Worse Is Better

Vijay Gill/Bikash Koley/Paul Schultz for Google Network Architecture
Google

June 14, 2010
NANOG 49, San Francisco

Worse is Better

You see, everybody else is too afraid of looking stupid because they just can't keep enough facts in their head at once to make multiple inheritance, … or multithreading, or any of that stuff work. So they sheepishly go along with whatever faddish ~~programming~~ network craziness has come down from the architecture astronauts who speak at conferences and write books and articles and are so much smarter than us that they don't realize that the stuff that they're promoting is too hard for us.

Guiding Principle

- Important not to try to be all things to all people
  - Clients might be demanding 8 different things
    - Doing 6 of them is easy
    - …handling 7 of them requires real thought
    - …dealing with all 8 usually results in a worse system
  - more complex, compromises other clients in trying to satisfy everyone
- Don't build infrastructure just for its own sake
  - Identify common needs and address them
  - Don't imagine unlikely potential needs that aren't really there
  - Best approach: use your own infrastructure (especially at first!)
    - (much more rapid feedback about what works, what doesn't)

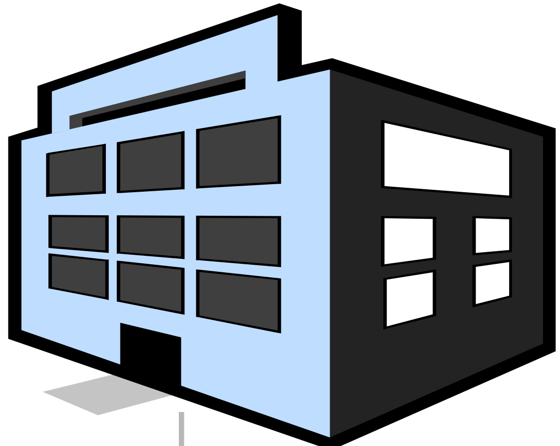
Google’s Mission

To organize the world's information and make it <u>UNIVERSALLY</u> accessible and useful.

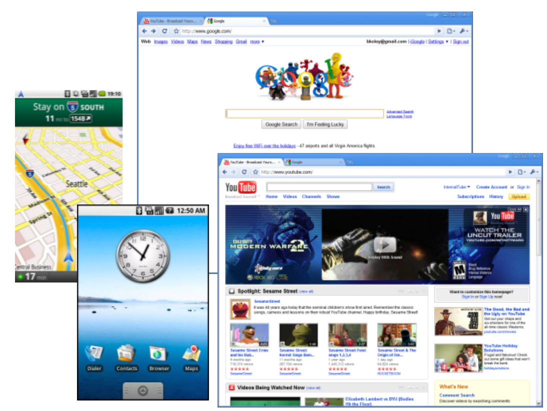
Scale Scale Scale

- **User base**
  - World population: 6.676 billion people (June'08, US Census est.)
  - Internet Users: 1.463 billion (>20%) (June'08, Nielsen/ITU)
  - Google Search: More than a billion searches daily
- **Geographical Distribution**
  - Google services are worldwide: over 55 countries and 112 languages
  - More than half our searches come from outside the U.S.
- **Data Growth**
  - Web expands/changes: billions of new/modified pages every month
  - Every few hours we crawl/refresh more than whole Library of Congress
  - YouTube gains over ~~13~~ ~~15~~ ~~18~~ 24 hours of video every minute, 1+ billion views a day
- **Latency Challenge**
  - Speed of Light in glass: 2 x 10^8 m/s = 2,000 km / 10 ms
  - “Blink of an eye response” = 100 ms

Warehouse Scale Computers



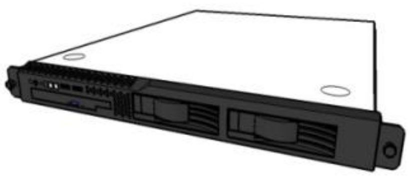
Consolidated Computing, Many UIs, Many Apps, Many Locations



Luiz André Barroso, Urs Hölzle, “The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines”, http://www.morganclaypool.com/doi/abs/10.2200/S00193ED1V01Y200905CAC006?prevSearch=allfield%253A%2528Urs%2529&searchHistoryKey=

