The Photonic Middleware Challenge

Cees de Laat
University of Amsterdam

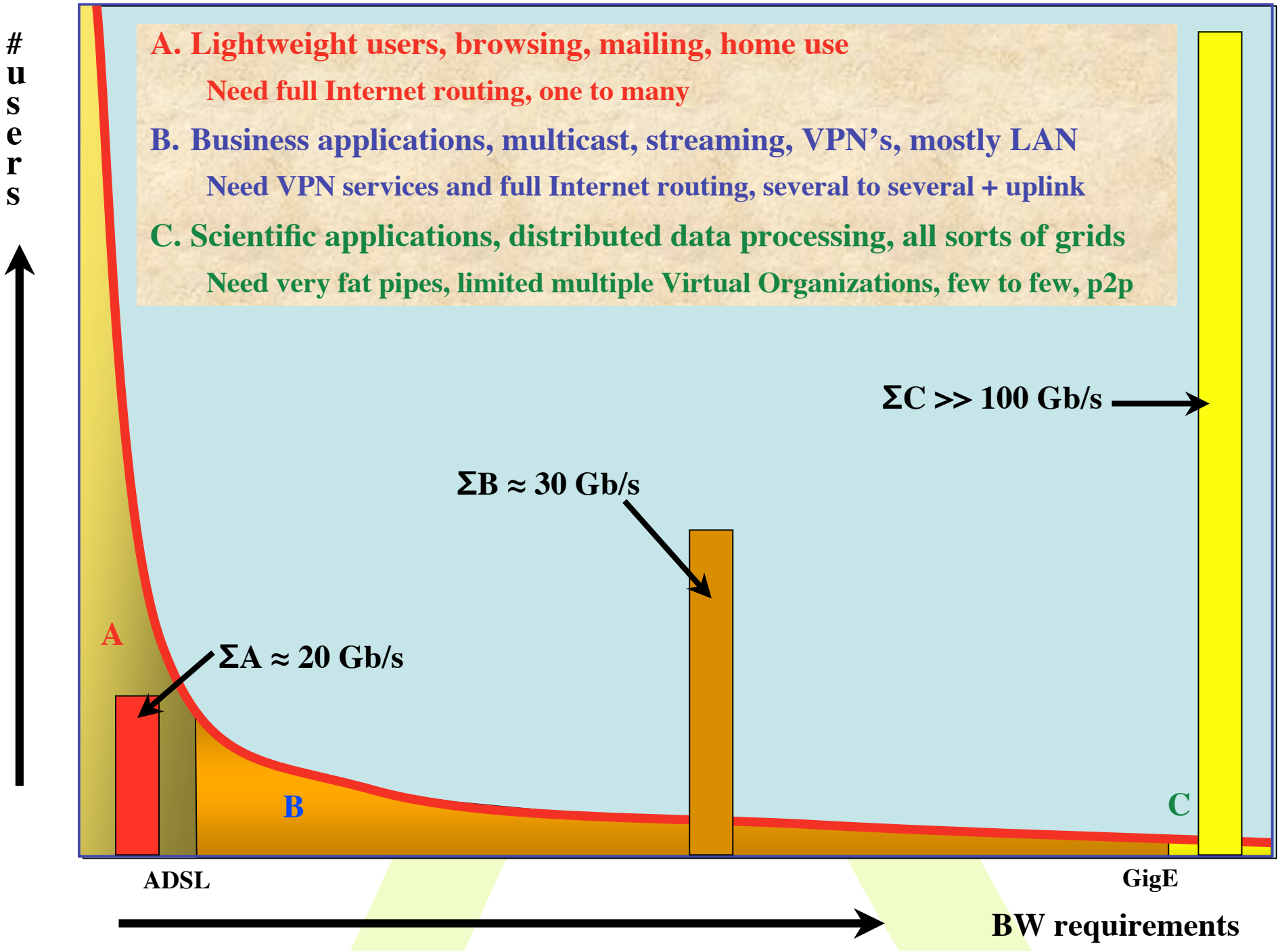
Contents

This page is intentionally left blank

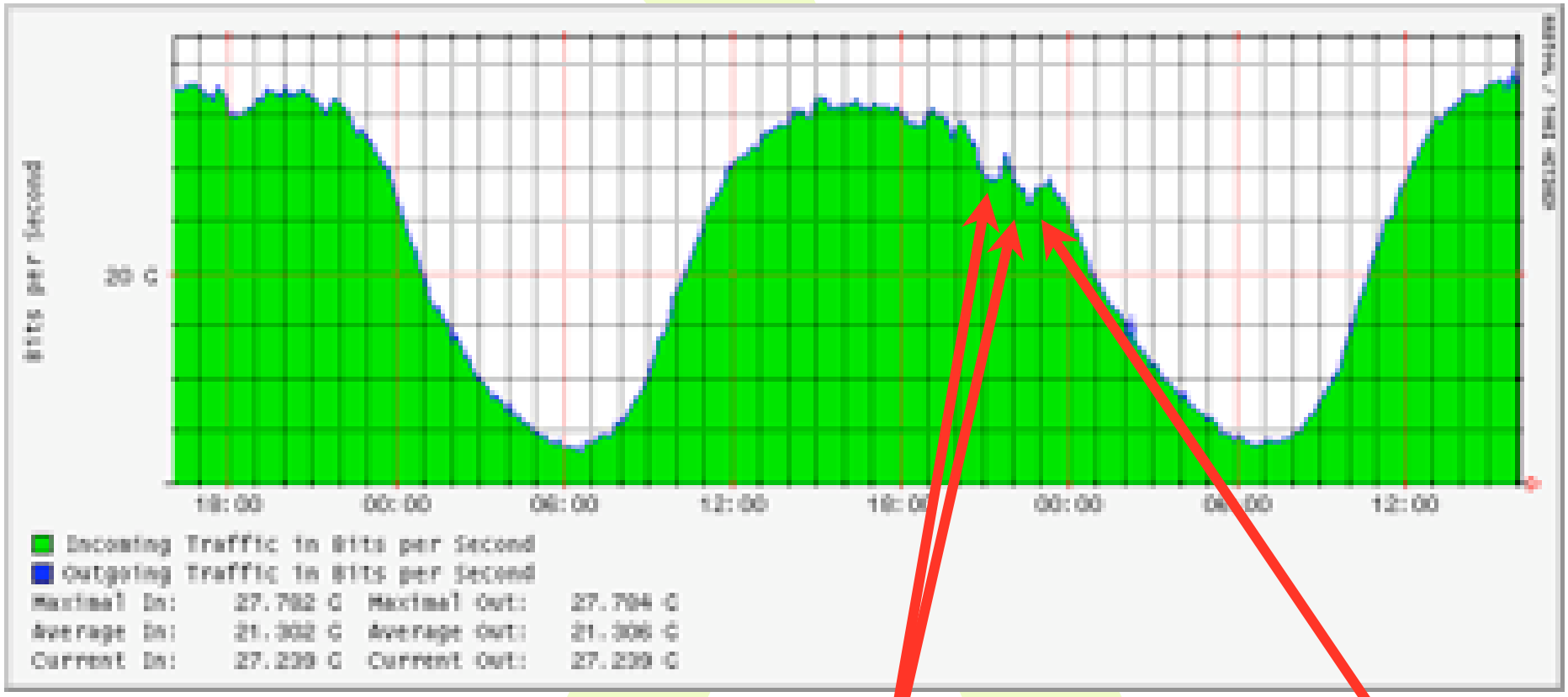
- Ref: www.this-page-intentionally-left-blank.org

The Playing Field

Visualization courtesy of
Bob Patterson, NCSA.


#users
A. Lightweight users, browsing, mailing, home use
Need full Internet routing, one to many
B. Business applications, multicast, streaming, VPN's, mostly LAN
Need VPN services and full Internet routing, several to several + uplink
C. Scientific applications, distributed data processing, all sorts of grids
Need very fat pipes, limited multiple Virtual Organizations, few to few, p2p
ΣC >> 100 Gb/s
ΣB ≈ 30 Gb/s
ΣA ≈ 20 Gb/s
A
B
C
ADSL
GigE
BW requirements

