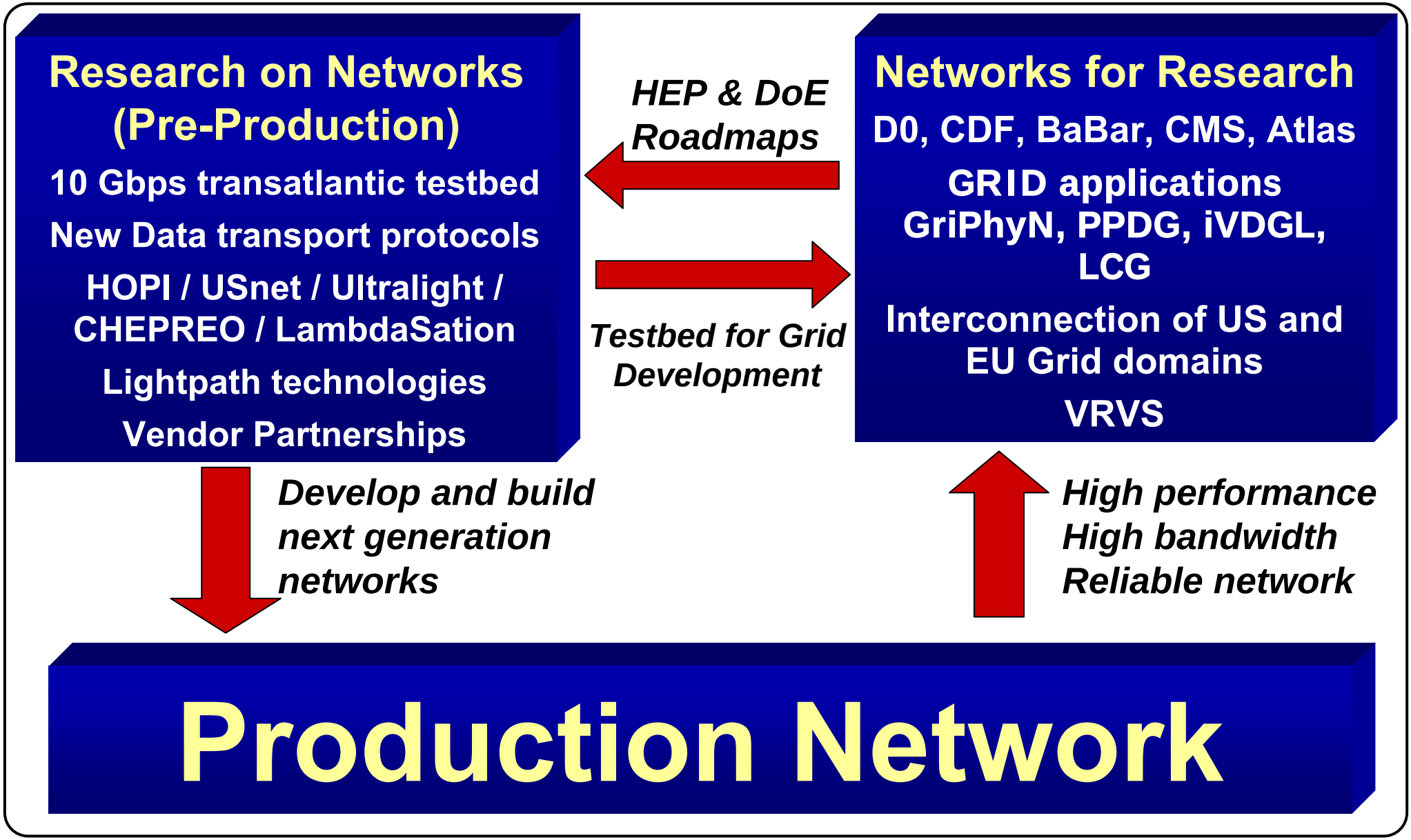
LHCNet topology and Operation

US T0-T1 Meeting

July 7, 2005

Sylvain Ravot

LHCNet mission



LHCNet services

"providing the right cost-effective transatlantic network infrastructure for HEP community's needs"

- Commodity Internet (at MANLAN)
- Access to US T1
  - FNAL, BNL
- Access from US T2s
  - ~10 Official, and some other large facilities (Cornell, Buffalo, etc.)
- Peerings with Abilene, ESNet and others
  - At Chicago (primary?) and New-York (secondary?)
- T1-T1 traffic?; T1-T2 traffic?
- Pre-Production
  - High speed disk-to-disk throughput
    - New end-systems (PCI-e; 64 bit cpu; New 10 GE NICs)
    - New data transport protocols (FAST and others)
  - Prototype the data movement services between CERN and the US
  - Monitoring, Command and Control (Monalisa: networks and Grids)
  - Circuit switching ("optical" control plane; light paths)
  - GLIF
  - UltraLight/LambdaStation; Access to HOPI (Internet2/NSF) and USNet (DOE) testbeds

Type of services

- Define needs and services:
  - Connection oriented
  - Packet switched vs circuit switched
  - Shared bandwidth vs dedicated bandwidth
  - L1, L2 and L3 services

Service	L1	L2	L3	Availability	Backup	Bandwidth in 2006	
Commodity Internet	No	No	Yes	99,999%	At L3	100 Mbps	
Connections to T1	Yes	Yes	Yes	99%	At ??	2x10Gbps	
Access to Abilene ESNet	No	??	Yes	99,999%	At L3	??	
Pre-production	Yes	Yes	Yes	90%	No	>10Gbps	

Major US-Partners

Chicago (Starlight)

- FNAL (10 Gbps; 6 x 10 Gbps this year)
- ESnet (10 Gbps)
- U. Michigan (10 Gbps; 3 x 10 Gbps this year)
- FIU/UF (10 Gbps via NLR & FLR)
- Caltech (10 Gbps via NLR)
- USnet (2 x 10 Gbps)
- HOPI (2 x 10 Gbps)
- U. Wisconsin Madison (10 Gbps via Starlight)
- TeraGrid (10 Gbps via Starlight)
- Abilene (10 Gbps via Starlight)

9 to 16 10GE waves in 2005

New-York (MANLAN)

- BNL (10 Gbps in 2006)
- ESnet, Abilene (10 Gbps via MANLAN)
- HOPI (2 x 10 Gbps)
- CANARIE (3 x 10 Gbps)
- NLR
- Buffalo (2 x 10 Gbps)
- Atlantic Wave (10 Gbps)

