch.org

RFC 2903, 2904, 2905, 2906, 3334

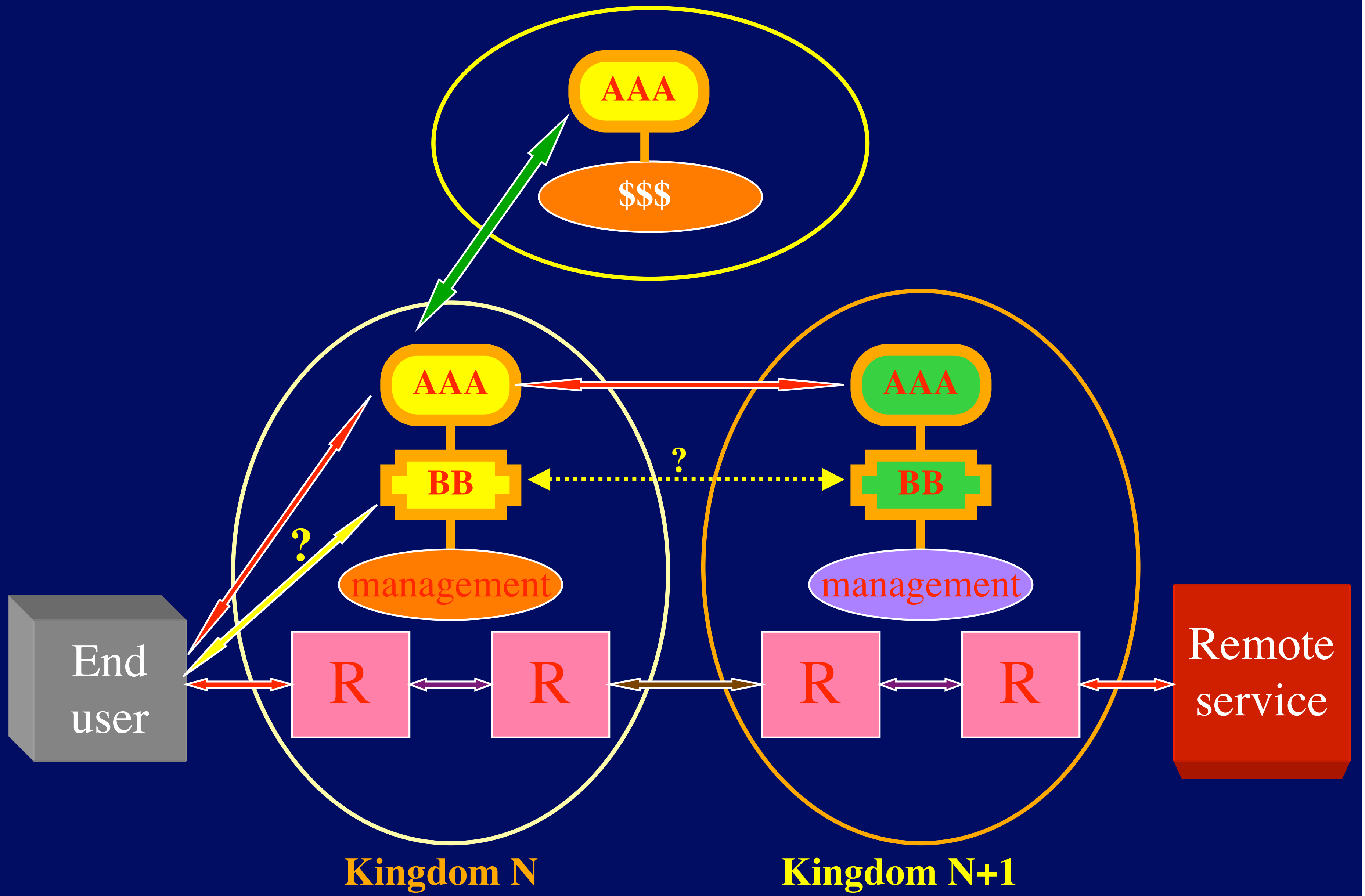
Authorization Models



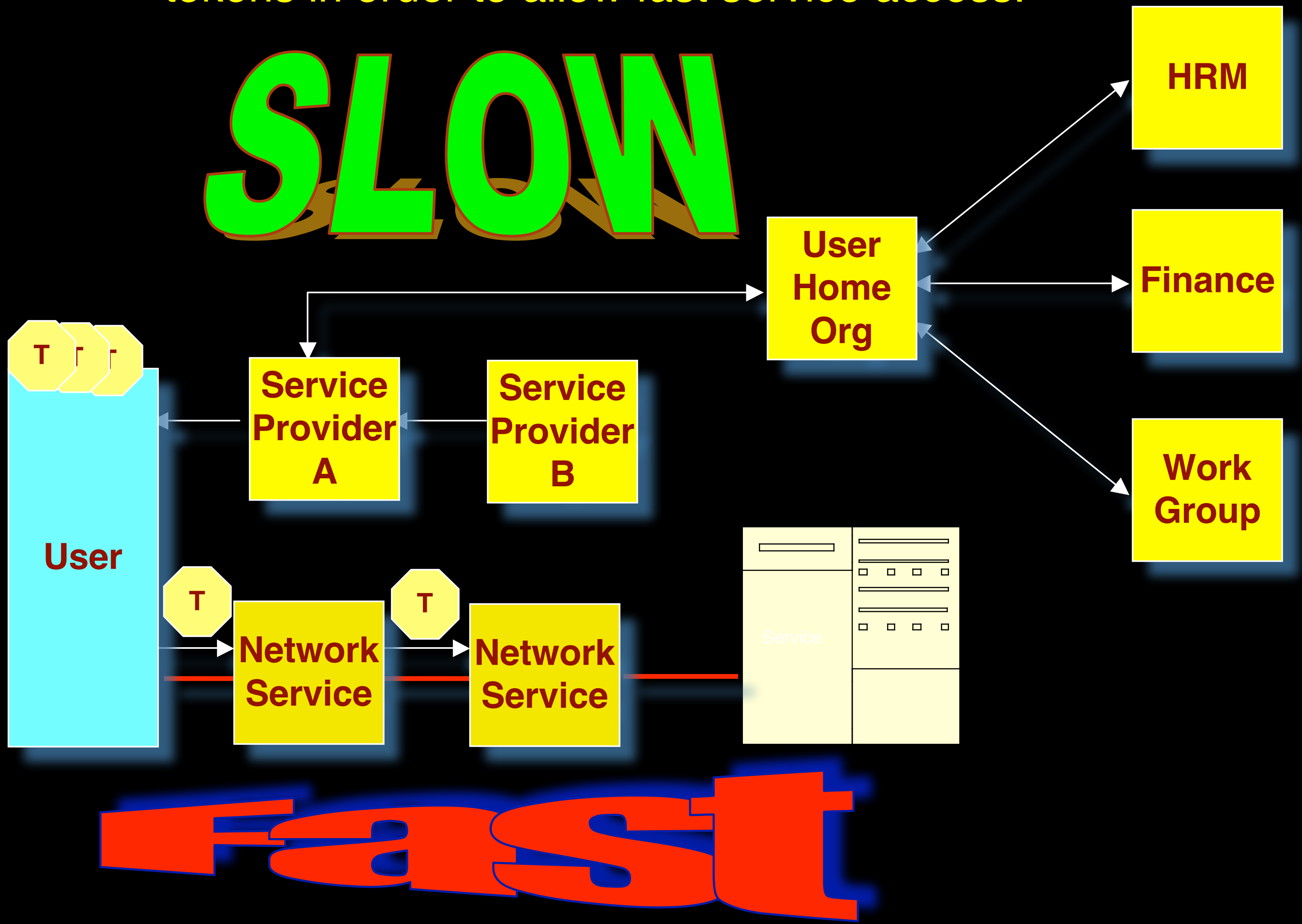


Starting point
1
Generic AAA server
Rule based engine
1
Policy
API
2
3
Data
PDP
4
Application Specific
Module
Policy
3
