

Open Exchanges

Where policy matters!

Cees de Laat

University of Amsterdam

Based on articles:

Freek Dijkstra, Cees de Laat, "Optical Exchanges", GRIDNETS conference proceedings, oct 2004,
<http://www.broadnets.org/2004/workshop-papers/Gridnets/DijkstraF.pdf>

And

Freek Dijkstra, Bas van Oudenaarde, Bert Andree, Leon Gommans, Paola Grosso, Jeroen van der Ham,
Karst Koymans and Cees de Laat, "A Terminology for Control Models at Optical Exchanges", LCNS, Volume
4543, july 2007, Page 49-60
<http://ext.delaat.net/pubs/2007-c-9.pdf>

See also:

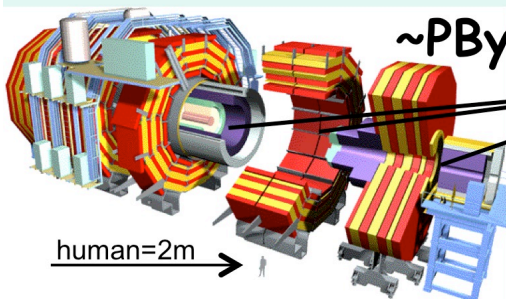
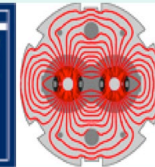
Freek Dijkstra, thesis "Framework for Path Finding in Multi-Layer Transport Networks", UvA, june 2009
<http://ext.delaat.net/pubs/2009-t-1.pdf>





LHC Data Grid Hierarchy

CMS as example, Atlas is similar



human=2m

~PByte/sec

Online System

Tier 0 + 1

~100 MBytes/sec

100000 flops/byte

10 Pflops/s

event simulation

event reconstruction

CMS detector: 15m X 15m X 22m
12,500 tons, \$700M.

Tier 1

~2.5 Gbits/sec

Italian Regional Center

German Regional Center

NIKHEF Dutch Regional Center

FermiLab, USA Regional Center

analysis

~0.6-2.5 Gbps

Tier2 Center

Center

Center

Center

Center

Tier 2

Tier 3

~0.6-2.5 Gbps

Institute

Institute

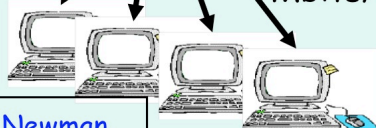
Institute

Institute

Physics data cache

100 - 1000 Mbits/sec

Tier 4



Workstations

CERN/CMS data goes to 6-8 Tier 1 regional centers, and from each of these to 6-10 Tier 2 centers.

Physicists work on analysis "channels" at 135 institutes. Each institute has ~10 physicists working on one or more channels.

2000 physicists in 31 countries are involved in this 20-year experiment in which DOE is a major player.

