Lambda Networking Research

www.science.uva.nl/~delaat

www.science.uva.nl/research/air

Cees de Laat

EU

SURFnet

University of Amsterdam

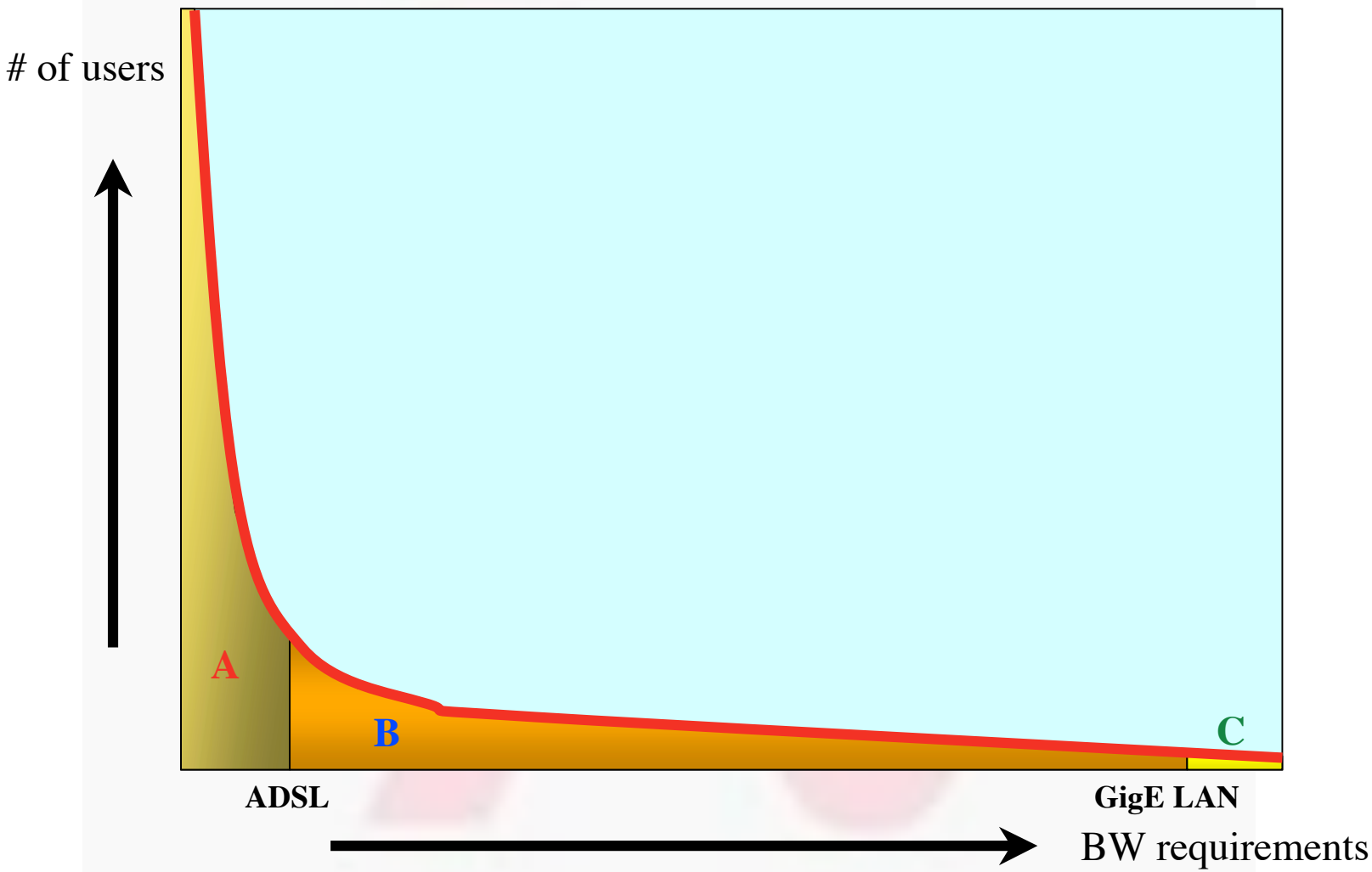
SARA

NikHef

Contents of this talk

- This space is intentionally left blank

Know the user



A -> Lightweight users, browsing, mailing, home use

B -> Business applications, multicast, streaming

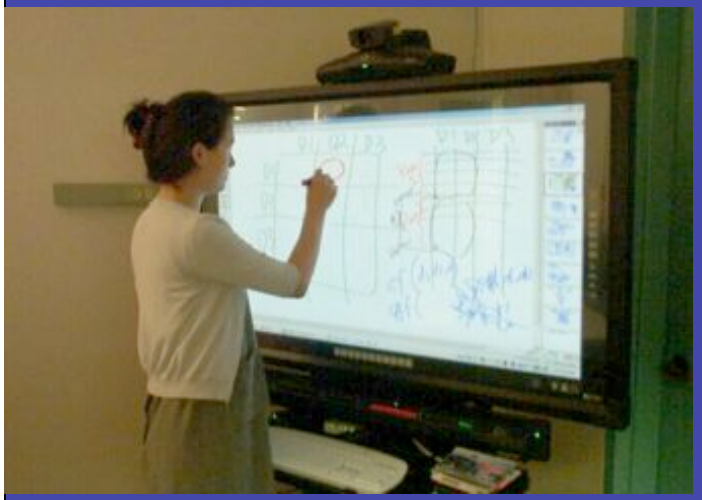
C -> Special scientific applications, computing, data grids, virtual-presence

