High Performance Networking for Grid Applications

www.science.uva.nl/~delaat

Cees de Laat

Faculty of Science

High Performance Networking for Grid Applications

www.science.uva.nl/~deλaat

Cees de λaat

EU

SURFnet

University of Amsterdam

SARA

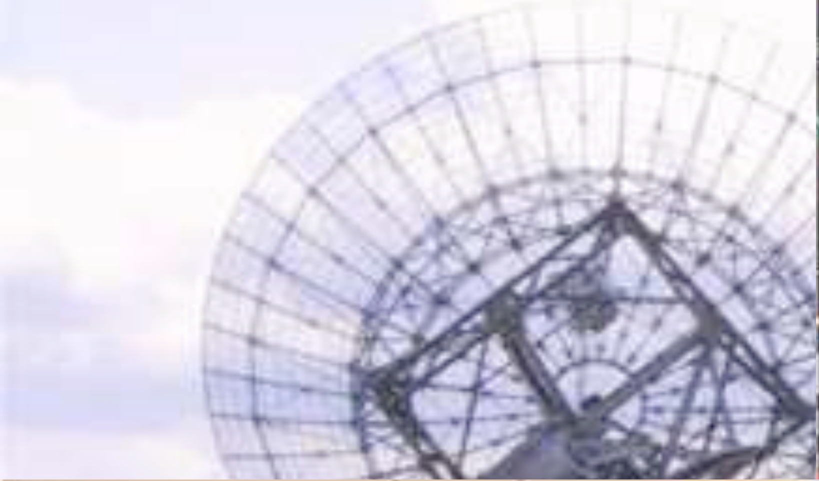
NIKHEF

NCF

Contents of this talk

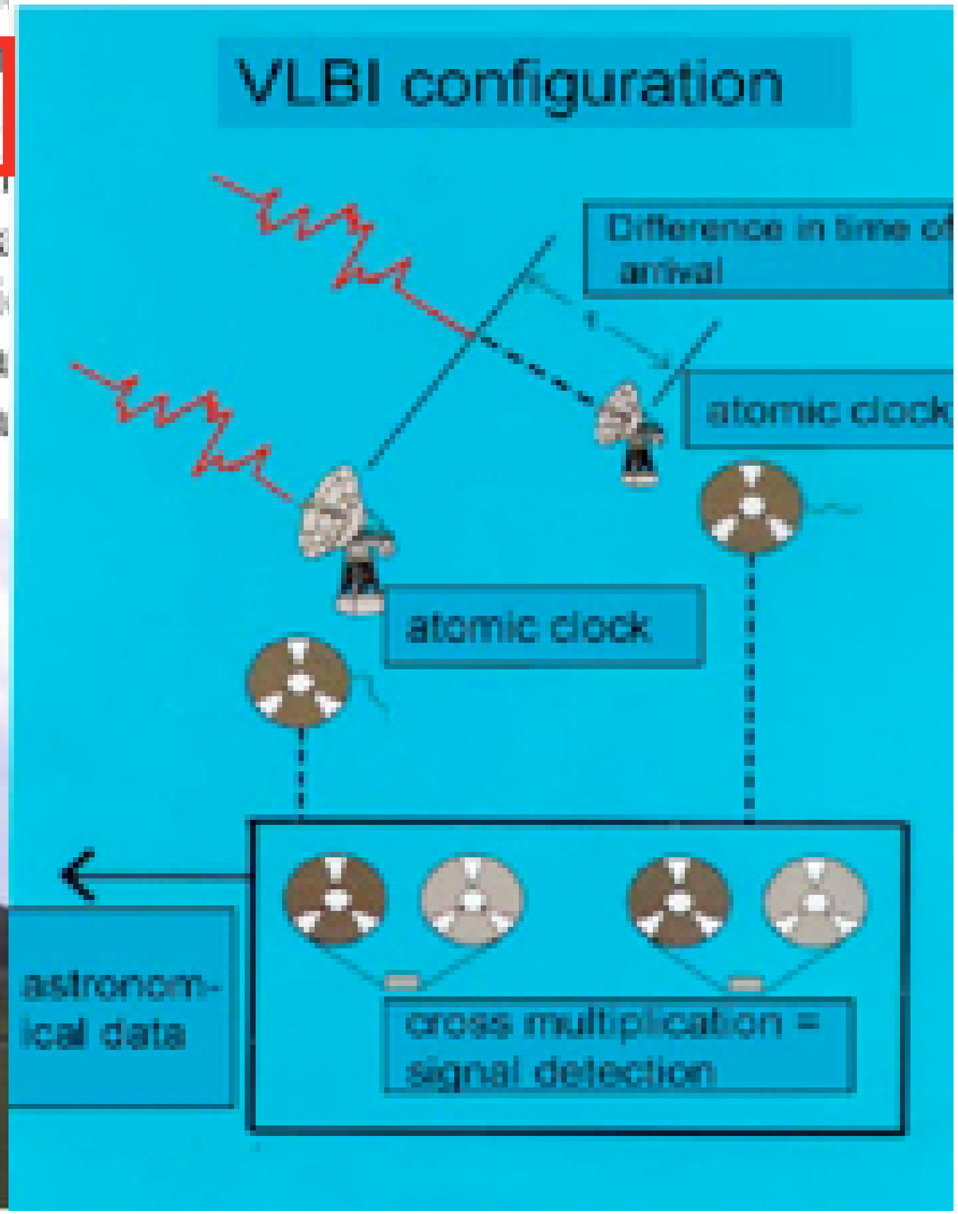
This slide is intentionally left blank

eVLBI



VLBI

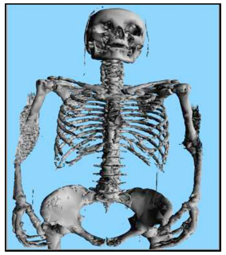
ger term VLBI is easily capable of generating many Gb of data per
The sensitivity of the VLBI array scales w
(~data-rate) and there is a strong push to i
. Rates of 8Gb/s or more are entirely feasibl
nder development. It is expected that paral
orrelator will remain the most efficient approa
s distributed processing may have an appli
lti-gigabit data streams will aggregate into la
or and the capacity of the final link to the da
tor.

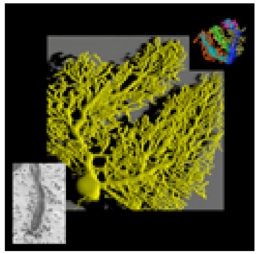
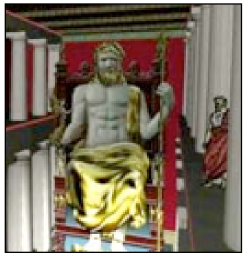
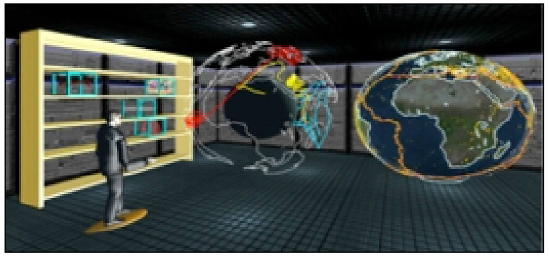
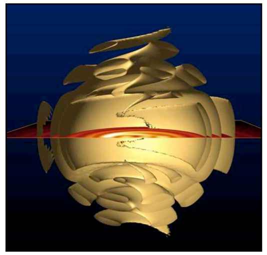