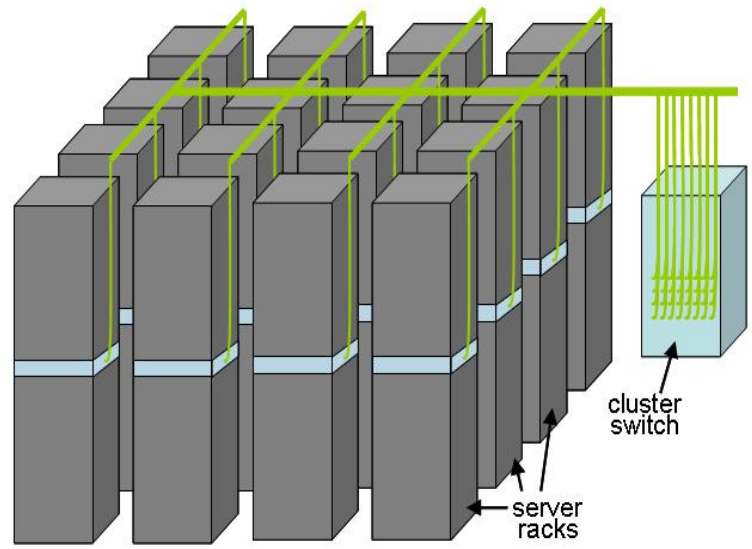

Warehouse-scale Computer

The Machinery



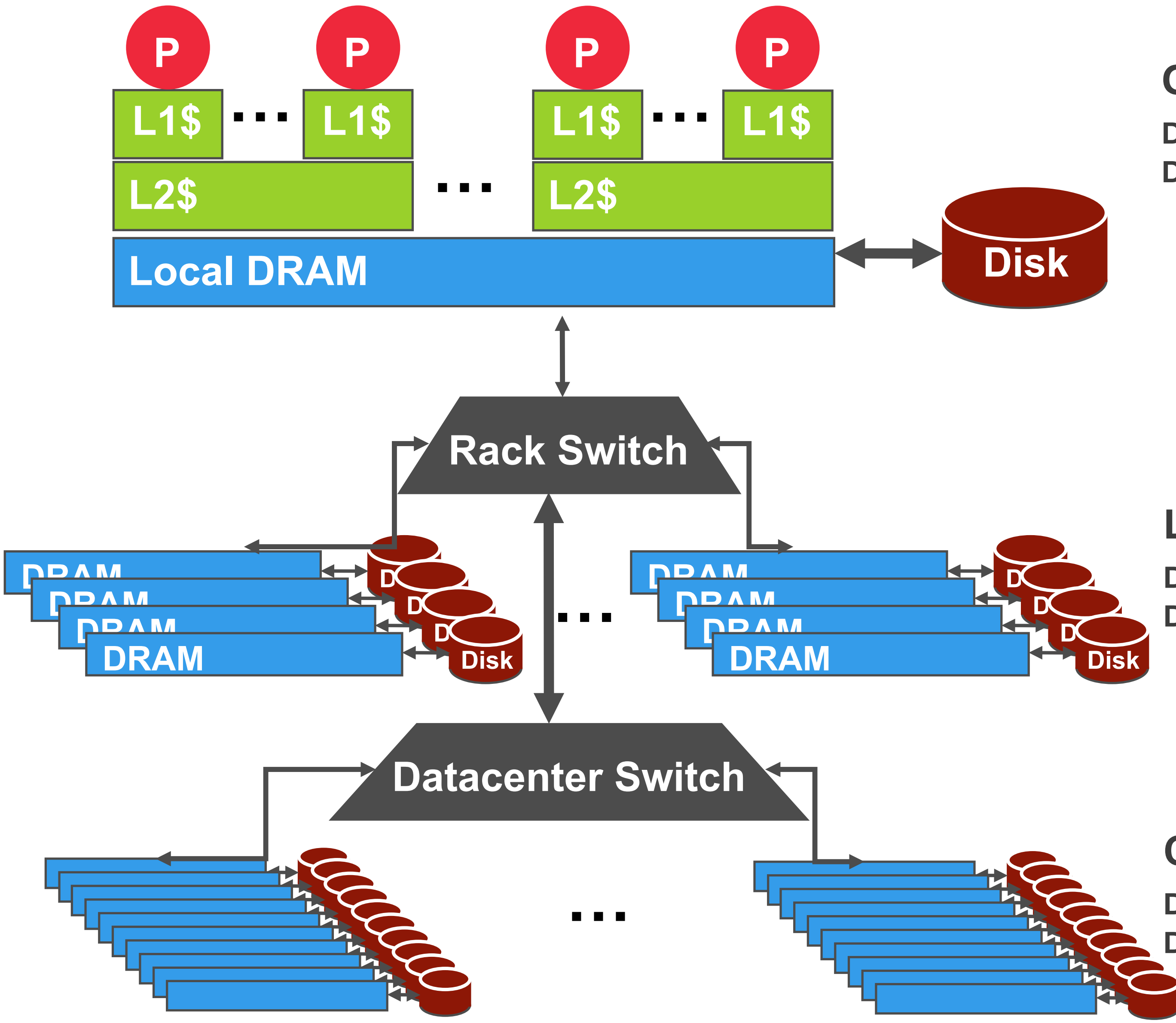
Servers
- CPUs
- DRAM
- Disks

Racks
- 40-80 servers
- Ethernet switch



Clusters

Example (small!) storage hierarchy



One server

DRAM: 16GB, 100ns, 20GB/s
Disk: 2TB, 10ms, 200MB/s

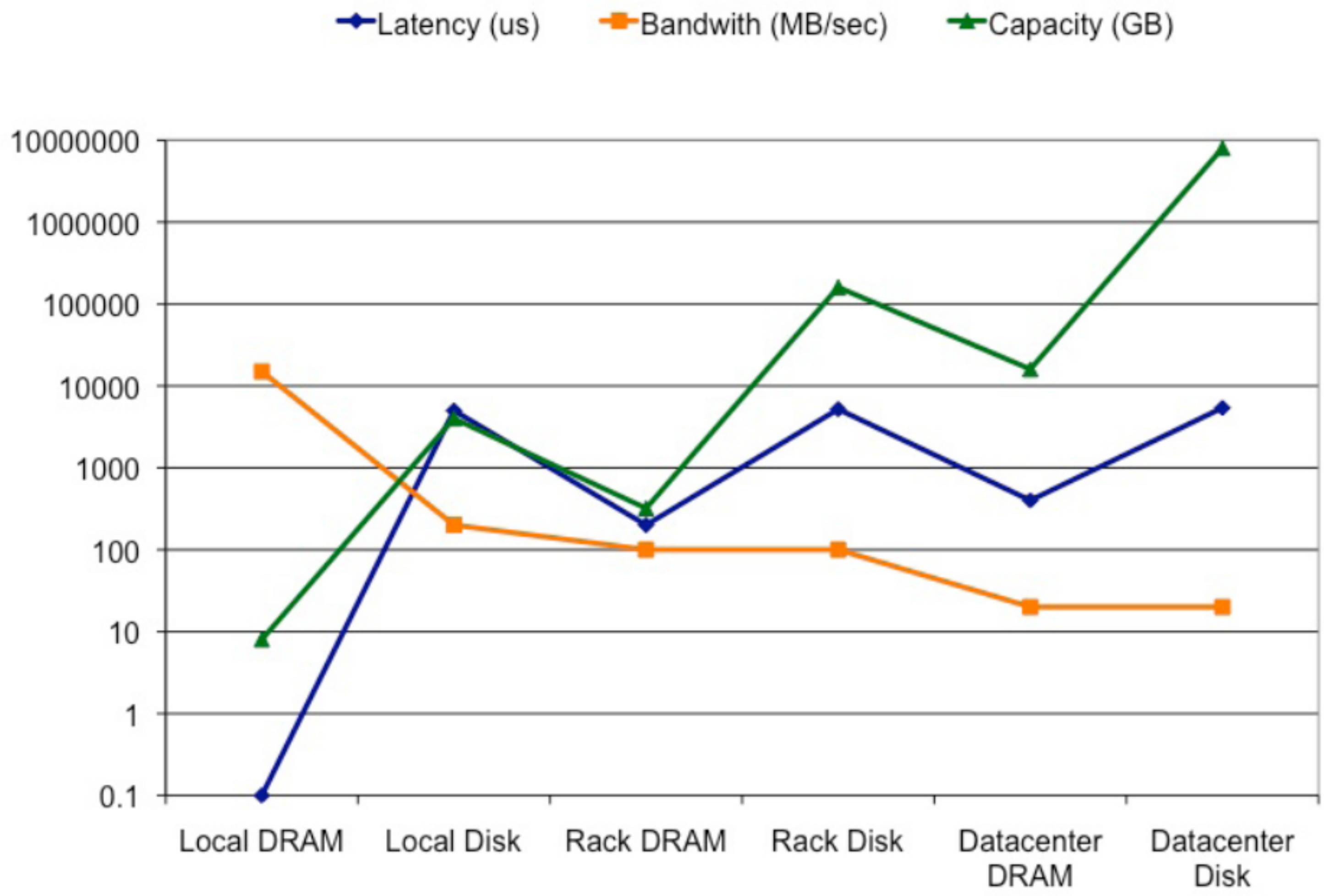
Local rack (80 servers)

DRAM: 1TB, 300us, 100MB/s
Disk: 160TB, 11ms, 100MB/s

Cluster (30 racks)

DRAM: 30TB, 500us, 10MB/s
Disk: 4.80PB, 12ms, 10MB/s

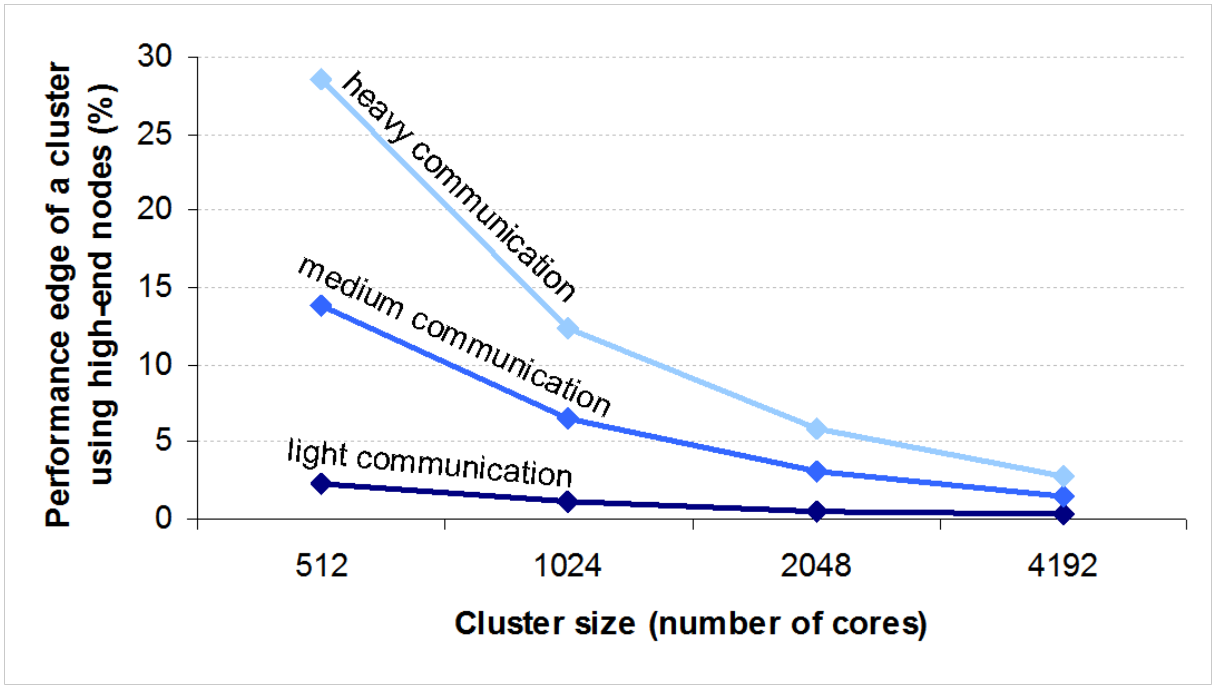
Storage Hierarchy (A different View)



A bumpy ride that has been getting bumpier over time

Communication performance?