8 to 10 10GE waves in 2005

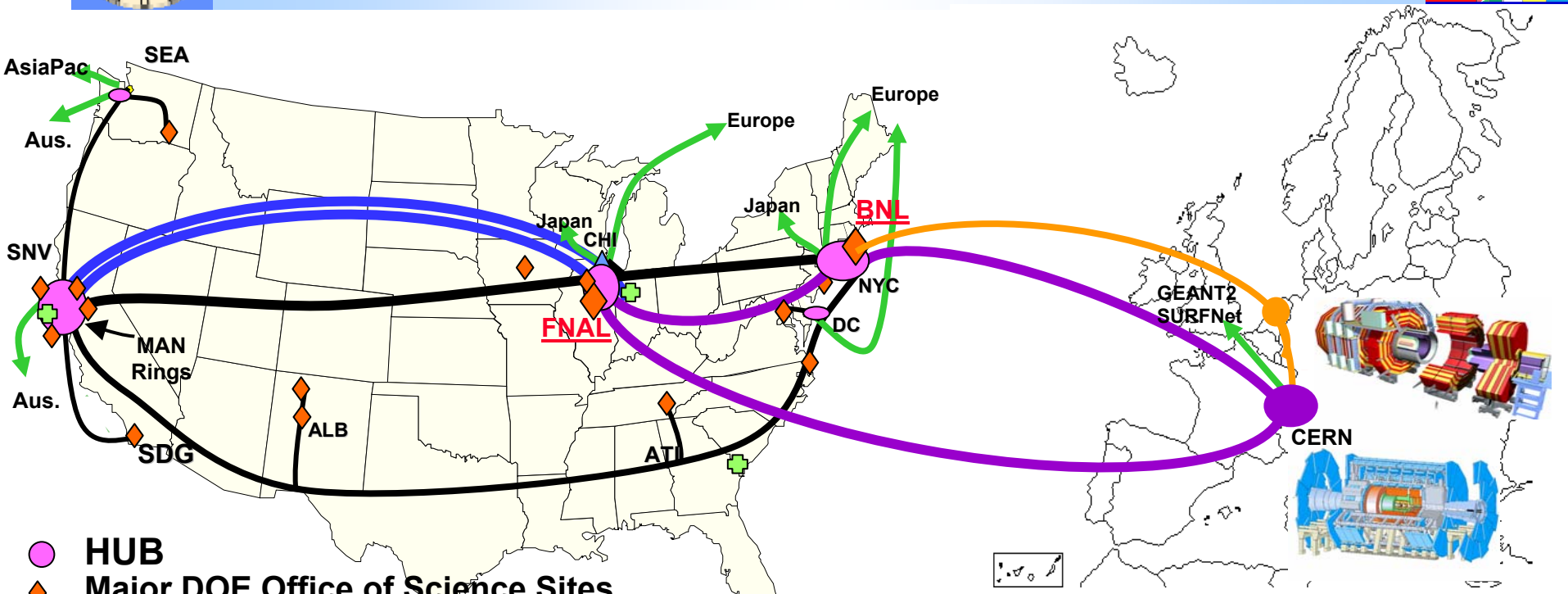
Transatlantic Network Requirements Estimates (in Gbps)

Year	2004	2005	2006	2007	2008	2009	2010
CERN-BNL (ATLAS)	0.5 (Via ESnet)	0.5	5	10	20	30	40
CERN-FNAL (CMS)	2.5	7.5	15	30	40	40	40
Other (ESNet, Tier2, Inter-Regional Traffic …)	1.5	2	10	10-20	20	40	40
TOTAL US-CERN BW	5	10	30	50-60	80	110	120
LHCNet BW	5	10	20	40	60	80	80
Other BW[1] (GEANT, Surfnet, IRNC, Gloriad…)	Backup[2]	Backup	10	10-20	20	30	40

[1] We assume that one-quarter to one-third of the bandwidth required is provided by other networks such as GN2 (the next-generation European backbone), Surfnet, Gloriad or the NSF IRNC link.

[2] The other networks are assumed to provide backup paths in case an LHCNet link goes down, or to provide “overflow” capacity if LHCNet is heavily subscribed.

LHCNet and ESNet Sept. 2005



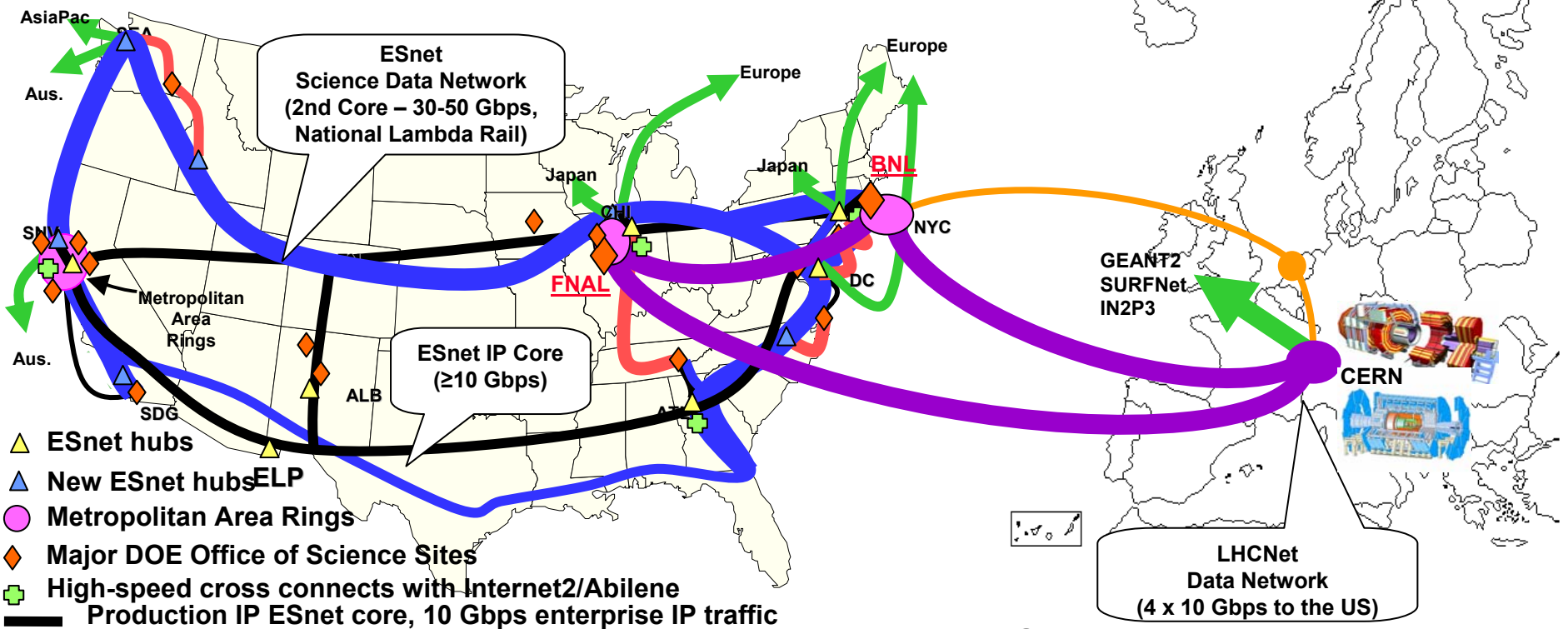
- HUB
- Major DOE Office of Science Sites
- High-speed cross connects with Internet2/Abilene
- Production IP ESnet core, 10 Gbps enterprise IP traffic
- USNet 10 Gbps circuit based transport. (DOE funded project)
- Major international
- LHCNet Data Network (10 Gb/s)
- NSF/IRNC circuit; GVA-AMS connection via Surfnet or Geant2 (10 Gb/s)