Courtesy Harvey Newman, CalTech and CERN

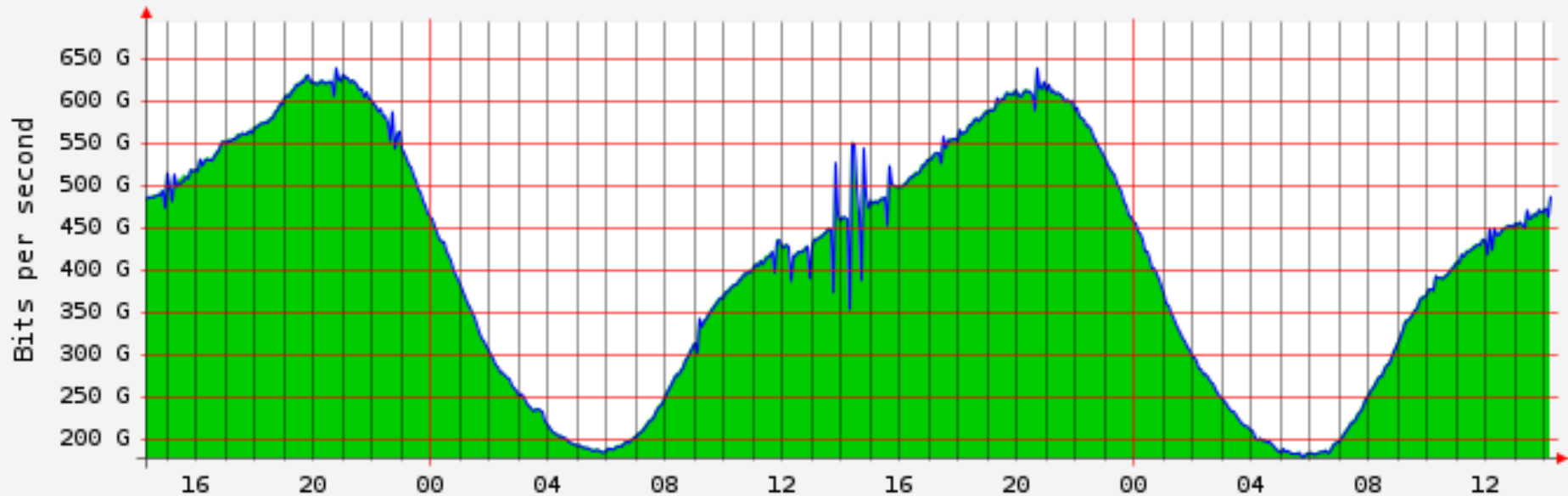
u
s
e
r
s

A. Lightweight users, browsing, mailing, home use

Need full Internet routing, one to all

B. Business/grid applications, multicast, streaming, VO's, mostly LAN

Need VPN services and full Internet routing, several to several + unlink to all



■ Input ■ Output

Peak In : 641.166 Gb/s Peak Out : 639.212 Gb/s

Average In : 413.749 Gb/s Average Out : 413.612 Gb/s

Current In : 488.105 Gb/s Current Out : 487.341 Gb/s

Copyright (c) 2009 AMS-IX B.V. [updated: 19-Feb-2009 14:15:20 +0100]

B

C

ADSL (12 Mbit/s)

BW **GigE**

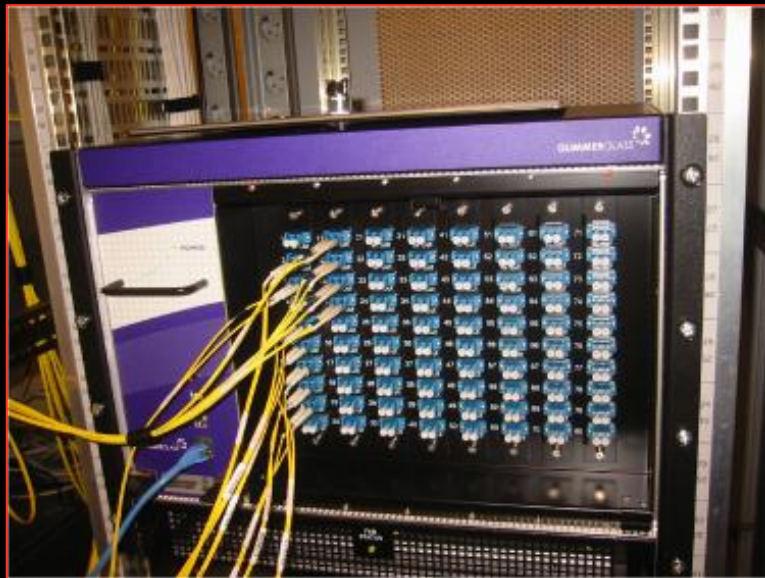
Ref: Cees de Laat, Erik Radius, Steven Wallace, "The Rationale of the Current Optical Networking Initiatives"
iGrid2002 special issue, Future Generation Computer Systems, volume 19 issue 6 (2003)