Integrating Distributed Collaborative Visualization


CAVE


AccessGrid

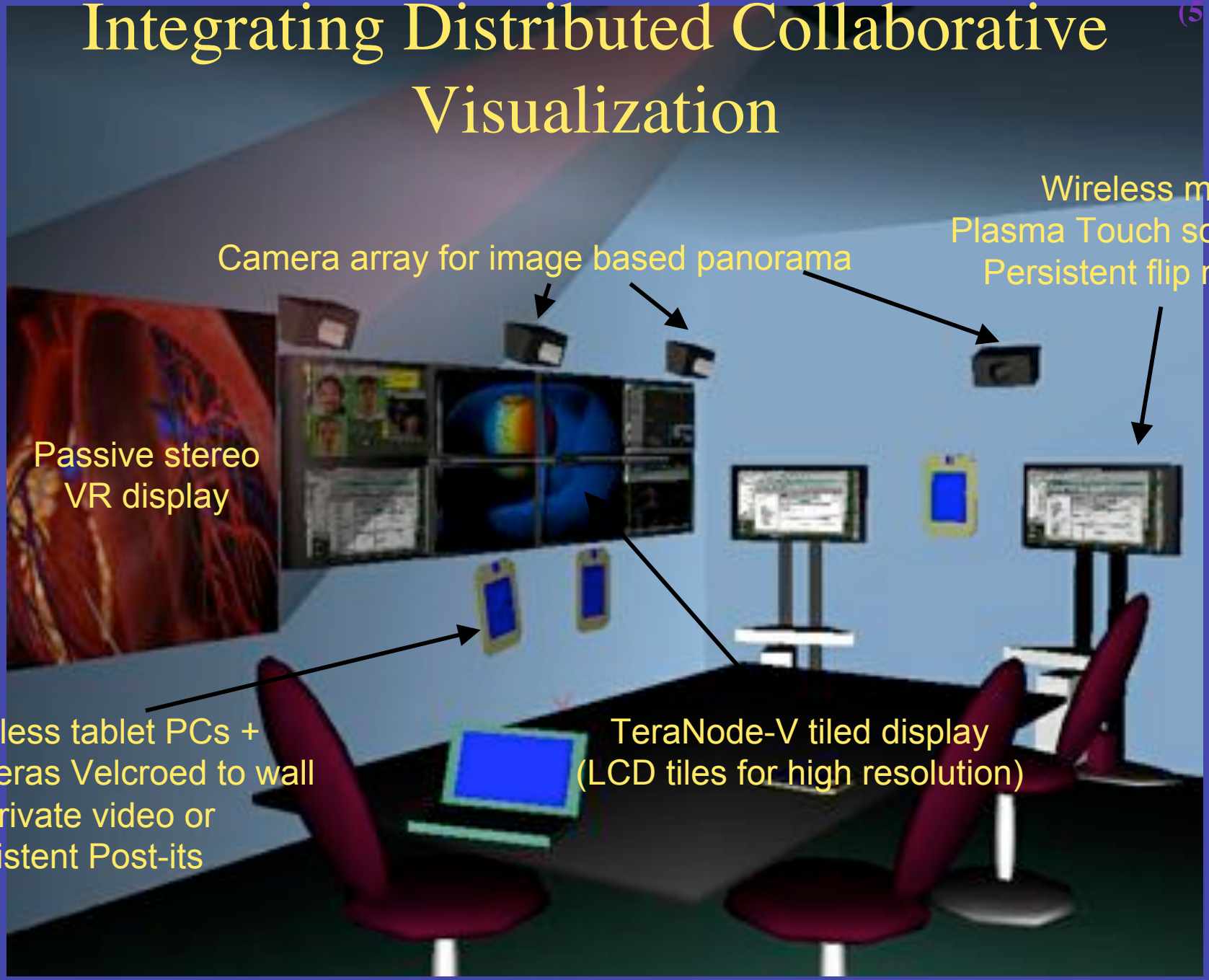
ImmersaDesk


Plasma Touch Screen


AGAVE: Passive Stereo Wall


PDAs, Tablet PCs, Laptops

Integrating Distributed Collaborative Visualization



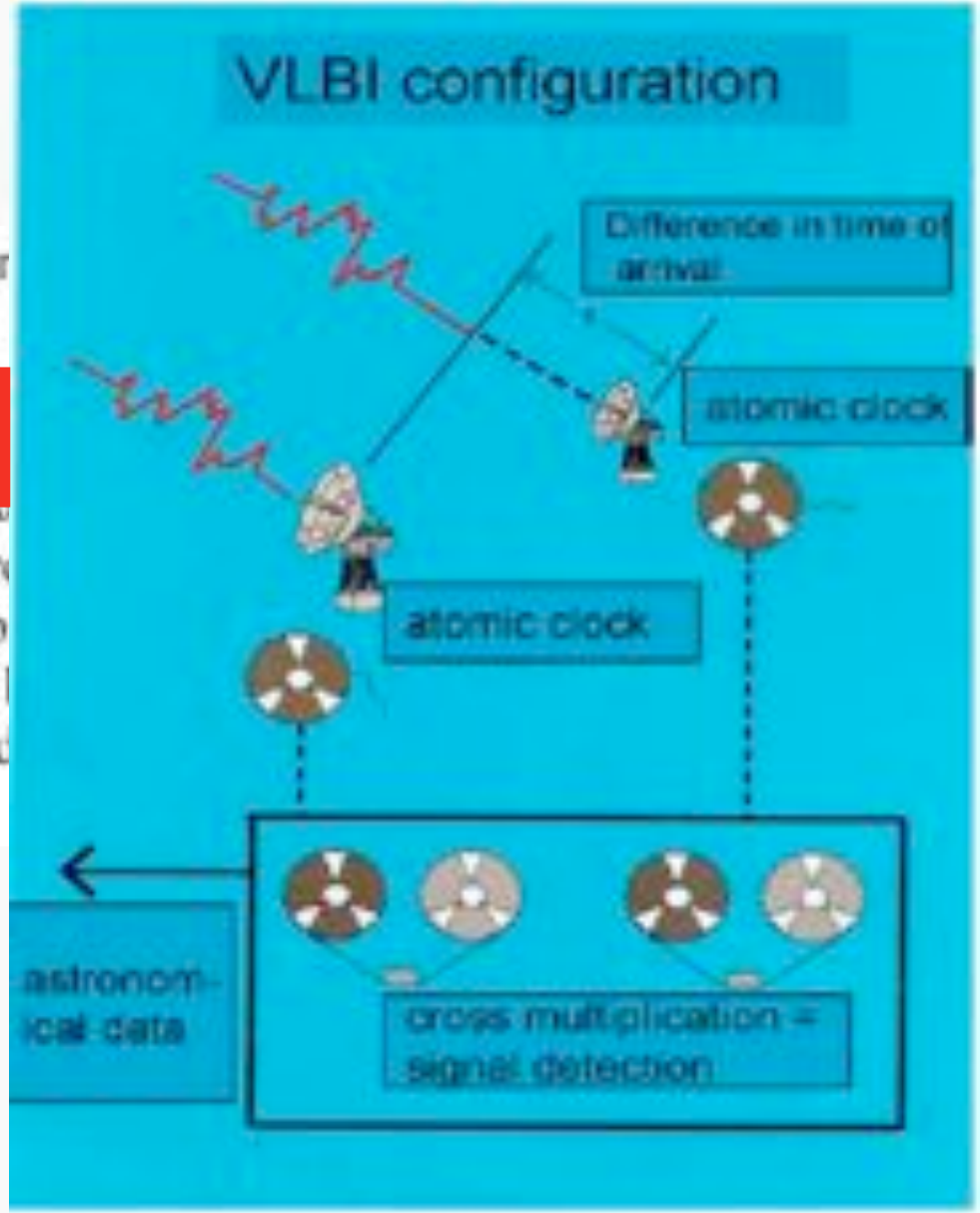
The pilot test described here aims to create a single baseline, real-time, radio interferometer between Jodrell Bank in the UK and Westerbork in the north of The Netherlands. The JIVE data processor in Dwingeloo, close to Westerbork, will be used to correlate the data.

Jodrell Bank Lovell Telescope - UK

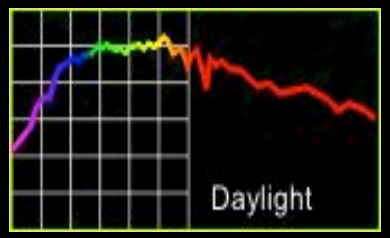
Westerbork Synthesis Radio Telescope - Netherlands

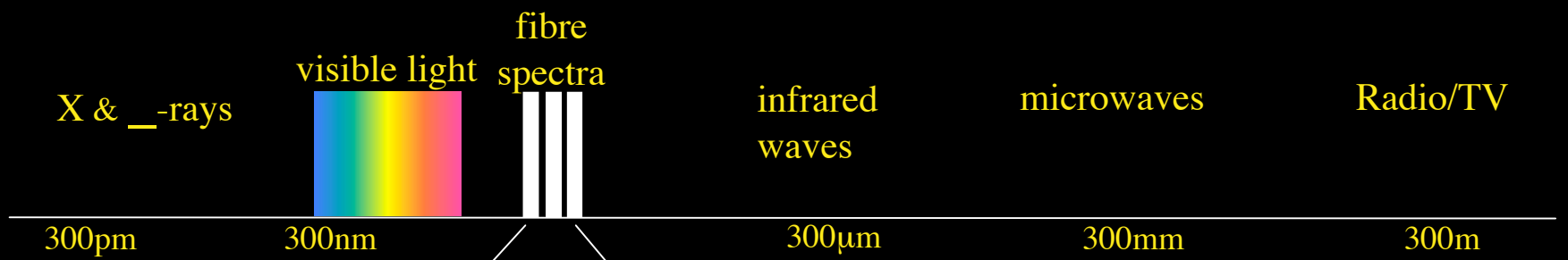
VLBI

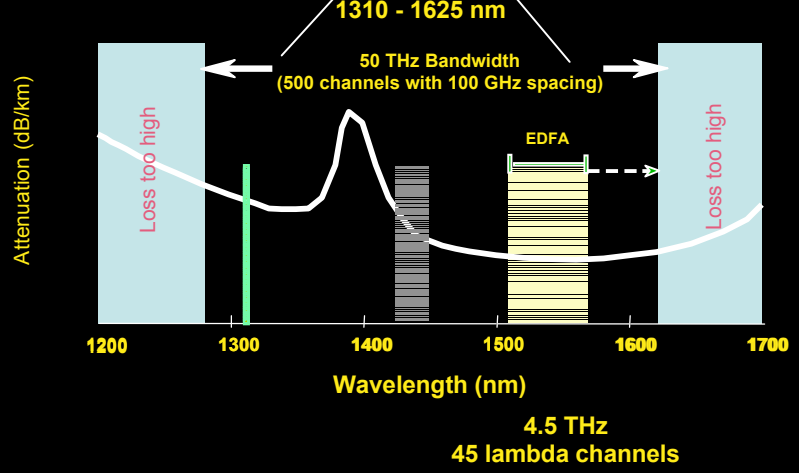
ger term VLBI is easily capable of gener
The sensitivity of the VLBI array scales
(~data-rate) and there is a strong push to
. Rates of 8Gb/s or more are entirely feasi
der development. It is expected that par
orrelator will remain the most efficient appr
s distributed processing may have an app
lti-gigabit data streams will aggregate into
or and the capacity of the final link to the
tor.