10Gb/s

≥ 2.5 Gb/s

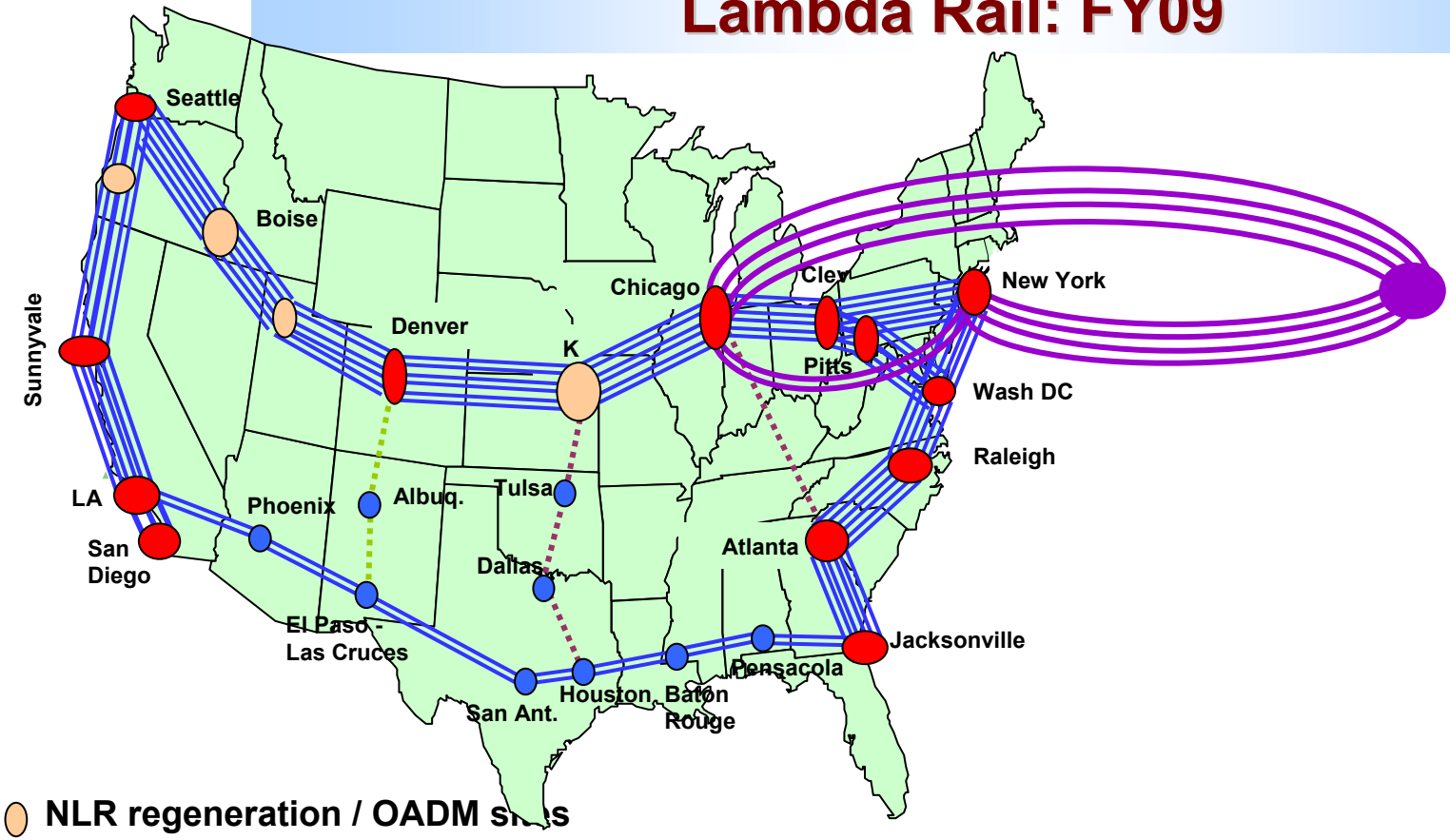
- Connections to ESnet Hubs in New-York and Chicago
- Multi-lambdas (10 Gbps) to BNL and FNAL
- Access to USNet for R&D

ESnet, NSF/IRNC and LHCNet – 2007/2008

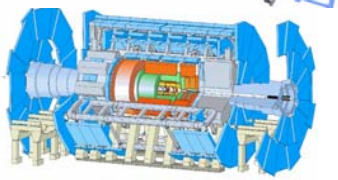


- End-to-end circuit provisioning across LHCNet and ESnet
- Provisioning will initially be provided by manual circuit configuration, managed on-demand in the future
- Technology collaboration with Glif, ESNet, Internet2 / Abilene…

LHCNet connection to Proposed ESnet Lambda Infrastructure Based on National Lambda Rail: FY09



CERN (Geneva)



NLR regeneration / OADM sites

NLR wavegear sites

ESnet via NLR (10 Gbps waves)

LHCNet (10 Gbps waves)