AMS-IX

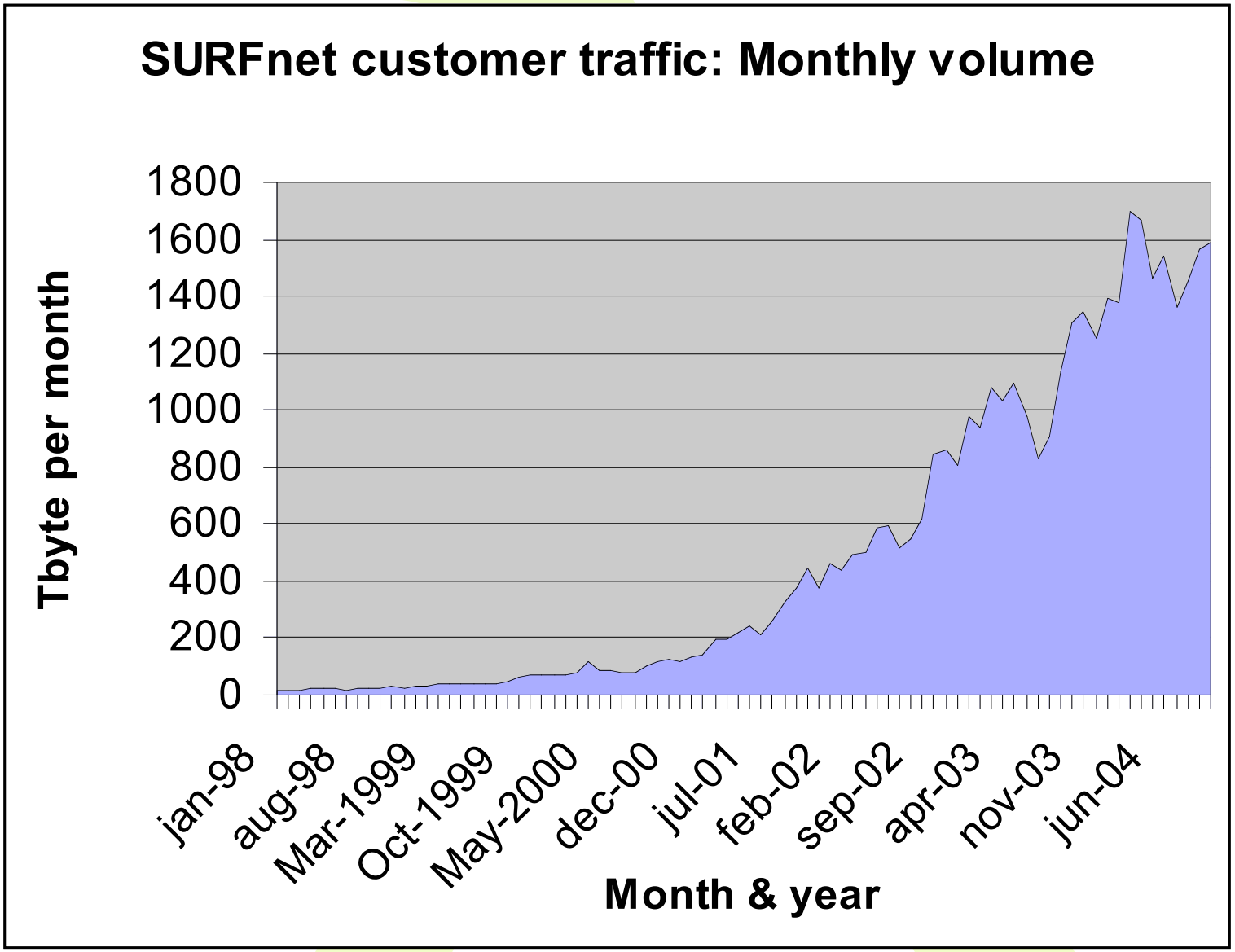


June 19th 2004

Lost :-(

European championship football **Holland -- Czech Republic**

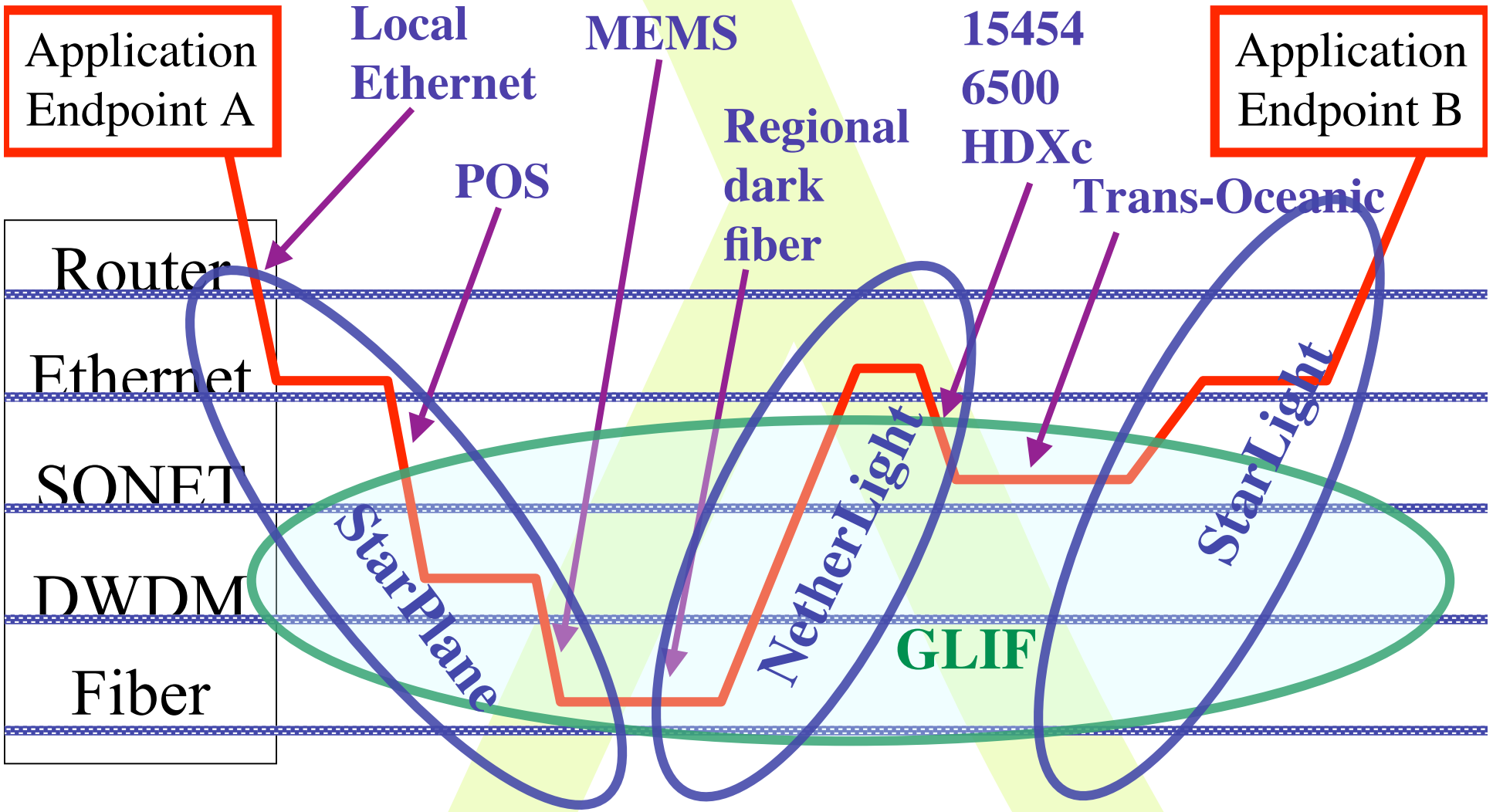
Routed L3 traffic growth



1600 Tbyte/month ≈ 5 Gbits/second

Slide courtesy Kees Neggers

How low can you go?



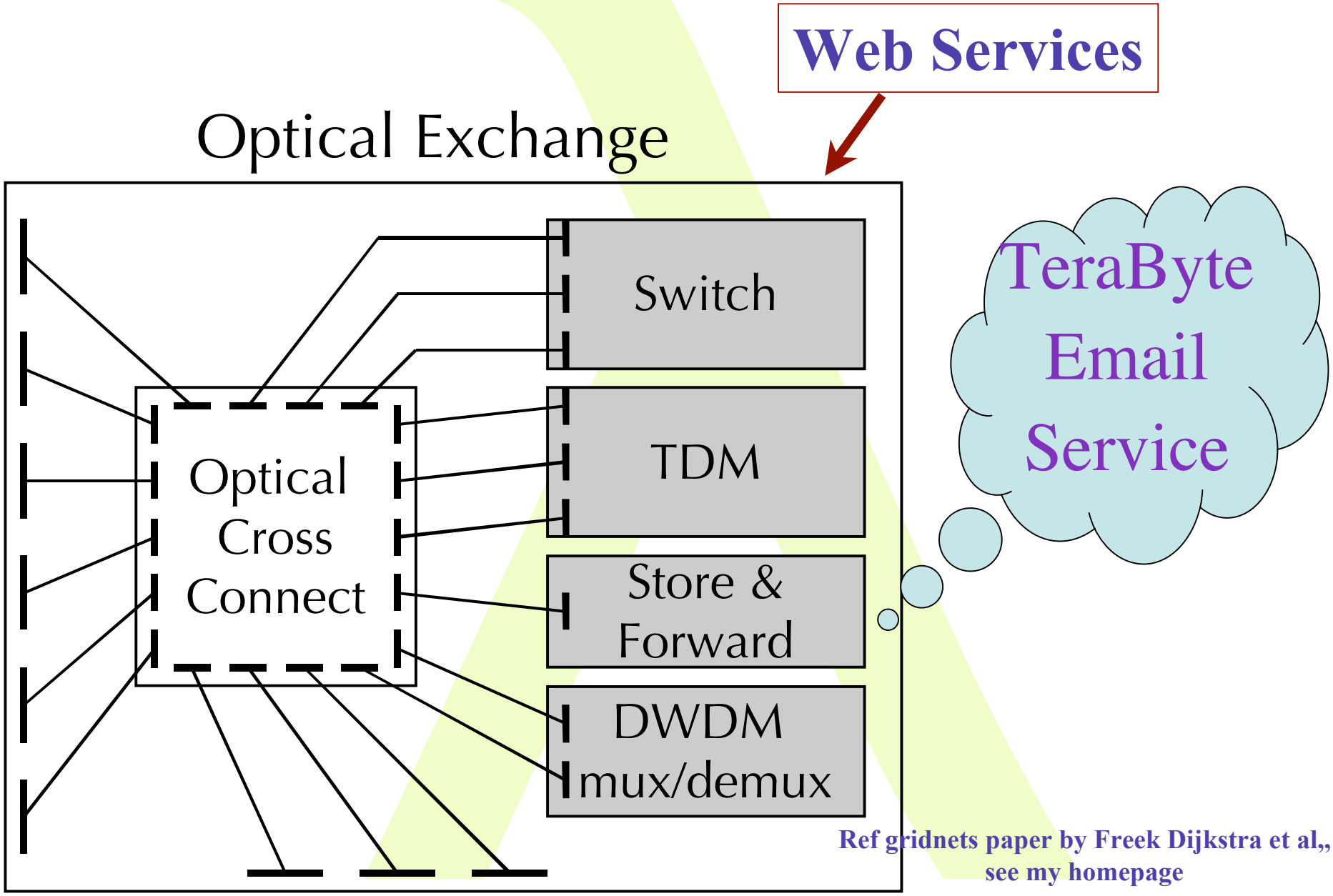
Services

SCALE / CLASS	2 Metro	20 National/ regional	200 World
A	Switching/ routing	Routing	ROUTER$
B	Switches + E-WANPHY VPN's	Switches + E-WANPHY (G)MPLS	ROUTER$
C	dark fiber DWDM MEMS switch	DWDM, TDM / SONET Lambda switching	Lambdas, VLAN's SONET Ethernet

Service Matrix

To / From	WDM (multiple λ)	Single λ, any bitstream	SONET/ SDH	1 Gb/s Ethernet	LAN PHY Ethernet	WAN PHY Ethernet	VLAN tagged Ethernet	IP over Ethernet
WDM (multiple λ)	cross-connect multicast, regenerate, multicast	WDM demux	WDM demux*	WDM demux *	WDM demux *	WDM demux *	WDM demux *	WDM demux *
Single λ, any bitstream	WDM mux	cross-connect multicast, regenerate, multicast	N/A *	N/A *	N/A *	N/A *	N/A *	N/A *
SONET/SDH	WDM mux	N/A *	SONET switch, +	TDM demux *	TDM demux[6]	SONET switch	TDM demux *	TDM demux *
1 Gb/s Ethernet	WDM mux	N/A *	TDM mux	aggregate, Ethernet conversion +	aggregate, eth. convert	aggregate, Ethernet conversion	aggregate, VLAN encap	L3 entry *
LAN PHY Ethernet	WDM mux	N/A*	TDM mux[6]	aggregate, Ethernet conversion	aggregate, Ethernet conversion +	Ethernet conversion	aggregate, VLAN encap	L3 entry *
WAN PHY Ethernet	WDM mux	N/A *	SONET switch	aggregate, Ethernet conversion	Ethernet conversion	aggregate, Ethernet conversion +	aggregate, VLAN encap	L3 entry *
VLAN tagged Ethernet	WDM mux	N/A *	TDM mux	aggregate, VLAN decap	aggregate, VLAN decap	aggregate, VLAN decap	Aggregate, VLAN decap & encap +	N/A
IP over Ethernet	WDM mux	N/A *	TDM mux	L3 exit *	L3 exit *	L3 exit *	N/A	Store & forward, L3 entry/exit+

Optical Exchange as Black Box



Ref gridnets paper by Freek Dijkstra et al,,
see my homepage

Network resources: management and monitoring

Motivation:
Users and applications should be able to:

- **monitor the performance of single network components,**
- **monitor the available resources in a single or multiple domains,**
- **monitor and setup dedicated light paths within an Optical Cross Connect**

through well defined interfaces.