iGrid 2002

September 24-26, 2002, Amsterdam, The Netherlands

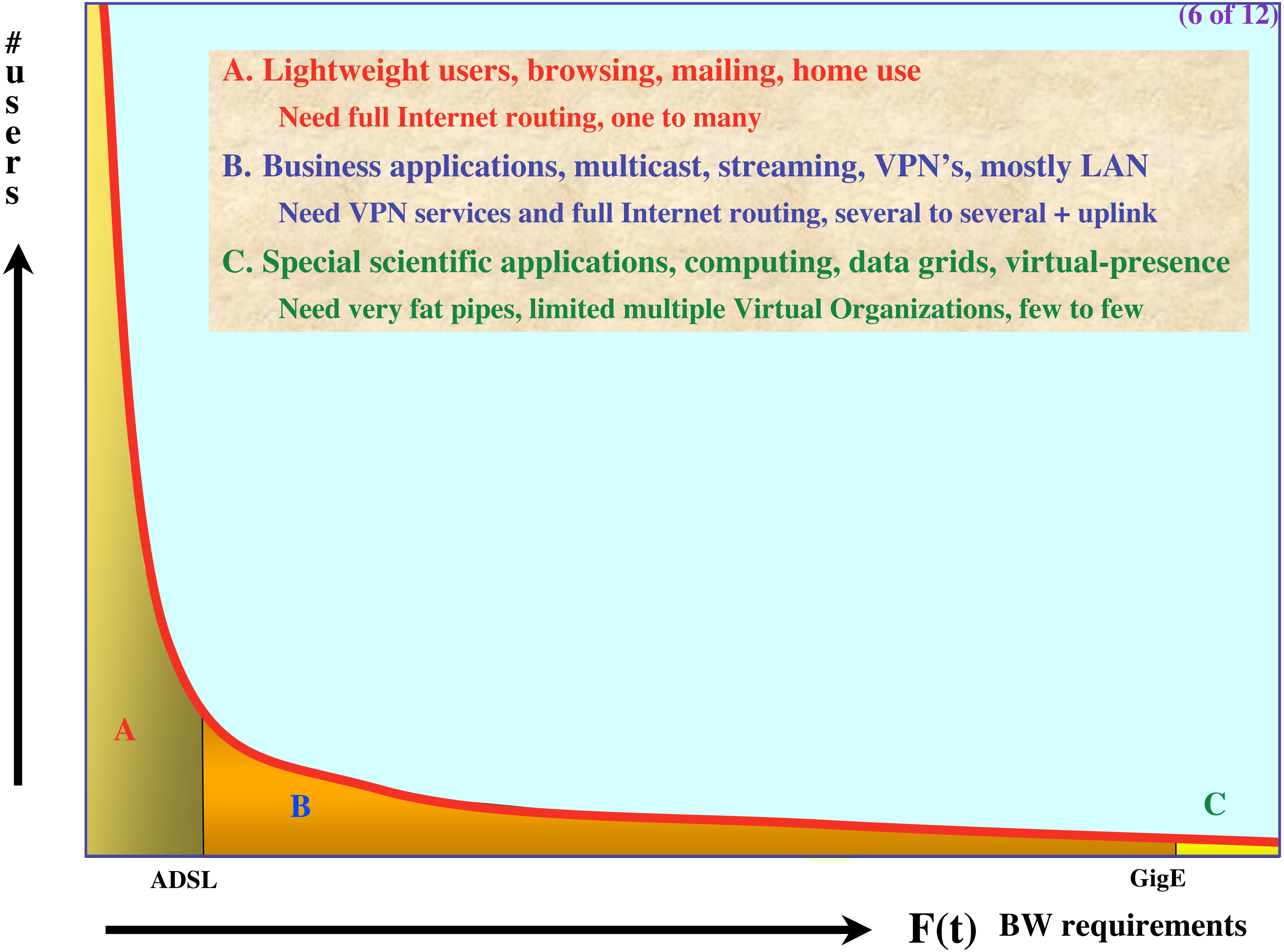
- 28 demonstrations from 16 countries: Australia, Canada, CERN, France, Finland, Germany, Greece, Italy, Japan, The Netherlands, Singapore, Spain, Sweden, Taiwan, United Kingdom, United States
- Applications demonstrated: art, bioinformatics, chemistry, cosmology, cultural heritage, education, high-definition media streaming, manufacturing, medicine, neuroscience, physics, tele-science



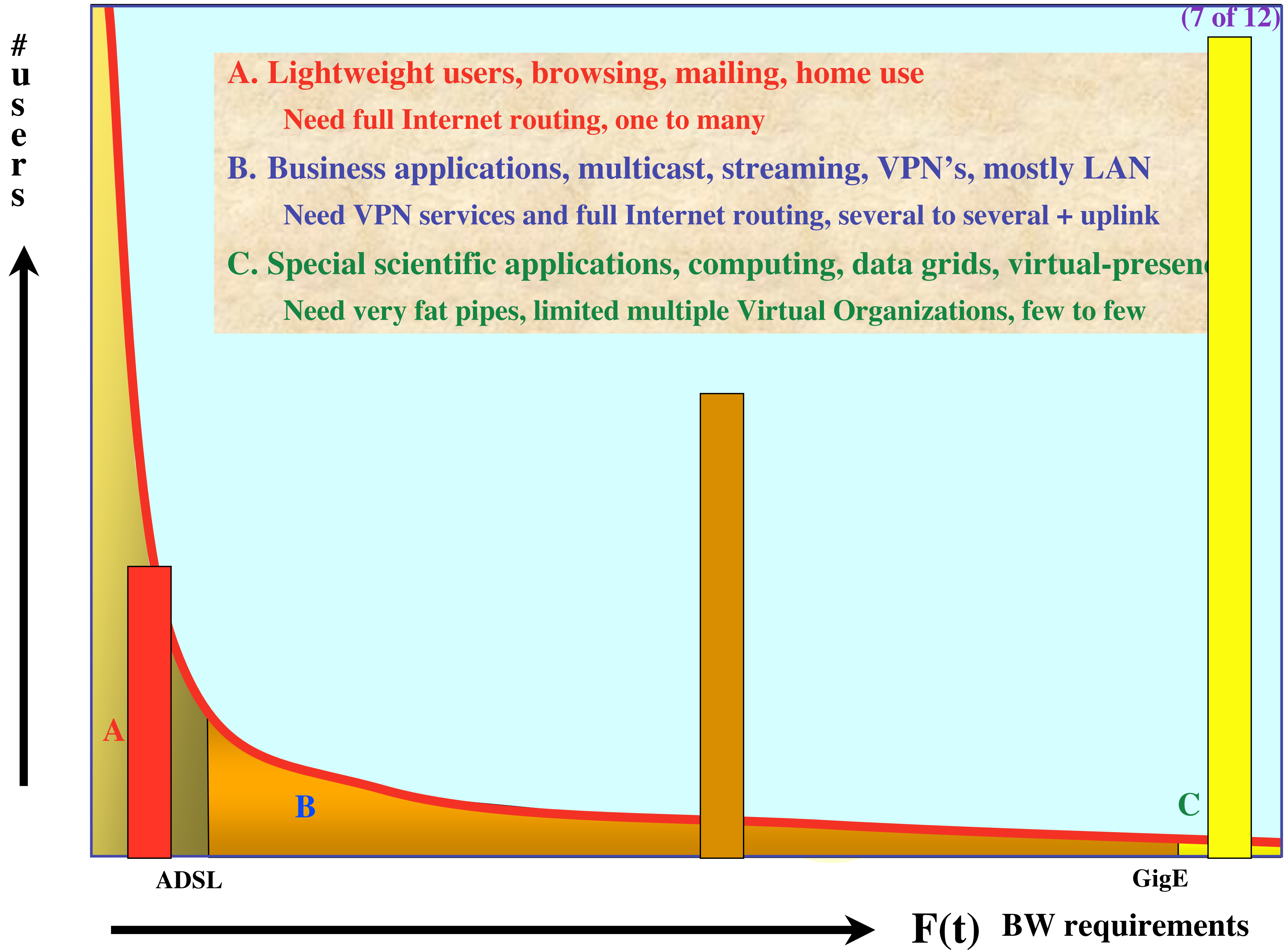





- Grid technologies demonstrated: Major emphasis on grid middleware, data management grids, data replication grids, visualization grids, data/visualization grids, computational grids, access grids, grid portals
- 25Gb transatlantic bandwidth (100Mb/attendee, 250x iGrid2000!)