- LHCNet: To ~80 Gbps by 2009
- Routing + Dynamic managed circuit provisioning

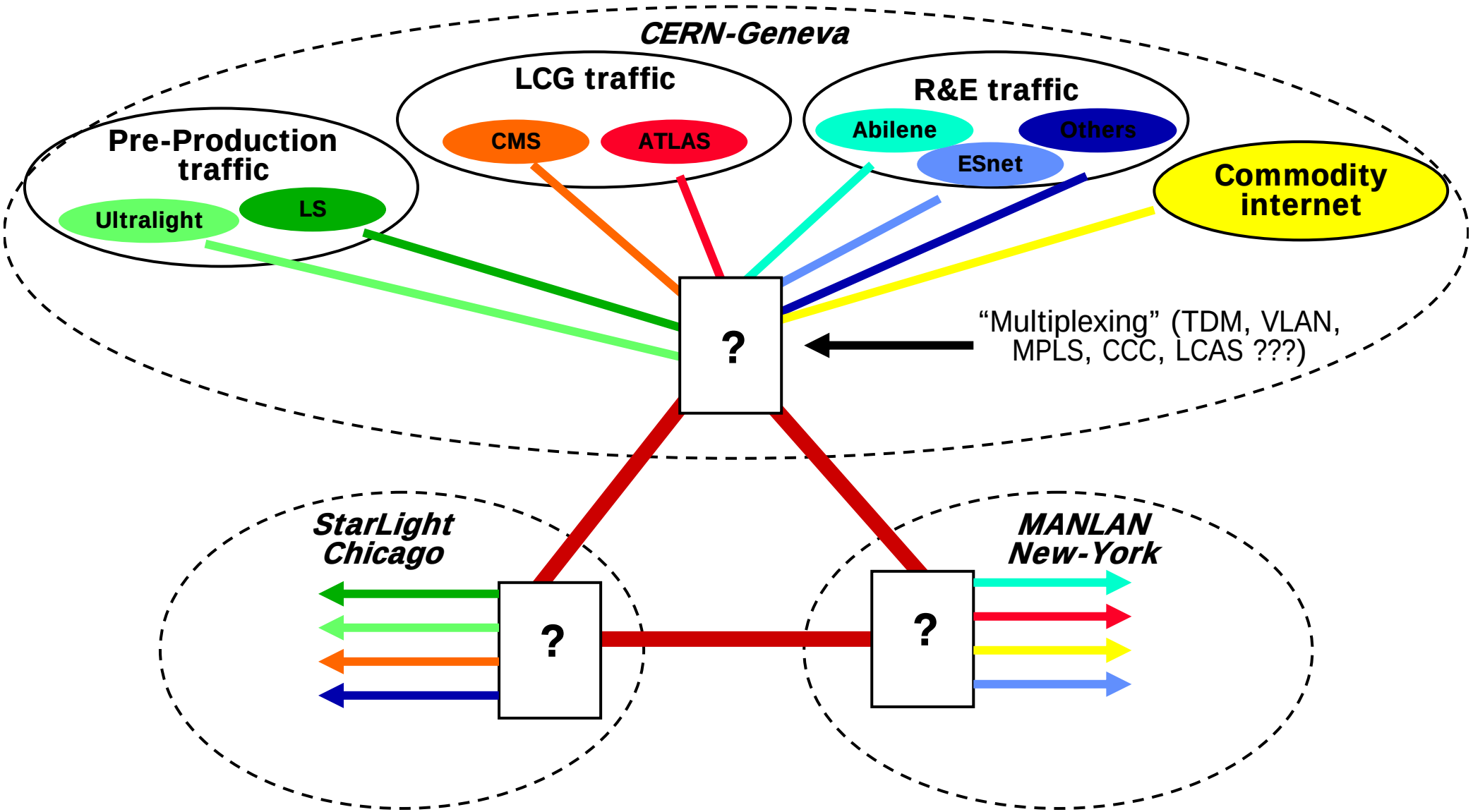
New Topology

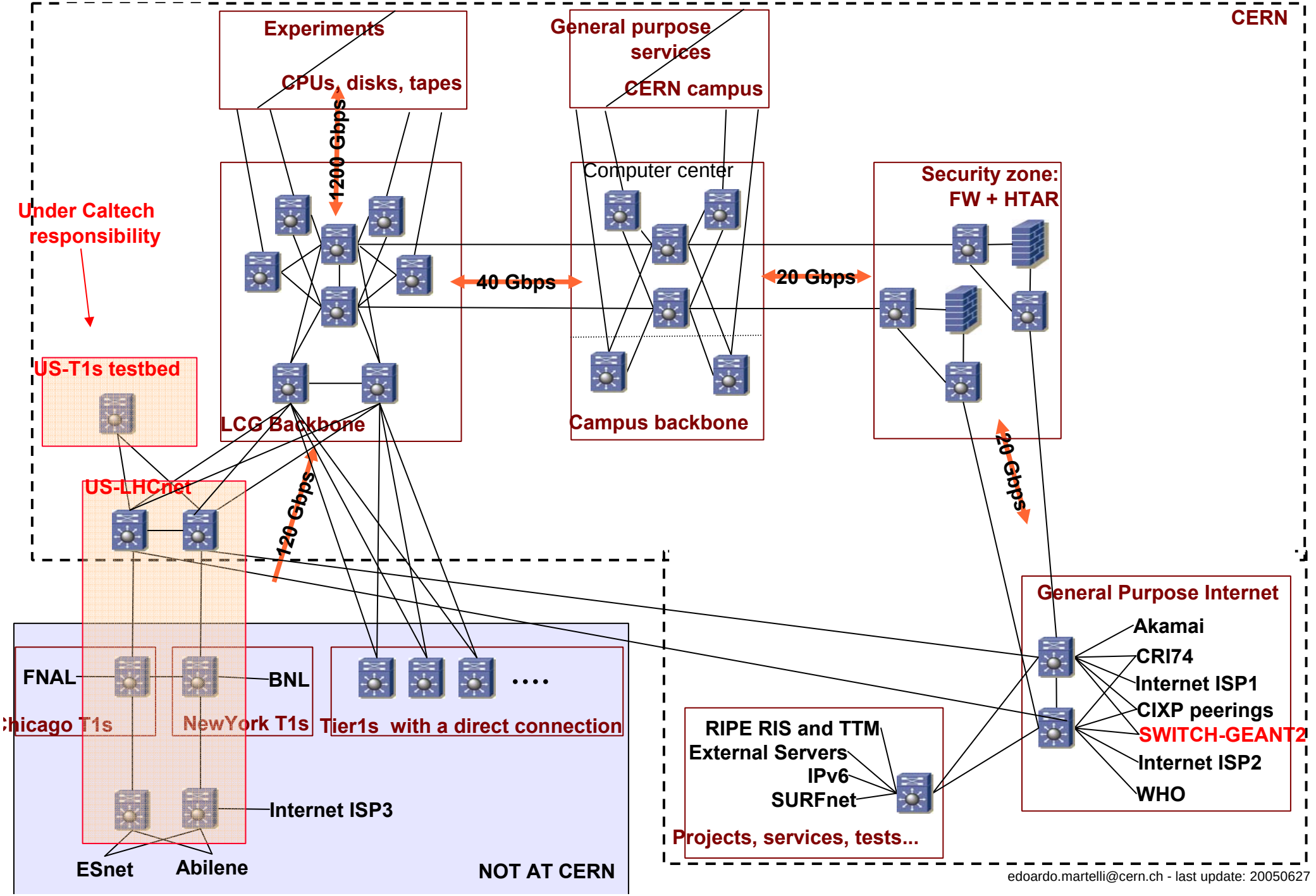
- Constraints
  - SONET Framing
  - Cost
  - “Symmetric” setup on both sides of the Atlantic
  - Redundancy
  - Flexibility at all layers
  - Scalable
- Technology reliability and interoperability
  - WAN-PHY
  - GFP capable products
- Pre-Production
  - Prepare each year for the production network of the following year
  - Testbed directly attached to the LHCNet
  - Enough resources to fully saturate 2 10 Gbps links now.
    4 10 Gbps links by (?) Fall 2006 ? Or Spring 2007 ?

New Technology Candidates: Opportunities and Issues

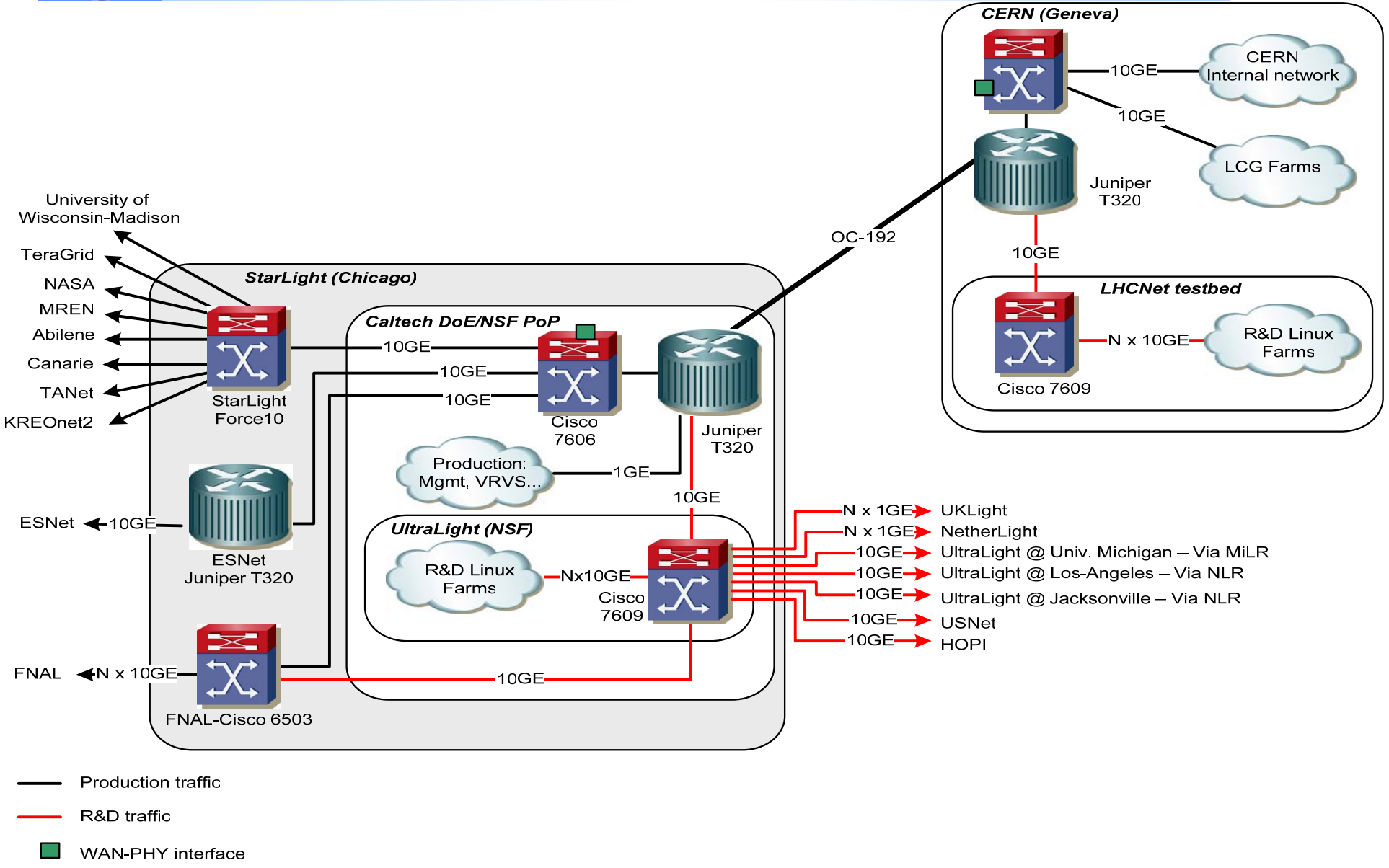
- **New standard for SONET infrastructures**
  - Alternative to the expensive Packet-Over-Sonet (POS) technology currently used
  - May change significantly the way in which we use SONET infrastructures
- **10 GE WAN-PHY standard**
  - Ethernet frames across OC-192 SONET networks
  - Ethernet as inexpensive linking technology between LANs & WANs
  - Supported by only a few vendors
- **LCAS/VCAT standards**
  - Point-to-point circuit-based services
  - Transport capacity adjustments according to the traffic pattern.
  - “Bandwidth on Demand” becomes possible for SONET network
- **NEW standards and NEW hardware**
  - Intensive evaluation and validation period
  - WAN-PHY tests in July 2005
  - LCAS/VCAT tests at the end of 2005 ?