- Big advantage in communication performance **if** computation fits in 1 system
- Performance edge of big iron systems deteriorates if used in very large clusters



Execution time = 1ms + f*(100ns*1/#nodes + 100us*(1 – 1/#nodes))

Reliability & Availability

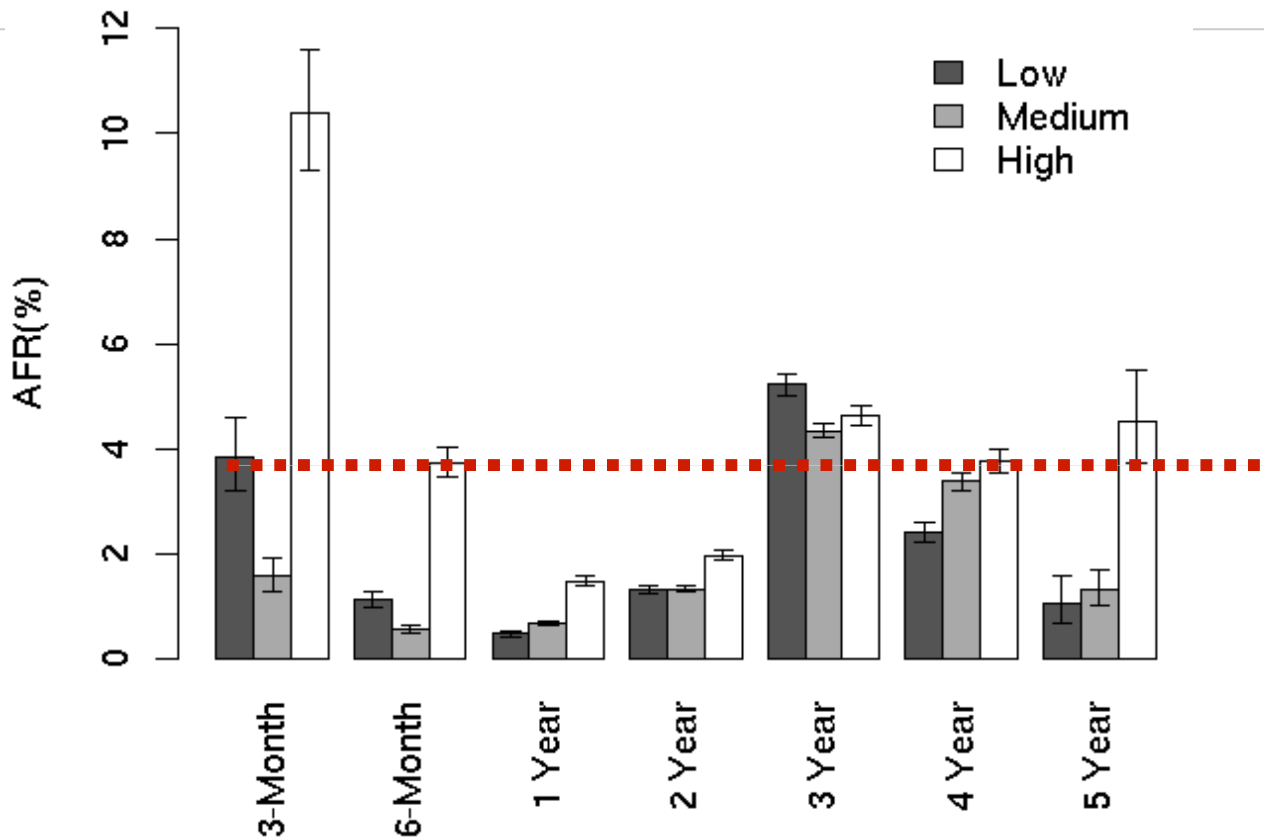
- Things will crash. Deal with it!
  - Take super reliable servers (MTBF of 30 years)
  - Build a machine with 10 thousand of those
  - Watch one fail per day
- Fault-tolerant software is inevitable
- Typical yearly flakiness metrics
  - 1-5% of your disk drives will die
  - Servers will crash at least twice (2-4% failure rate)
- Internet is not a “five nines” fabric
  - ~99% availability, varying significantly with geography
  - Wild dogs, sharks, dead horses, thieves, blasphemy, drunken hunters
  - Very strange bugs are common when you have lots of gear

The Joys of Real Hardware

Typical first year for a new cluster:

~0.5 overheating (power down most machines in <5 mins, ~1-2 days to recover)
~1 PDU failure (~500-1000 machines suddenly disappear, ~6 hours to come back)
~1 rack-move (plenty of warning, ~500-1000 machines powered down, ~6 hours)
~1 network rewiring (rolling ~5% of machines down over 2-day span)
~20 rack failures (40-80 machines instantly disappear, 1-6 hours to get back)
~5 racks go wonky (40-80 machines see 50% packetloss)
~8 network maintenances (4 might cause ~30-minute random connectivity losses)
~12 router reloads (takes out DNS and external vips for a couple minutes)
~3 router failures (have to immediately pull traffic for an hour)
~dozens of minor 30-second blips for dns
~1000 individual machine failures
~thousands of hard drive failures
slow disks, bad memory, misconfigured machines, flaky machines, etc.

Hard drive failures



(48000 disks)(0.04 fail/y)(6 h)/(8766 h/y) =

1.3 disk failures per run

Must design assuming drives will fail!

Graph source: Pinheiro, Weber, Barroso: FAST'07

Internet Landscape is Changing

Google

2009 Internet Observatory Report

The "ATLAS 10" Today

Rank	Provider	Percentage
1	Level(3)	5.77
2	Global Crossing	4.55
3	ATT	3.35
4	Sprint	3.2
5	NTT	2.6
6	Cogent	2.77
7	Verizon	2.24
8	TeliaSonera	1.82
9	Savvis	1.35
10	AboveNet	1.23

(a) Top Ten 2007

Rank	Provider	Percentage
1	Level(3)	9.41
2	Global Crossing	5.7
3	Google	5.2
4		
5		
6	Comcast	3.12
7		
8	*Intentionally omitted*	
9		
10		

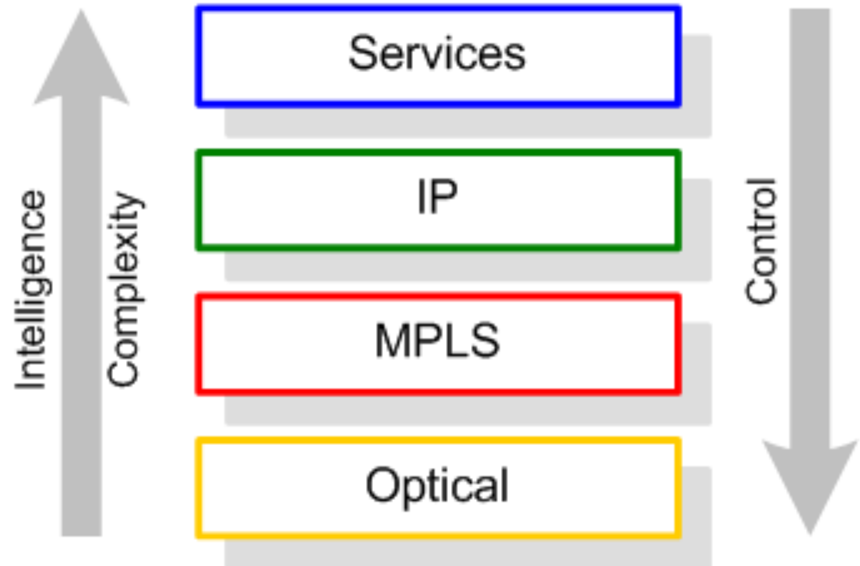
(b) Top Ten 2009

- Based on analysis of anonymous ASN (origin/transit) data
  - Weighted average percentage
- Top ten has NO direct relationship to Observatory participation
- Tier1s still carry significant traffic volumes (and profitable services)
- But Comcast and Google join the top ten