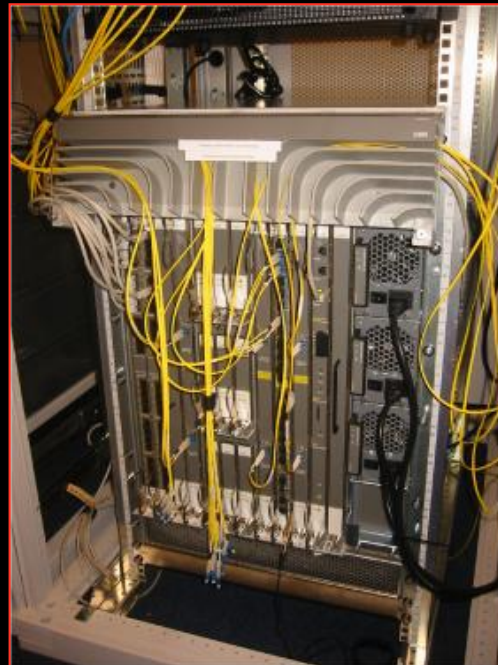
Towards Hybrid Networking!

- Costs of photonic equipment 10% of switching 10 % of full routing
 - for same throughput!
 - Photonic vs Optical (optical used for SONET, etc, 10-50 k\$/port)
 - DWDM lasers for long reach expensive, 10-50 k\$
- Bottom line: look for a hybrid architecture which serves all classes in a cost effective way
 - map A -> L3 , B -> L2 , C -> L1 and L2
- Give each packet in the network the service it needs, but no more !