Design

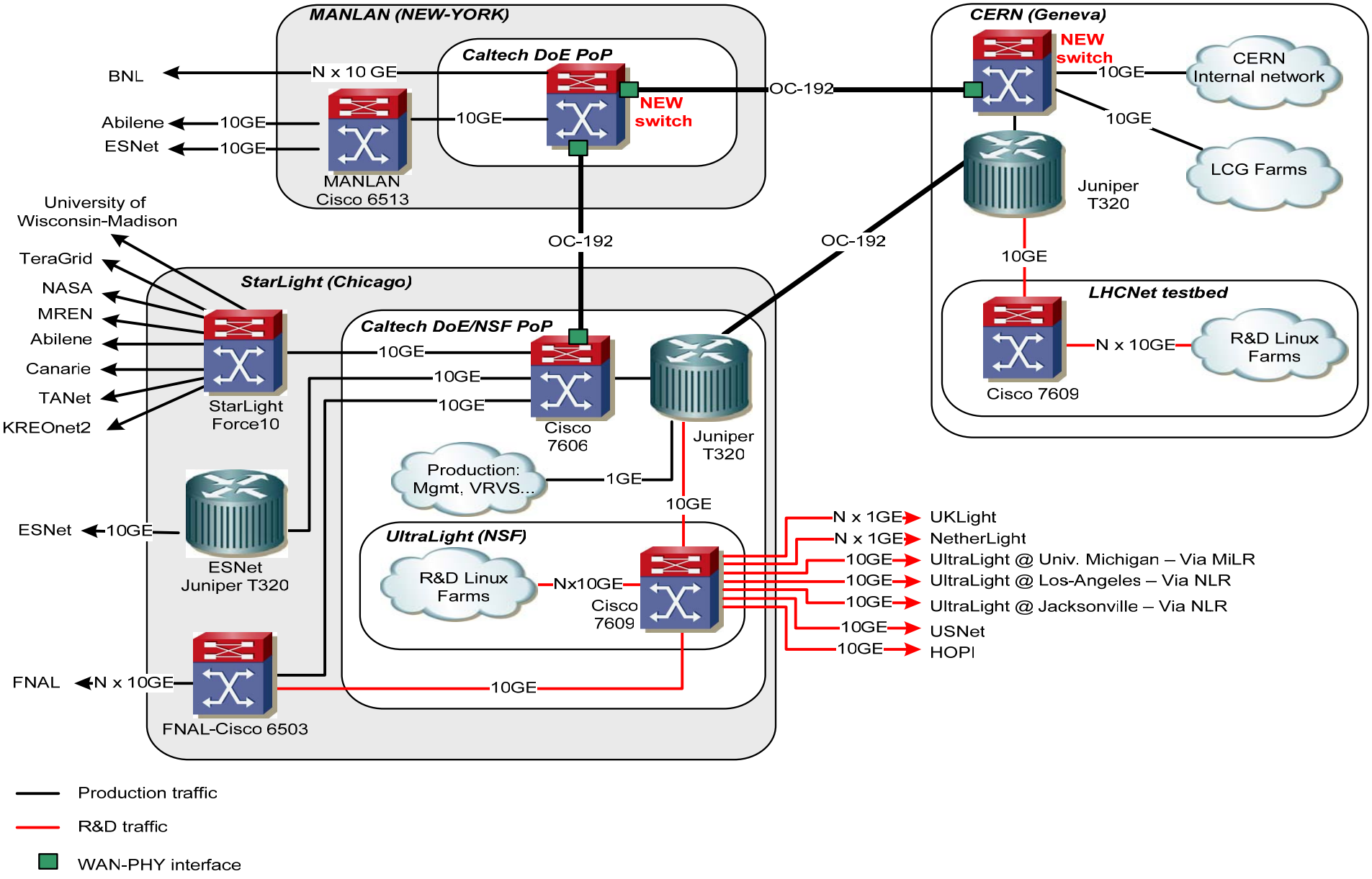

CERN-Geneva
LCG traffic
CMS
ATLAS
R&E traffic
Abilene
Others
ESnet
Pre-Production traffic
Ultralight
LS
Commodity internet
?
"Multiplexing" (TDM, VLAN, MPLS, CCC, LCAS ???)
StarLight Chicago
?
MANLAN New-York
?


CERN
Experiments
CPUs, disks, tapes
General purpose services
CERN campus
1200 Gbps
Computer center
Security zone:
FW + HTAR
Under Caltech responsibility
40 Gbps
20 Gbps
US-T1s testbed
LCG Backbone
Campus backbone
20 Gbps
US-LHCnet
120 Gbps
General Purpose Internet
Akamai
CRI74
Internet ISP1
CIXP peerings
SWITCH-GEANT2
Internet ISP2
WHO
FNAL
BNL
Chicago T1s
NewYork T1s
Tier1s with a direct connection
....
RIPE RIS and TTM
External Servers
IPv6
SURFnet
Projects, services, tests...
Internet ISP3
ESnet
Abilene
NOT AT CERN
edoardo.martelli@cern.ch - last update: 20050627