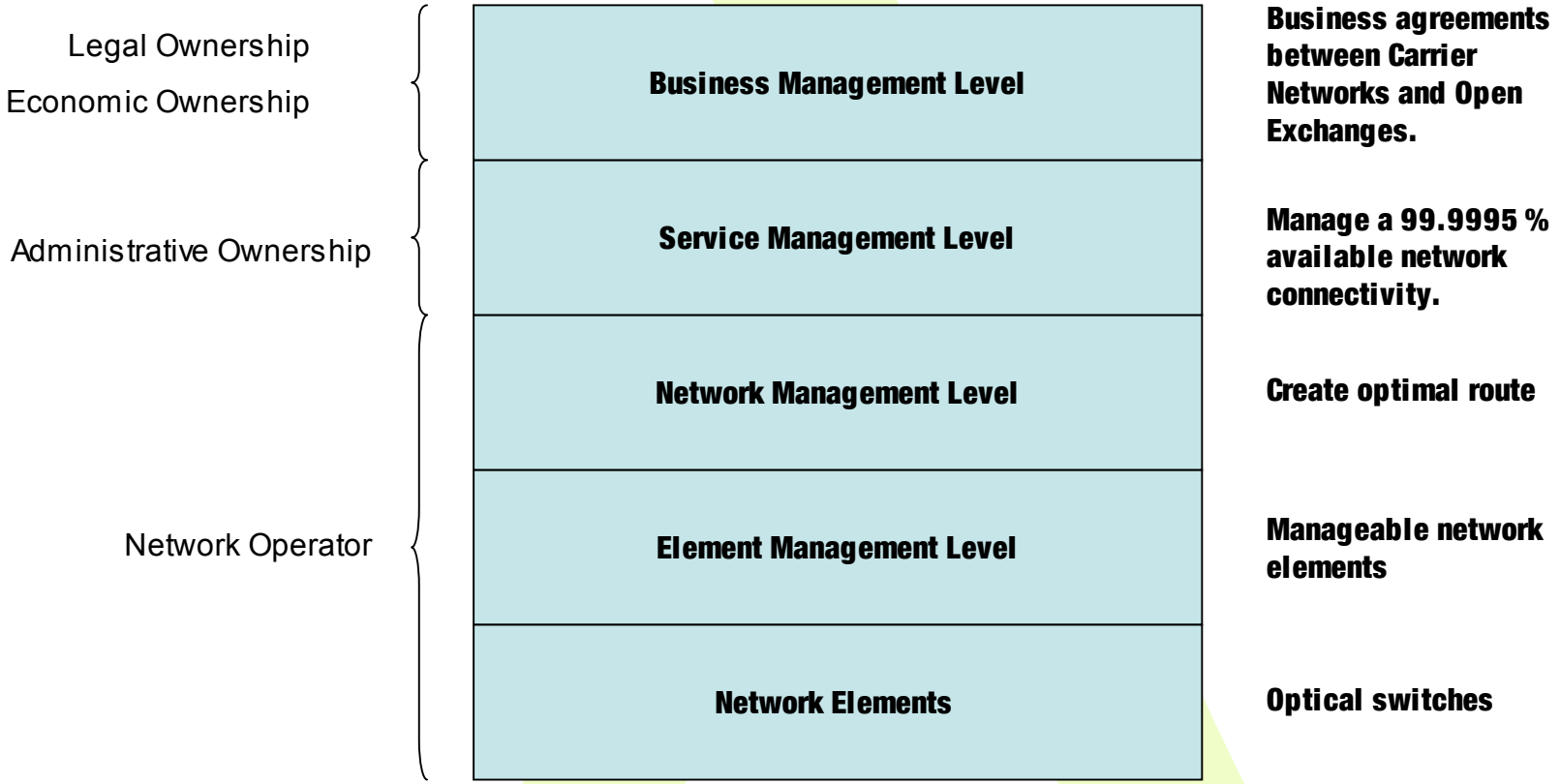
Objective:
Provide access to authorized users and applications to network resources through Web Services.

Current work:

- **definition of models for network components;**
- **definition of models for resource brokers;**
- **publication of available interfaces via WSDL;**
- **implementation of Web Services;**
- **integration with AAA for user authentication and authorization to use the service.**

More information:
http://vangogh0.uva.netherlight.nl/AIRWebServices/doc/NetherLightWS.htm

ISO Telecommunications Management Networks (TMN) reference model

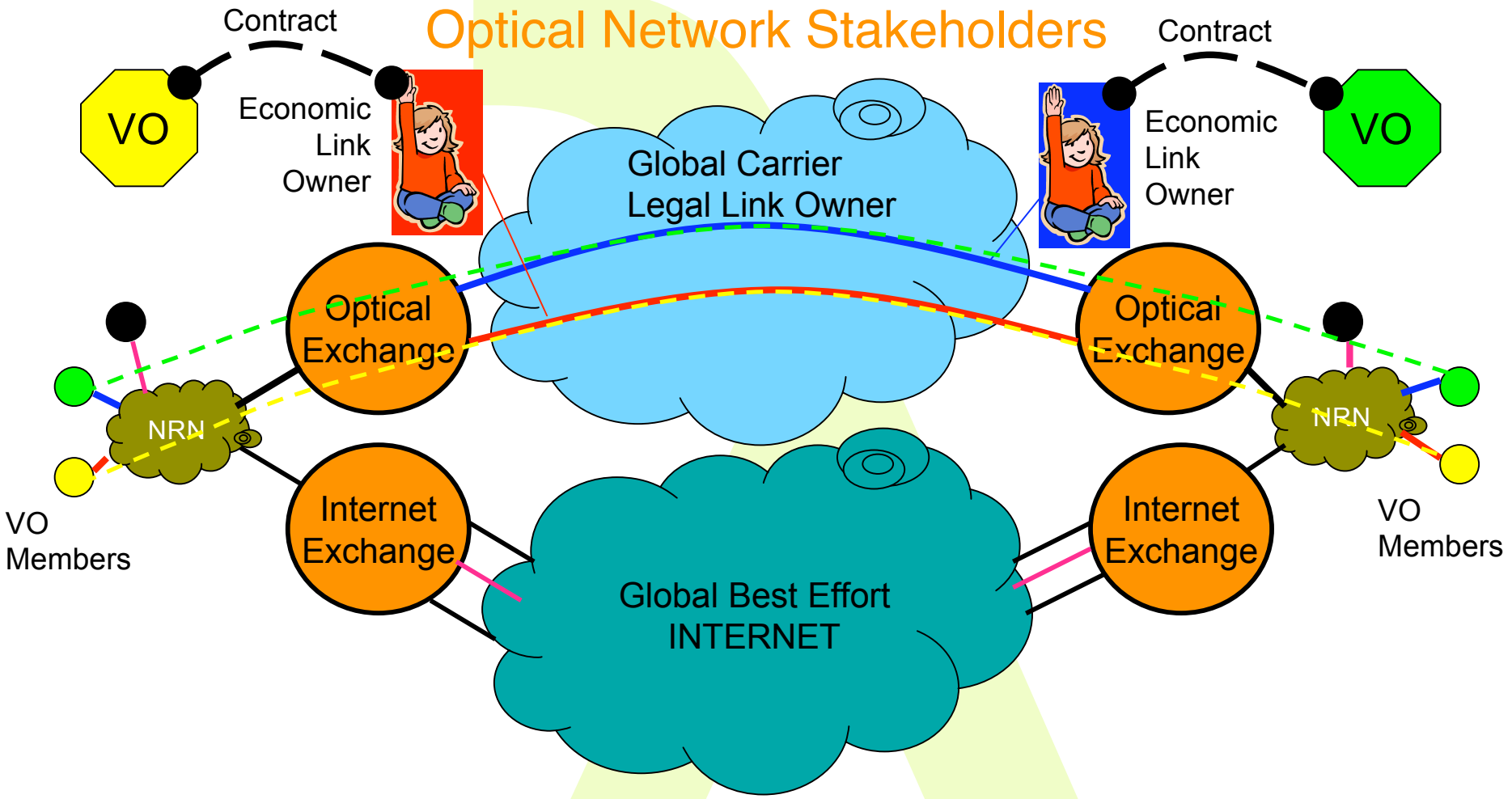


TMN is based on the OSI management framework and uses an object-oriented approach, with managed information in network resources modeled as attributes in managed objects. TMN is defined in ITU-T M.3000 series recommendations

Ownership of resources

- **Legal Owner:**
  - Organization that legally owns a resource.
  - A legal owner may sell the right to economically use the resource.
- **Economic Owner:**
  - Acquires economic resource usage right a from legal resource owner.
  - A contract details terms by which a resource may be used.
  - Economic owners may outsource resource management to an Administrative Owner by means of a service level agreement.
- **Administrative Owner:**
  - Technically implements the terms of a service level agreement
  - Signals requests to other AO's and handles responses.
  - Collects accounting information.
- **Relationship between owners:**
  - Legal, economic and administrative owners may or may not be independent organizations.
  - Economic owners may acquire resources from different legal owners.
  - Administrative owners may serve different economic owners.
  - Economic owners may establish contracts with other economic owners to create more elaborate services. Technical details are delegated and implemented by Administrative Owners.

Optical Network Stakeholders



Problem

In order to enable a dynamic, cost effective VO business operation, Economic Link Owners Red and Blue need to create and have the ability to implement link usage contracts with VO's leading to the creation of **Optical Private Network (OPN)** between VO members.

Role definitions

- **Legal Link Owner (LLO):** Sells the right to use a link to an ELO's

- **Economic Link Owner (ELO):** Acquires the right to use a link and creates agreements with Economic VO's about the usage of its links.ELO's will terminate a link at an optical exchange based on a contract with an EPO.