Optical bandwidth: room to grow







Optical transport windows:

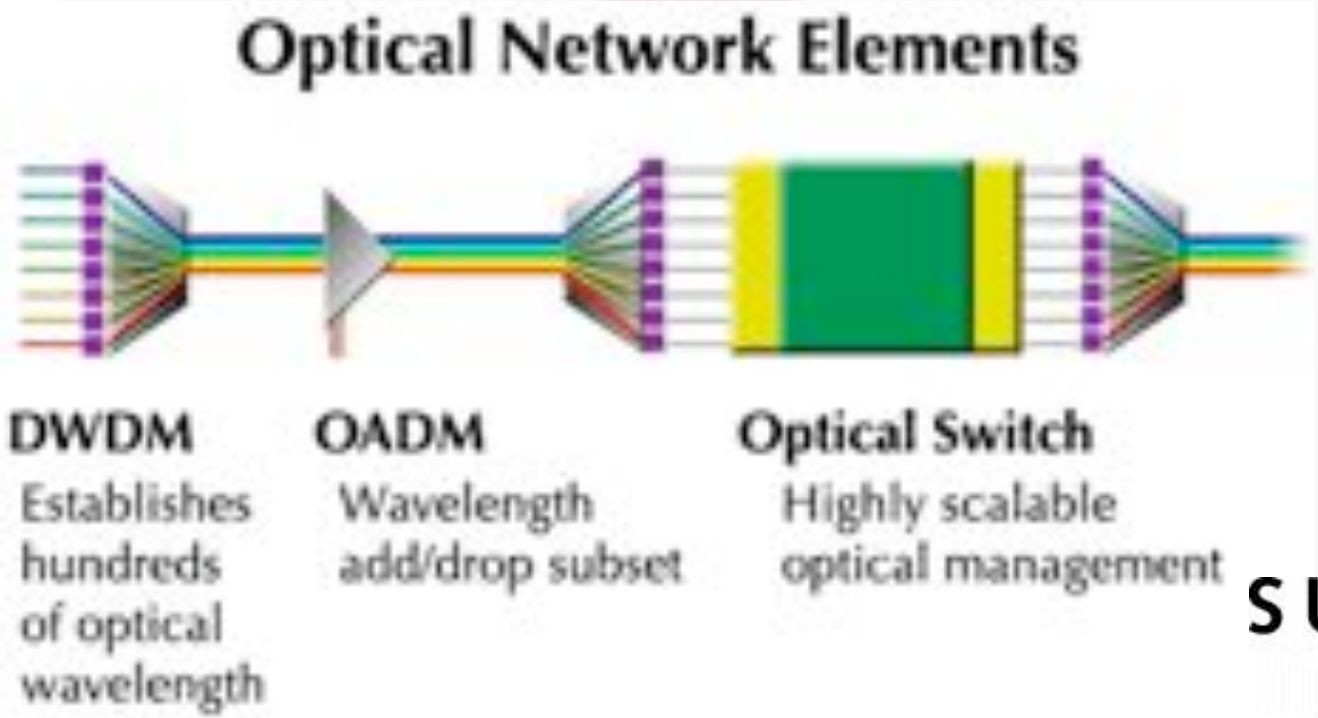
1: 850nm	short reach
2: 1310nm	**intermediate reach**
3a: 1550nm	**long reach**
	1530-1565: C-band
3b: 1600nm	**1565-1620: L-band**

Source: Gert Nieveld, Global Crossing

SURFnet

Lambda networking

- WDM: Wavelength Division Multiplexing:
  - multiple colors (*lambdas*) on a single fiber
- OADM: Add/drop traffic in optical domain
- OXC: Optical Cross-Connect



SURFnet

So, what's up doc

Suppose:

- **Optical components get cheaper and cheaper**
- **Dark (well, dark?) fibers abundant**
- **Number of available λ/user -> ∞**
- **Speeds of 10, 100, 1000 Gbit/s make electrical domain packet handling physically difficult**

Then:

- **λ provisioning for grid applications becomes feasible**
- **Long term view ---> full optical**

Why optical networking and IP ?

- Well established IP world of applications
- Provides worldwide addressing scheme, DNS, URL's, Routing, etc.
- Optical networks are supposed to bring speed
- Only Lambda's may look like a telephone system
- How to marry both worlds ?

Connection less versus connection oriented

- Connection less
  - postal office
  - mail
  - internet (IP)
  - datagram delivery
- Connection oriented
  - telephone system
  - 3 phases: establish, use, release
  - order preserved
  - file transfer
  - waste of resources
  - TCP
- role can change in each layer

Example of connection less communication (int)

Optical networking, 3 scenarios

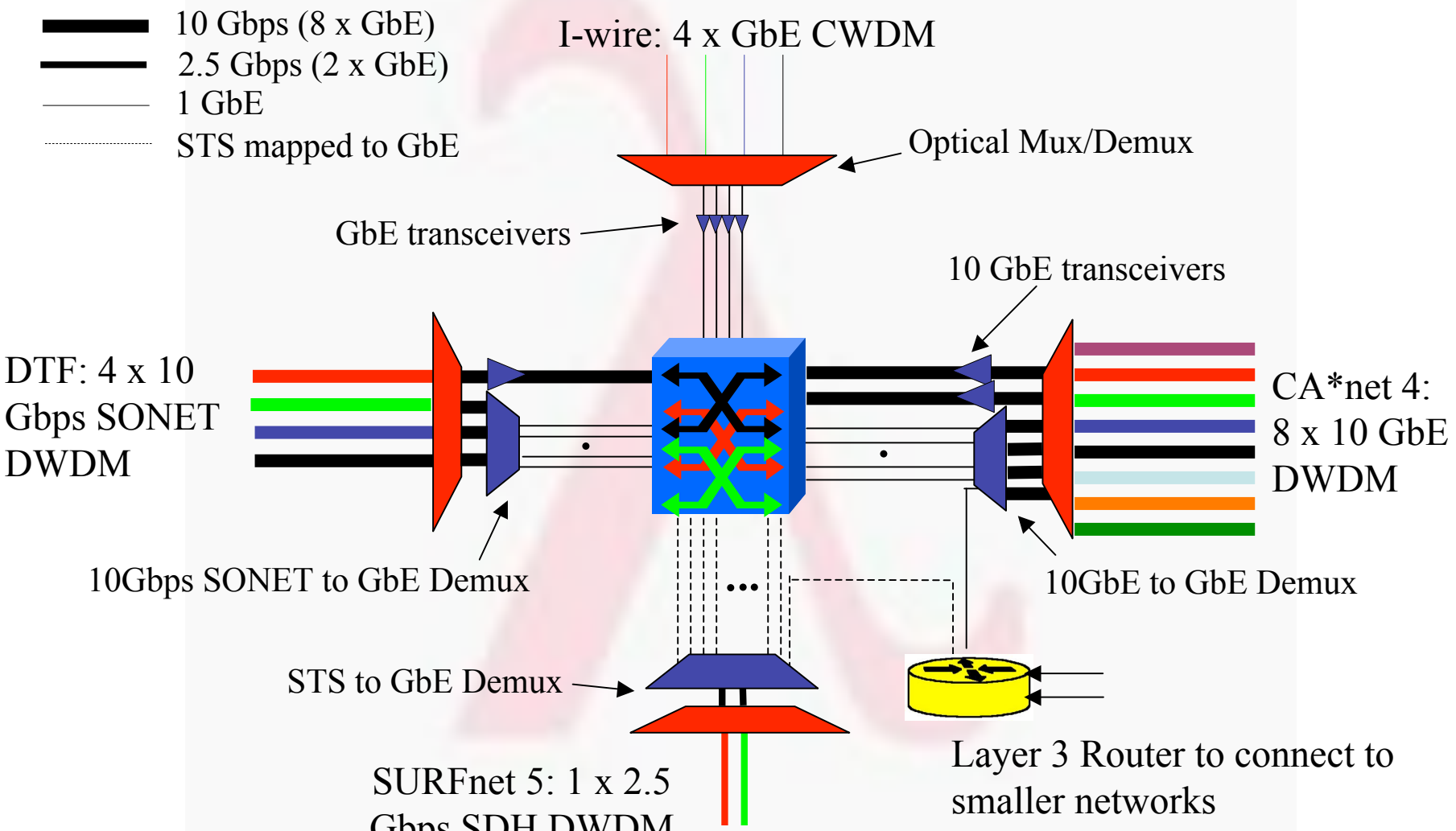
- **Lambdas for internal ISP bandwidth provisioning**
  - **An ISP uses a lambda switching network to make better use of its (suppliers) dark fibers and to provision to the POP's. In this case the optical network is just within one domain and as such is a relatively simple case.**
- **Lambda switching as peering point technology**
  - **In this use case a layer 1 Internet exchange is build. ISP's peer by instantiating lambdas to each other. Is a N*(N-1) and multi domain management problem.**
- **Lambda switching as grid application bandwidth provisioning**
  - **This is by far the most difficult since it needs UNI and NNI protocols to provision the optical paths through different domains.**

Current technology + (re)definition

- Current (to me) available technology consists of SONET/SDH switches
- DWDM+switching coming up
- Starlight uses for the time being VLAN's on Ethernet switches to connect [exactly] two ports
- So redefine a λ as:

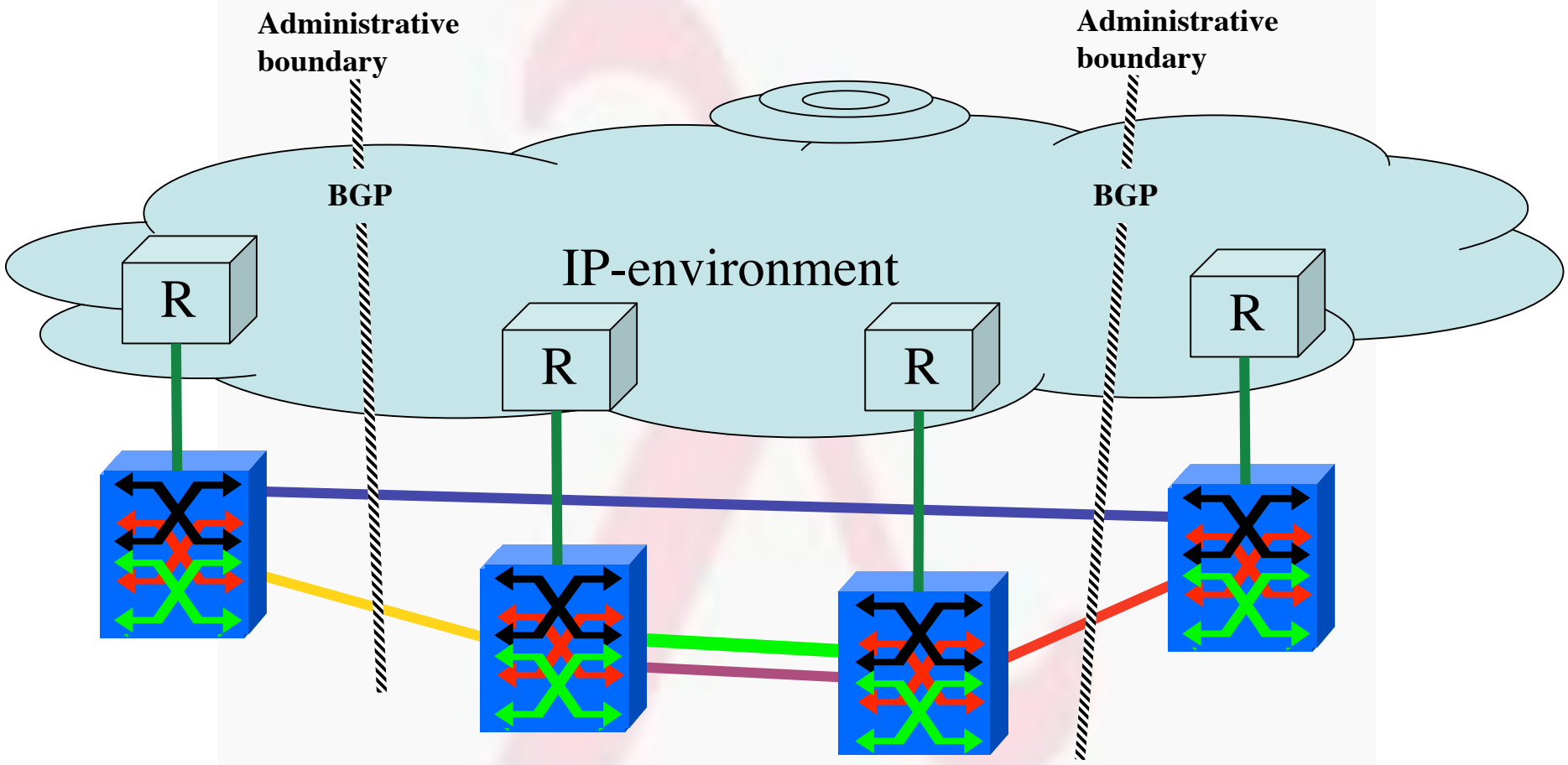
"a λ is a pipe where you can inspect packets as they enter and when they exit, but principally not when in transit. In transit one only deals with the parameters of the pipe: number, color, bandwidth"