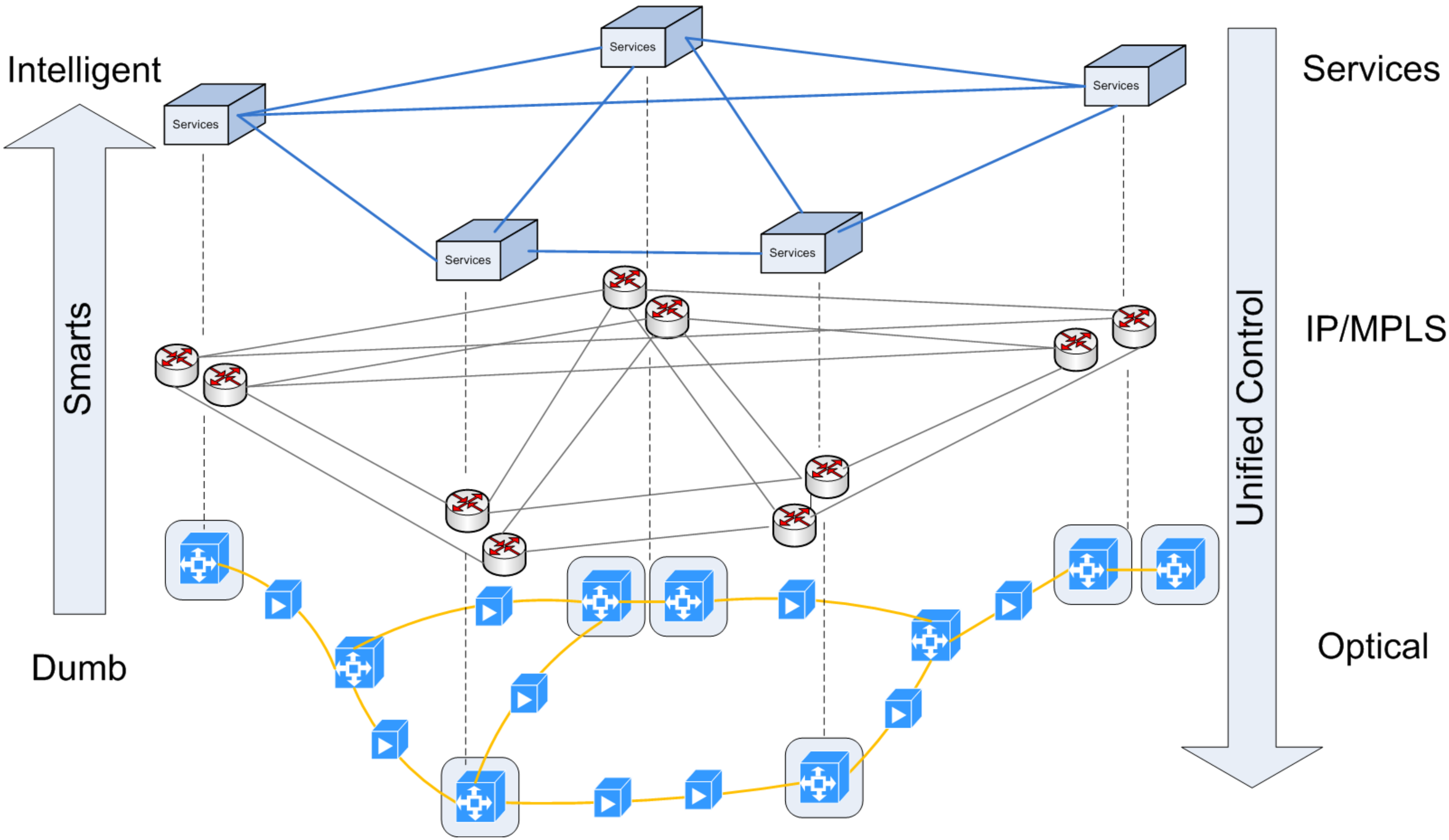
http://www.nanog.org/meetings/nanog47/presentations/Monday/Labovitz_ObserveReport_N47_Mon.pdf

Layer Cake

- Service Layer – Massively Scalable, Highly Dynamic. Services Drive the Network. Application Layer Control Pushed Down Into Network.
- IP Layer – Standardized, Resilient and Universal Compute Interconnect and Service Delivery
- MPLS Layer – Forwarding, TE, Fast Restoral
- Optical Layer – Cost-effective simple, high-BW, point-to-point connectivity



Topology



Intelligent
Smarts
Dumb
Services
Services
Services
Services
Services
Services
IP/MPLS
Optical
Unified Control

Traffic Classes

Google

- **Two types of traffic**
- **1. User (“I”)**

  Predominantly user initiated
- **2. Machine-to-machine (“M”)**

  Bulk, predominantly machine to machine
- Different SLA and requirements
- Different scale of traffic

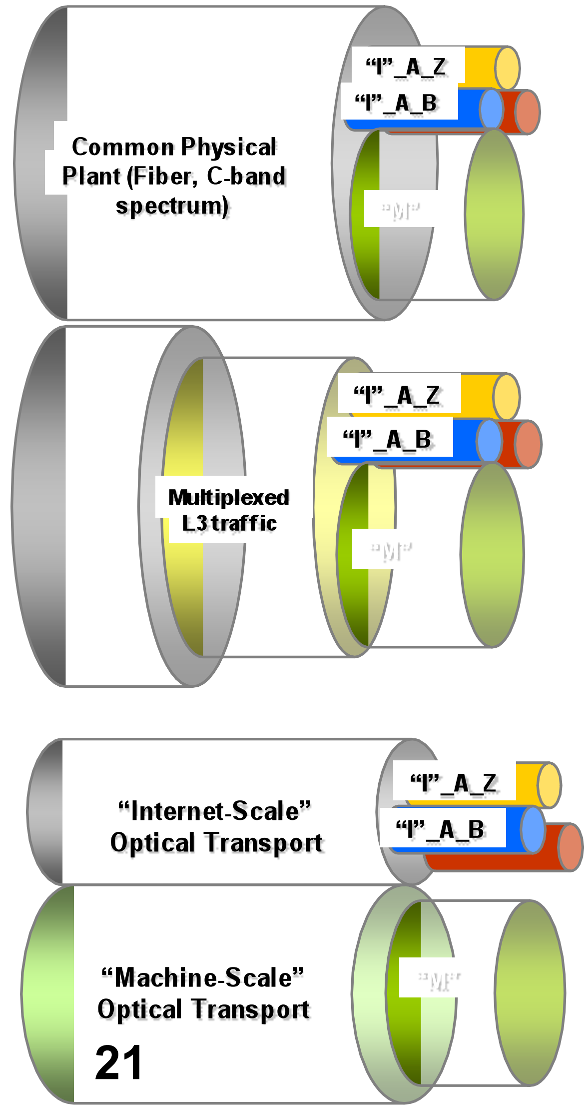
Converged vs Overlay?