Data
5
5
4’
Service
Accounting
Metering
3
Acct Data
PEP

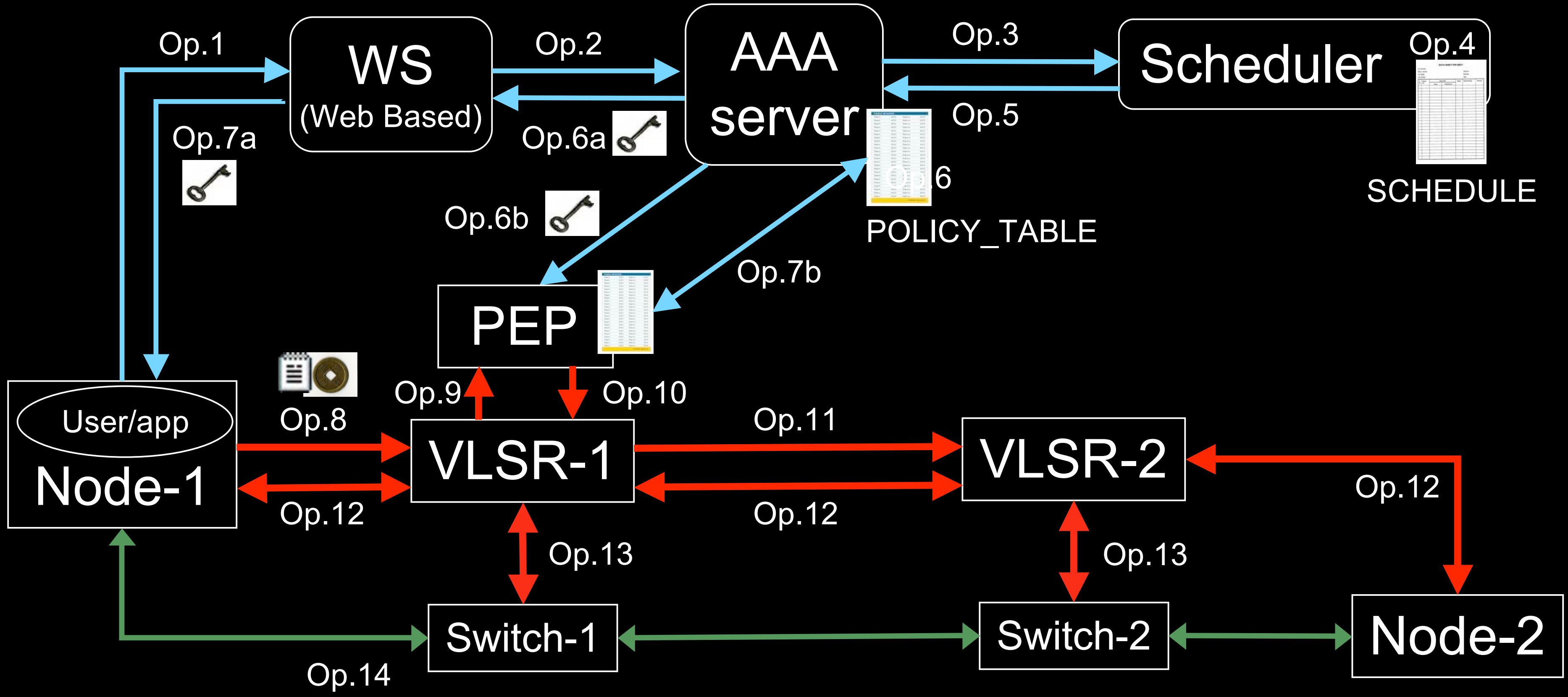
The need for AAA



Use AAA concept to split (time consuming) service authorization process from service access using secure tokens in order to allow fast service access.