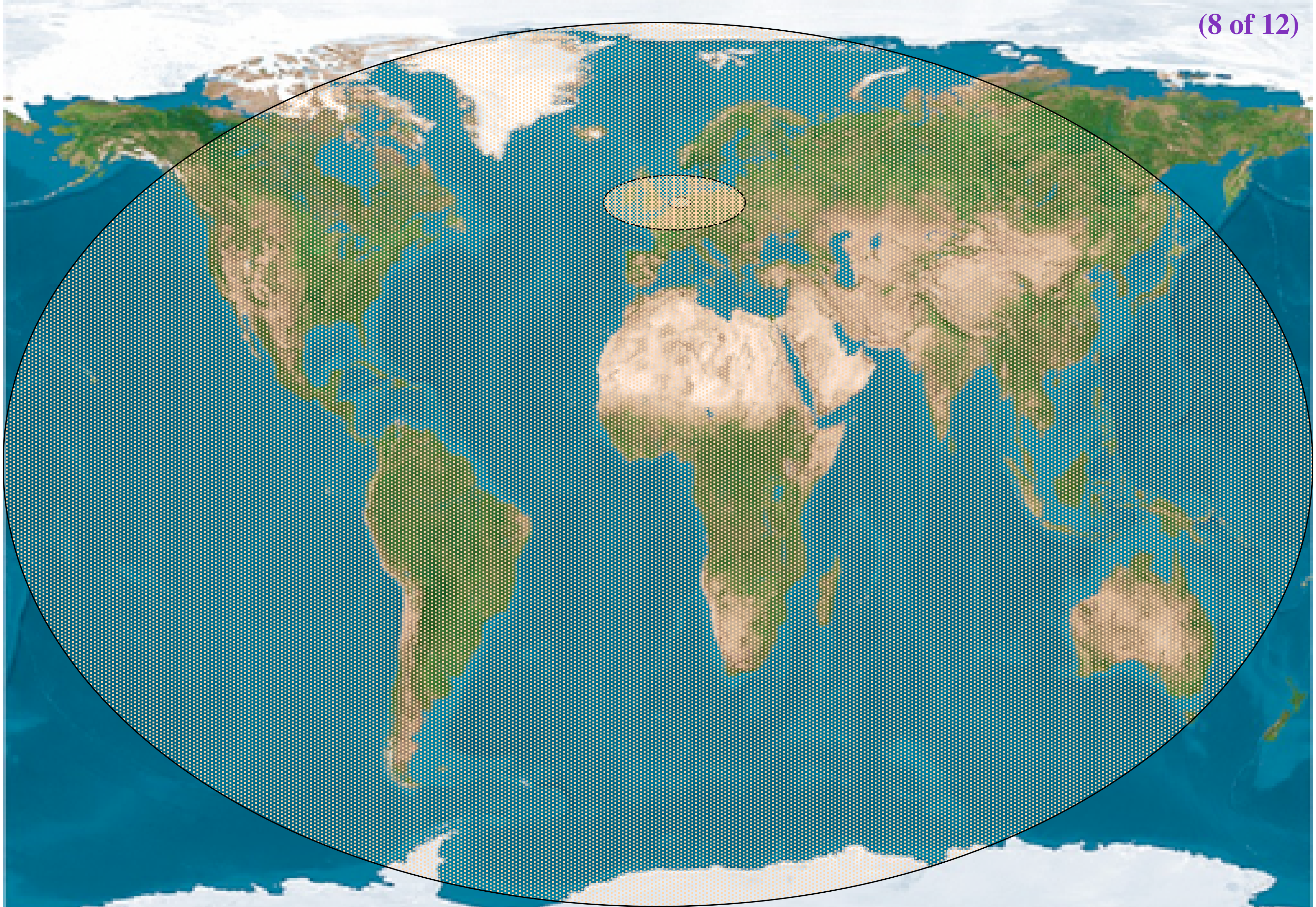
www.igrid2002.org



(6 of 12)
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. Special scientific applications, computing, data grids, virtual-presence
Need very fat pipes, limited multiple Virtual Organizations, few to few
users
A
B
C
ADSL
GigE
F(t) BW requirements



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. Special scientific applications, computing, data grids, virtual-presenc
Need very fat pipes, limited multiple Virtual Organizations, few to few
users
A
B
C
ADSL
GigE
F(t) BW requirements



Scale 2-20-200

The only formula's

$$\#\,\lambda(rtt) \approx \frac{200 * e^{(t-2002)}}{rtt}$$

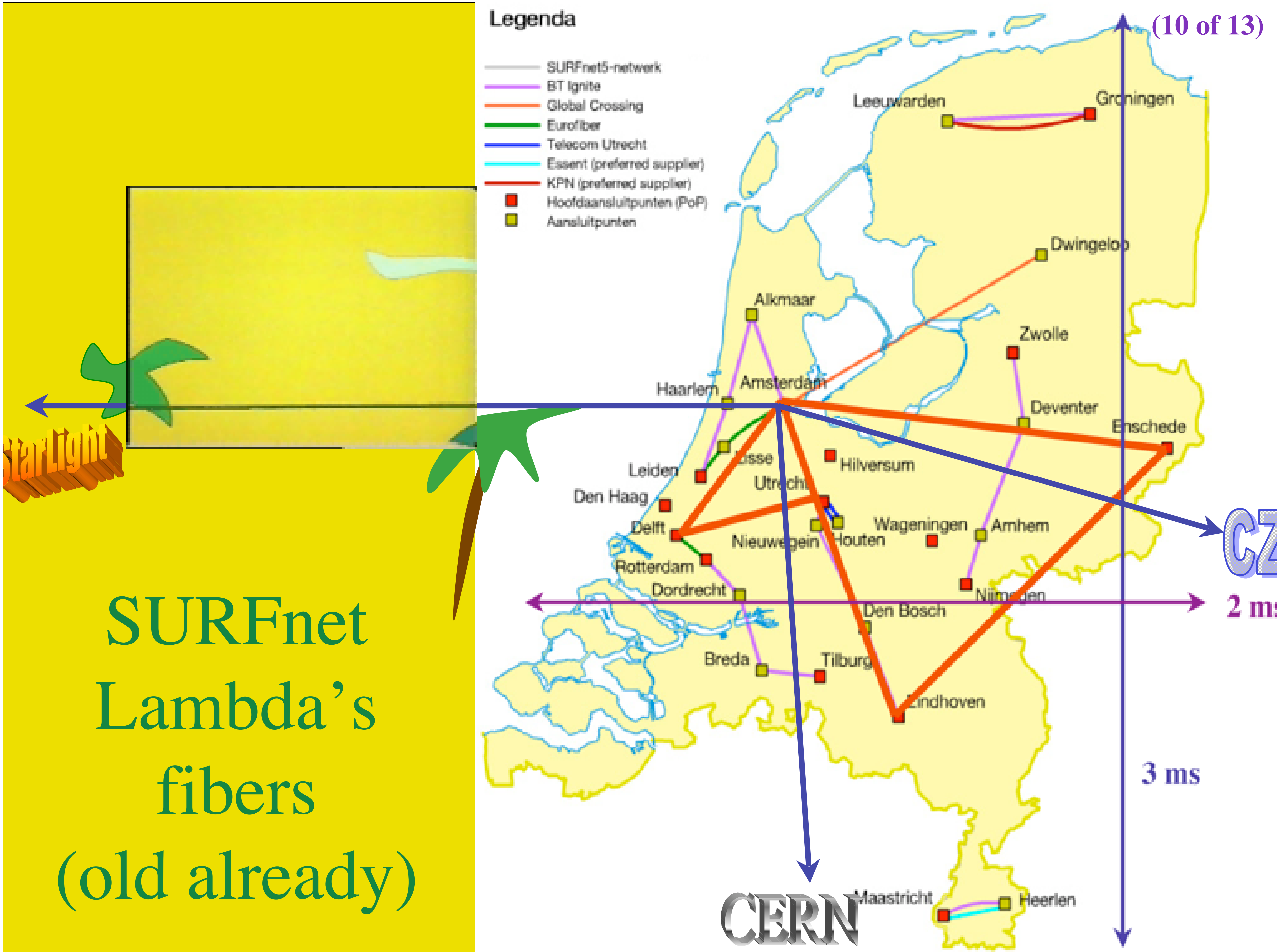
Now, having been a High Energy Physicist we set