1. Converged

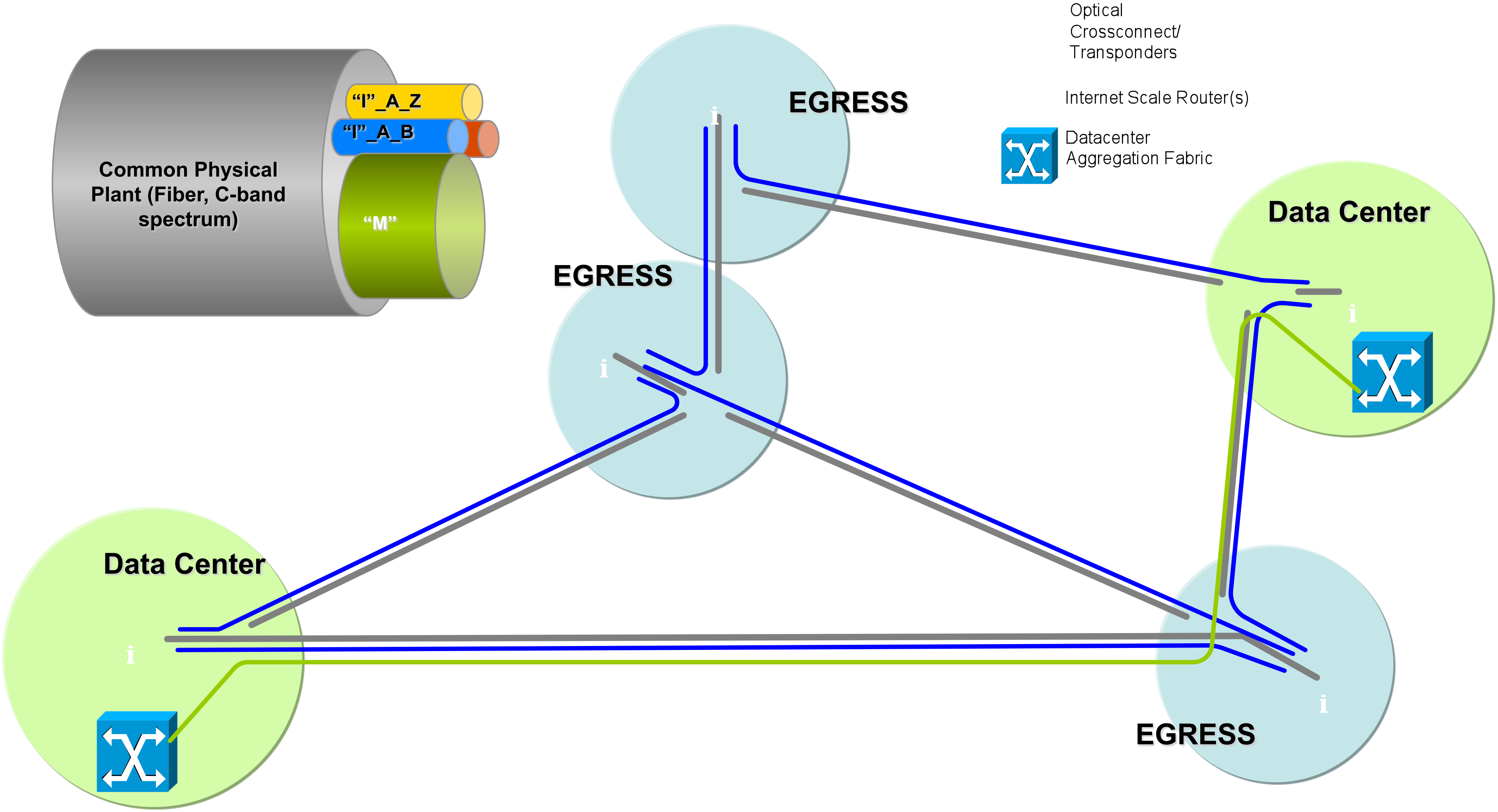
- Layer-1 (common transport infrastructure)
- Layer-3 (common routing and transport infrastructure)

2. Overlay

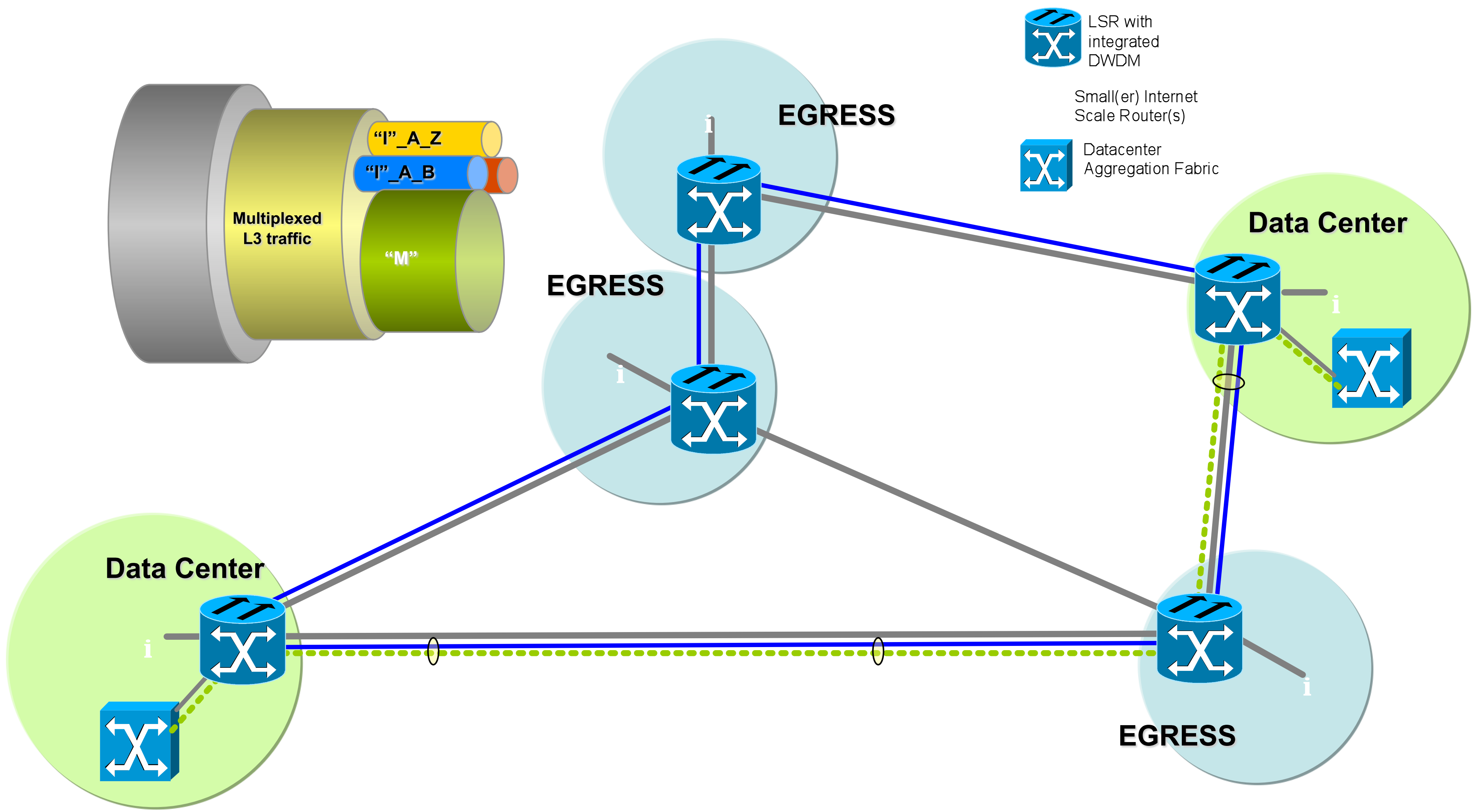
Same fiber path, different transport with appropriate optimization for each layer-3