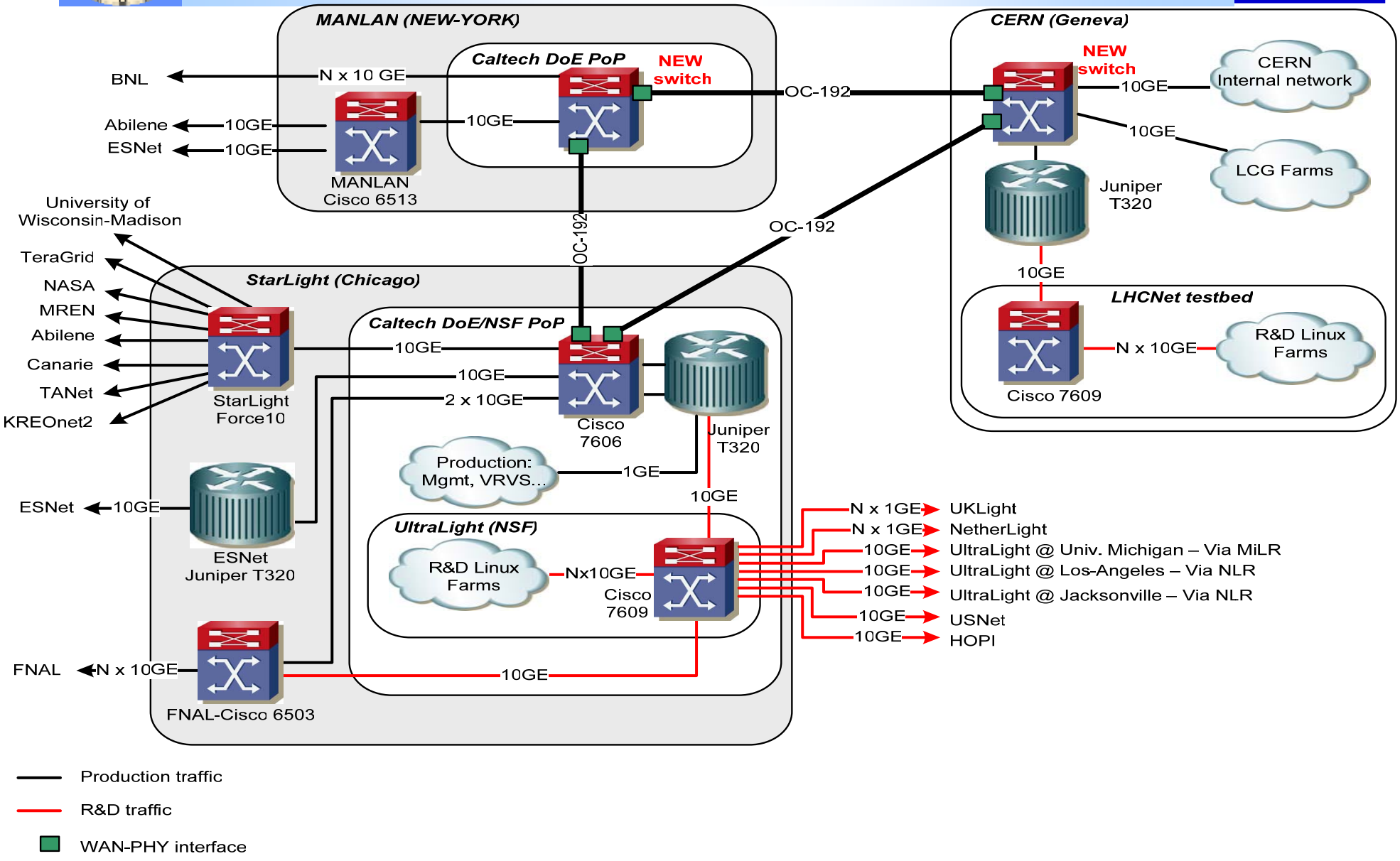
Current setup



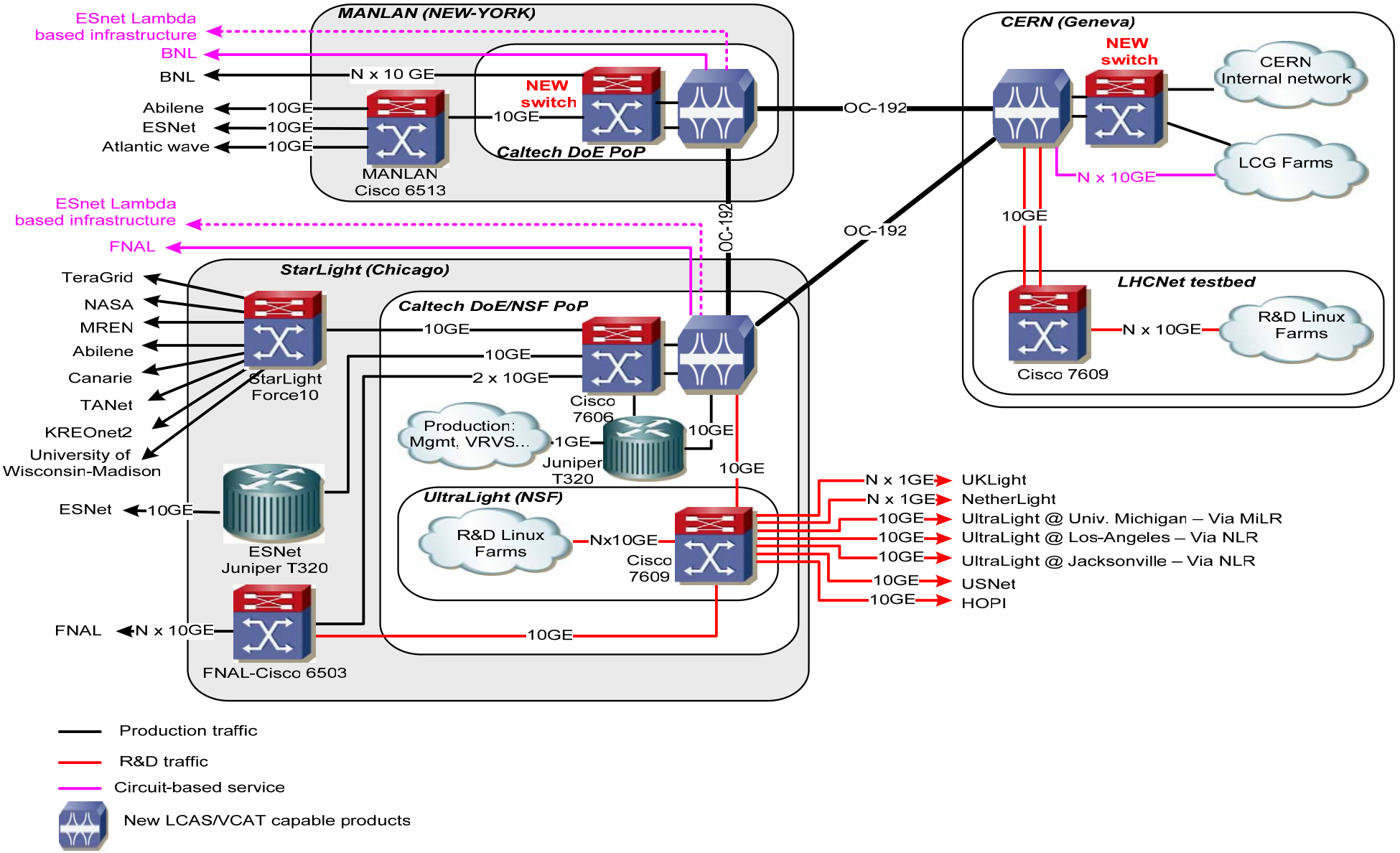
September 2005 (WAN-PHY option)



End of 2005


MANLAN (NEW-YORK)
Caltech DoE PoP
NEW switch
BNL
N x 10 GE
Abilene
10GE
ESNet
10GE
MANLAN Cisco 6513
OC-192
CERN (Geneva)
NEW switch
CERN Internal network
10GE
10GE
LCG Farms
Juniper T320
10GE
LHCNet testbed
N x 10GE
R&D Linux Farms
Cisco 7609
OC-192
University of Wisconsin-Madison
TeraGrid
NASA
MREN
Abilene
Canarie
TANet
KREOnet2
StarLight (Chicago)
StarLight Force10
Caltech DoE/NSF PoP
10GE
10GE
2 x 10GE
Cisco 7606
Juniper T320
Production: Mgmt, VRVS...
1GE
10GE
ESNet
10GE
ESNet Juniper T320
UltraLight (NSF)
R&D Linux Farms
Nx10GE
Cisco 7609
FNAL
N x 10GE
FNAL-Cisco 6503
10GE
N x 1GE
UKLight
N x 1GE
NetherLight
10GE
UltraLight @ Univ. Michigan – Via MiLR
10GE
UltraLight @ Los-Angeles – Via NLR
10GE
UltraLight @ Jacksonville – Via NLR
10GE
USNet
10GE
HOPI
Production traffic
R&D traffic
WAN-PHY interface

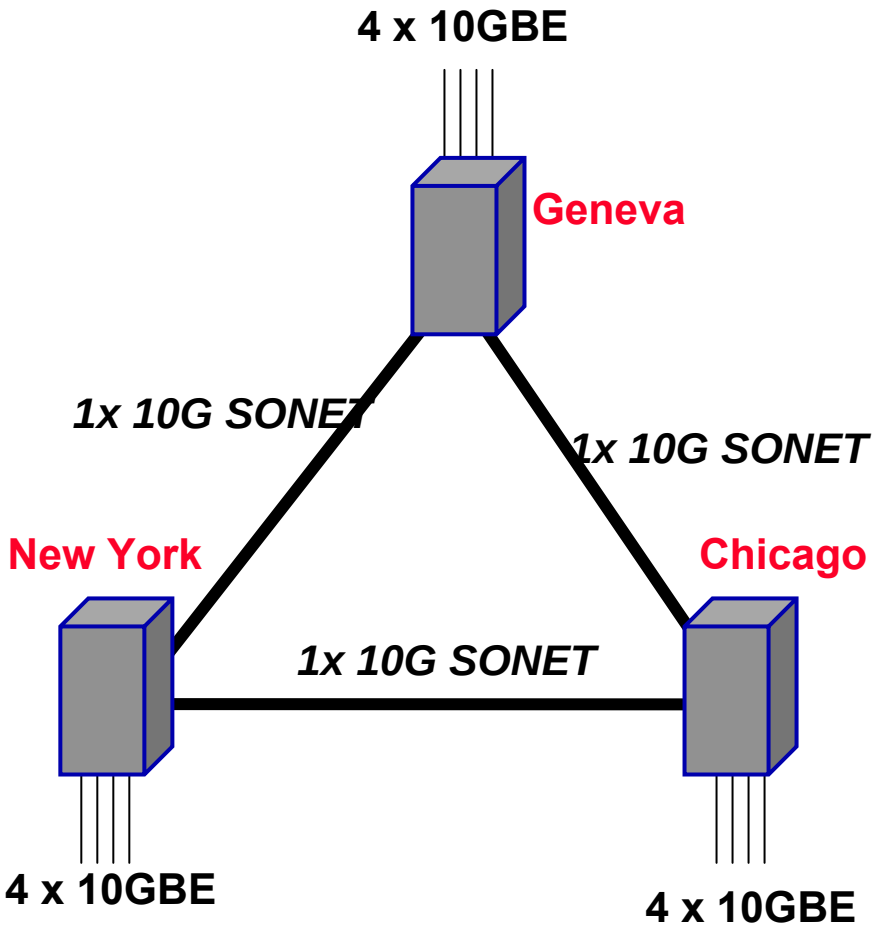
2007-2008 (Circuit Oriented Services?)