c = 1

e = 1

$\bar{h}$ = 1

and the formula reduces to: $\#\,\lambda(rtt) \approx \frac{200 * e^{(t-2002)}}{rtt}$



SURFnet Lambda's fibers (old already)

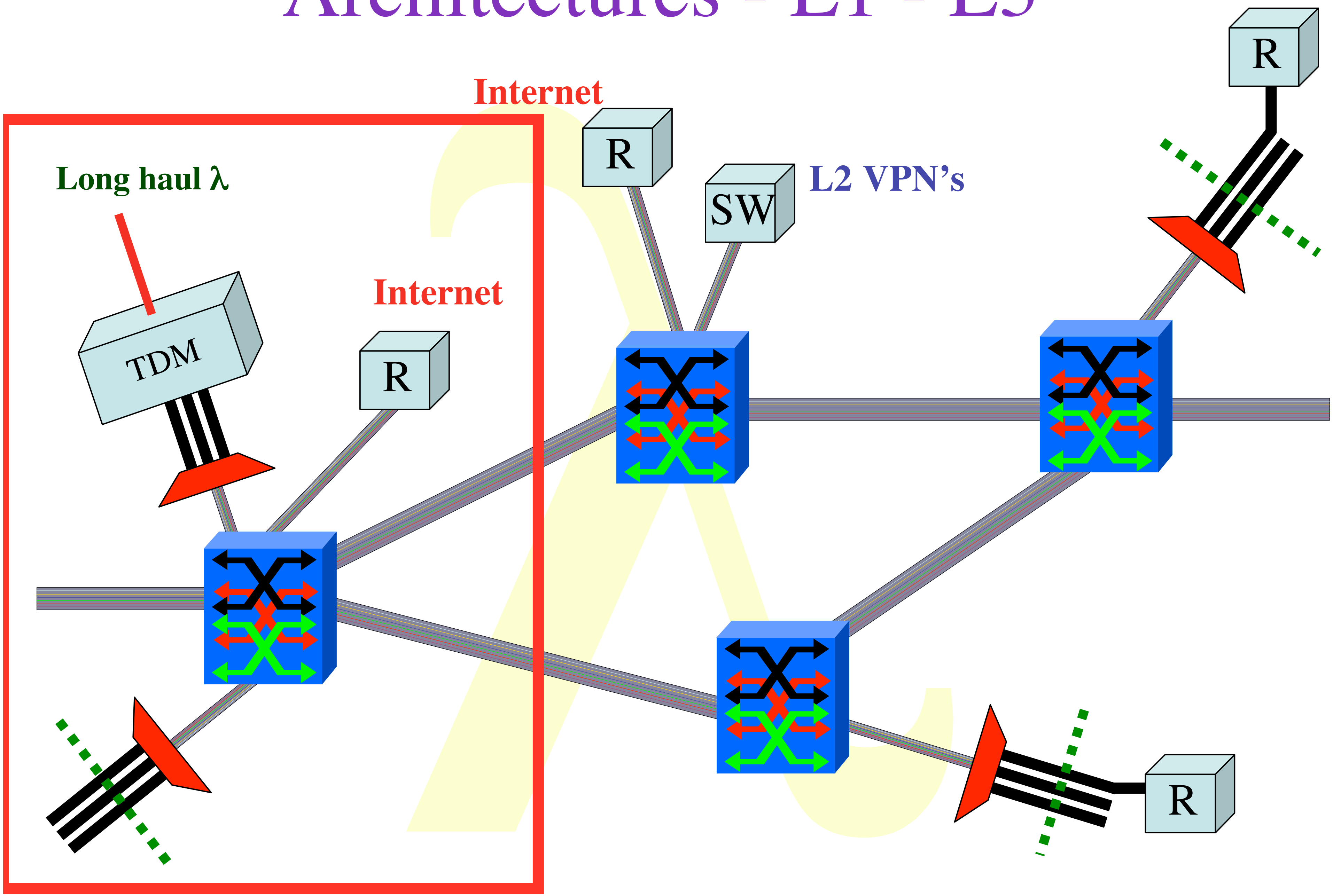
Services

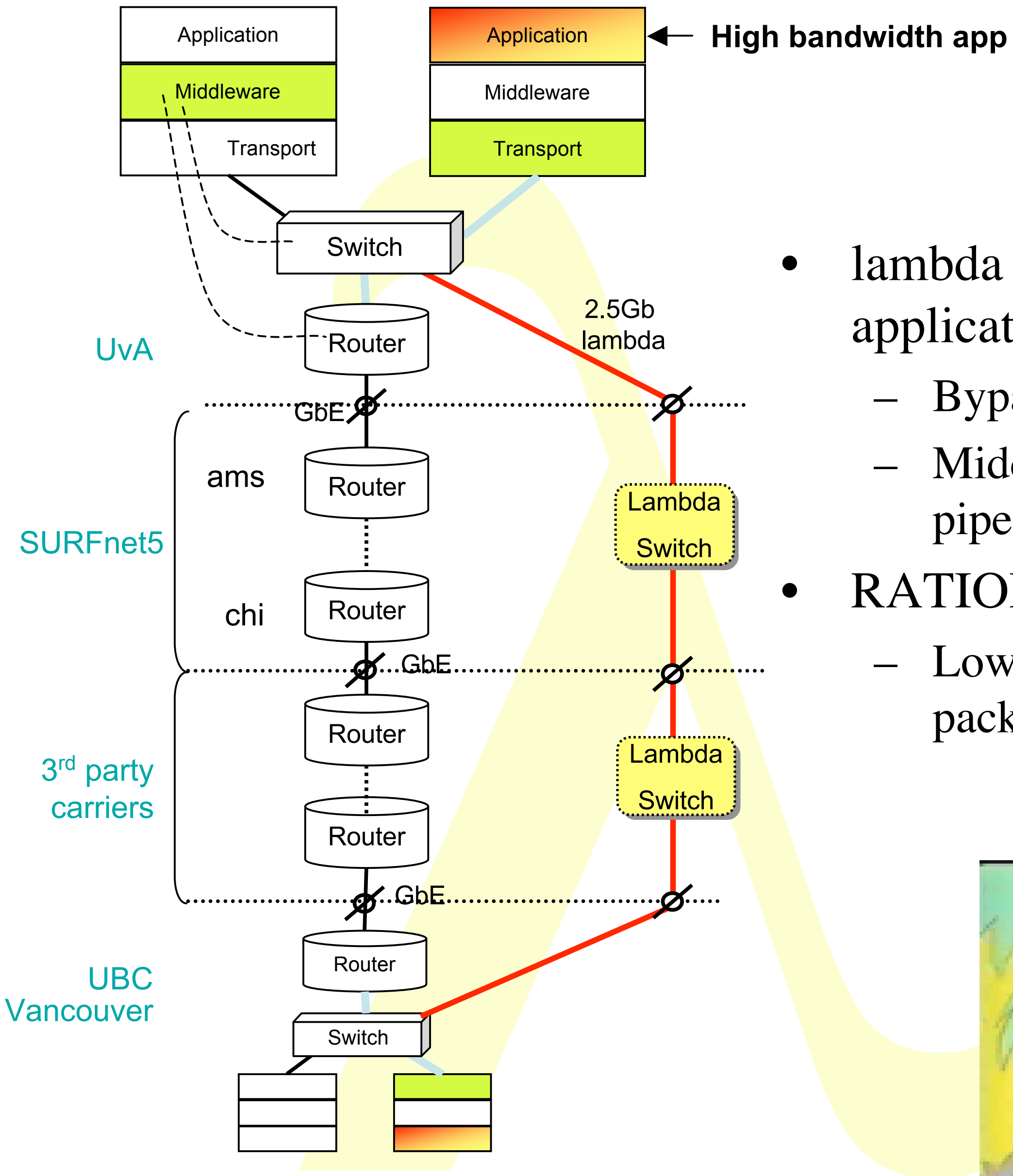
SCALE / CLASS	2 Metro	20 National/ regional	200 World
A	Switching/ routing	Routing	ROUTER$
B	VPN's, (G)MPLS	VPN's Routing	ROUTER$
C $\# \lambda(rtt) \approx \frac{200 * e^{(t-2002)}}{rtt}$	dark fiber Optical switching	Lambda switching	Sub-lambdas, ethernet-sdh