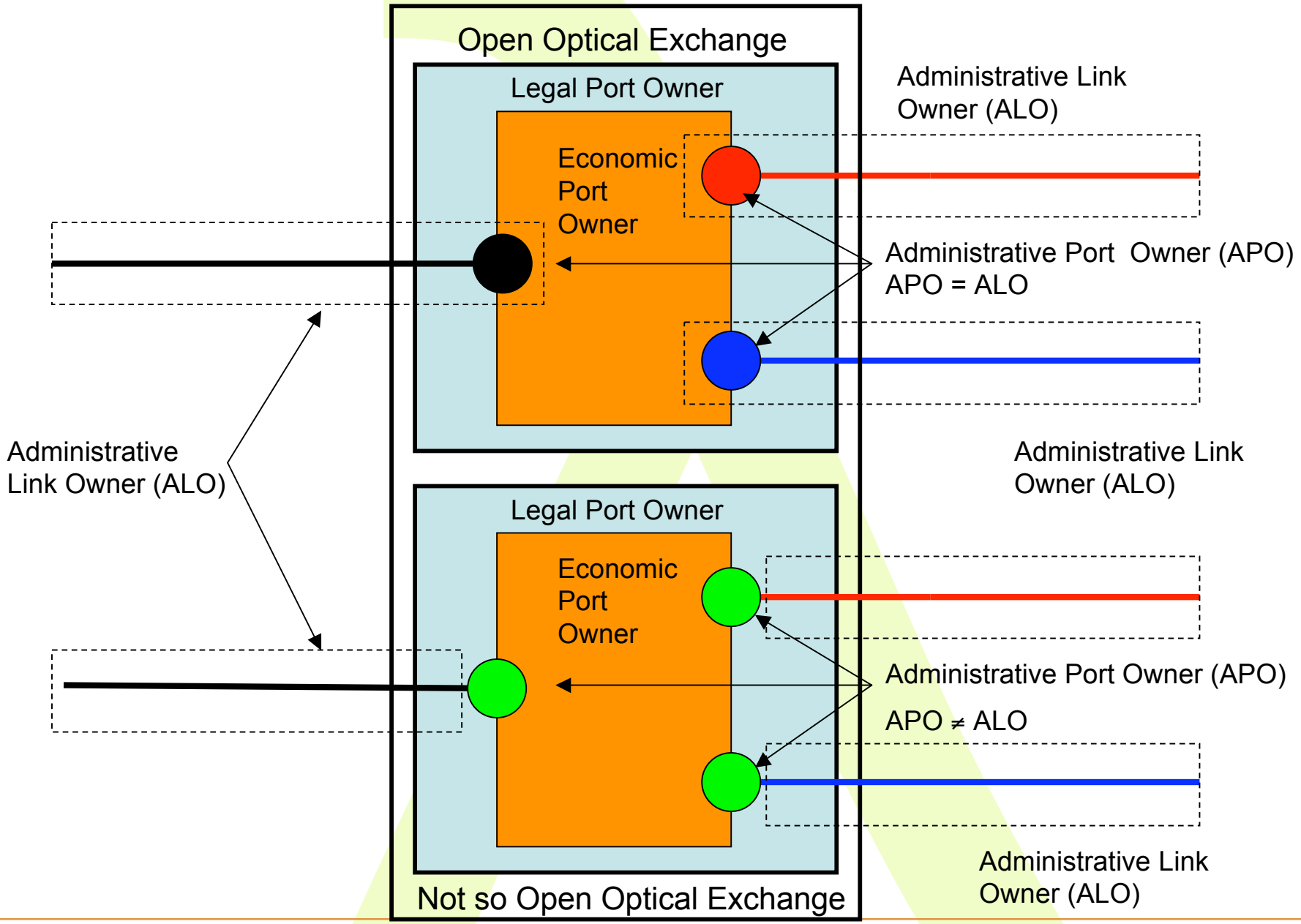
- **Administrative Link Owner (ALO):** Translates the ELO defined business rules governing link access to technical rules that are subsequently pushed to the APO for enforcement (optical link fibers have no electronic control).

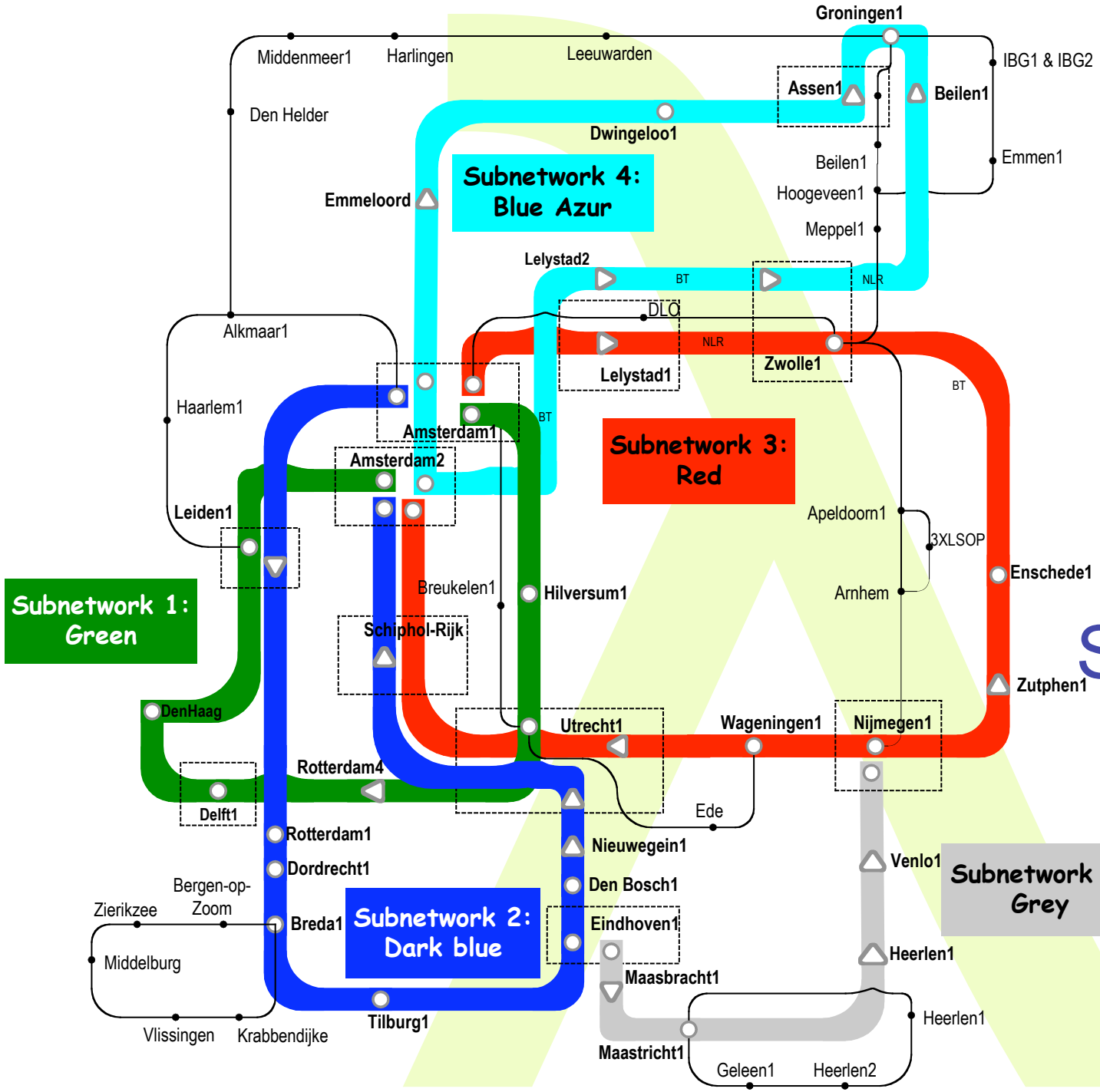
- **Legal Port Owner (LPO):** Owns optical switch-ports. Usage rights are sold to EPO's. Multiple LPO's may be present within an Optical Exchange.

- **Economic Port Owner (EPO):** Acquires the usage right from one or more LPO's for one or more ports on the Optical Exchange. EPO's establishes contracts to allow peering with own or other EPO ports on behalf of ELO's.

- **Administrative Port Owner (APO):** an entity that accepts peering policies from ALO's. Peering policies are based on the agreements between ELO and a VO. Creates connections with own ports or other ports from different APO's based on requests with credentials from VO's members or its proxy .

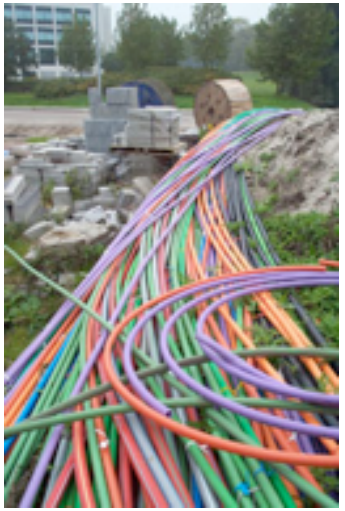
Optical Exchange Stakeholders





Common Photonic Layer (CPL) in SURFnet6

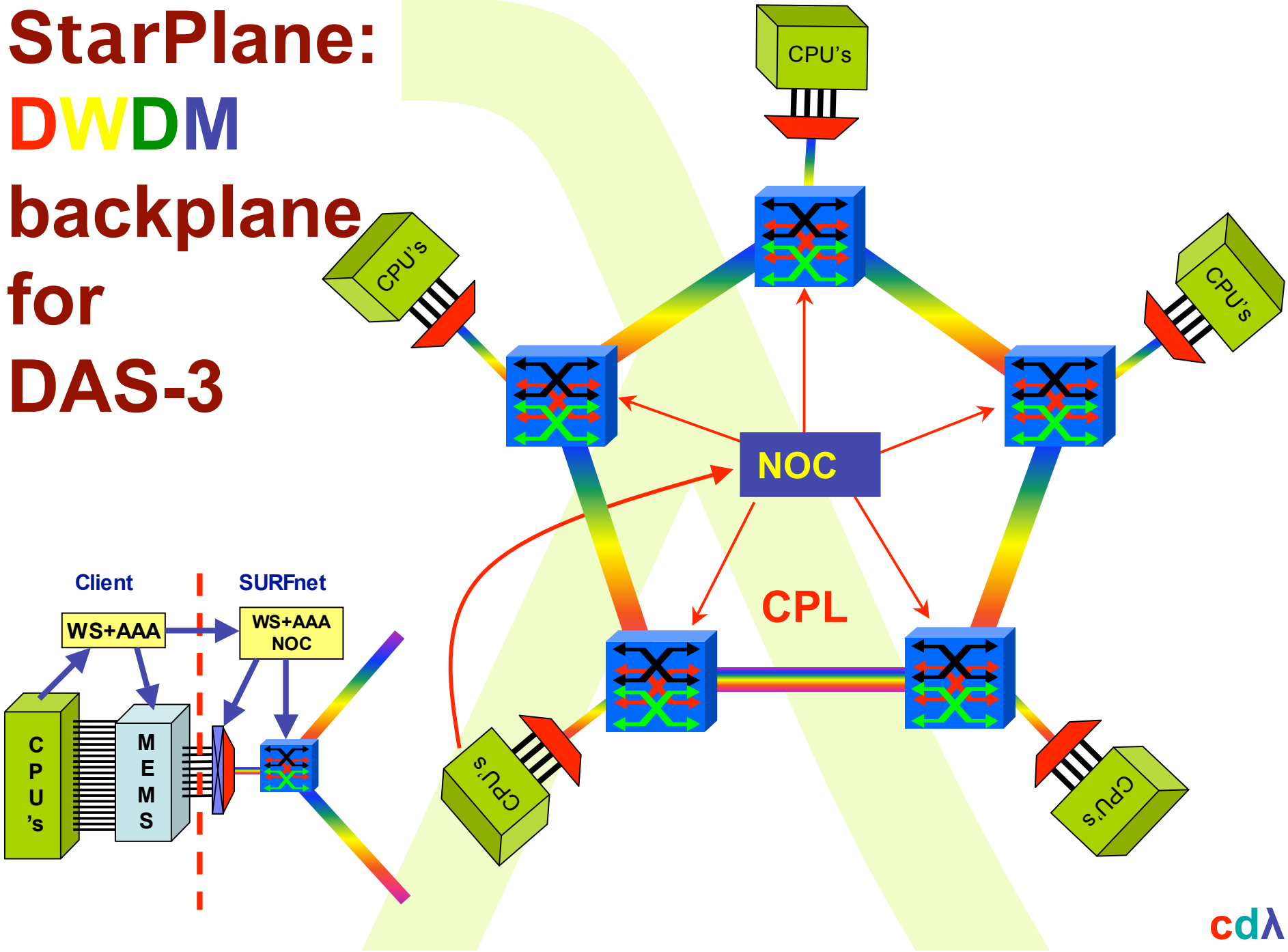
Laying of fiber near/at Science Park Amsterdam