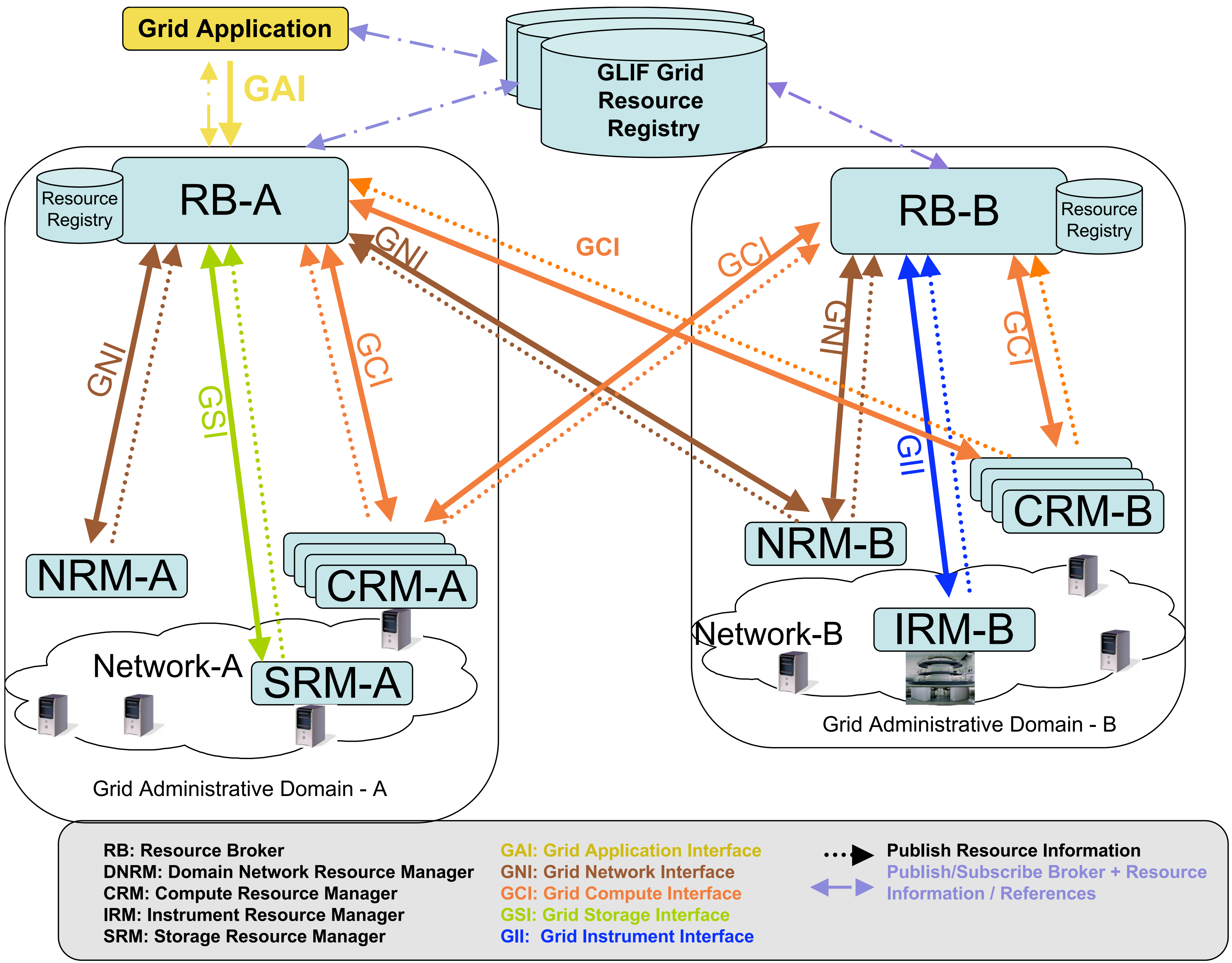
DRAGON GMPLS & TBN Demo, SC06 Tampa

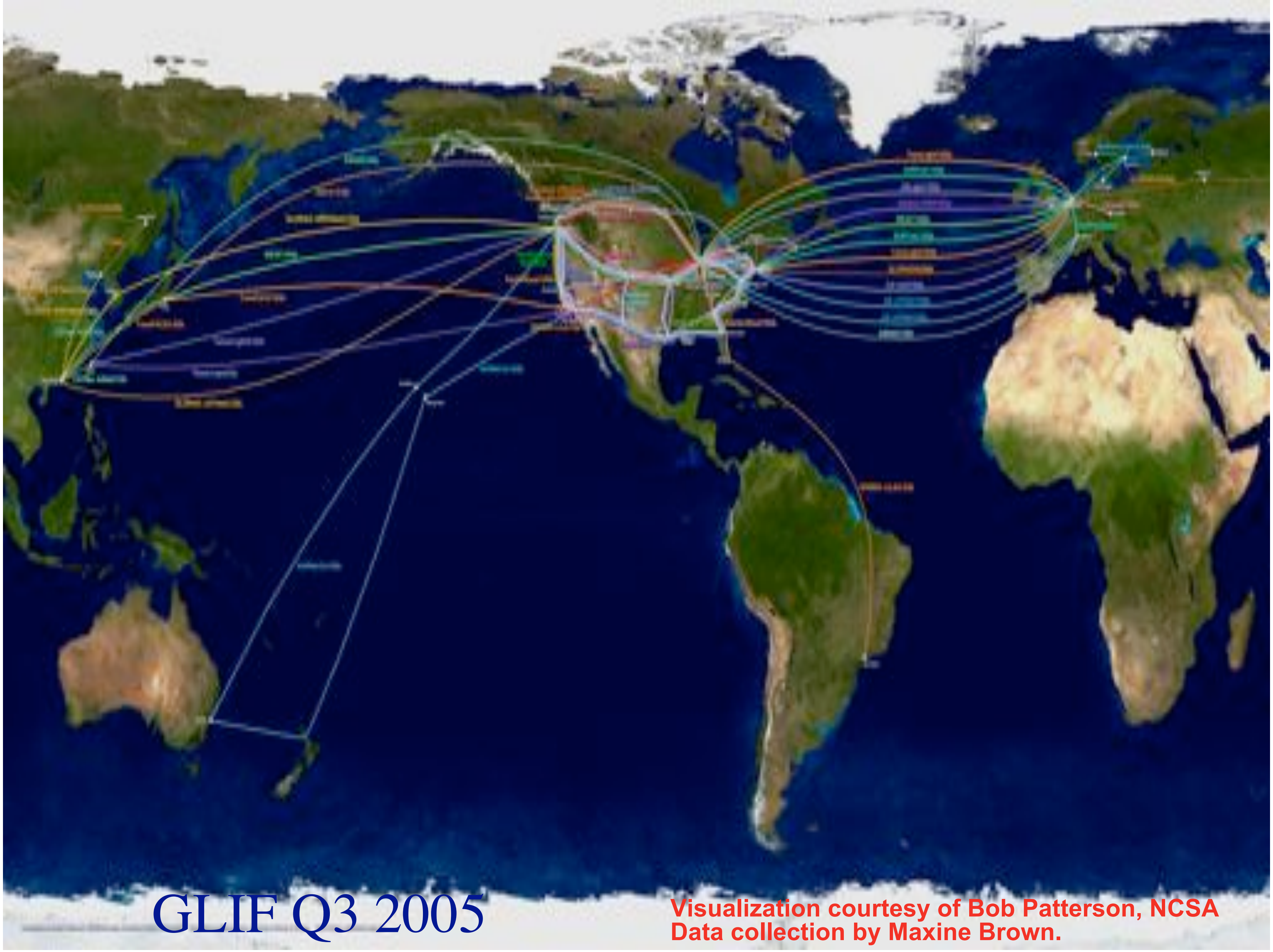


1. User (on Node1) requests a path via web to the WS.
2. WS sends the XML requests to the AAA server.
3. AAA server calculates a hashed index number and submits a request to the Scheduler.
4. Scheduler checks the SCHEDULE and add new entry.
5. Scheduler confirms the reservation to the AAA.
6. AAA server updates the POLICY_TABLE.