So what are the facts

- **Costs of fat pipes (fibers) are one/third of cost of equipment to light them up**
  - **Is what Lambda salesmen tell me**
- **Costs of optical equipment 10% of switching 10 % of full routing equipment for same throughput**
  - **100 Byte packet @ 40 Gb/s -> 20 ns to look up in 140 kEntries routing table (light speed from me to you!)**
- **Big sciences need fat pipes**
- **Bottom line: look for a hybrid architecture which serves all users in a cost effective way**

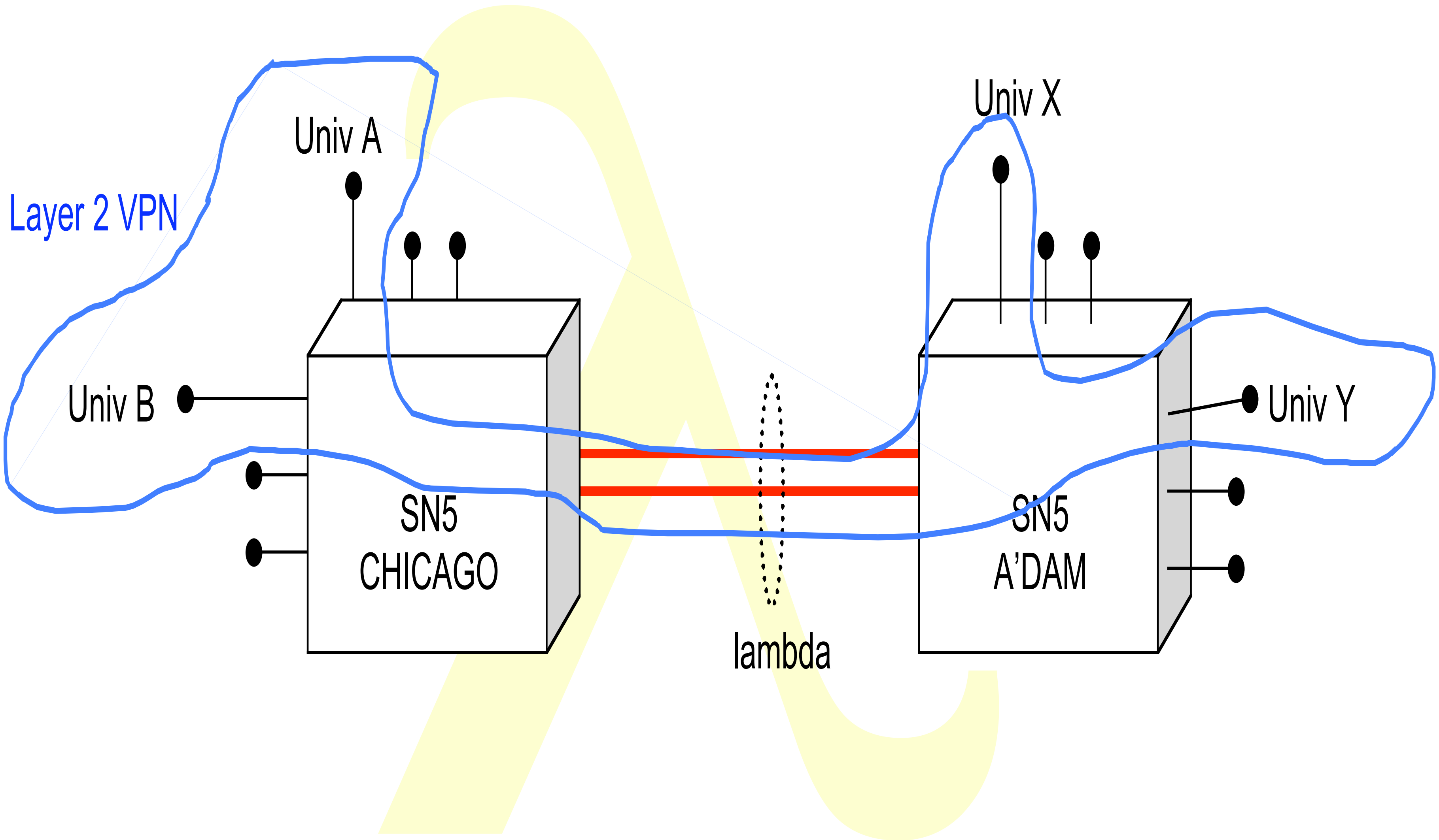
Architectures - L1 - L3

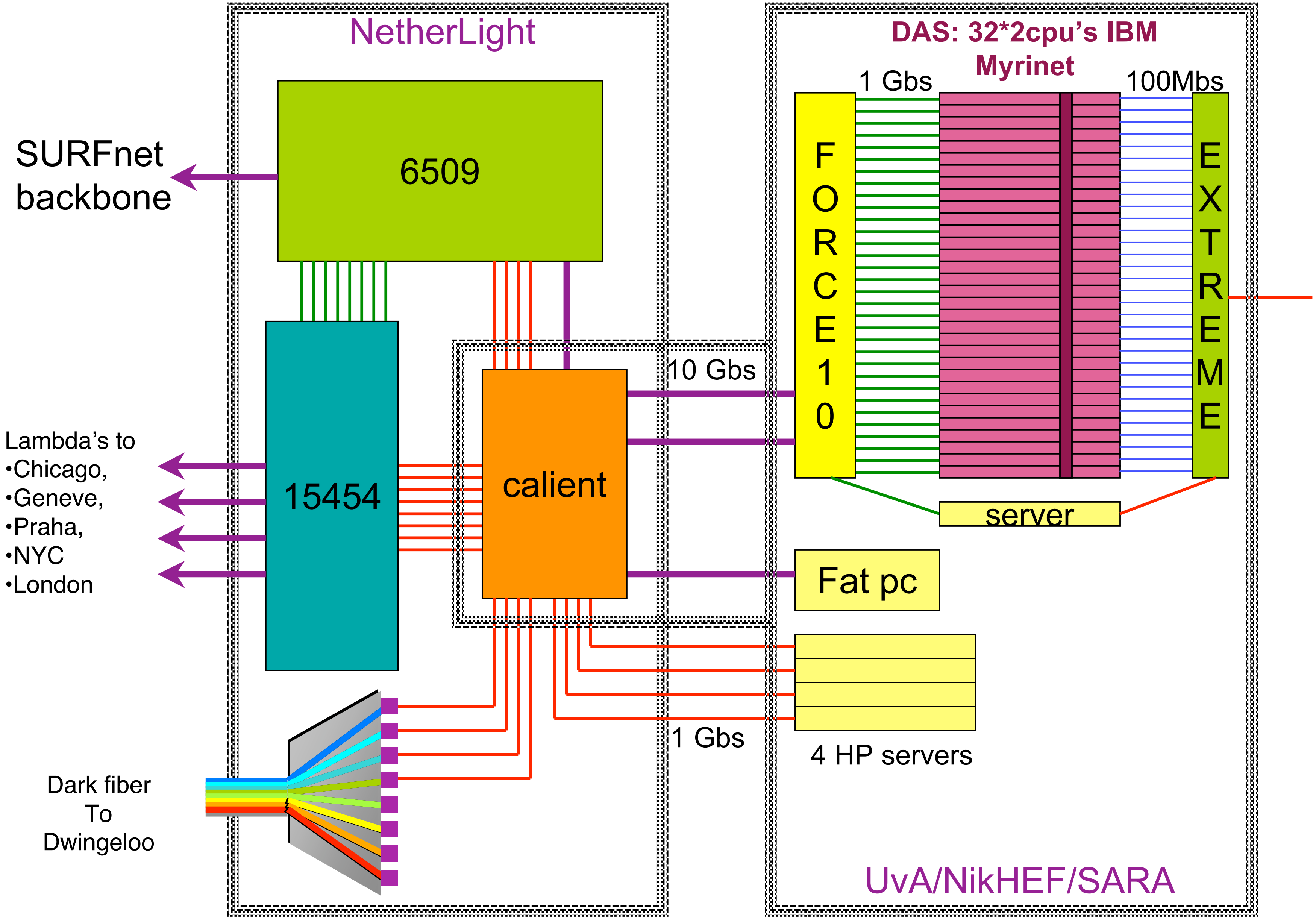




- lambda for high bandwidth applications