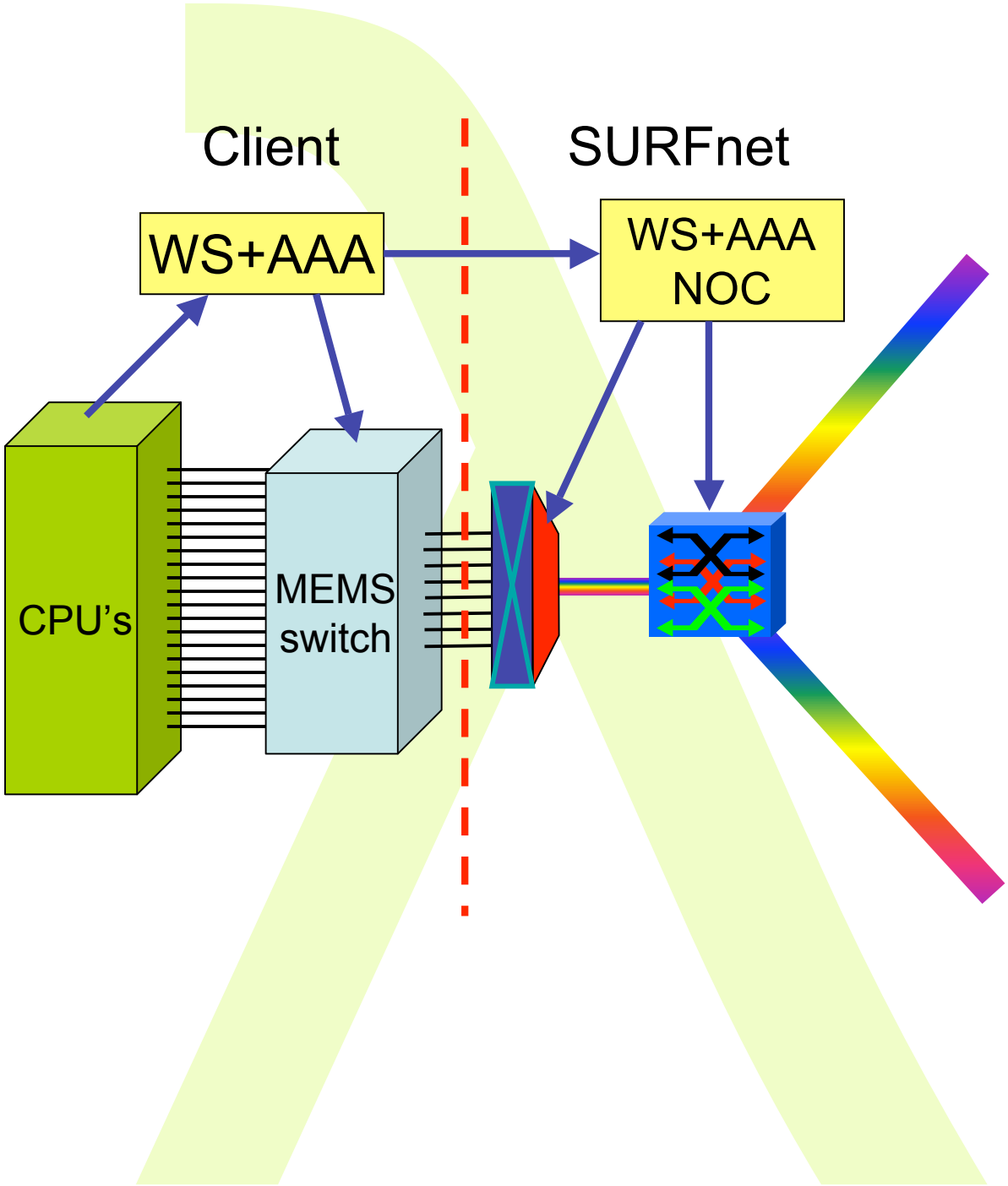
Pictures by Yuri Demchenko


SURFnet on Lambda inspection in Science Park Amsterdam :-)

StarPlane: DWDM backplane for DAS-3

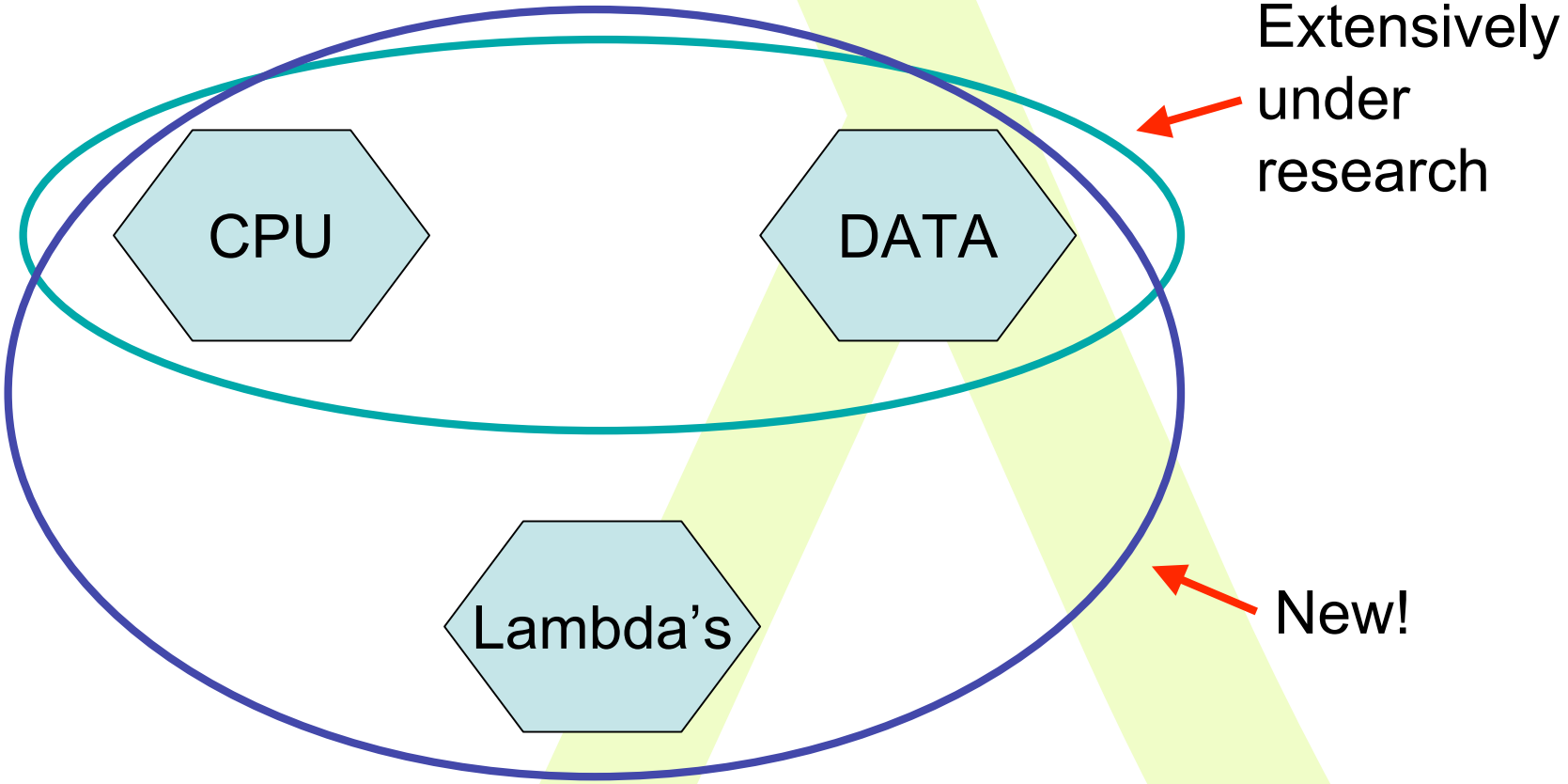


cdλ



CdL

GRID-Colocation problem space



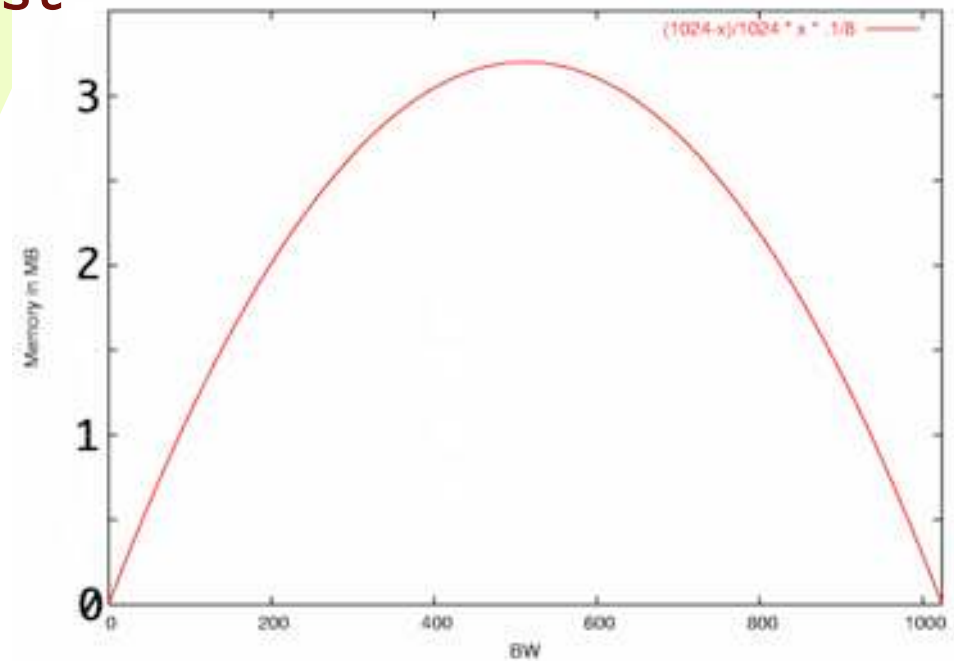
Layer - 2 requirements from 3/4

TCP is bursty due to sliding window protocol and slow start algorithm.