2005 Workplan

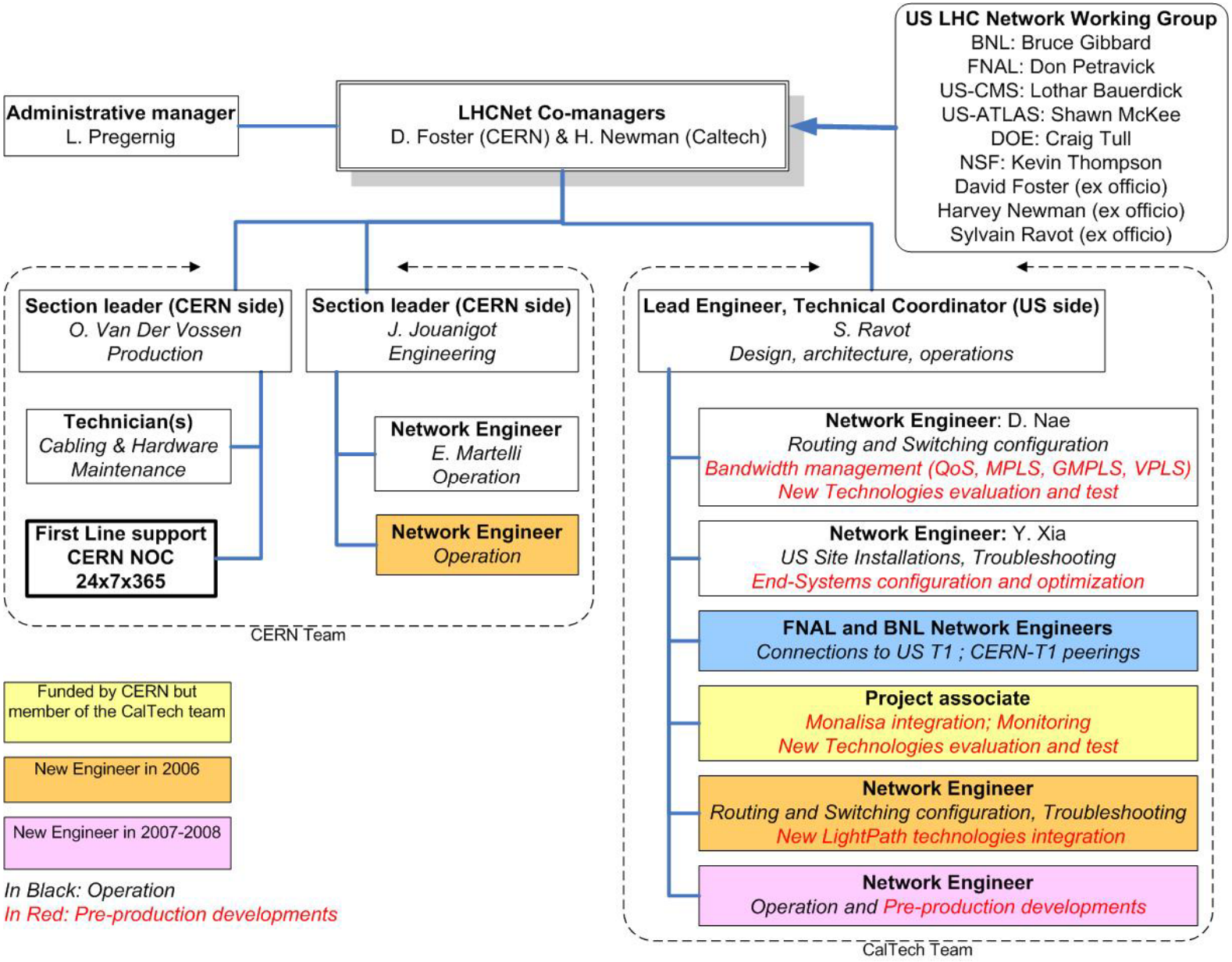
- June 2005: Call for tender; CERN financial committee
- July 2005: Negotiations with vendors
  - Splitting the service between two suppliers
- July 2005: Topology & Equipment selection
  - WAN-PHY tests ?? Cisco, F10?
- September - October: 2005 Circuits delivery
- January 2006: Direct path to BNL