Converged L-1 Network

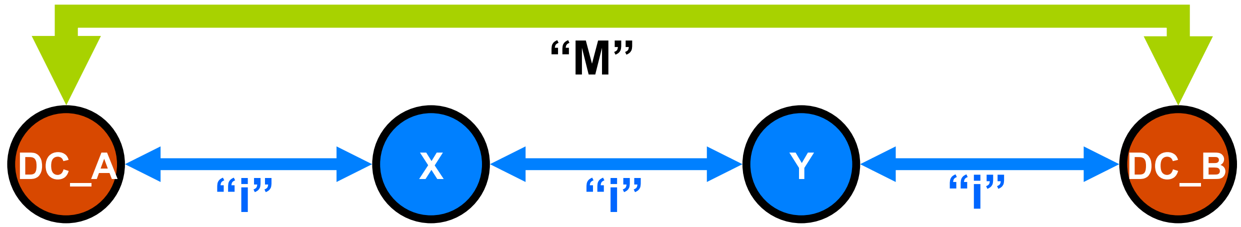


Converged L-1+L-3

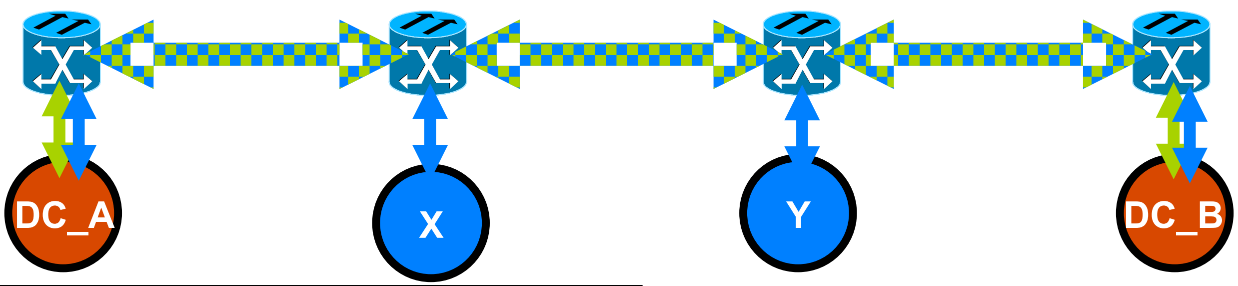


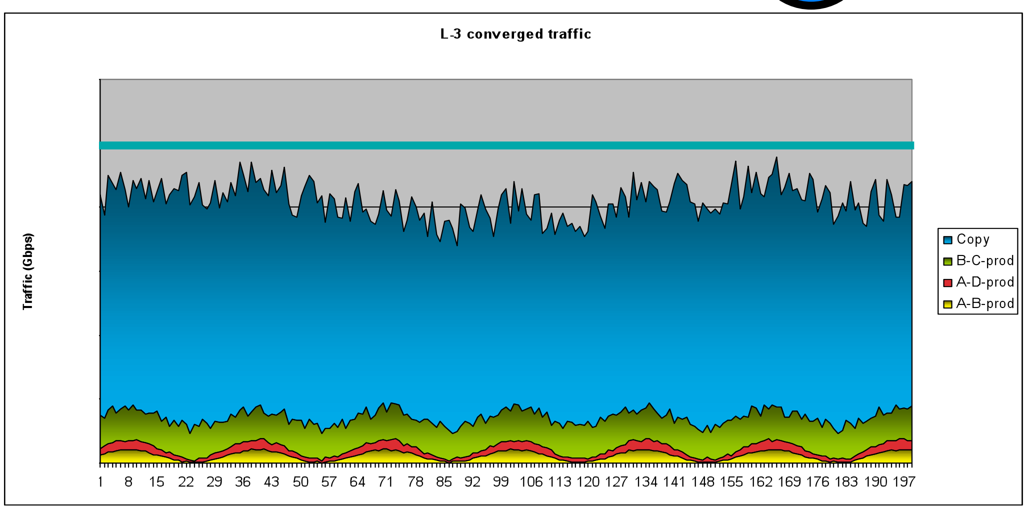
Converged L3 Explained

Separate L3 for "I" class and "M" class traffic



Converged L3 for "I" class and "M" class traffic

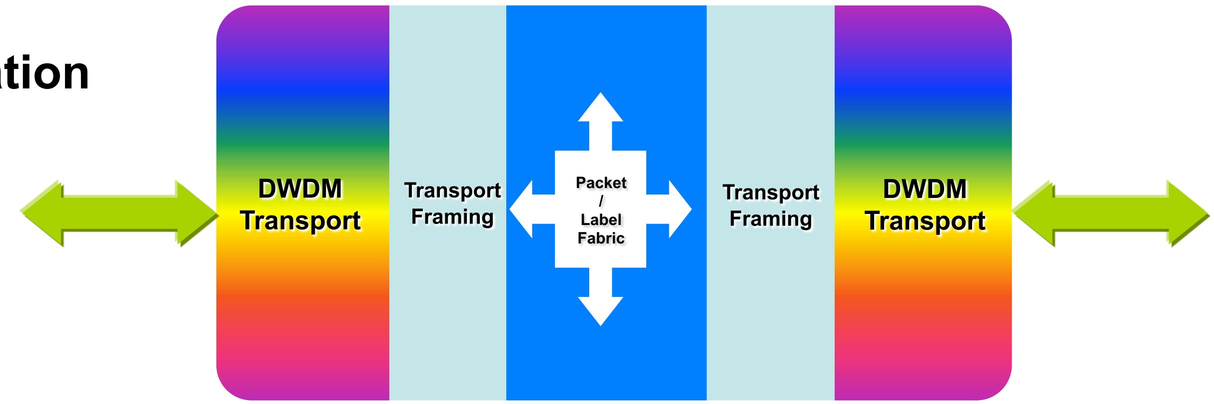




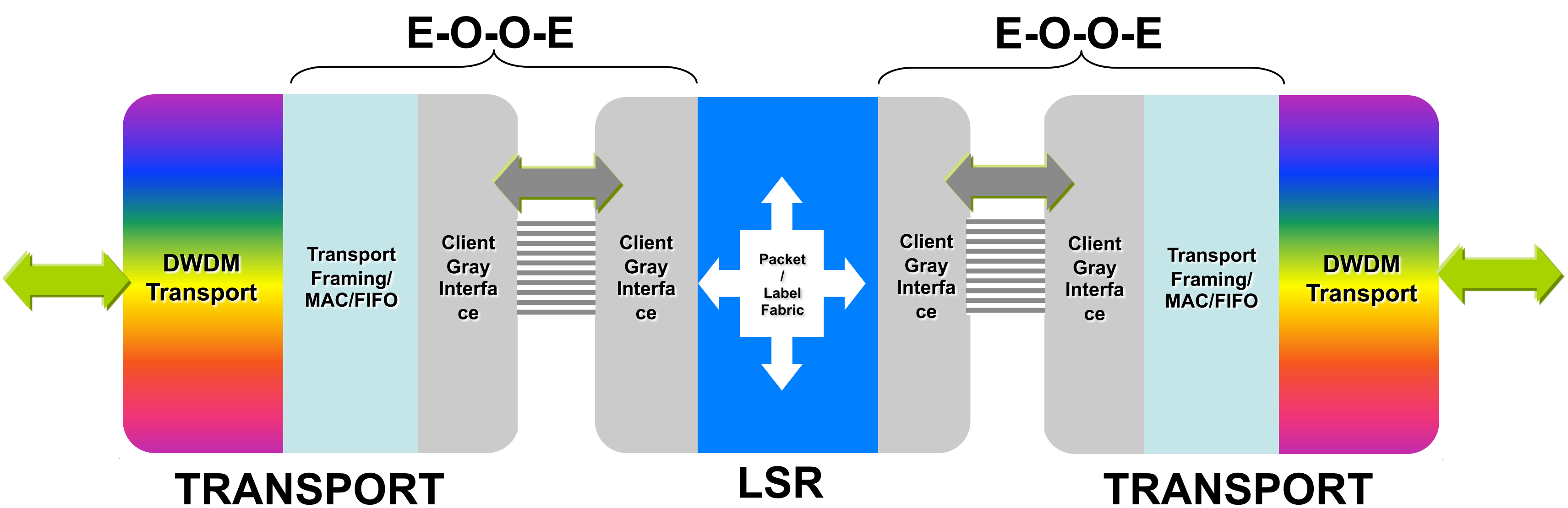
- **Better utilization of transport pipes through stat-muxing of "I" class and "M" class traffic**