Possible STAR LIGHT configuration

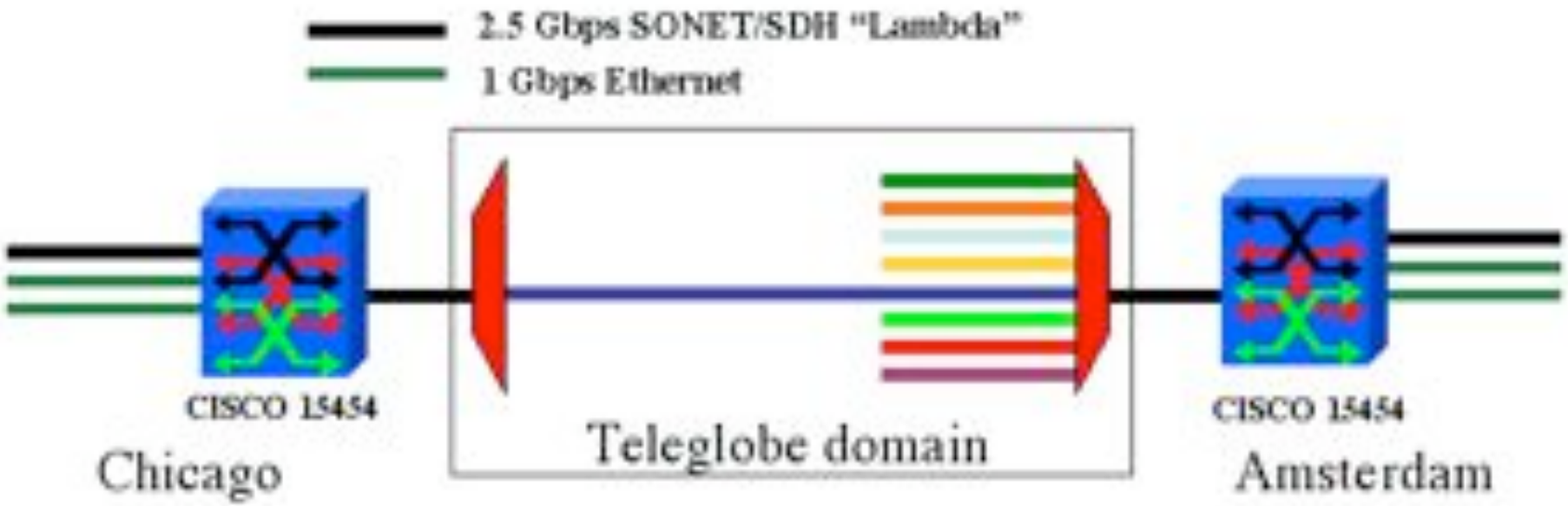


Courtesy Bill St.Arnaud

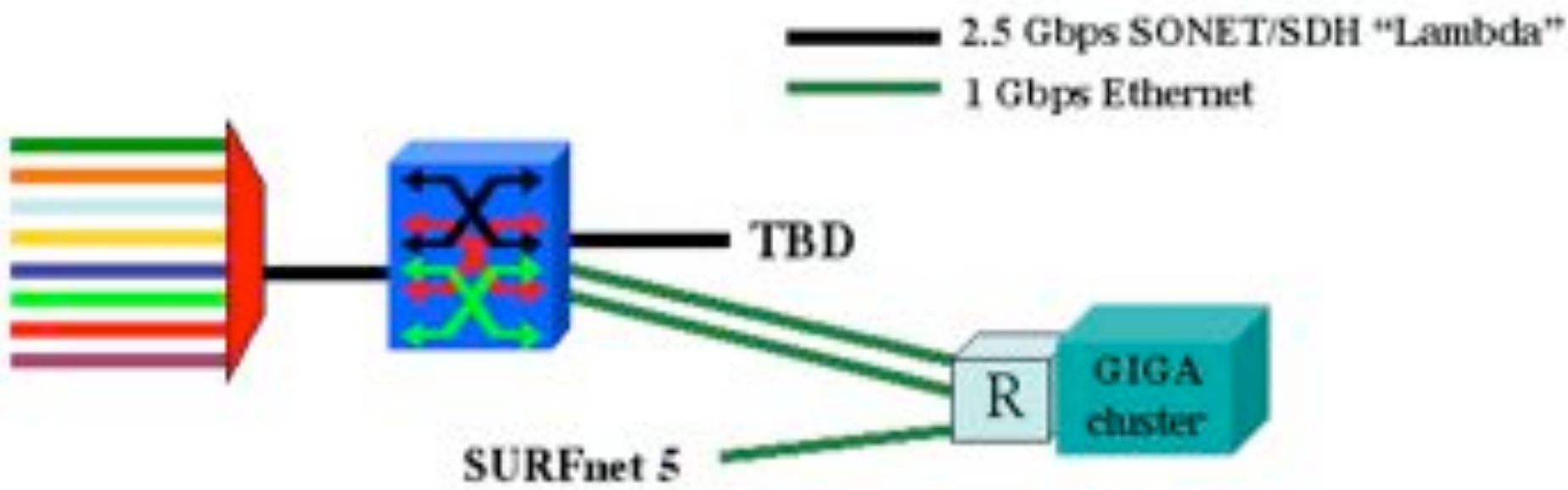
Multi domain IP controlled



Basic phase 1

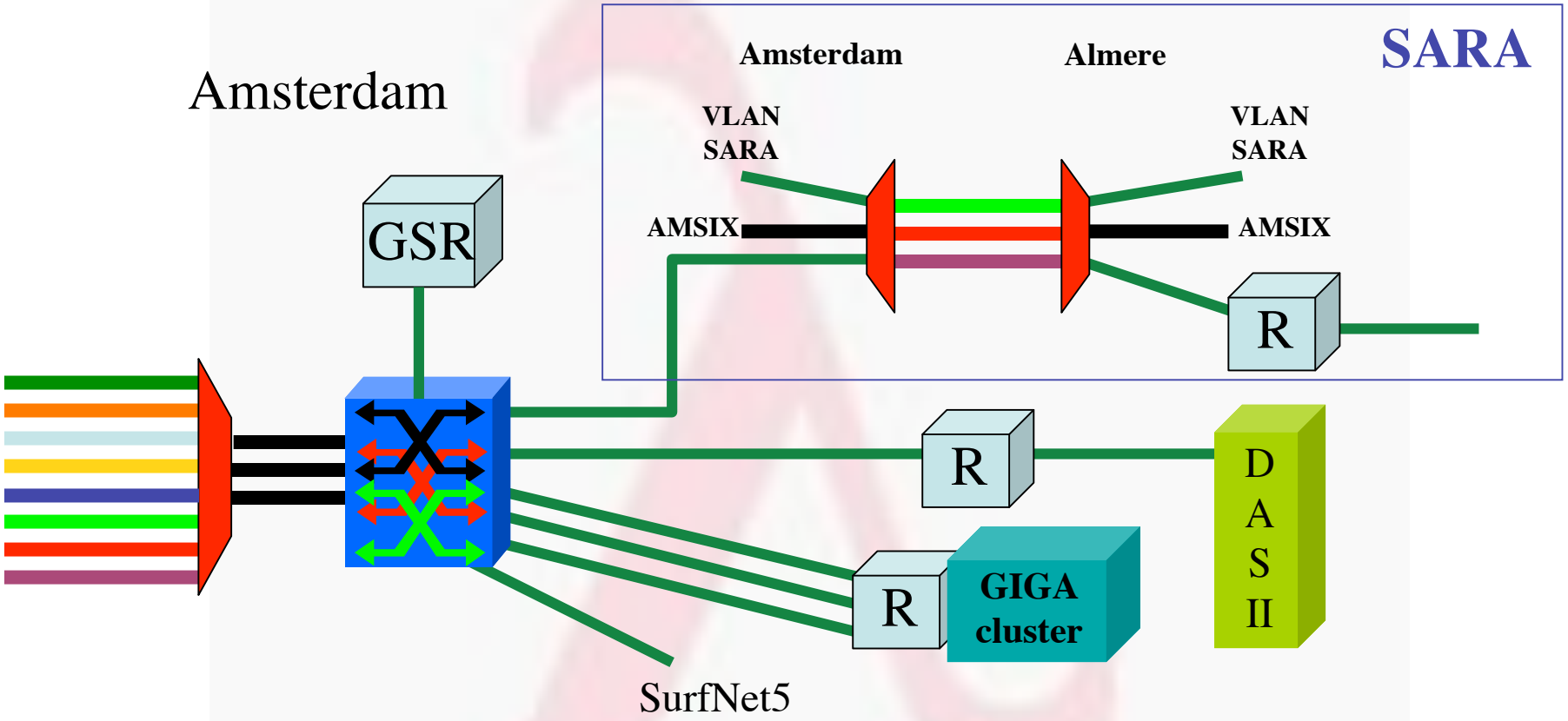


Amsterdam 1st phase



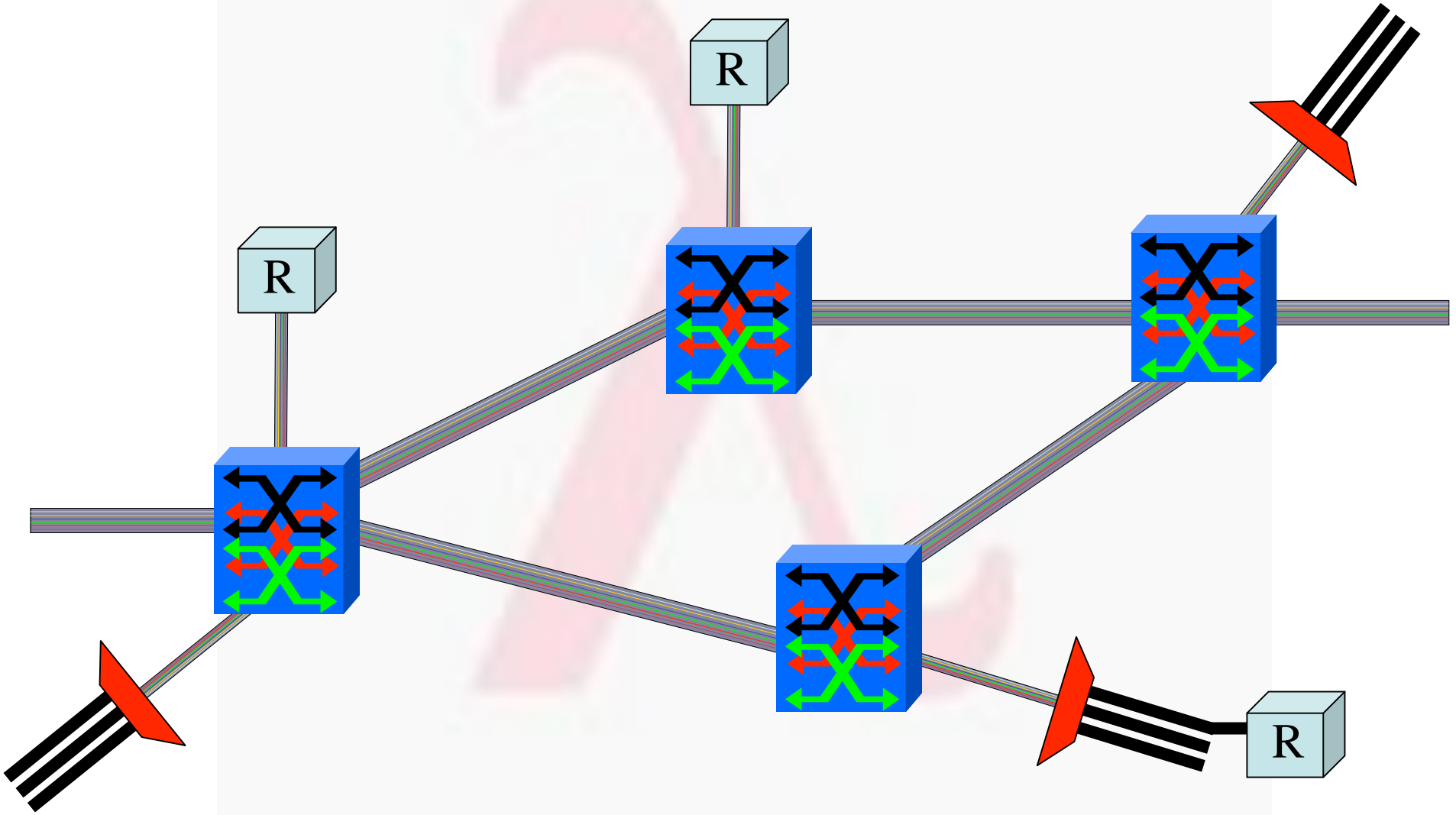
2.5 Gbps SONET/SDH "Lambda"