  - Bypass of production network
  - Middleware may request (optical) pipe
- RATIONALE:
  - Lower the cost of transport per packet

Virtual Organization on L2



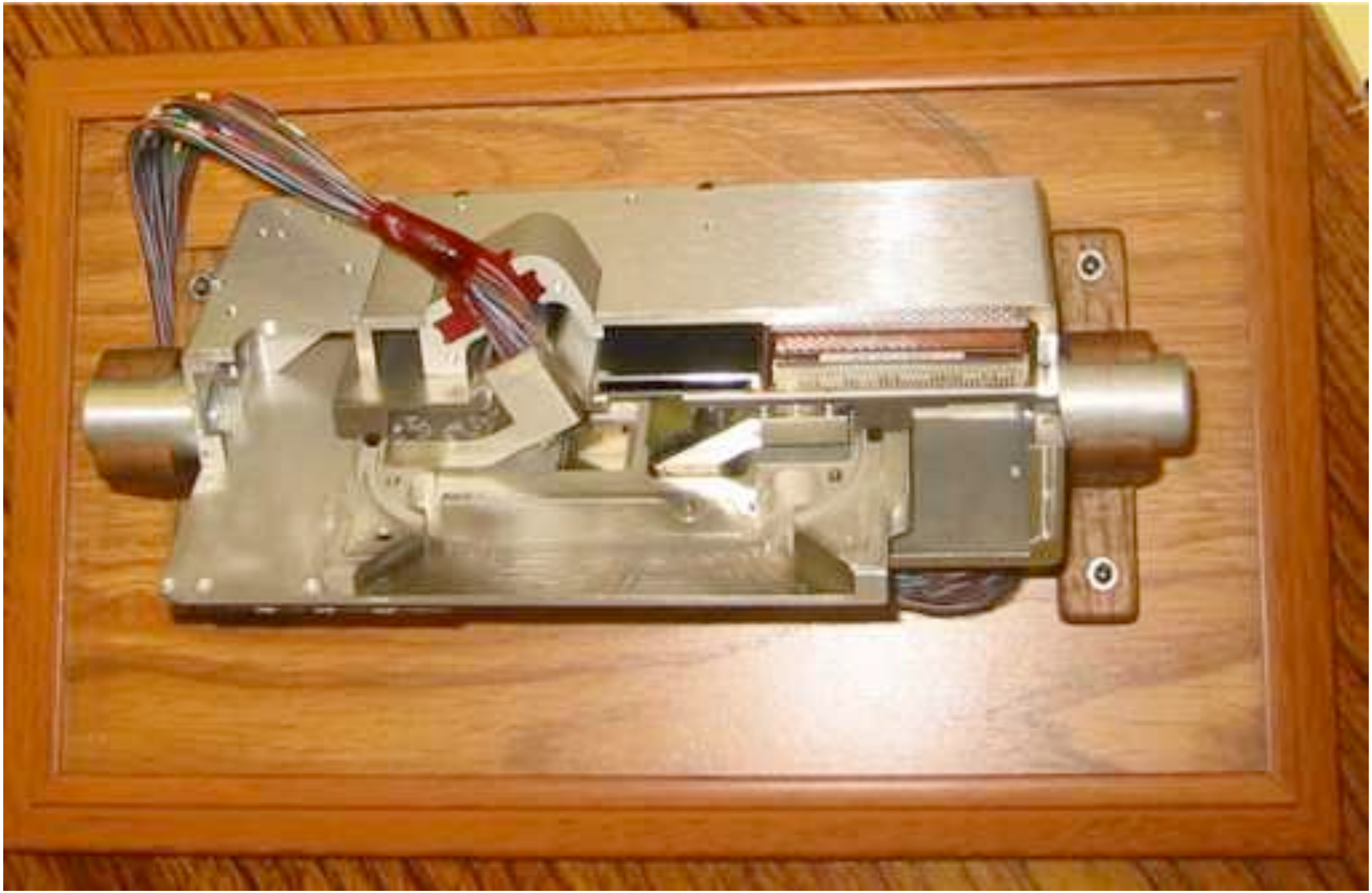


NetherLight
SURFnet backbone
6509
15454
calient
Lambda's to
•Chicago,
•Geneve,
•Praha,
•NYC
•London
Dark fiber
To
Dwingeloo
10 Gbs
1 Gbs
DAS: 32*2cpu's IBM
Myrinet
1 Gbs
100Mbs
FORCE10
EXTREME
server
Fat pc
4 HP servers
UvA/NikHEF/SARA

UVA/EVL's
64*64
Optical Switch
@ NetherLight
in SURFnet POP
@ SARA

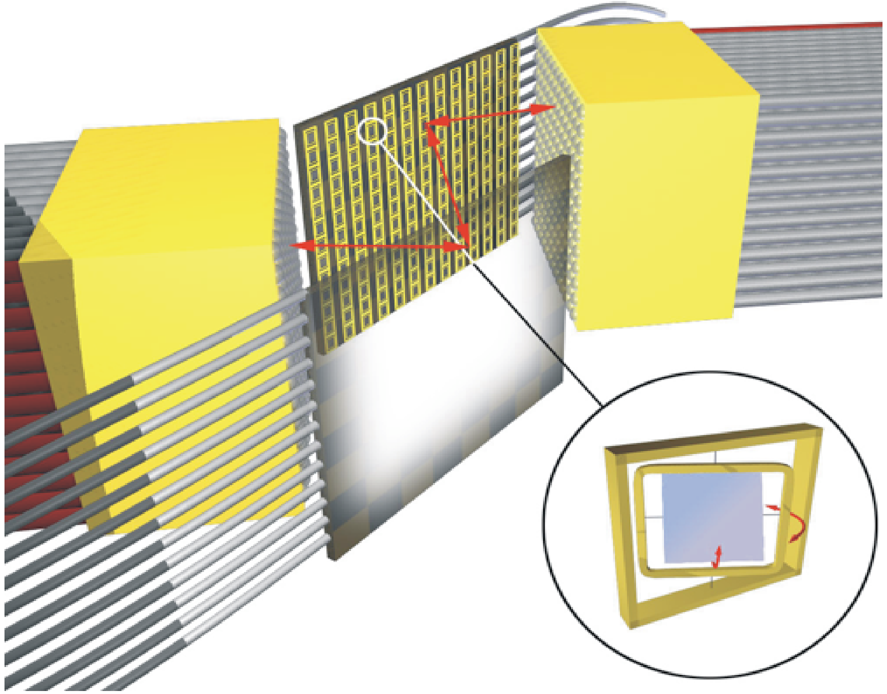
Costs 1/100th of a similar throughput router but with specific services!