L1 \approx 2-3 k\$/port



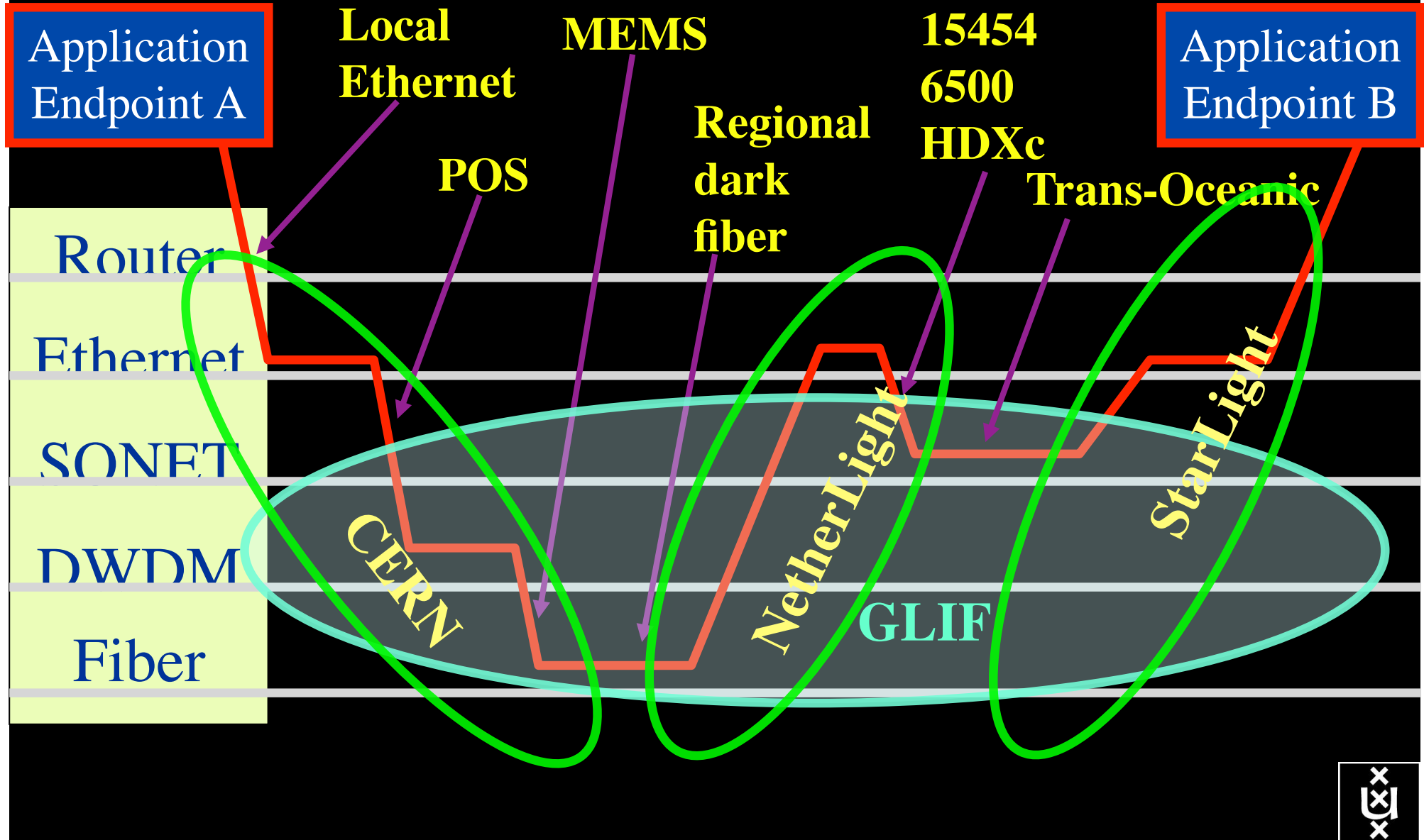
L2 \approx 5-8 k\$/port



L3 \approx 75+ k\$/port



How low can you go?