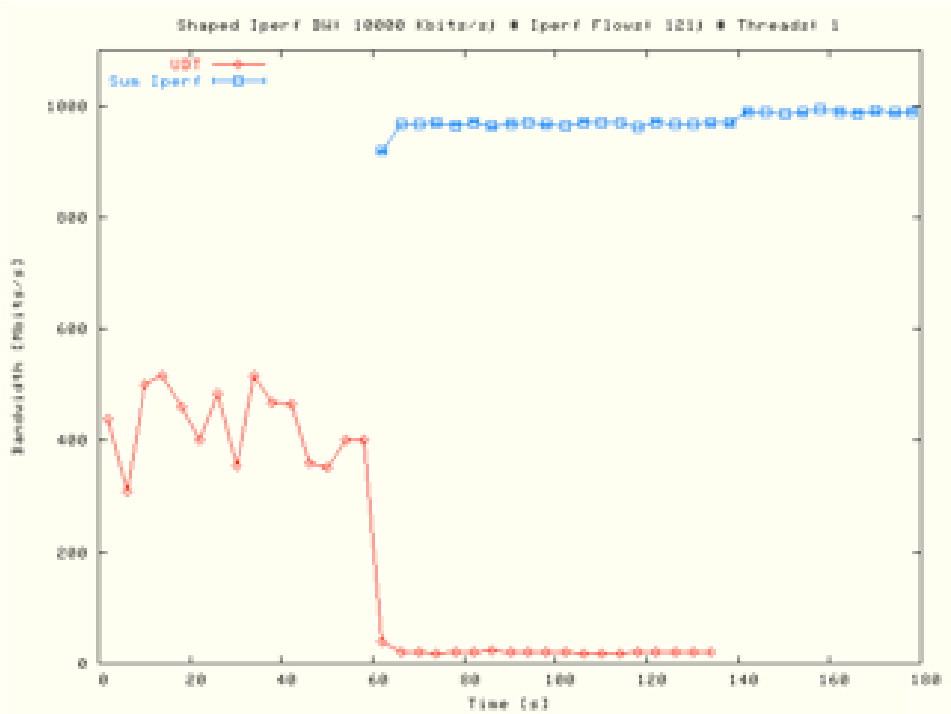
```
Window = BandWidth * RTT     &    BW == slow

                                fast - slow
Memory-at-bottleneck = ------------ * slow * RTT
                                  fast
```

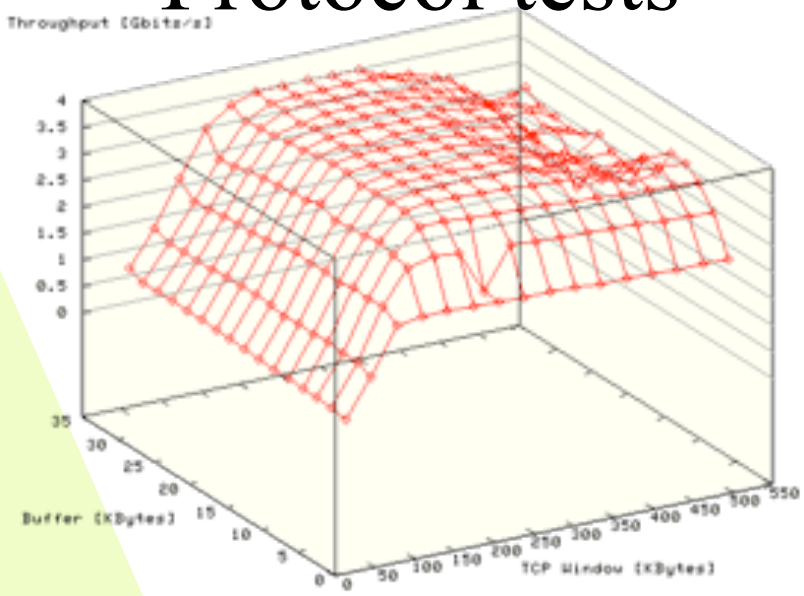
So pick from menu:

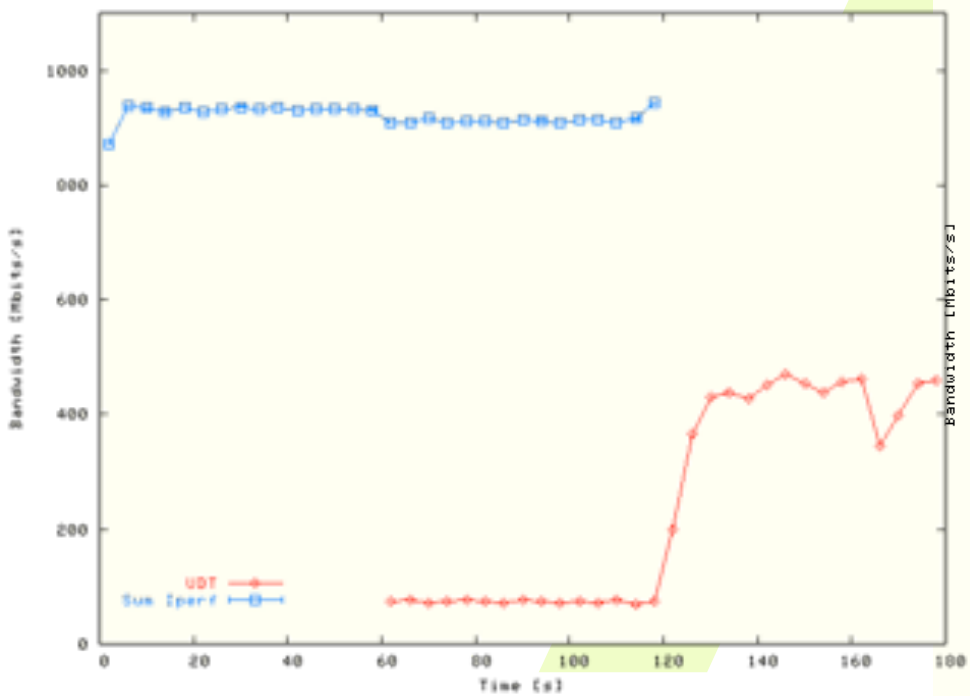
- *Flow control*
- *Traffic Shaping*
- *RED (Random Early Discard)*
- *Self clocking in TCP*
- *Deep memory*

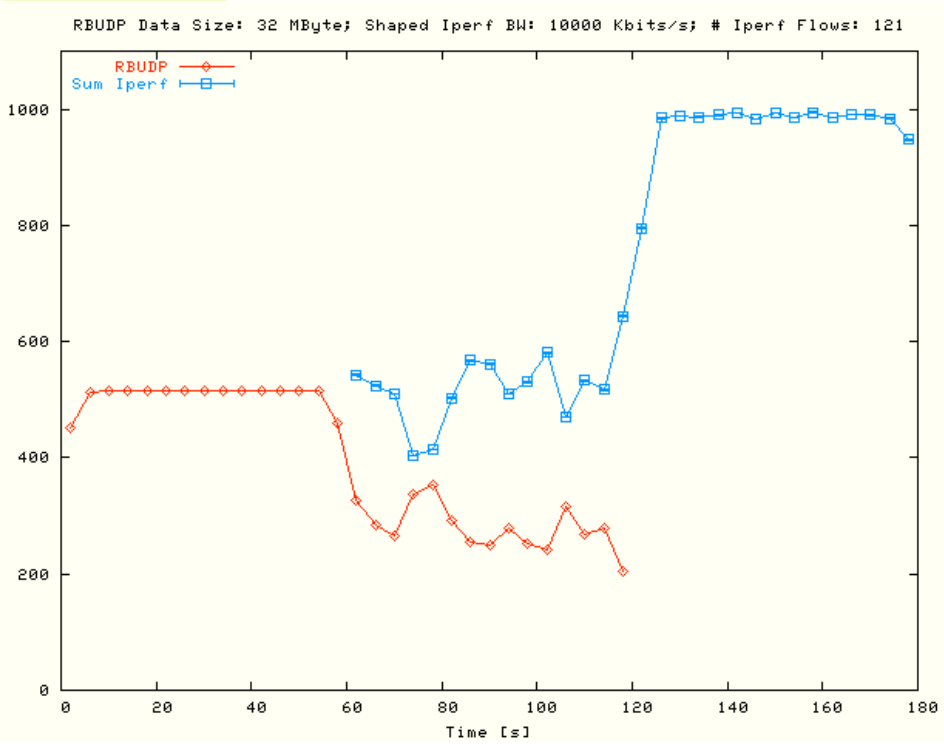


Protocol tests









Grid and network tests

Motivation:

As more and more Grids are being built and deployed we expect that in some cases network tests and measurements will have to be conducted on such infrastructures.