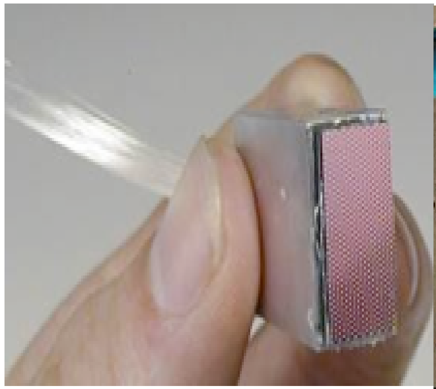
MEMS optical switch (CALIENT)

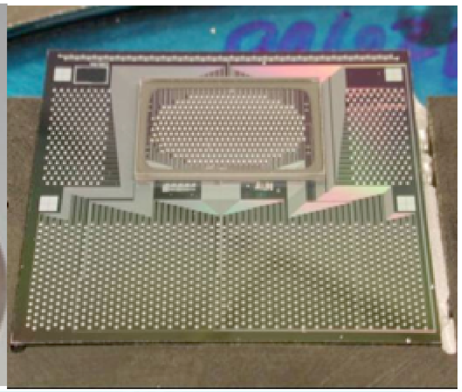


Core Switch Technology

(Intermezzo-3)







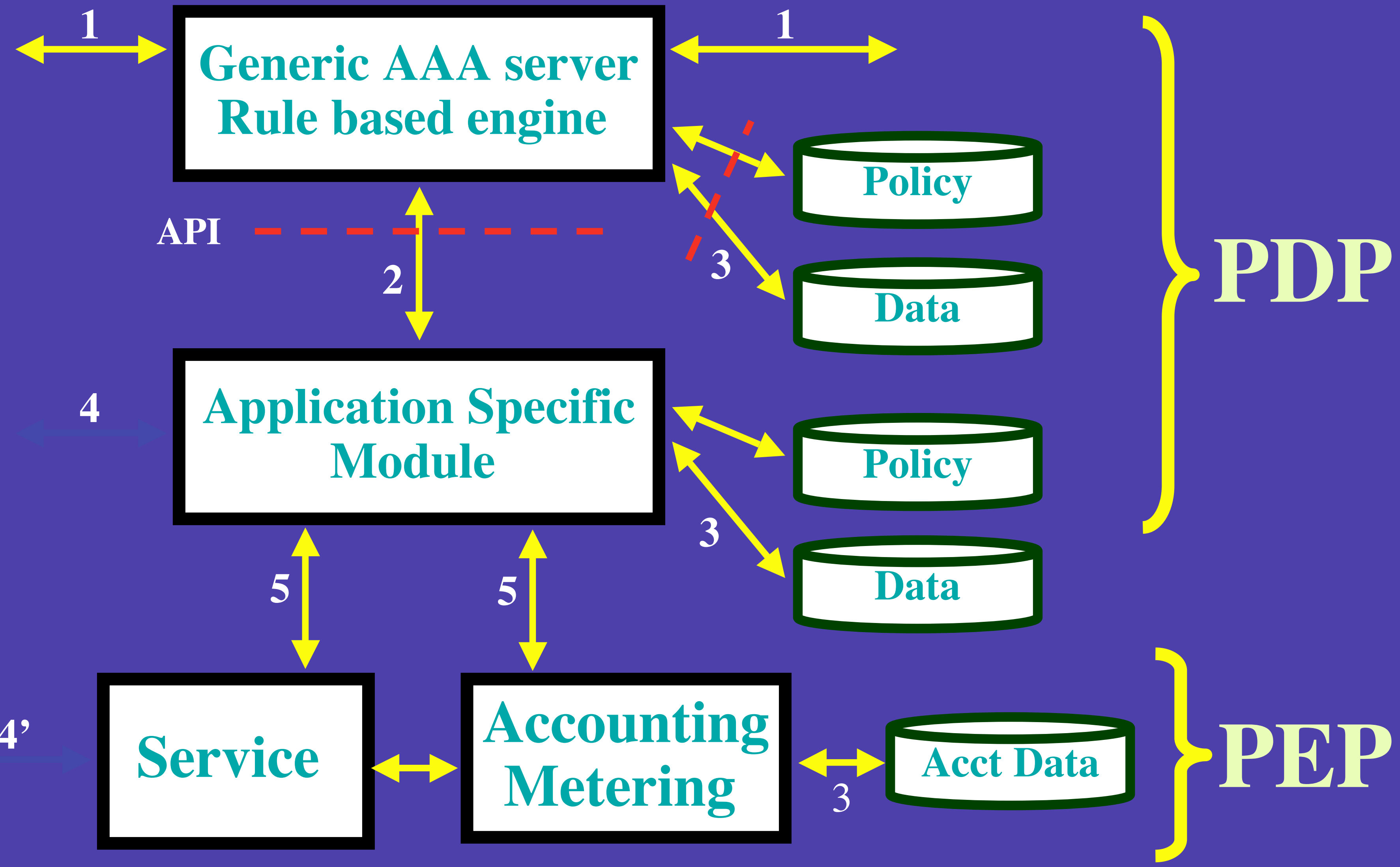
- **3D MEMS structure**
  - Bulk MEMS – High Density Chips
  - Electrostatic actuation
  - Short path length (~4cm)
  - <1.5 dB median loss
- **Completely Non-blocking**
  - Single-stage up to 1Kx1K
  - 10 ms switching time
- **Excellent Transparency**
  - Polarization
  - Bit rate
  - Wavelength

[where innovation comes to light™]