Two LSR Junction Options

1. LSR with ITU integration

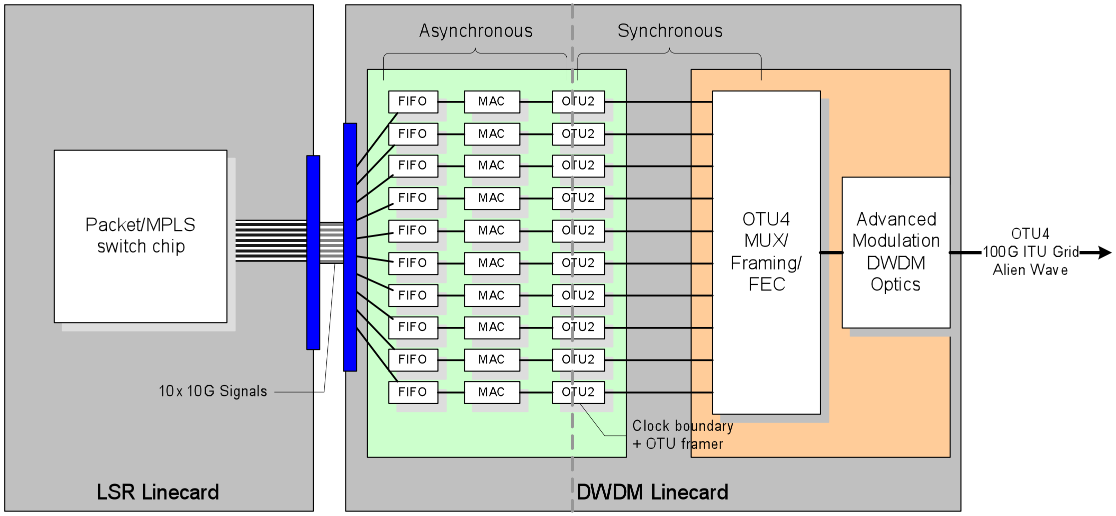


2. LSR without ITU integration

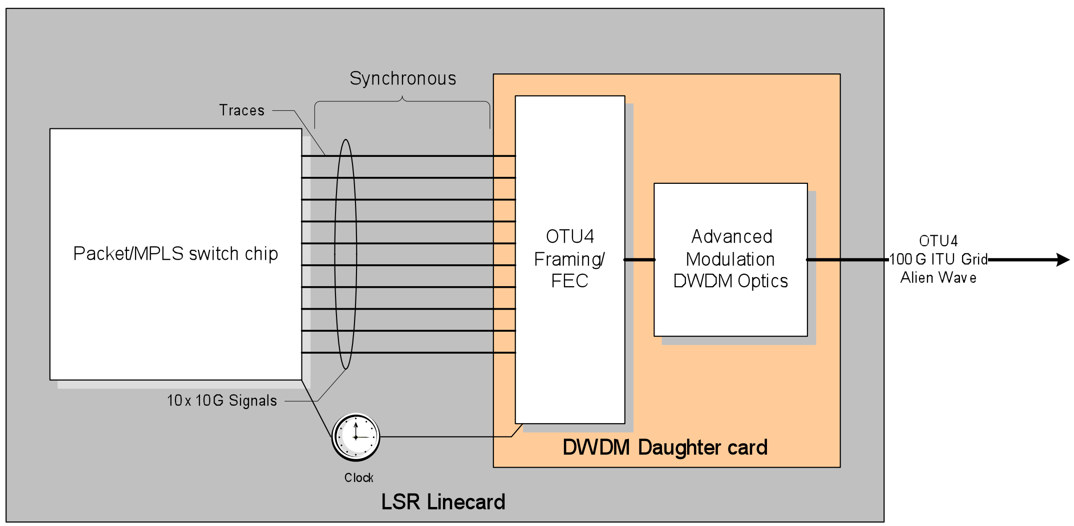


Integrated LSR: 2 types

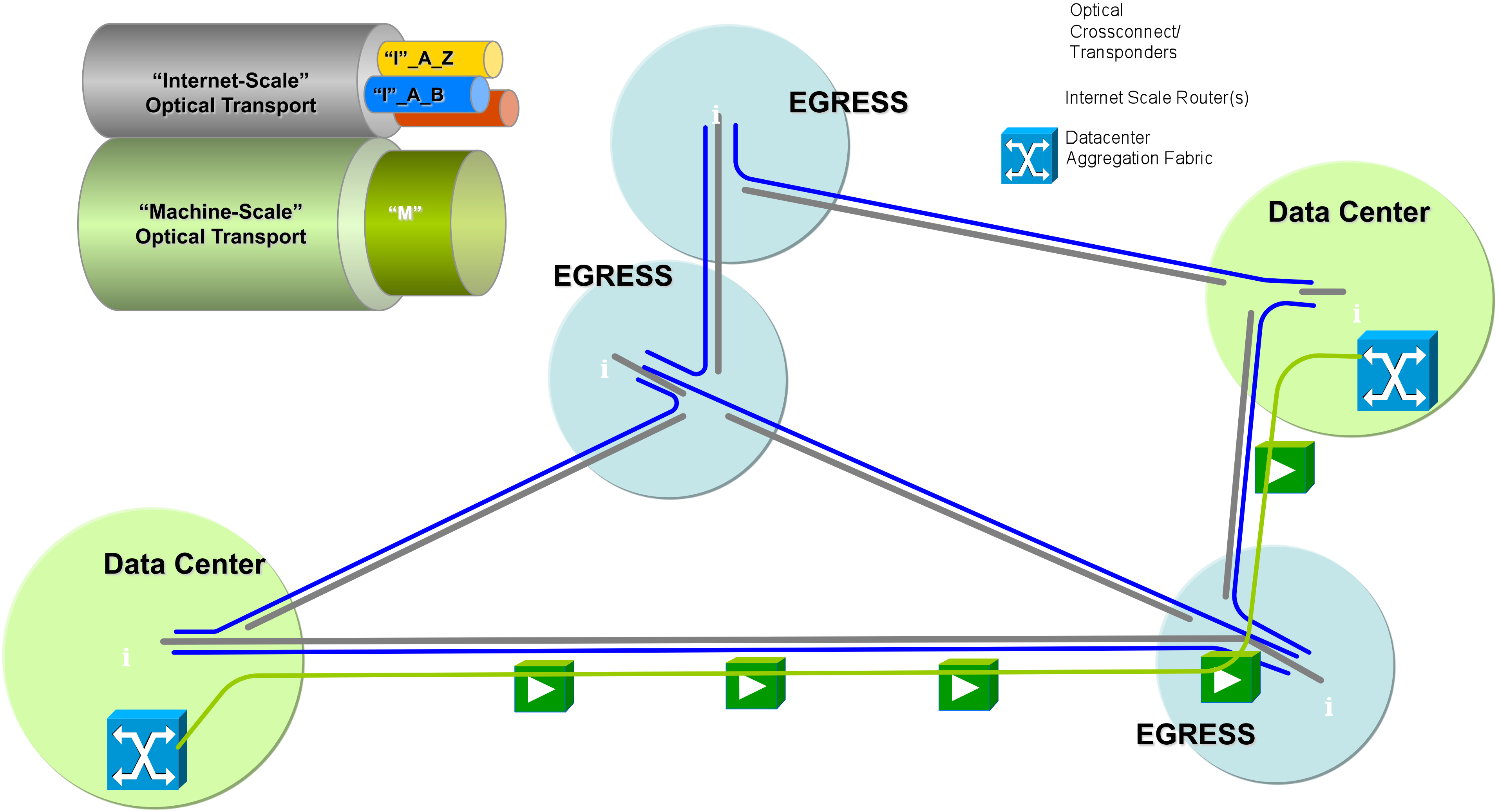
"Loose Integration": LSR + standalone DWDM equipment; cheap client optics integration



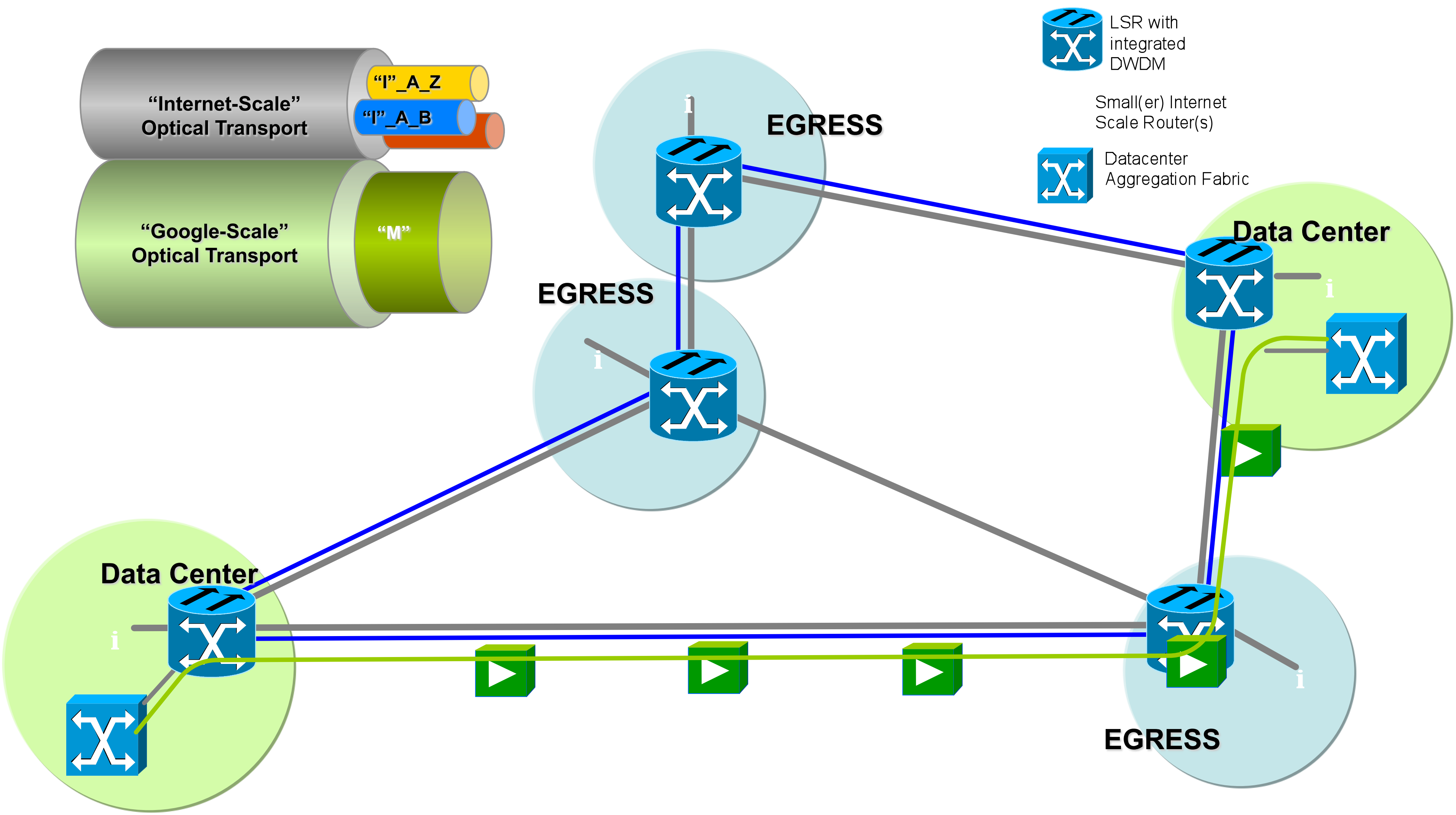
"Tight Integration": LSR with Integrated DWDM; Cost of buffering, sync and clocking eliminated