10/100/1000 Mbps Ethernet

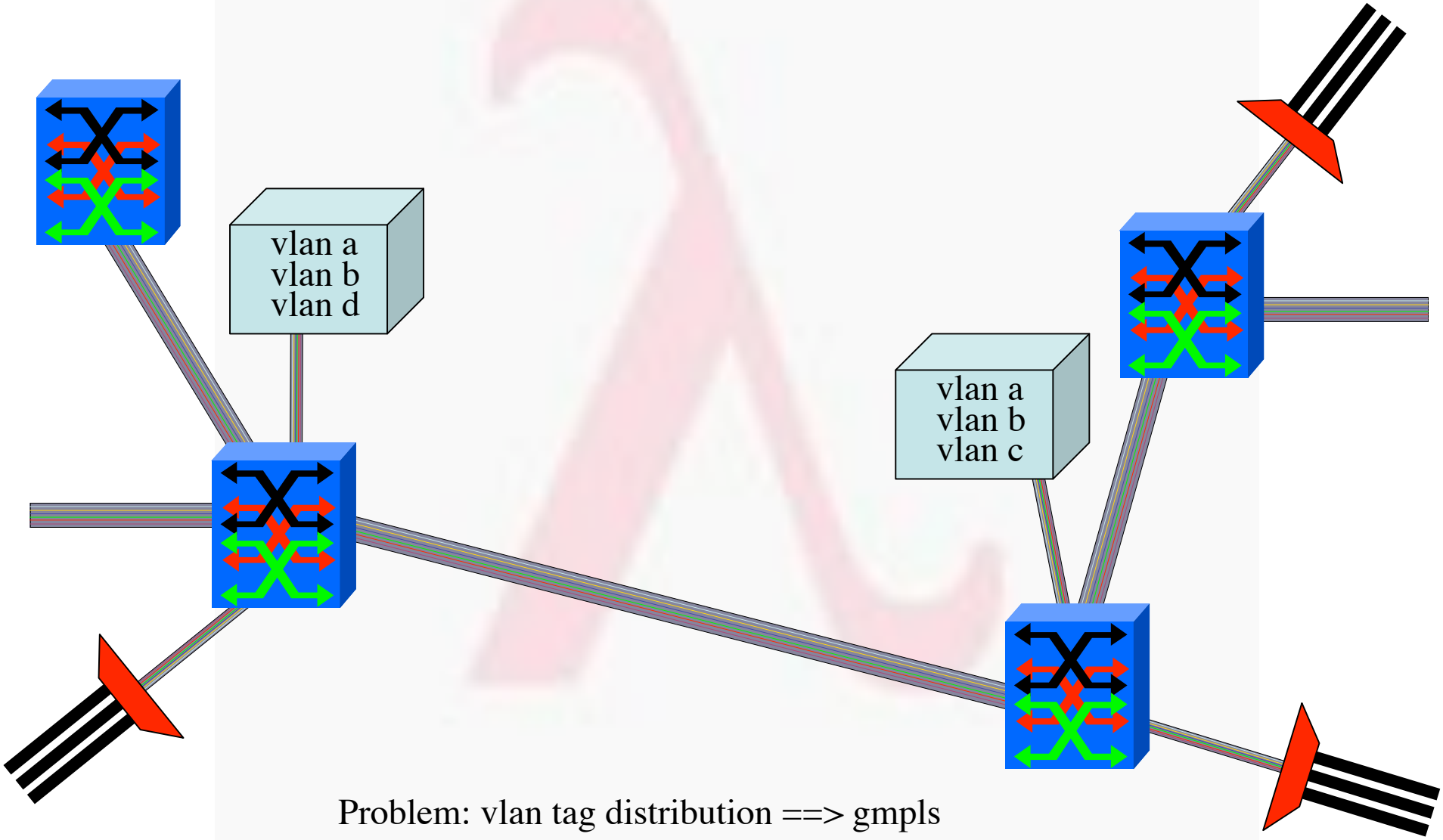


Amsterdam 2nd phase

Other architectures - L1 - 3

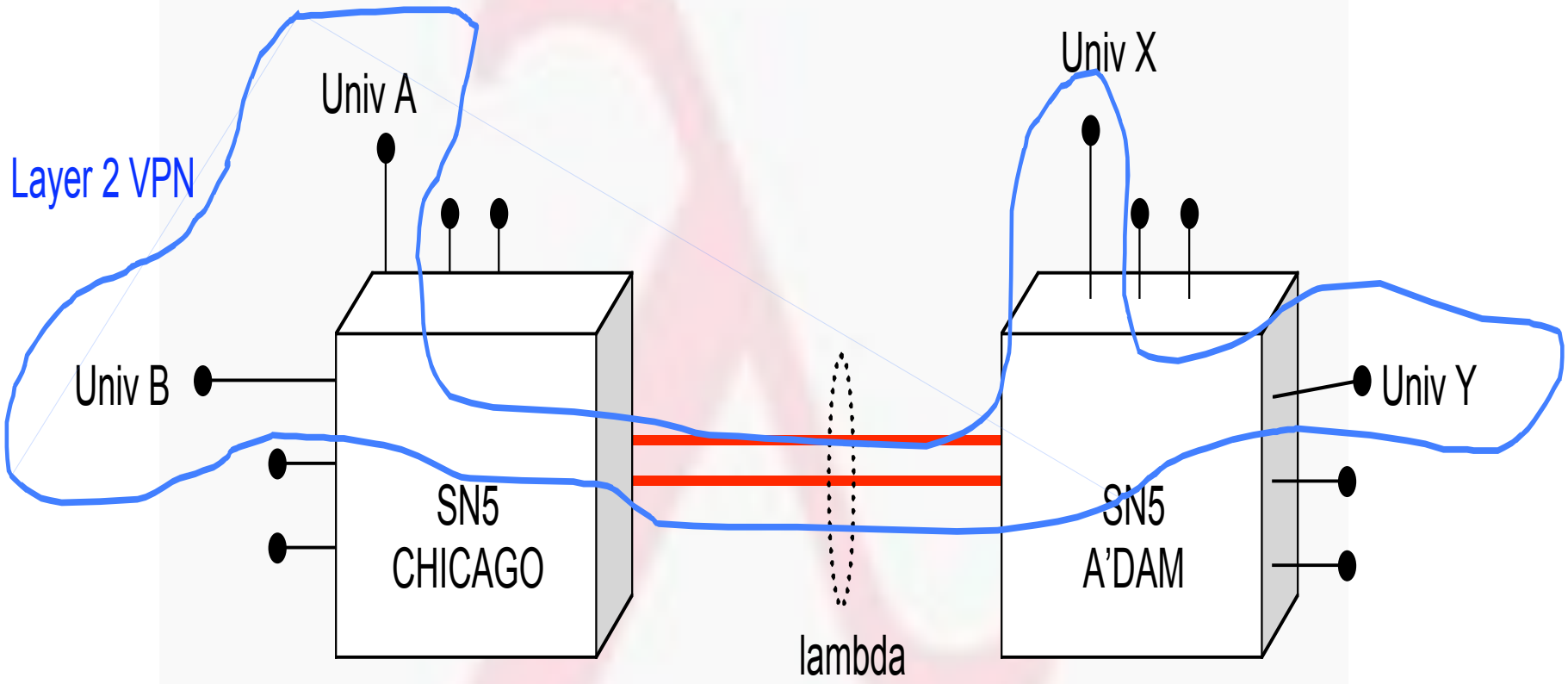


Other architectures - Distributed virtual IEX'es

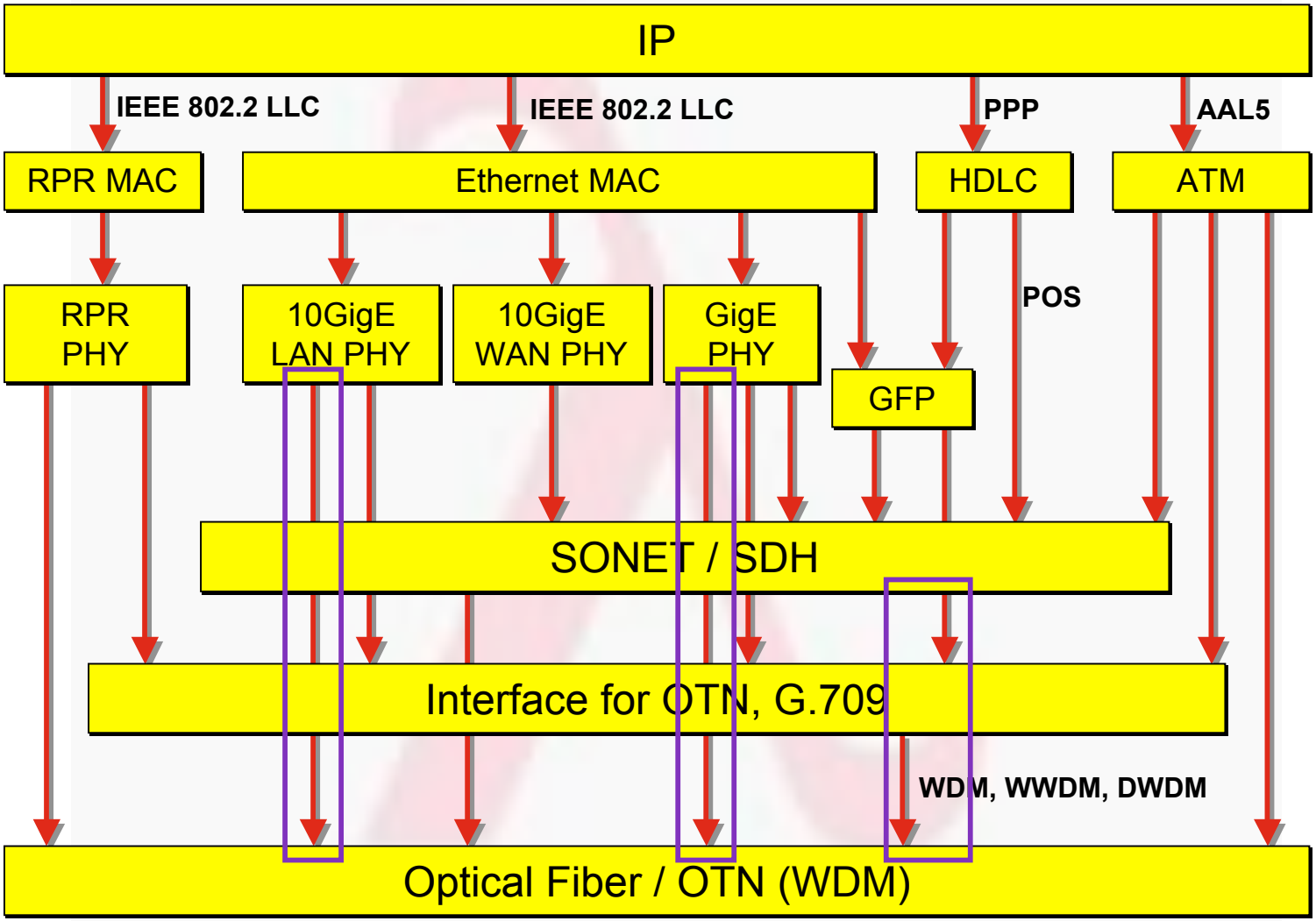


Problem: vlan tag distribution ==> gmpls

Distributed L2



IP networking over Optics