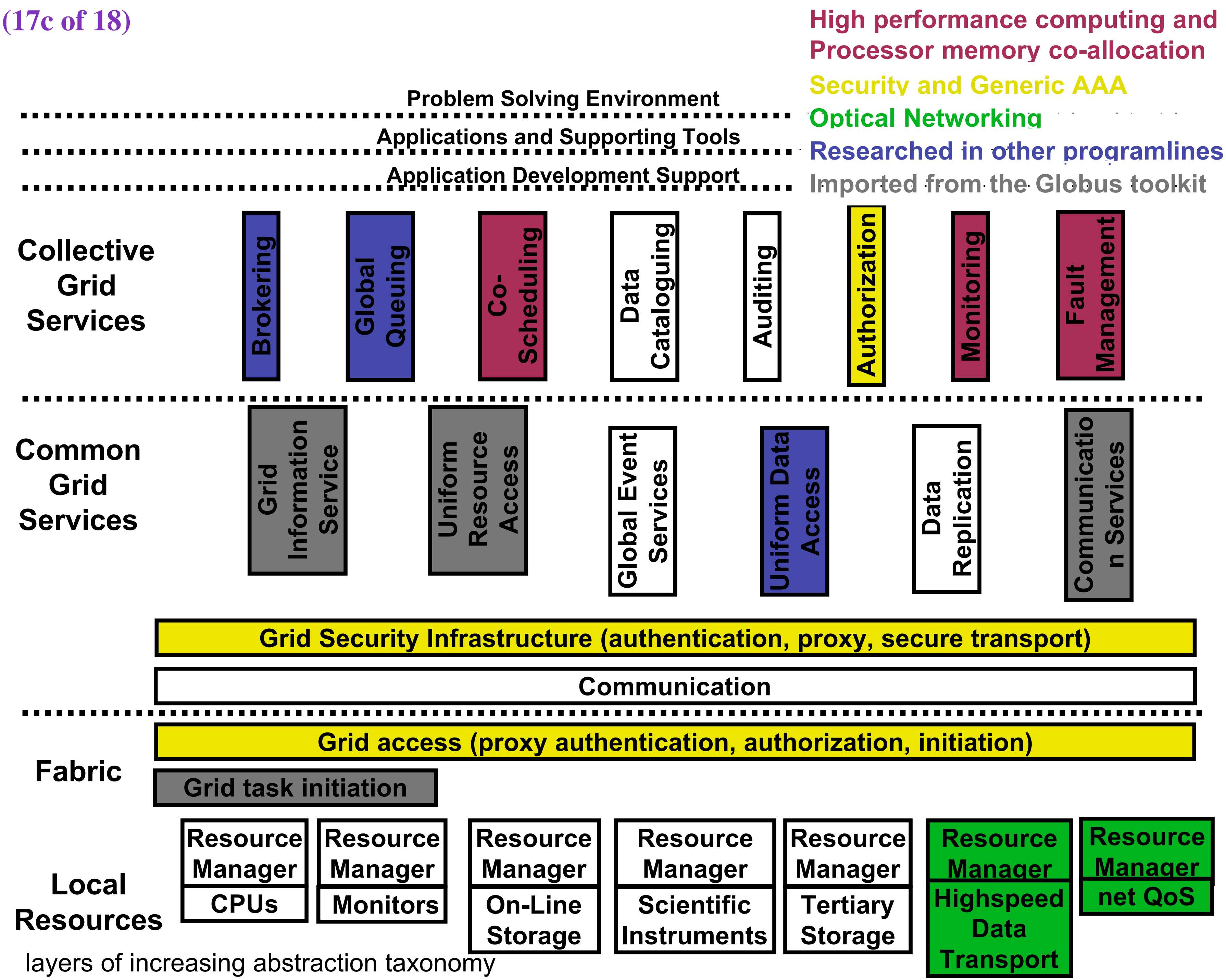
What is a LambdaGrid?

- A *grid* is a set of networked, middleware-enabled computing resources.
- A *LambdaGrid* is a grid in which the lambda networks themselves are resources that can be scheduled, like computing, data, visualization, etc resources. The ability to schedule and provision lambdas provides *deterministic* end-to-end network performance for real-time or time-critical applications, which cannot be achieved on today's grids.

Starting point



RFC 2903 - 2906 , 3334 , policy draft



layers of increasing abstraction taxonomy

(Future) Projects

- National:
  - NCF Grid project
  - VLE
  - GigaPort-NG
  - LOFAR
  - SARA-UVA LG
- European
  - DataGrid & DataTAG
    - UvA, NikHEF, SARA
  - DEISA
- International
  - NetherLight
  - StarLight
  - UKLight
  - AnyLight, LowLight, BackLight
  - Optiputer

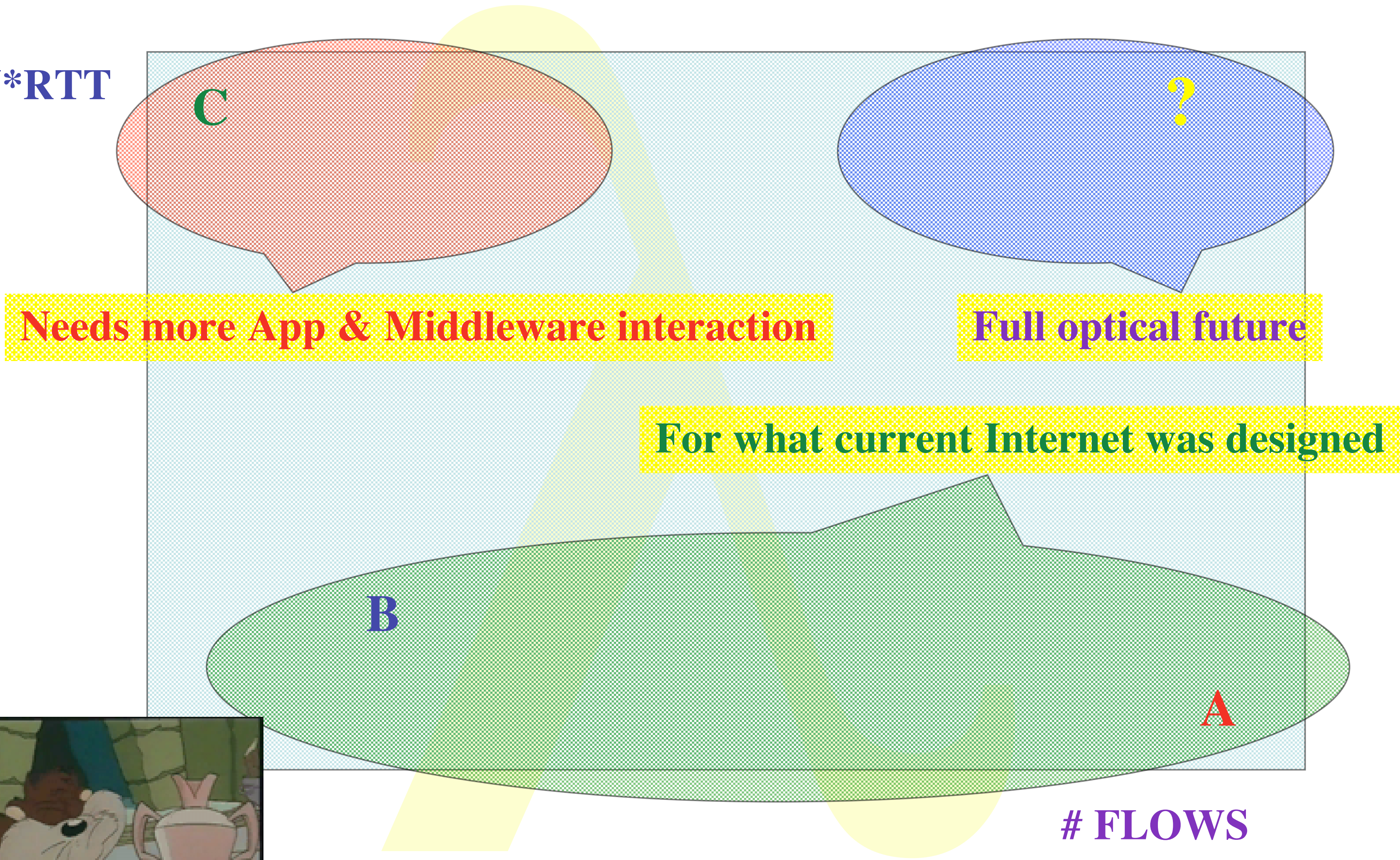
Research:

Models of Lambda networking

Transport

AAA

Transport in the corners



The END

Thanks to

SURFnet: Kees Neggers,UIC&iCAIR: Tom DeFanti, Joel Mambretti, CANARIE: Bill St. Arnaud

This work is supported by:SURFnet, EU-IST project DATATAG, SARA