



Technologies, Architecture and Services for the Next-Generation Core Optical Networks

Adel A. M. Saleh

DARPA

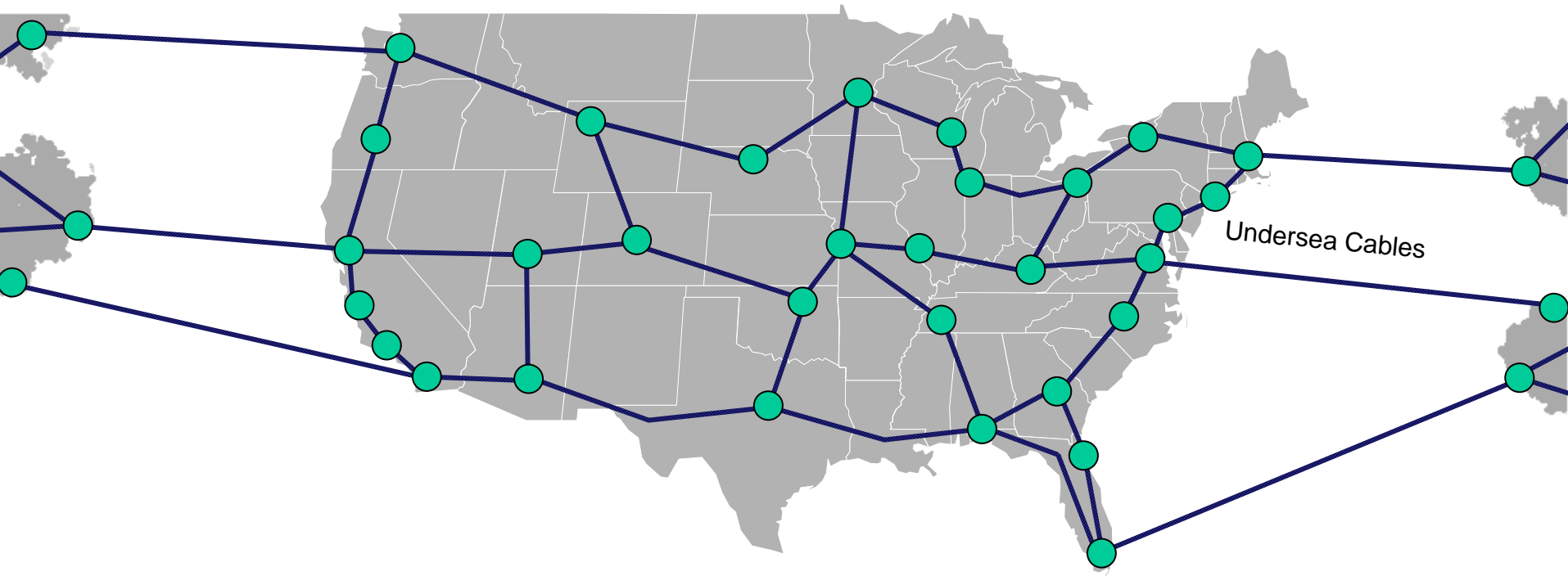
Strategic Technology Office

Arlington, Virginia, USA

Adel.Saleh@darpa.mil

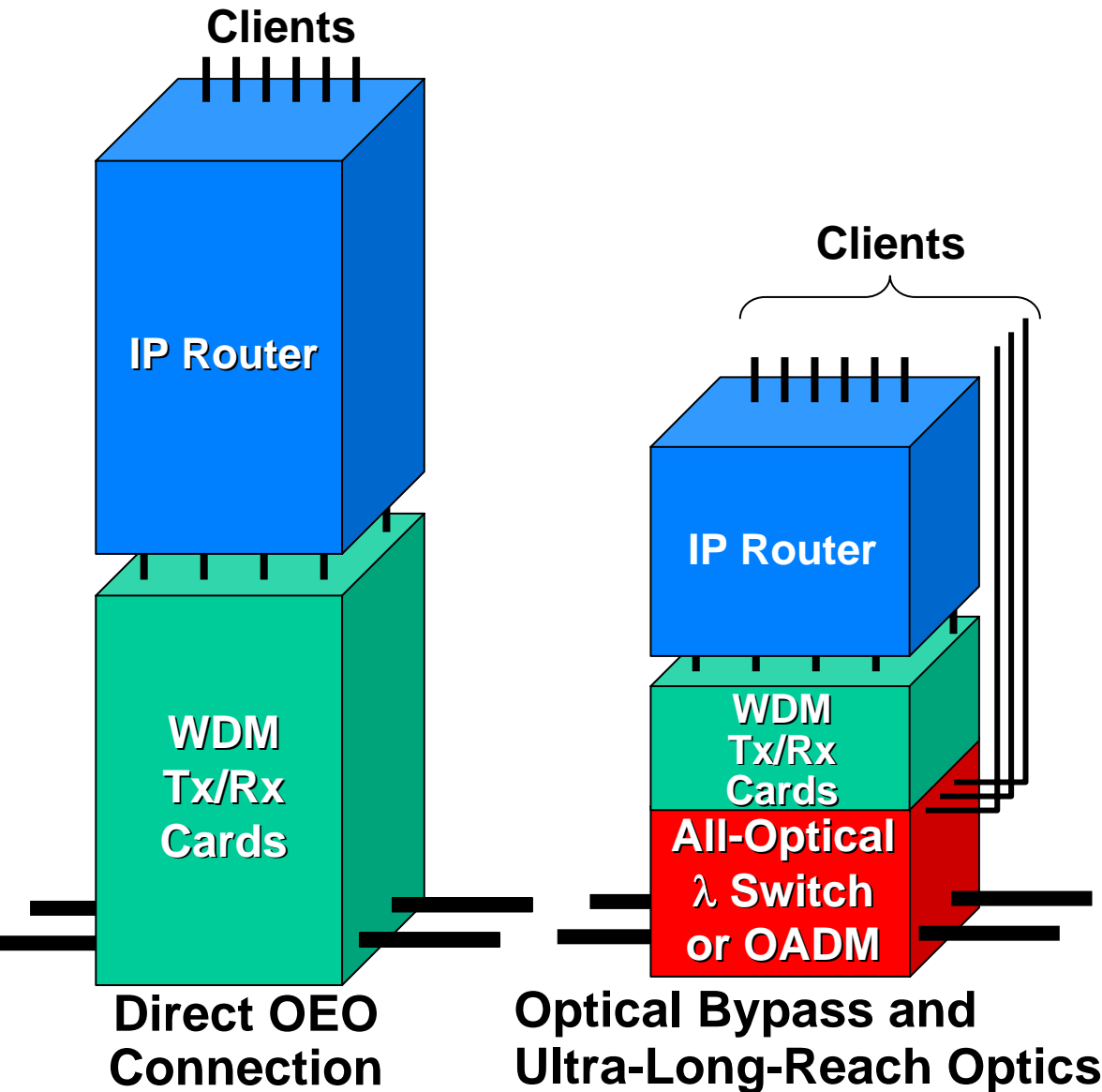


Vision: Increased Throughput, Robustness and Security with Reduced Latency, Capital Cost and Operational Cost