Services

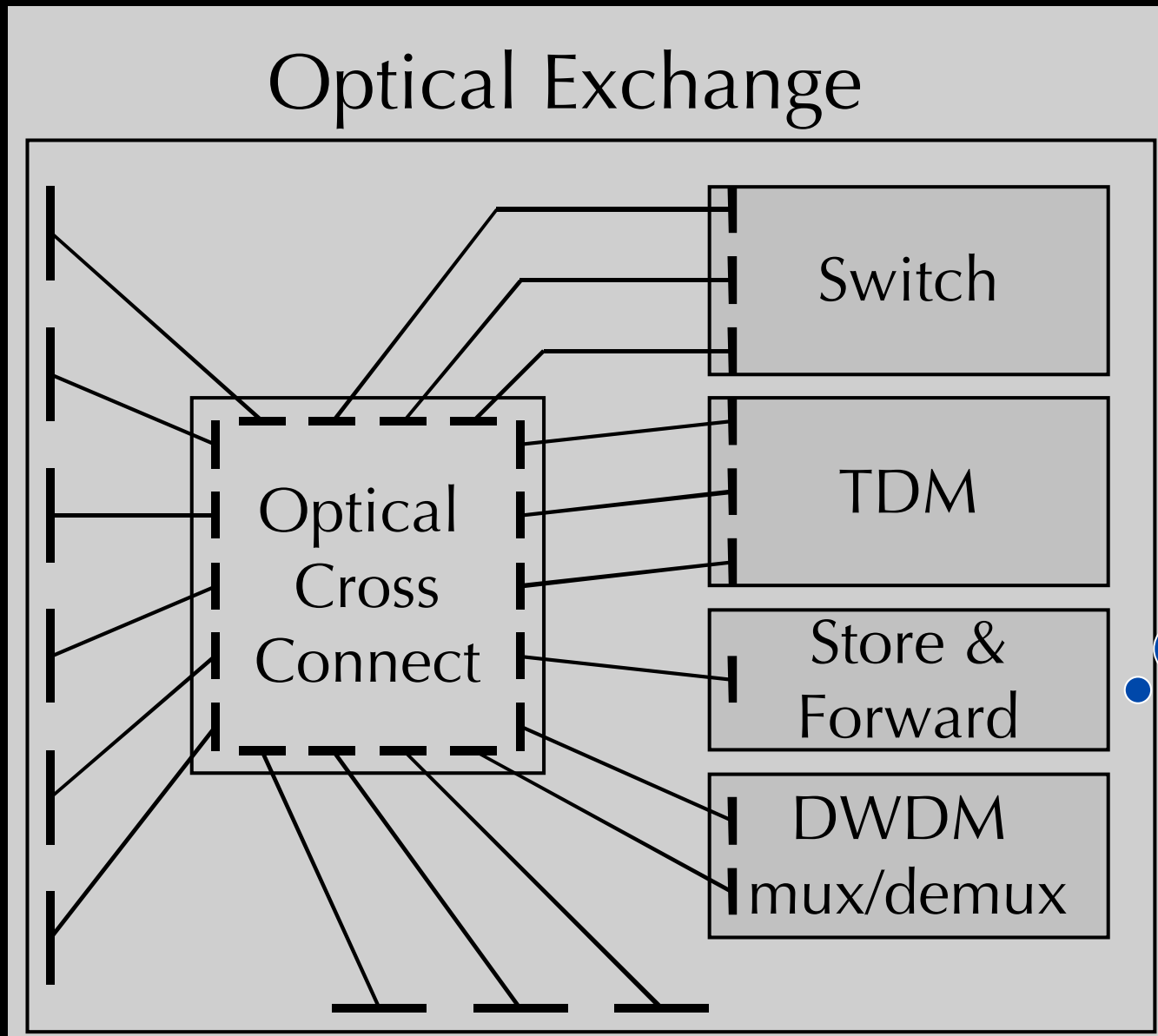
<div style="text-align: center;">SCALE</div> <div style="text-align: center;">CLASS</div>	<div style="text-align: center;">2 Metro</div>	<div style="text-align: center;">20 National/ regional</div>	<div style="text-align: center;">200 World</div>
<div style="text-align: center;">A</div>	<div style="text-align: center;">Switching/ routing</div>	<div style="text-align: center;">Routing</div>	<div style="text-align: center;">ROUTER\$</div>
<div style="text-align: center;">B</div>	<div style="text-align: center;">Switches + NGE VPN's</div>	<div style="text-align: center;">Switches + NGE (G)MPLS</div>	<div style="text-align: center;">ROUTER\$</div>
<div style="text-align: center;">C</div>	<div style="text-align: center;">dark fiber DWDM MEMS switch</div>	<div style="text-align: center;">DWDM, TDM / SONET Lambda switching</div>	<div style="text-align: center;">Lambdas, VLAN's SONET Ethernet</div>