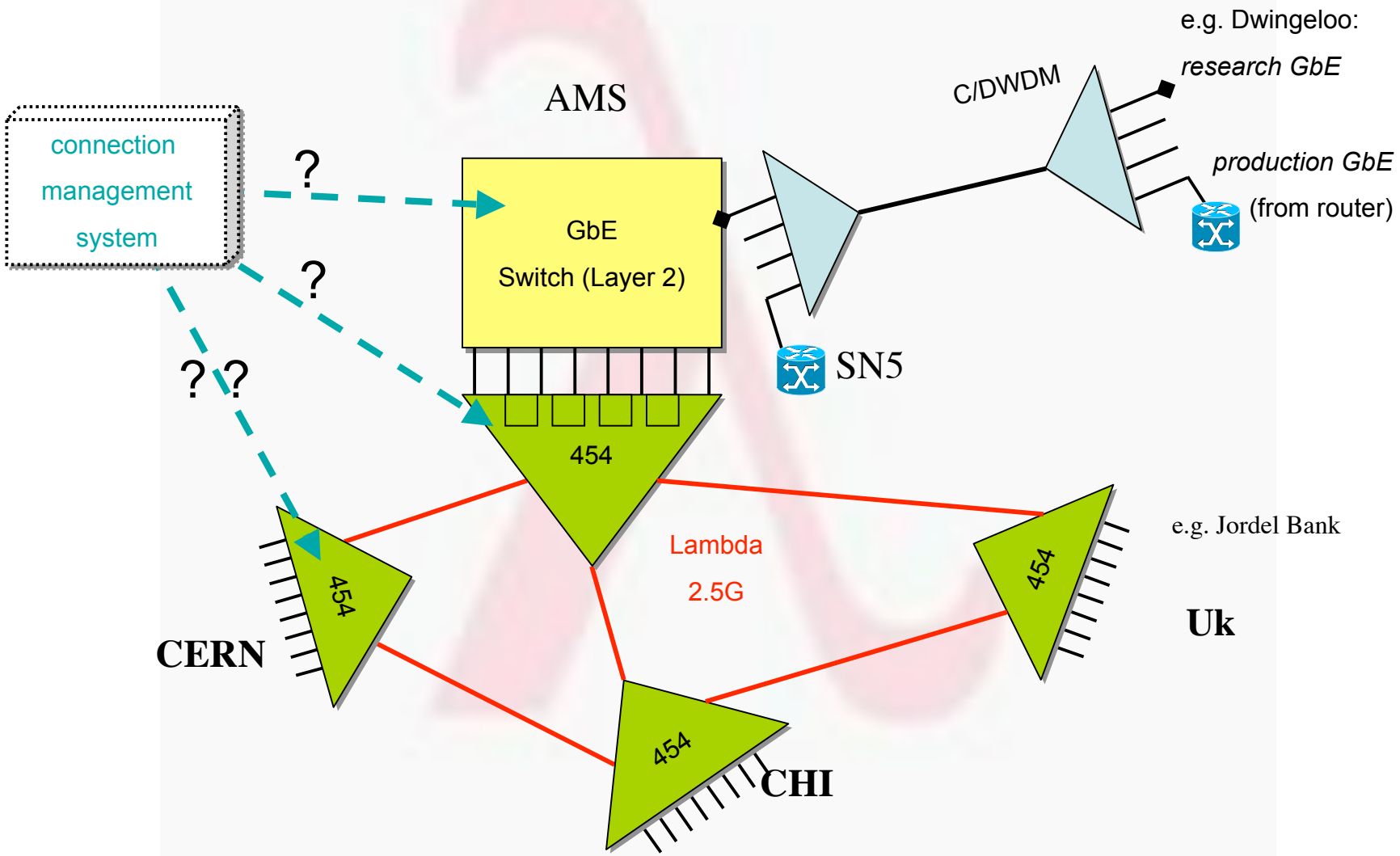


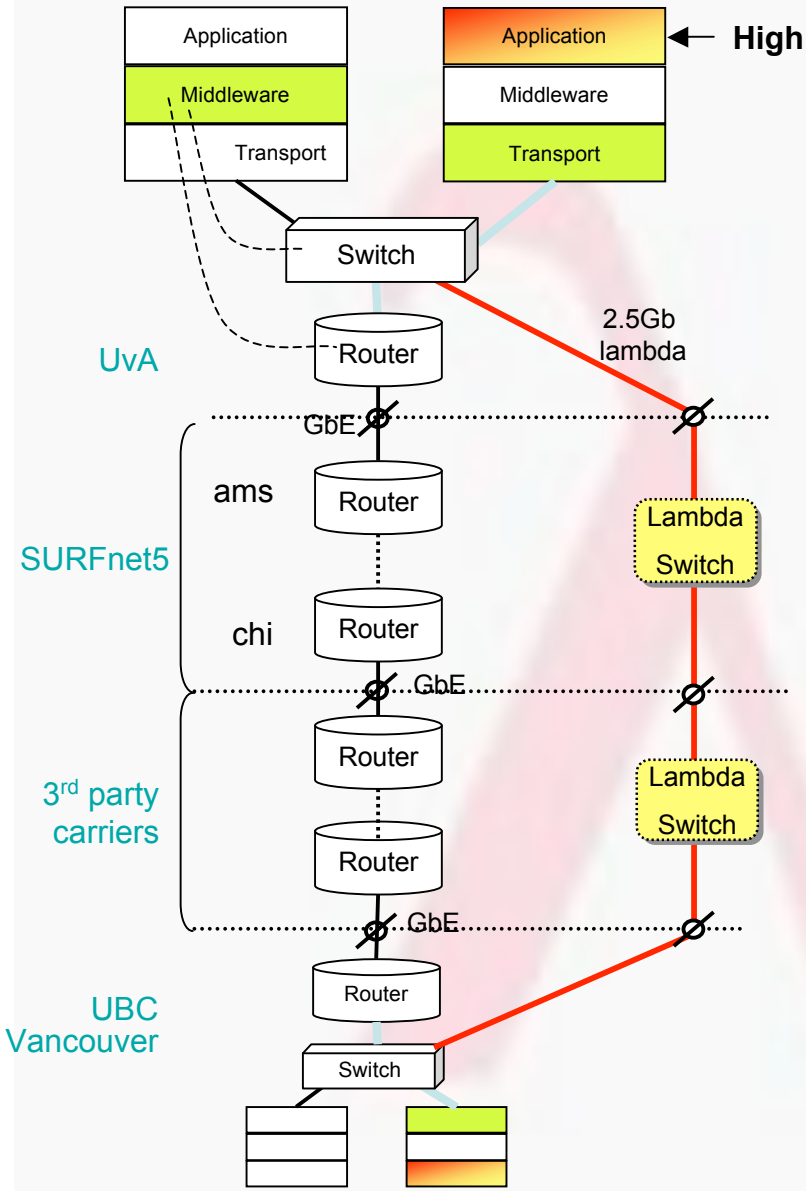
RPR = Resilient Packet Ring, IEEE 802.17
HDLC = High-Level Data-Link Control
POS = Packet over SONET/SDH
GFP = Generic Framing Procedure (ANSI T1 X1-driven standard)

OTN = Optical Transport Network
WDM = Wavelength Division Multiplexing
WWDM = Wide WDM
DWDM = Dense WDM