Overlay L-1/L-3 Networks



“I-scale”LSR and “M-scale” Overlay



Assumptions/ Parameters Used

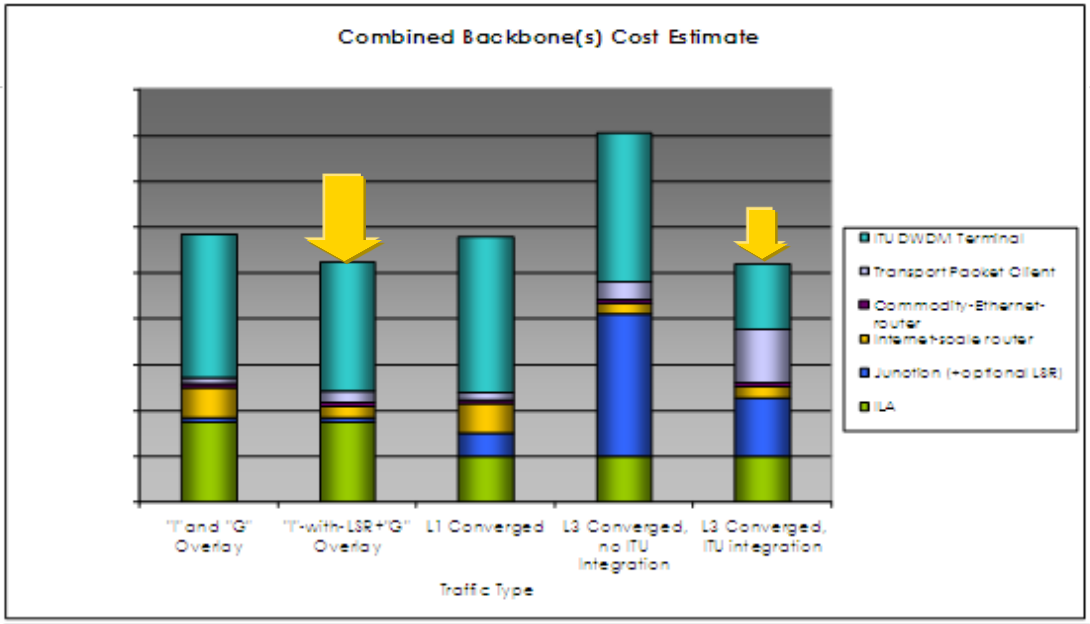
- This analysis takes into account of CapEx only
- Projected internet-scale traffic based on current traffic pattern is used
- Data_center-pair Machine-scale demand is a parameter in this analysis
- Availability of 100Gbps transport technology is assumed
- Availability of cost-effective LSR with/without optical integration is also assumed

- **Cost models: (normalized on a $ Per 10Gbps basis)***

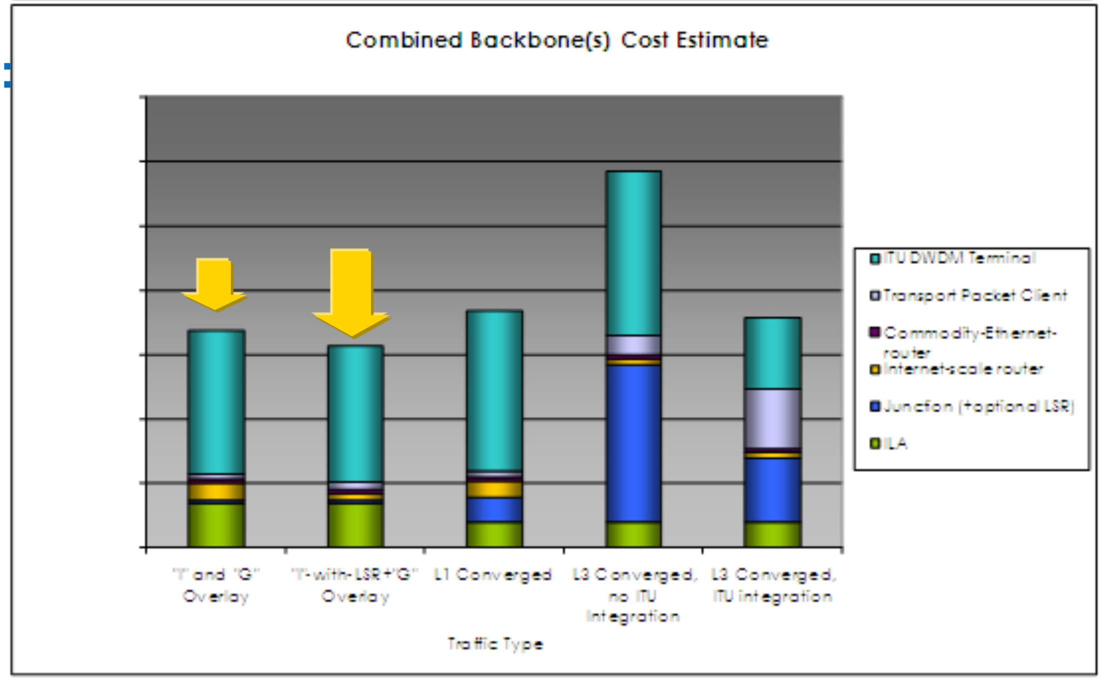
10G Port Type	Port Cost
Internet Scale Router	600
Merchant Silicon Scale Router	30
Merchant Silicon Scale LSR	30
Internet Scale LSR (no DWDM)	120
Internet Scale LSR with Transport Framing (ITU lambda not included)	180
DWDM Terminal Interface	400
DWDM Regen Interface	150
Gray Client Interface	10

***Pricing exercise done using average street price for the equipment where available**

Different scales for each graph (absolute cost)



"Machine" Traffic: X Tbps, "User" Traffic: T Tbps



"Machine" Traffic: 2*X Tbps, "User" Traffic: T Tbps

Conclusions

- Google's web-services generate large amounts of Traffic (*More Internal Than External*)
- Traffic volumes make cost efficiency a primary metric
- Cost-effective LSRs provide the simplest way to effectively stat-mux traffic and add/drop
- However, <u>DWDM HAS to be integrated with LSR functionality</u>
- Depending on add/drop locations and relative scale of "I" and "M" traffic, either a flat DWDM-LSR based network or an LSR based network with optical express are the ways to go
- A simple unified end-to-end control plane at the highest network layer (IP/MPLS) is a MUST
- <u>No Need For Additional Control Plane At Optical Layer – Other Layers Provide Necessary Functionality</u>

Thank You!