Objective:

- **to determine if and how Grids are suitable for network tests**
- **to defining the requirements for the applications and the resources available through the Grid.**

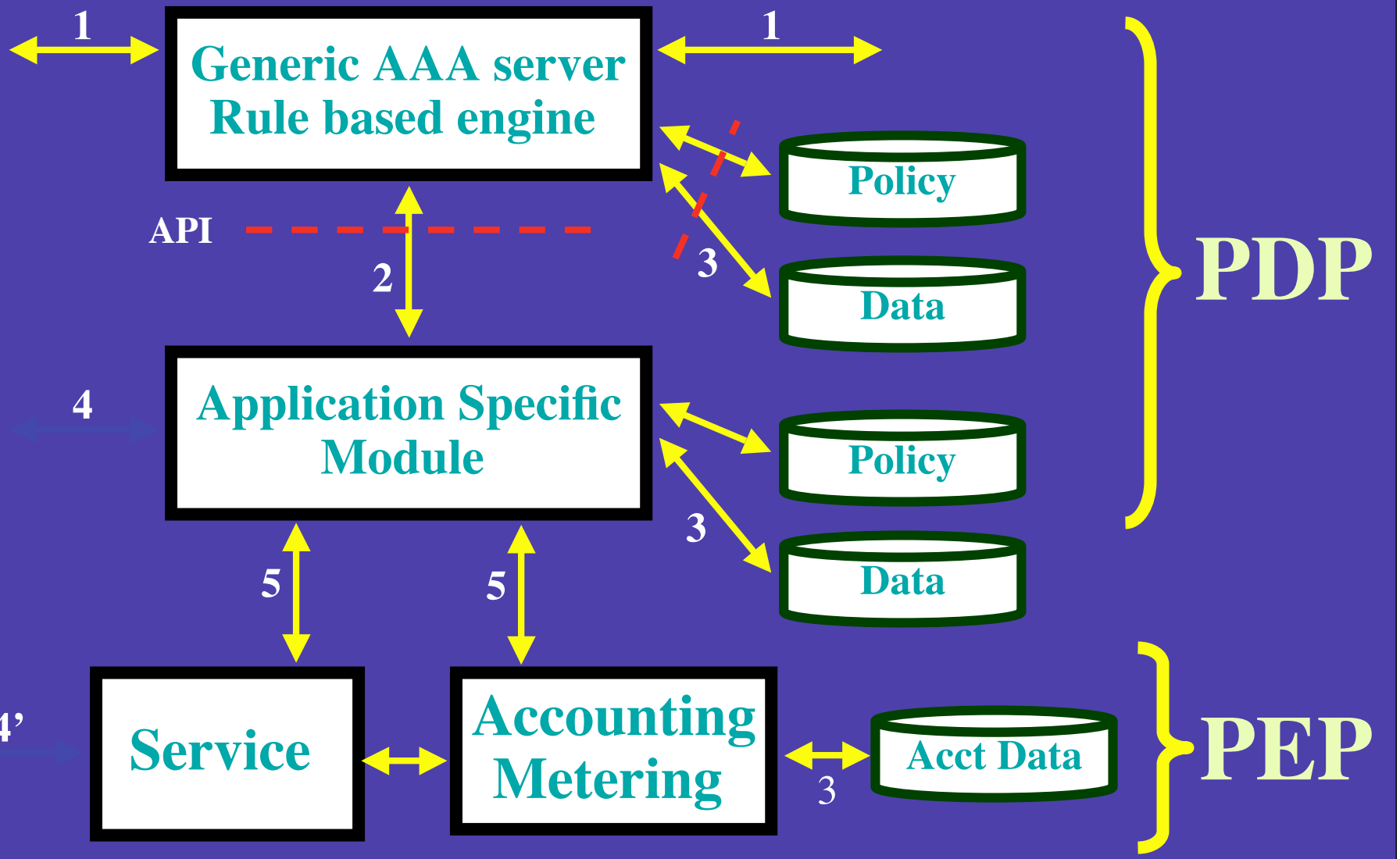
Current work:

- **deployment of standard test tools on Grids**
- **evaluation of Grid tools as network test tools (i.e GridFTP)**
- **design of measurement infrastructure**
- **implementation on the DAS-2 cluster, with Globus and MPICH-G2**
- **ongoing analysis of test results.**

More information:

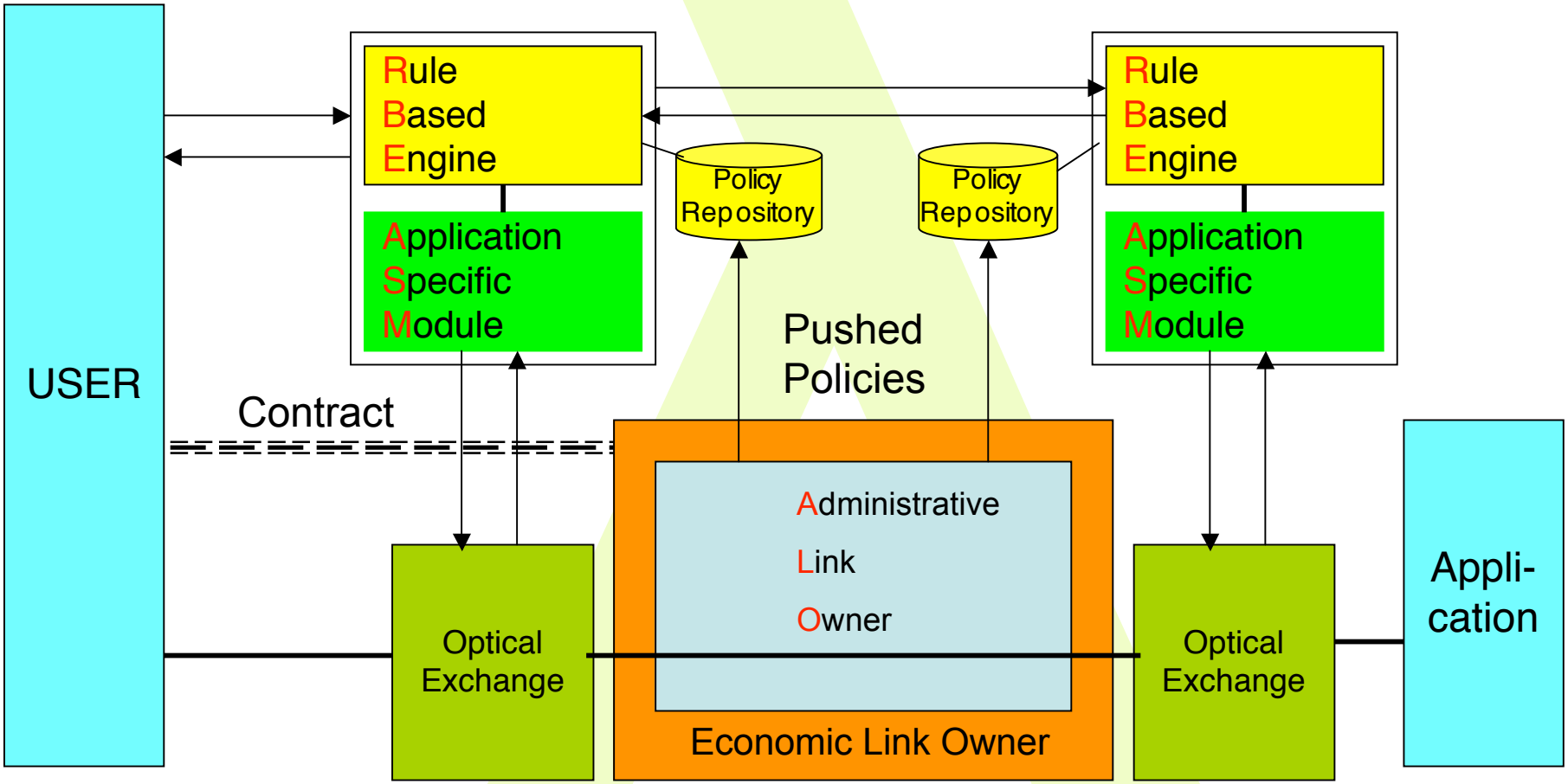
http://vangogh0.uva.netherlight.nl/GridFTP-tests/Intro.php