Cost Estimation

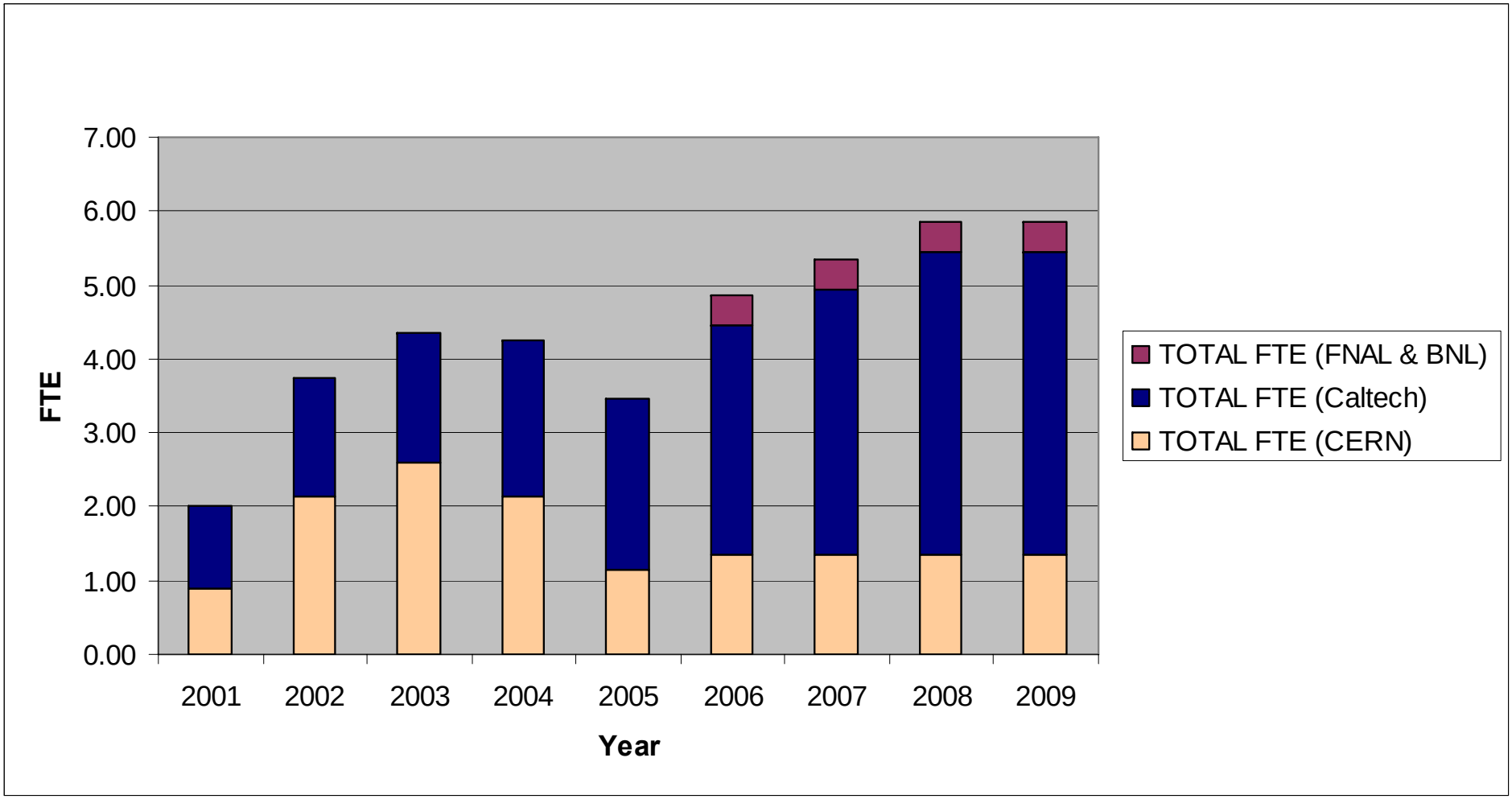


	Cost Estimation (including discount)	Remarks
Packet-over-Sonet	1,500 M$	Only two 10GE LAN-PHY modules Routing
WAN-PHY	330 M$	Routing and Switching
GFP/ LCAS	793 M$	Circuit-oriented

Organization Proposal

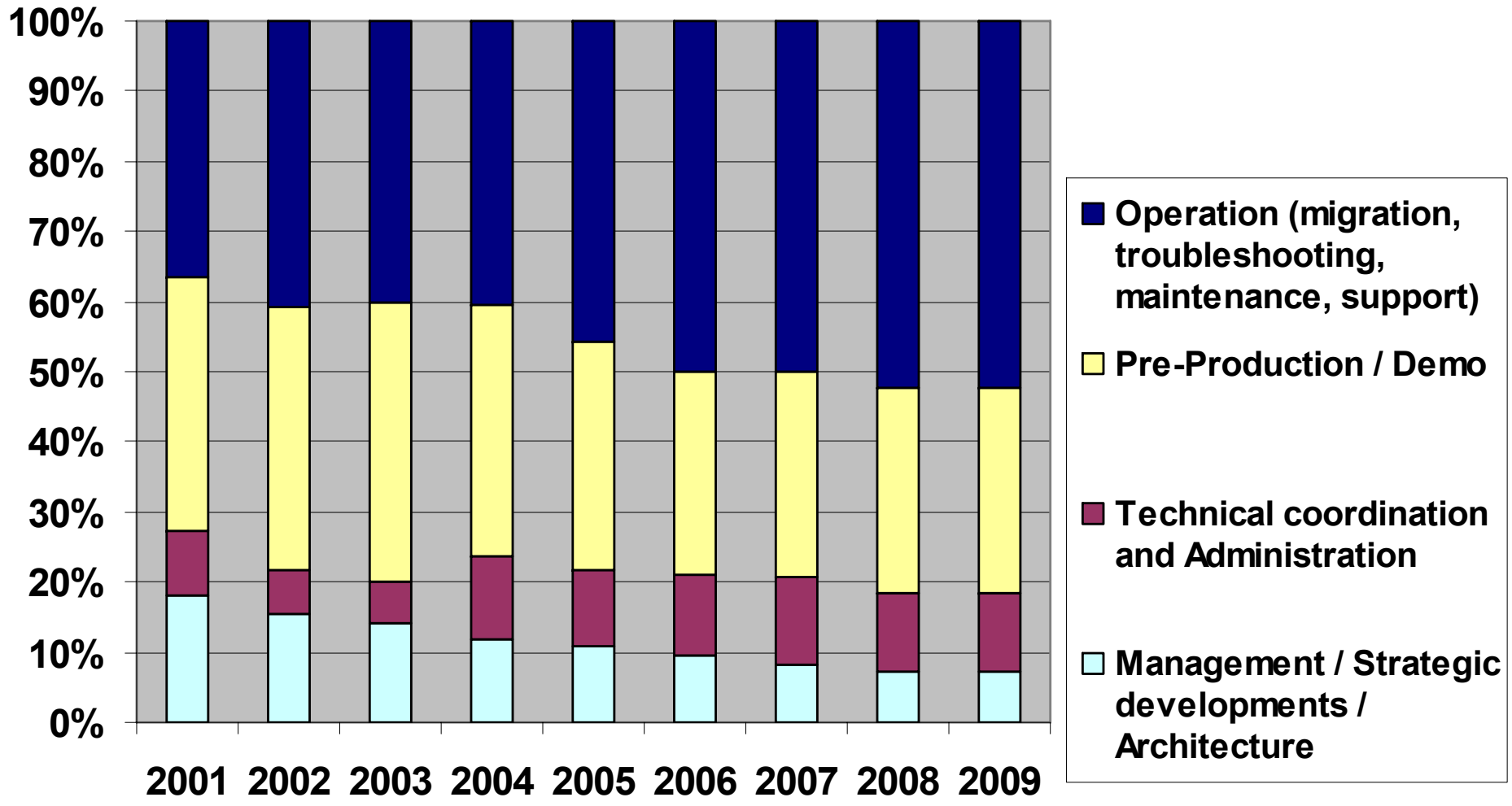


Manpower



- **2005-2008: 0.6 CERN FTE dedicated to Pre-Production**

Caltech Staff Activity Distribution



GFP/LCAS cost estimation

CERN Proposal - Config #2							
Partcode	Description	End User Price (USD)	Geneva	NYC	Chicago	Total Module	Extended price
Chassis Options -							
CD-SDH-ST-PKG-1-B	New COREDIRECTOR CD STARTER PACKAGE inc Rack	$82,775	0	1	0	1	$82,775
CI-SDH-ST-PKG-1-B	New COREDIRECTOR CI STARTER PACKAGE exc Rack	$35,000	1	0	1	2	$70,000
500-8001-001	KIT,INSTALL,FIBER,INTERBAY MANAGEMENT PANEL,CD	$4,150	0	1	0	1	$4,150
134-0069-902	Universal LM Blank	$425	3	26	3	32	$13,600
114-0106-100	OM2 Blank - universal connector	$200	0	0	0	0	$0
Line Modules	**Line Modules**						
134-0039-950	Line Module, LM-2 (134-0039-950)	$8,750	1	2	1	4	$35,000
	PALM Module 10G	$34,000	4	4	4	12	$408,000
	Optical Modules						
134-0107-900	STM64 Modules, SR2, LC Con (1510nm I-64.2 and I-64.2r)	$11,375	2	**2**	2	**6**	**$ 67800**
134-94xx-900	STM-64 DWDM Tunable 3rd Party Interop	$39,600	0	0	0	0	$0
134-0171-900	SFP-SX 50M, LC CONN	$140	0	0	0	0	$0
134-0172-900	SFP-LX, 10KM, LC CONN	$280	0	0	0	0	$0
	10G XFP	$2,190	4	4	4	12	$26,280
	Software						
S21-0004-302	Right to use Licence - Line Module Mesh - R3.0.2	$3,250	1	2	1	4	$13,000
S21-0006-302	Right to use Licence - 10G Optical Module R3.0.2	$1,500	2	4	2	8	$12,000
	Right to use Licence -PALM LM	$3,680	4	4	4	12	$44,160
	Element Licence fees						
009-2002-378	EMS Licence fee CI	$5,830	1	0	1	2	$11,660
009-2002-377	EMS Licence fee CD	$11,660	0	1	0	1	$11,660
			Total				**$ 793,000**