SURFnet

Lambda/GbE exchange





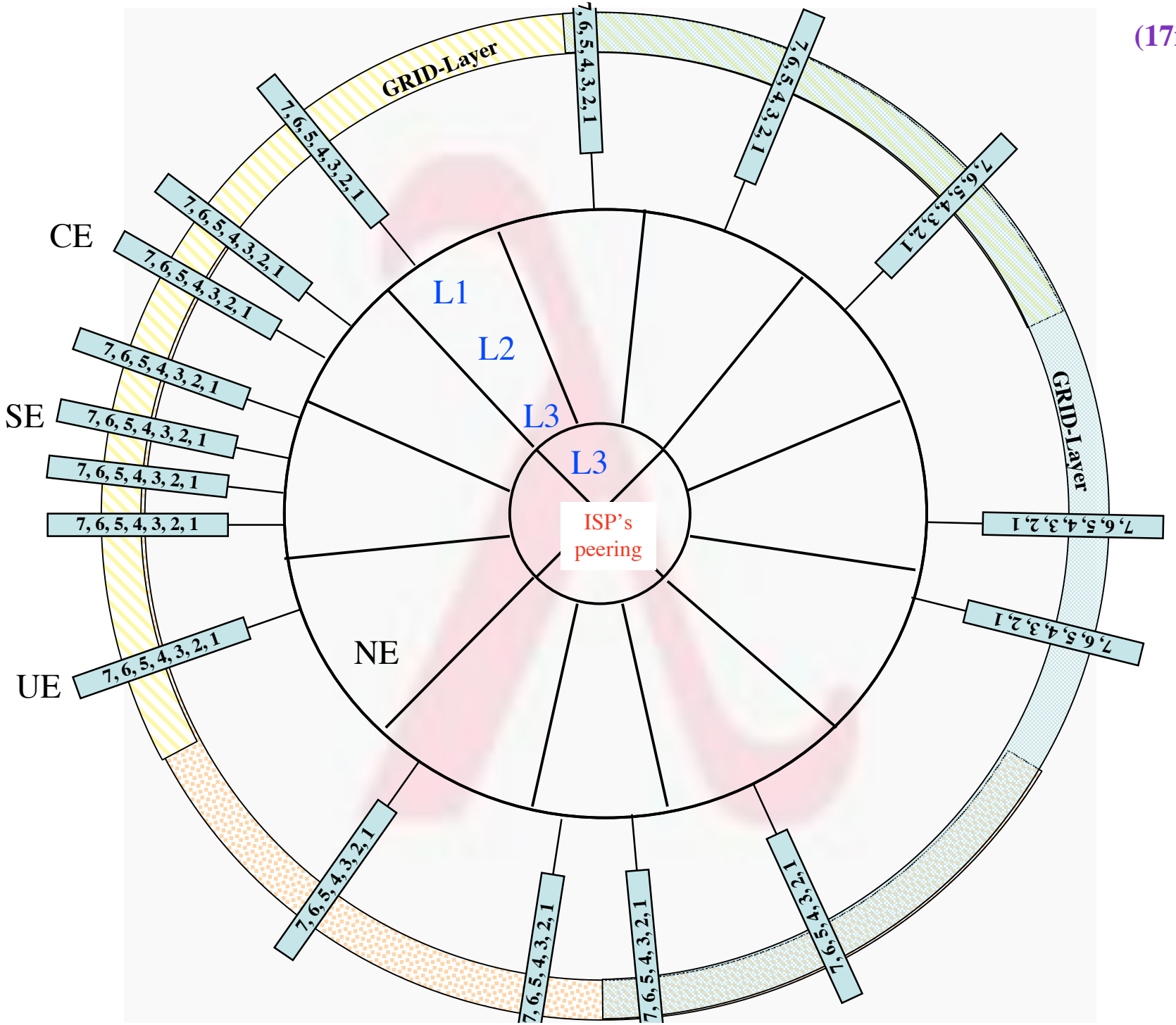
- lambda for high bandwidth applications
  - Bypass of production network
  - Middleware may request (optical) pipe

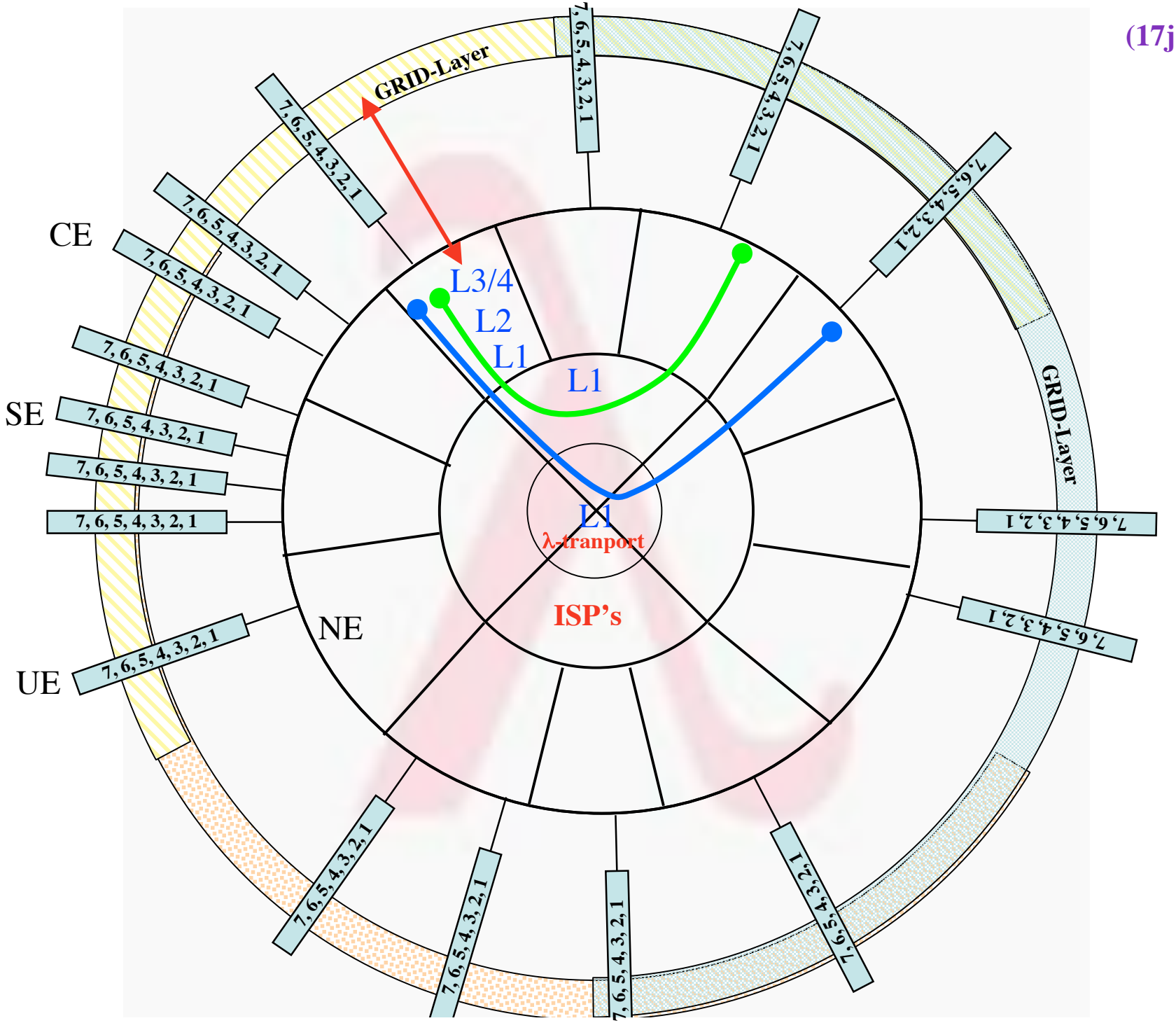
research on λ's

- how to get traffic in and out of lambdas
- how to map load on the network to a map of lambdas
- how to deal with lambdas at peering points
- how to deal with provisioning when more administrative domains are involved
- how to do fine grain near real time grid application level lambda provisioning

Research with λ's

- High speed TCP (high rtt and BW)
- Routing stability
- Routing responsibility
- Extremely multihomed Networks
- Roles, organizational issues
- SLA's
- Models (Connection less versus oriented)
- Discreet versus continuous in time


GRID-Layer
GRID-Layer
CE
SE
UE
NE
L1
L2
L3
L3
ISP's peering
7, 6, 5, 4, 3, 2, 1


GRID-Layer
7, 6, 5, 4, 3, 2, 1
CE
SE
NE
UE
L3/4
L2
L1
L1
L1
λ-tranport
ISP's
GRID-Layer

Example projects

- Do a few well defined proof of concepts
- Aim at IGRID2002
- Look into dynamics in lower layer
- Define interfaces
- Make Lambda setup take 300 ms in stead of 3 months
- Look into roles of different players
  - Who is owner of what part (GSP)
- Multi domain, policy, BW-brokering, AAA
- Optimal?
  - make 10 % protected and run the rest in scavenger mode