Service Matrix

From	To	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+

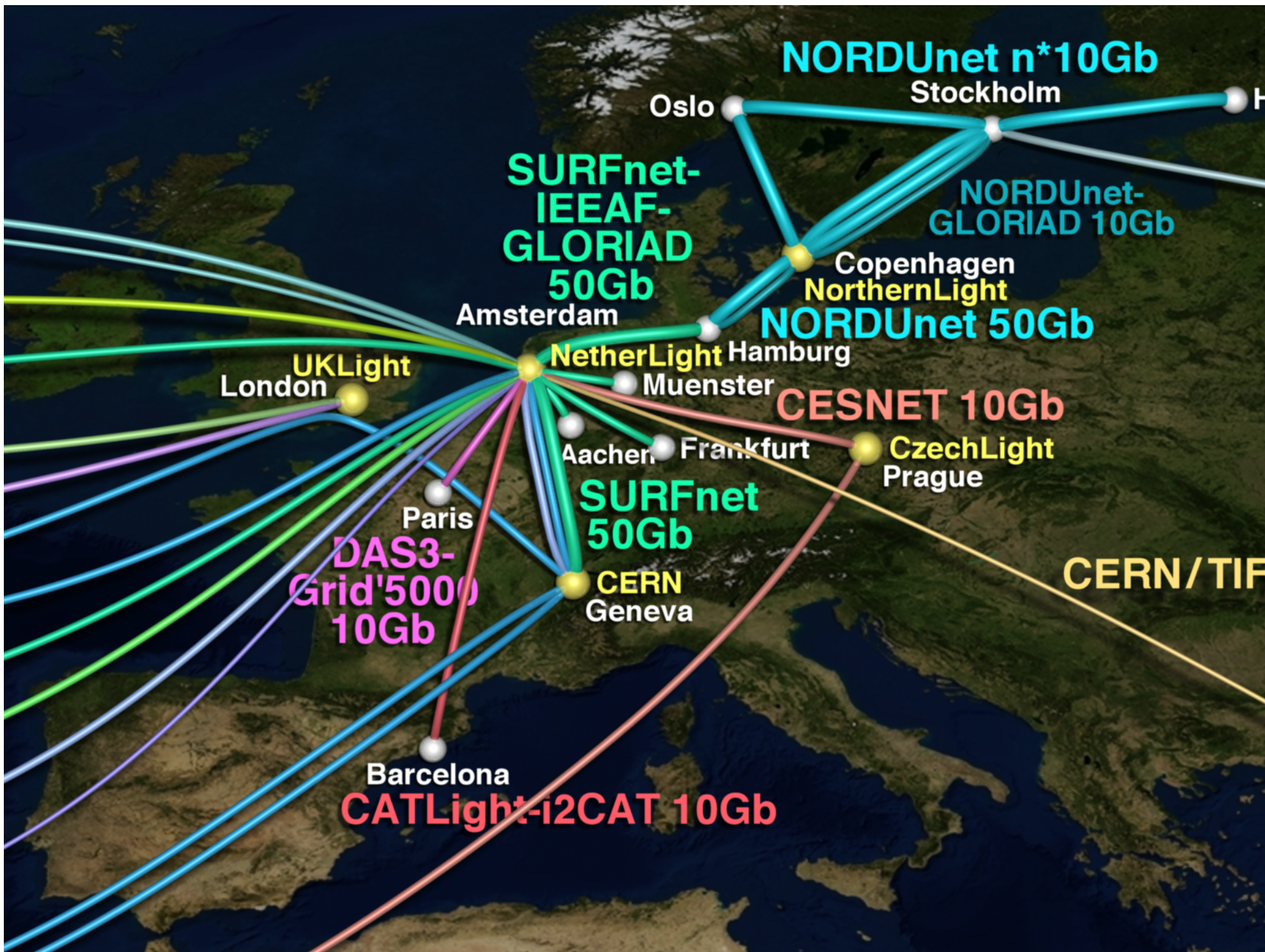
Optical Exchange as Black Box



- TeraByte
- Email
- Service

Ref: gridnets paper by Freek Dijkstra, Cees de Laat





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:**

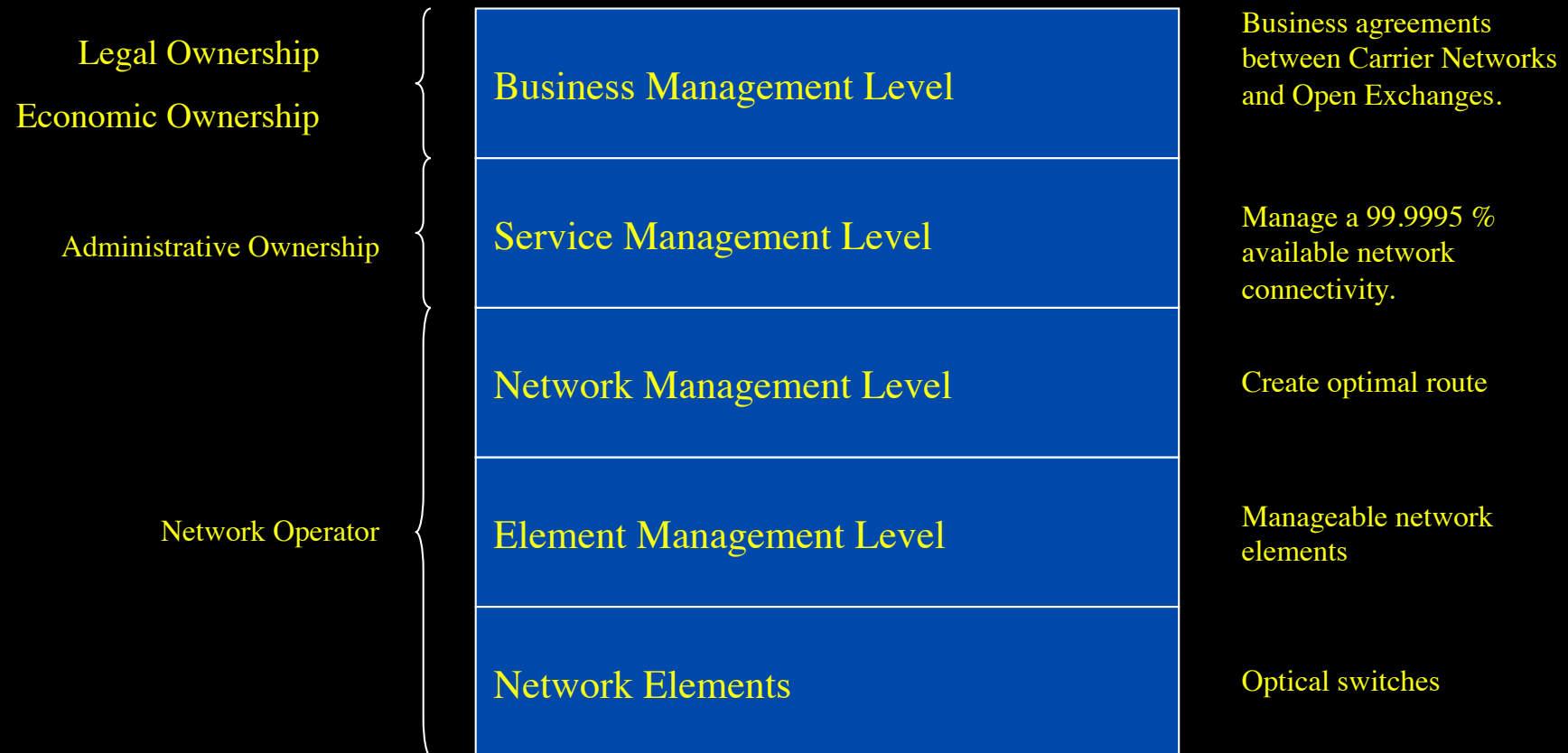
- Technically implements the terms of a service level agreement
- Signals requests to other AO's and handles responses.
- Collects accounting information.

- **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.



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



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 .



Possible roles and objectives

- **VO:**

- Administrative entity that administers VO memberships
- Administers technical implementations of contracts with ELO's regarding link usage rights.
- Technically delegates ELO link usage rights to VO members.

- **VO Member:**

- needs on-demand high-volumes traffic exchanges with their peers via one or more links owned by and ELO.
- is connected via a NRN to an Optical Exchange.
- sends connection requests to NRN containing delegated VO rights

- **National Research Network:**

- Operates a hybrid network infrastructure that allows re-direction of VO member traffic destined to ELO links.
- Is LLO and ELO and ALO.
- Non-ELO traffic is directed towards the regular Internet.
- May re-advertise link availability obtained from Optical Exchange.
- Acts as proxy for VO member connection requests for ELO links.