Starting point AAA

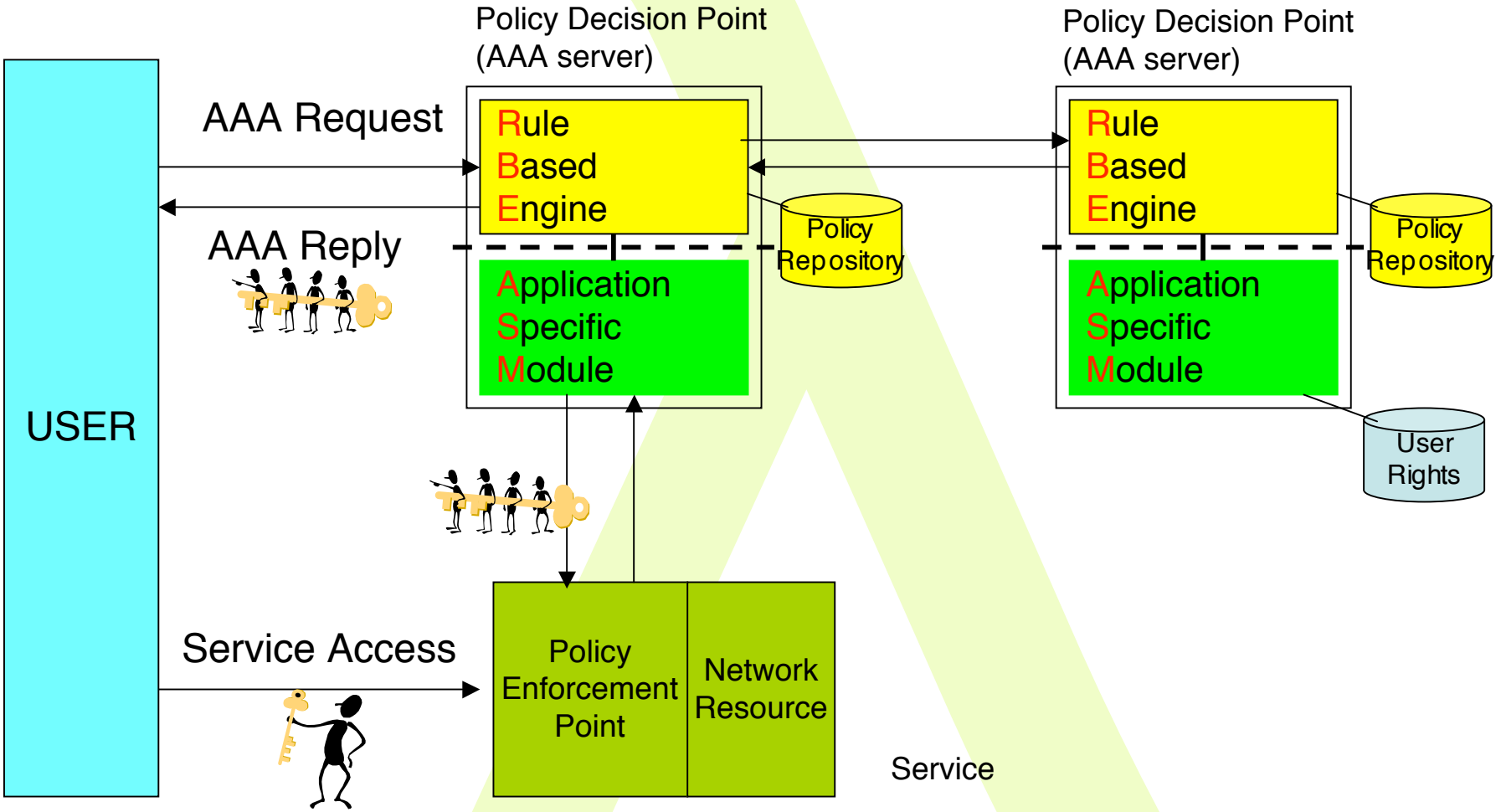


RFC 2903 - 2906 , 3334 , policy draft

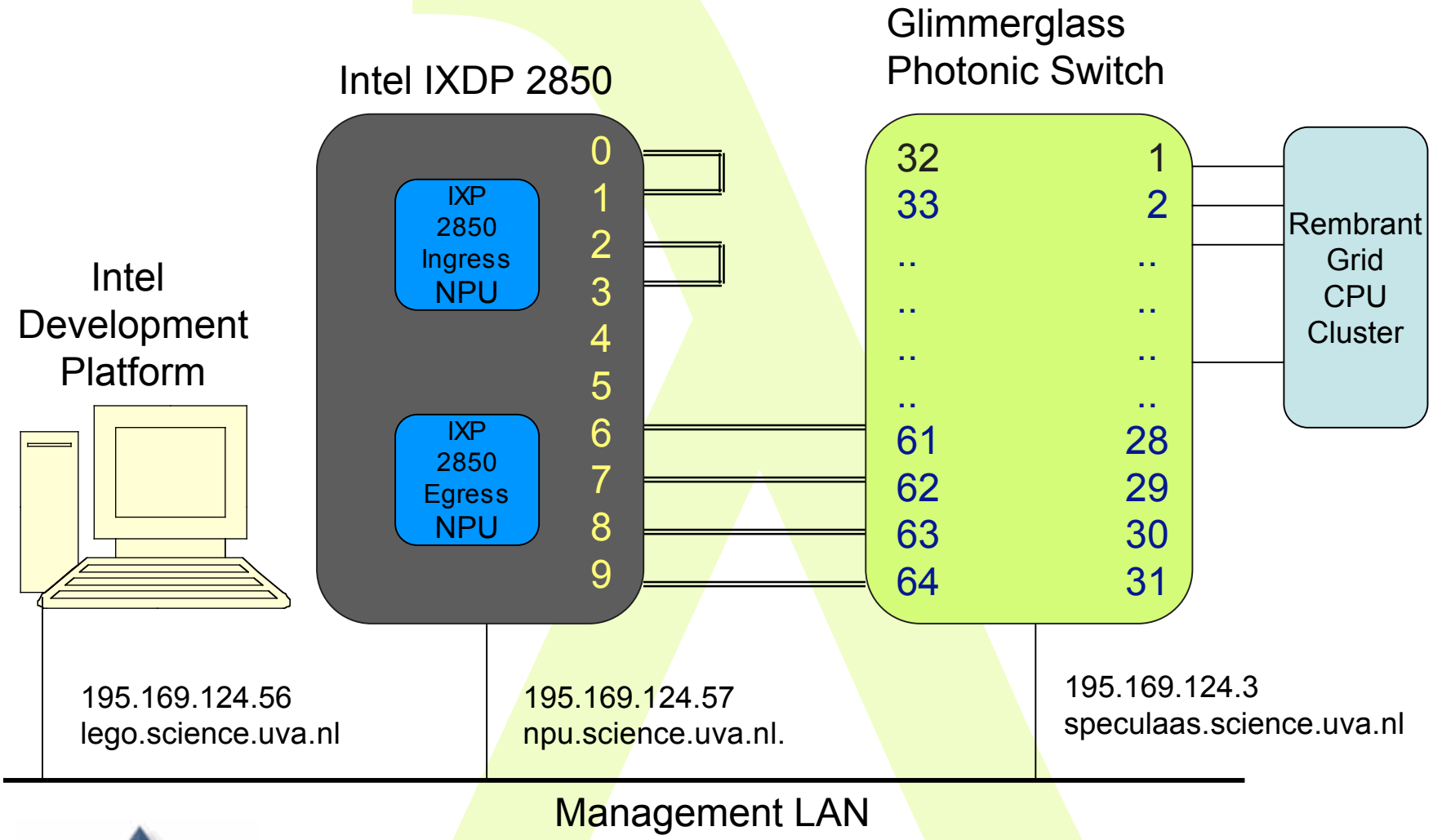
Optical Exchange Control



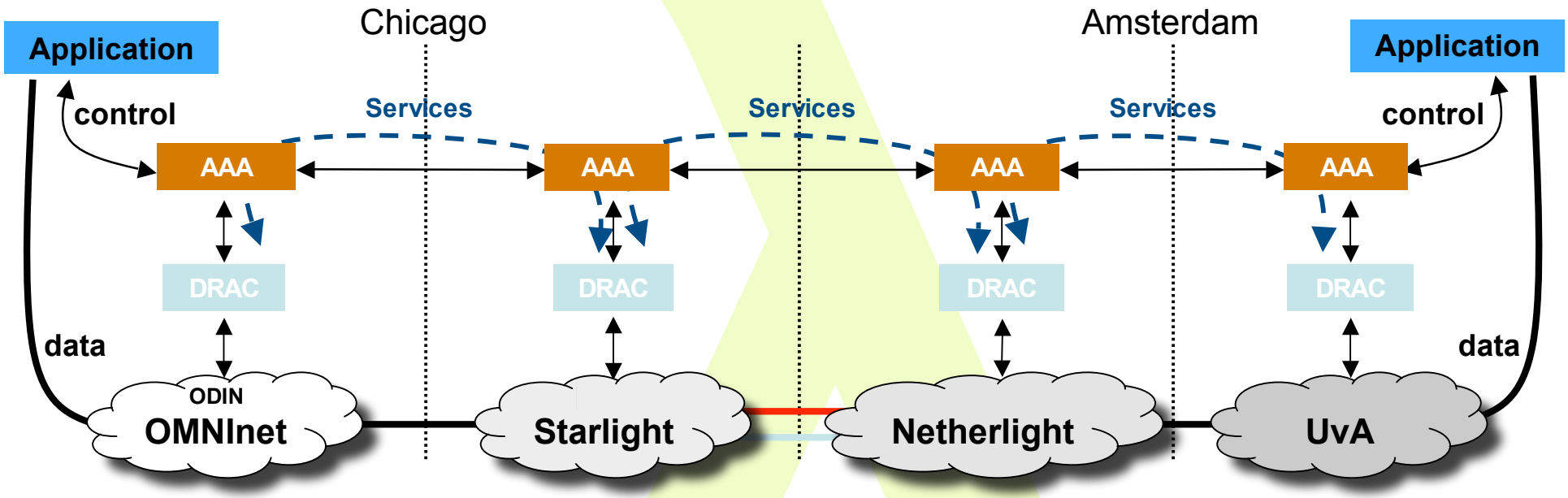
Token based networking



Configuration Intel NPU experiment.



SC2004 CONTROL CHALLENGE

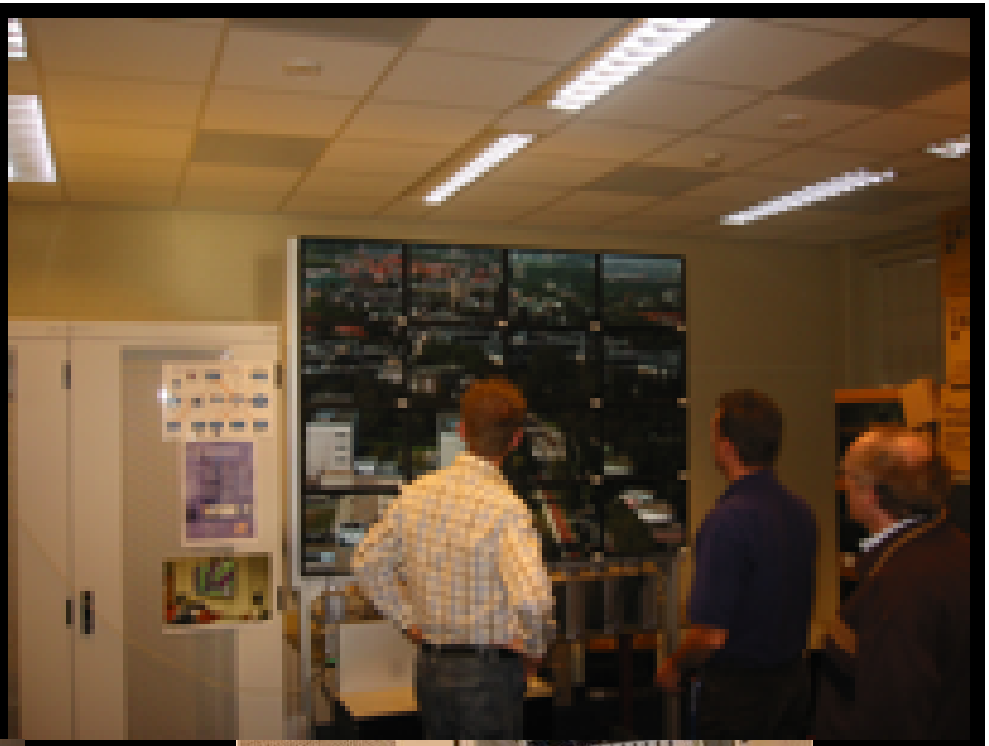


- finesse the control of bandwidth across multiple domains
- while exploiting scalability and intra- , inter-domain fault recovery
- thru layering of a novel SOA upon legacy control planes and NEs


GOT CONTROL TOO?
Visit Booth #1333

Highlights Generic AAA work

- Development of Generic AAA Toolkit V1.0
- Toolkit demo available via web.
- Hooks into EGEE LCAS/LCMAPS and Globus V4.0
- Deployment GAAA toolkit in:
  - Optical Exchange Control plane
  - Collaboratory.nl for job-centric access control of devices.
  - Chaining network control planes (UCLP, Nortel DRAC).
  - Token based networking.
  - Gridftp via EGEE.



SC2004 “Dead Cat” demo

SuperComputing 2004,
Pittsburgh,
Nov. 6 to 12, 2004

Produced by:
Michael Scarpa
Robert Belleman
Peter Sloot

Many thanks to:
AMC
SARA
GigaPort
UvA/AIR
Silicon Graphics, Inc.
Zoölogisch Museum

Some Thoughts

- Photonic, SONET/SDH, Ethernet, GMPLS, VLAN, Routed networks
- Optical/Photonic Exchanges, Grid resources, sensor grids
- Workflow support
- Monitoring and testing
- Different scales (national, continental, trans oceanic)
- Various set-up / tear down times [days - subseconds]
- Multiple administrative domains
- Transport dependency on properties of lower layers
- USER FRIENDLY !!!!!!

Not The END

Thanks to
SURFnet: Kees Neggers,UIC&iCAIR: Tom DeFanti, Joel Mambretti, CANARIE: Bill St. Arnaud
Freek Dijkstra, Hans Blom, Leon Gommans, Bas van oudenaarde, Arie Taal, Pieter de Boer, Bert Andree, Martijn de Munnik, Antony Antony, Rob Meijer, VL-team.

Partially complete list:

- Caas
- Chase
- Cess
- Kess
- Case