6a. AAA server issues an encrypted key to the WS.
6b. AAA server passes the same key to the PEP.
7a. WS passes the key to the user.
7b. AAA server interacts with PEP to update the local POLICY_TABLE on the PEP.

8. User constructs the RSVP message with extra Token data by using the key and sends to VLSR-1.
9. VLSR-1 queries PEP whether the Token in the RSVP message is valid.
10. PEP checks in the local POLICY_TABLE and return YES.
11. When VLSR-1 receives YES from PEP, it forwards the RSVP message.
12. All nodes process RSVP message(forwarding/response)
13. The Ethernet switches are configured
14. LSP is set up and traffic can flow


Grid Application
GAI
GLIF Grid Resource Registry
Resource Registry
RB-A
GNI
GSI
GCI
NRM-A
CRM-A
Network-A
SRM-A
Grid Administrative Domain - A
GCI
GNI
GCI
RB-B
Resource Registry
GNI
GCI
GII
CRM-B
NRM-B
Network-B
IRM-B
Grid Administrative Domain - B
RB: Resource Broker
DNRM: Domain Network Resource Manager
CRM: Compute Resource Manager
IRM: Instrument Resource Manager
SRM: Storage Resource Manager
GAI: Grid Application Interface
GNI: Grid Network Interface
GCI: Grid Compute Interface
GSI: Grid Storage Interface
GII: Grid Instrument Interface
Publish Resource Information
Publish/Subscribe Broker + Resource Information / References

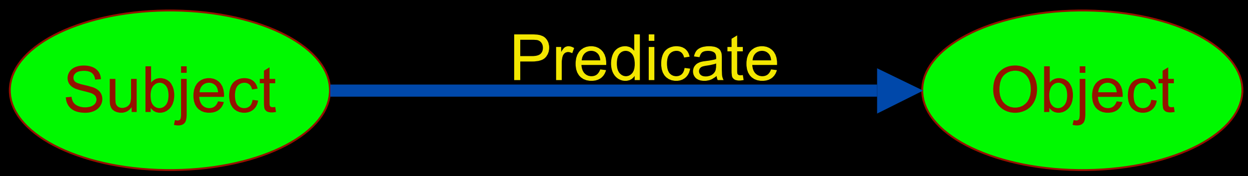


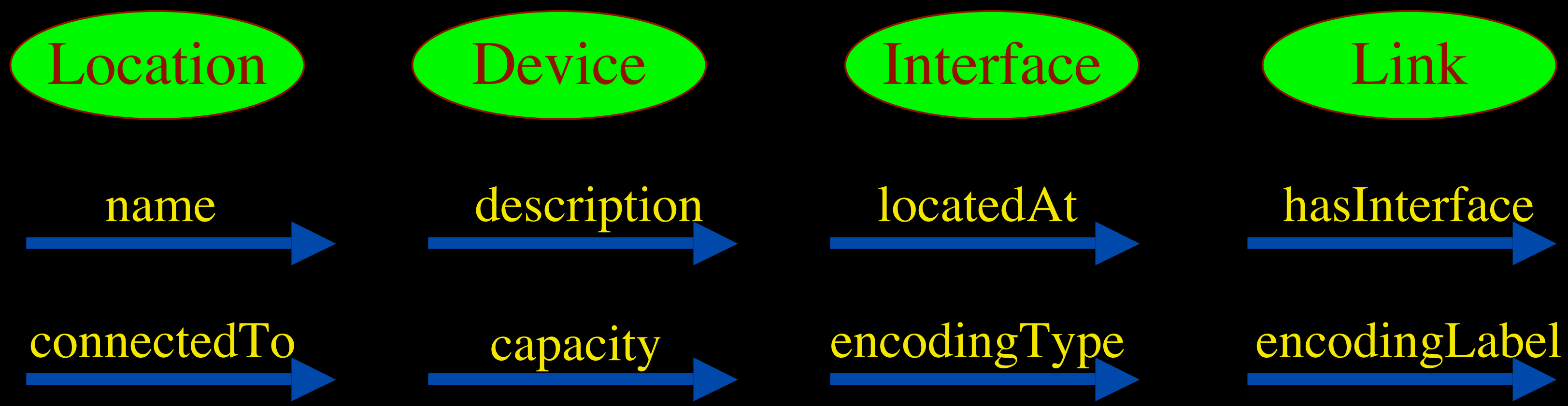
GLIF Q3 2005

Visualization courtesy of Bob Patterson, NCSA
Data collection by Maxine Brown.

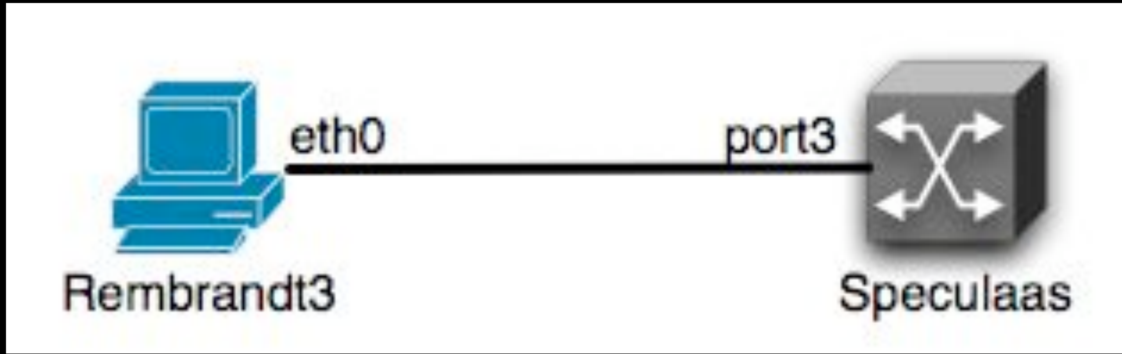
Network Description Language

- From semantic Web / Resource Description Framework.
- The RDF uses XML as an interchange syntax.
- Data is described by triplets:





NDL Example



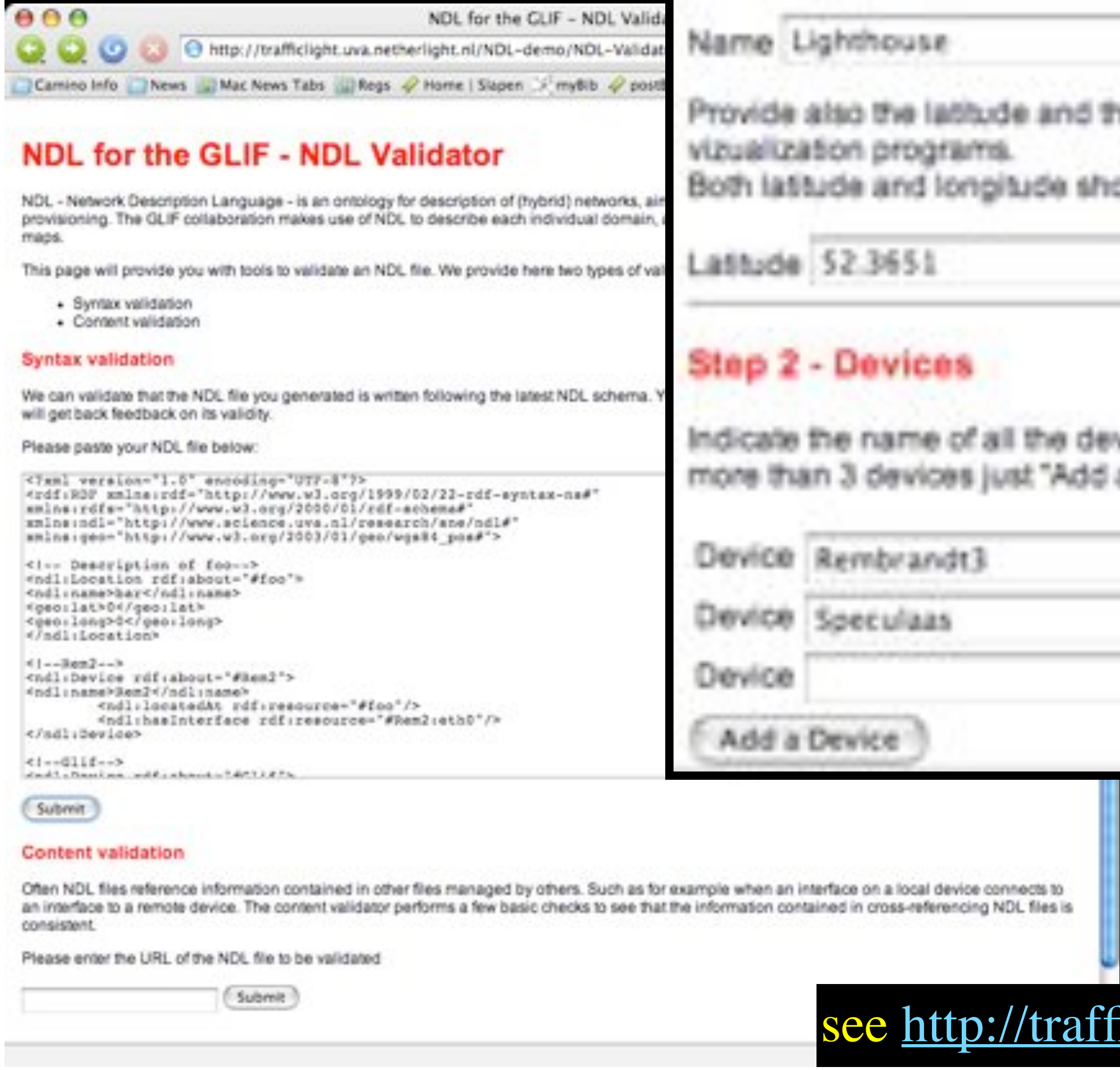

```xml
<ndl:Device rdf:about="#Rembrandt3">
    <ndl:name>Rembrandt3</ndl:name>
    <ndl:locatedAt rdf:resource="#Lighthouse"/>
    <ndl:hasInterface rdf:resource="#Rembrandt3:eth0"/>
</ndl:Device>
<ndl:Interface rdf:about="#Rembrandt3:eth0">
    <ndl:name>Rembrandt3:eth0</ndl:name>
    <ndl:connectedTo rdf:resource="#Speculaas:port3"/>
</ndl:Interface>
```

NetherLight in RDF

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:ndl="http://www.science.uva.nl/research/air/ndl#">
<!-- Description of Netherlight -->
<ndl:Location rdf:about="#Netherlight">
      <ndl:name>Netherlight Optical Exchange</ndl:name>
</ndl:Location>
<!-- TDM3.amsterdam1.netherlight.net -->
<ndl:Device rdf:about="#tdm3.amsterdam1.netherlight.net">
      <ndl:name>tdm3.amsterdam1.netherlight.net</ndl:name>
      <ndl:locatedAt rdf:resource="#amsterdam1.netherlight.net"/>
      <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/1"/>
      <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/3"/>
      <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:501/4"/>
      <ndl:hasInterface rdf:resource="#tdm3.amsterdam1.netherlight.net:503/1"/>
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
      <ndl:hasInterface rdf:resour
```

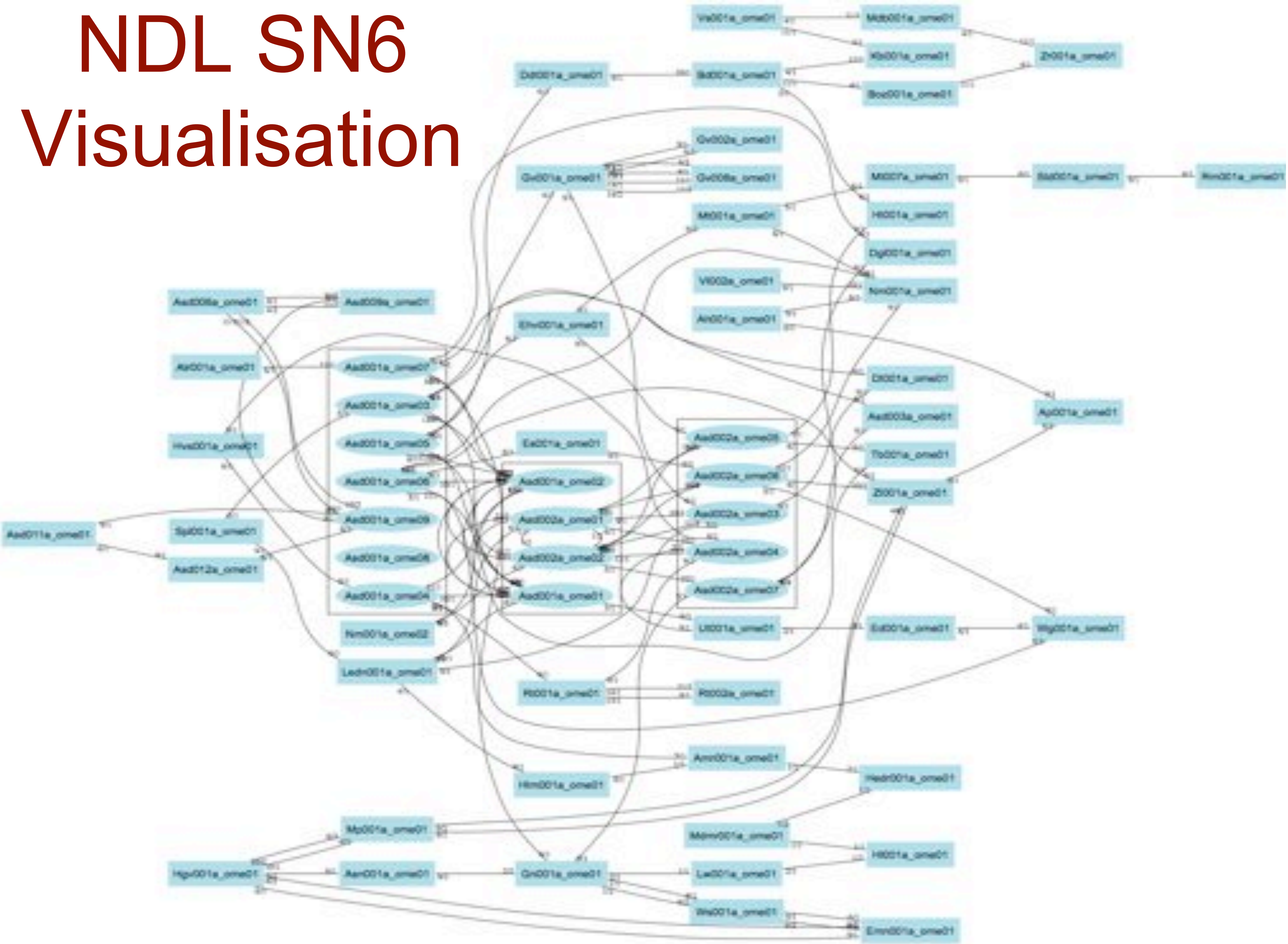
```
<!-- all the interfaces of TDM3.amsterdam1.netherlight.net -->

<ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/1">
            <ndl:name>tdm3.amsterdam1.netherlight.net:POS501/1</ndl:name>
            <ndl:connectedTo rdf:resource="#tdm4.amsterdam1.netherlight.net:5/1"/>
</ndl:Interface>
<ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/2">
            <ndl:name>tdm3.amsterdam1.netherlight.net:POS501/2</ndl:name>
            <ndl:connectedTo rdf:resource="#tdm1.amsterdam1.netherlight.net:12/1"/>
</ndl:Interface>
```

NDL Generator and Validator



see http://trafficlight.uva.netherlight.nl/NDL-demo/

NDL SN6 Visualisation



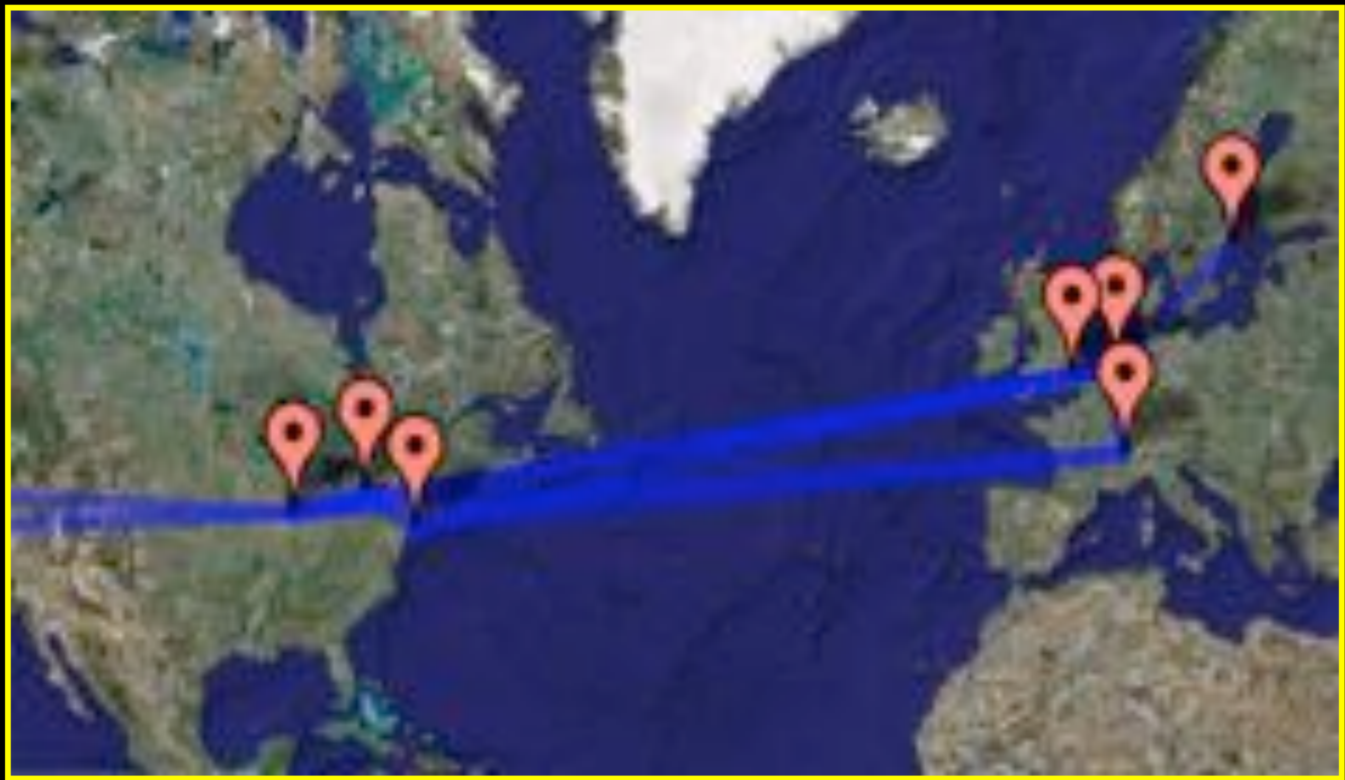
Current status: NDL

NDL - Network Description Language - an RDF based model for hybrid network descriptions.

It leverages all the semantic web tools, to provide:

- parsing of the RDF files
- graphs and visualization of connections and lightpaths
- lightpath provisioning support at inter and intra domain level.

Latest development s were presented at the GLIF meeting in Sep. '06.

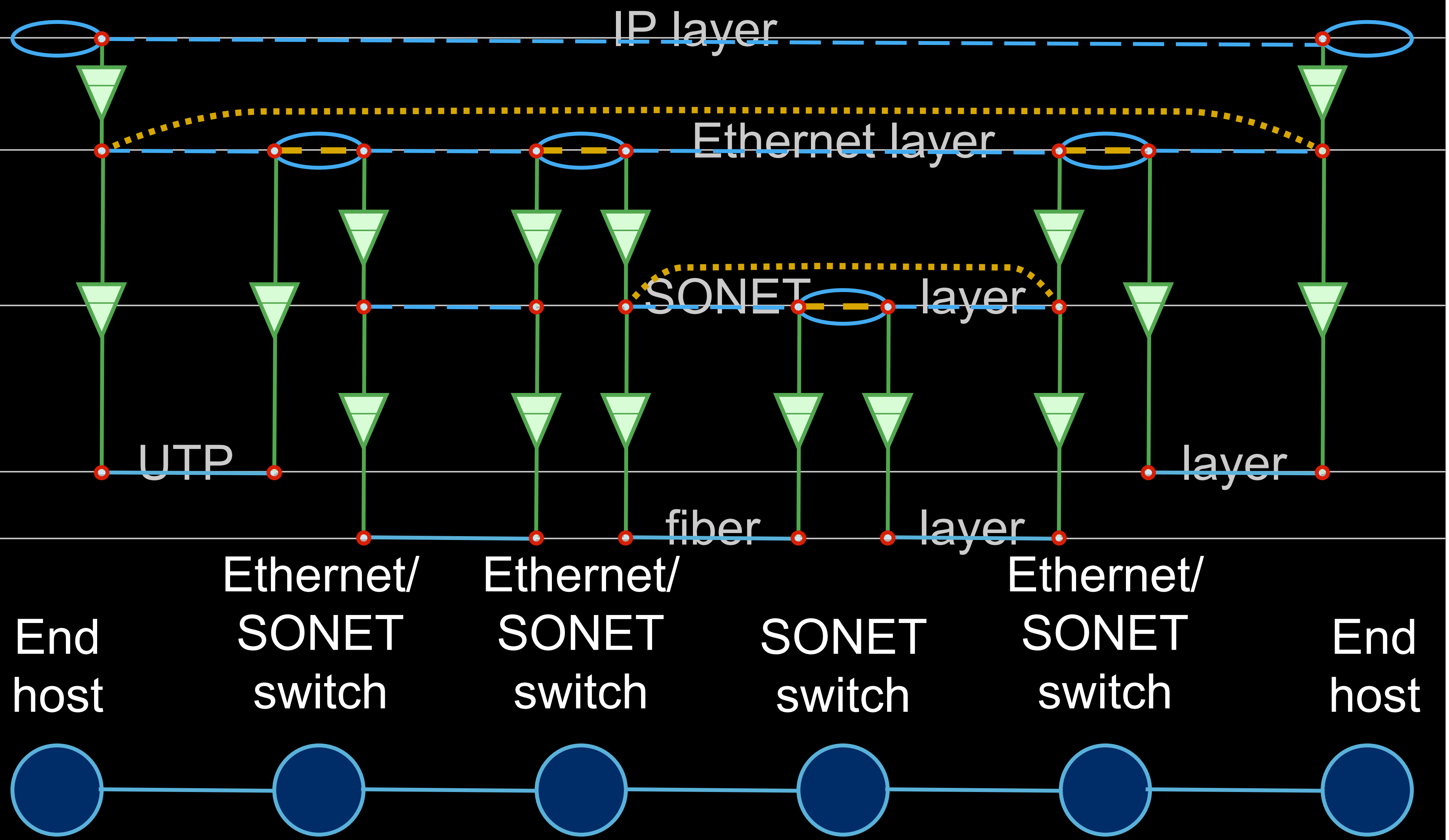


Google map and NDL…

…the GLIF connections described by NDL.

Multi-layer extensions to NDL

Layer schema based on G.805



OGF NML-WG

Open Grid Forum - Network Markup Language workgroup

Chairs:

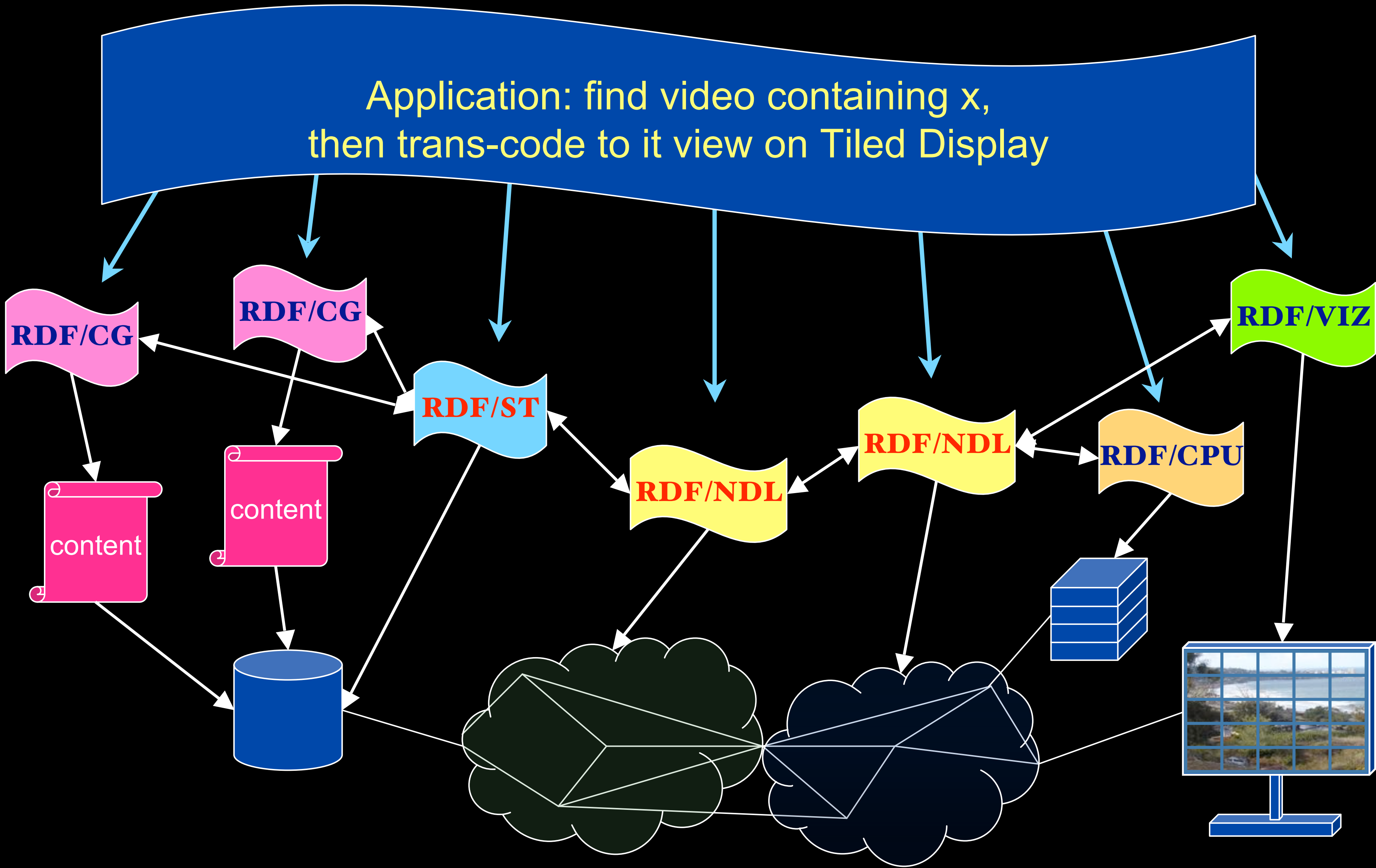
- Paola Grosso – Universiteit van Amsterdam
- Martin Swany – University of Delaware

Purpose:

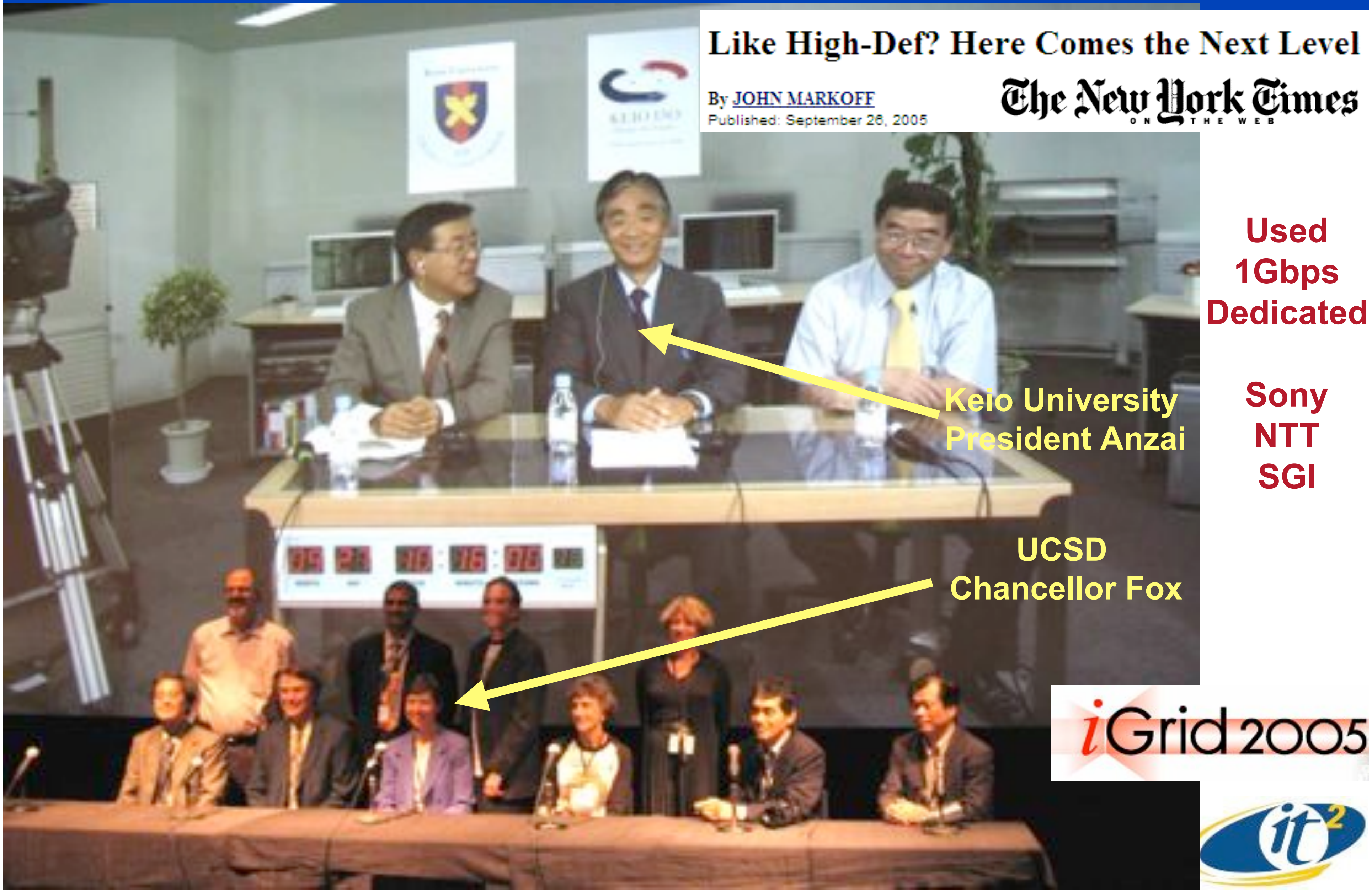
To describe network topologies, so that the outcome is a standardized network description ontology and schema, facilitating interoperability between different projects.

https://forge.gridforum.org/sf/projects/nml-wg

RDF describing Infrastructure



Keio/Calit2 Collaboration: Trans-Pacific 4K Teleconference



Used 1Gbps Dedicated

Sony
NTT
SGI

CineGrid@SARA

TeraThinking

- What constitutes a Tb/s network?
- CALIT2 has 8000 Gigabit drops ?->? Terabit Lan?
- look at 80 core Intel processor
  - cut it in two, left and right communicate 8 TB/s
- think back to teraflop computing!
  - MPI makes it a teraflop machine
- massive parallel channels in hosts, NIC's
- TeraApps programming model supported by
  - TFlops -> MPI / Globus
  - TBytes -> OGSA/DAIS
  - TPixels -> SAGE
  - TSensors -> LOFAR, LHC, LOOKING, CineGrid, ...
  - Tbit/s -> ?

ref Larry Smarr & CdL

Questions ?

Computing & Networking Services

Advanced School for Computing and Imaging