Possible roles and objectives

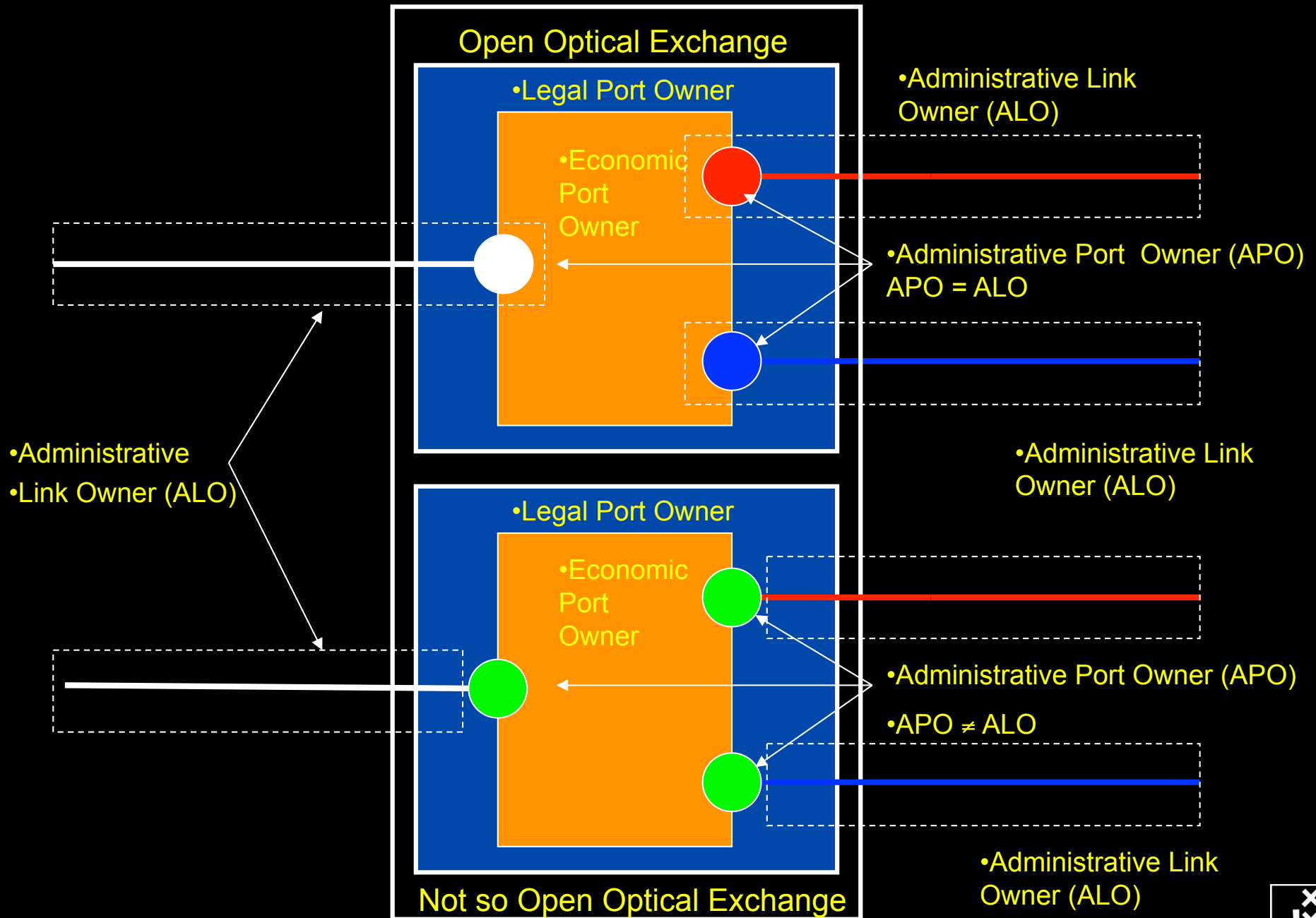
- **Optical Exchange:**

- Legal: Organization that facilitates LPO's and allow LLO's to terminate their links.
- Economic: Sells housing, rack space, termination facilities etc.
- Administrative: administers assets. Advertises topology info to other Optical Exchanges & NRNs ?

- **Open Optical Exchange:** (Part of) an optical exchange where ALO's also play the role of an APO.



Optical Exchange Stakeholders



Bottom line

In an open exchange the exchange is not involved in the decision for two peers to exchange traffic!



Questions ?