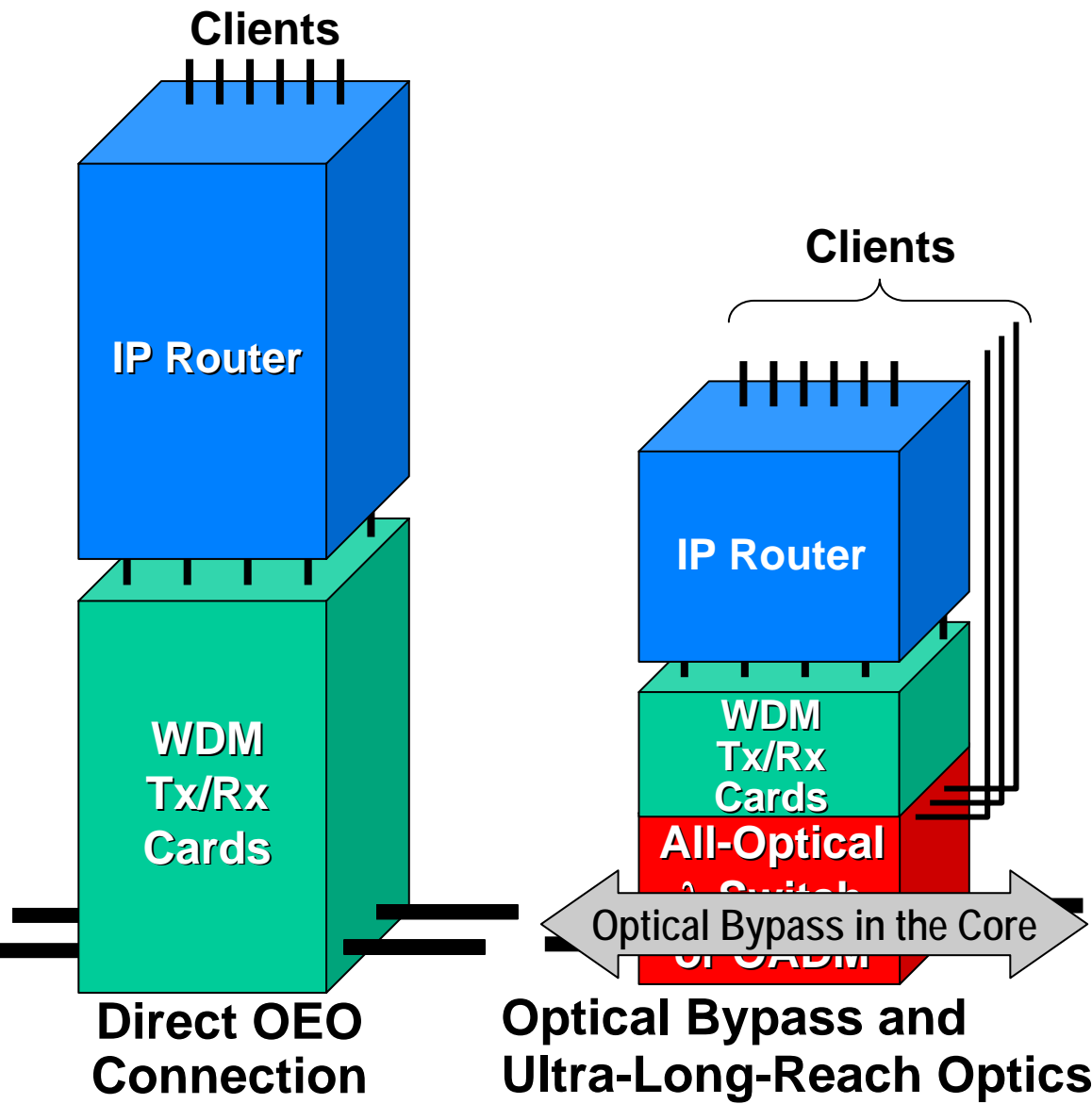


Current State of IP/WDM Network Node Architecture



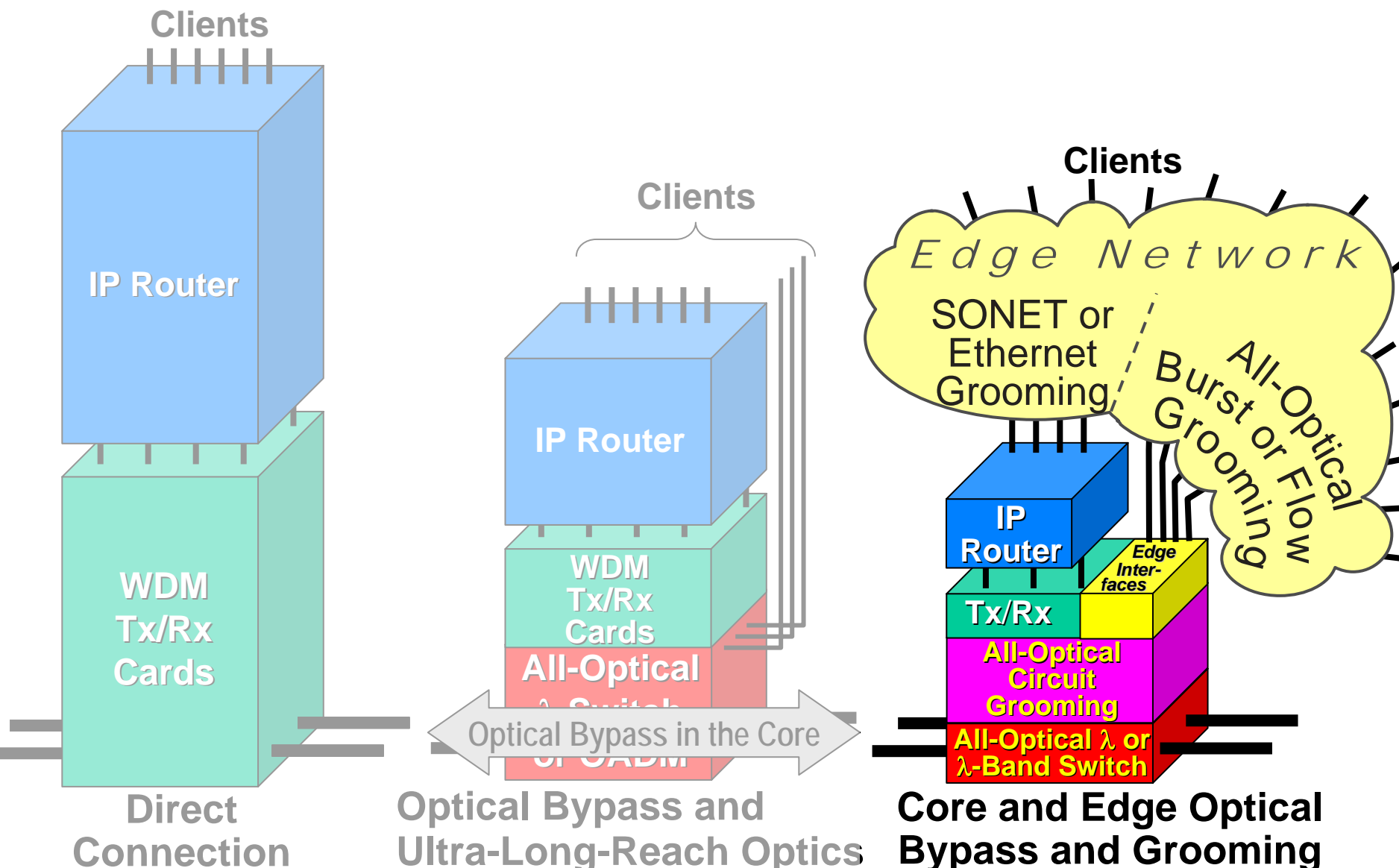


Current State of IP/WDM Network Node Architecture





Possible Evolution of IP/WDM Network Node Architecture



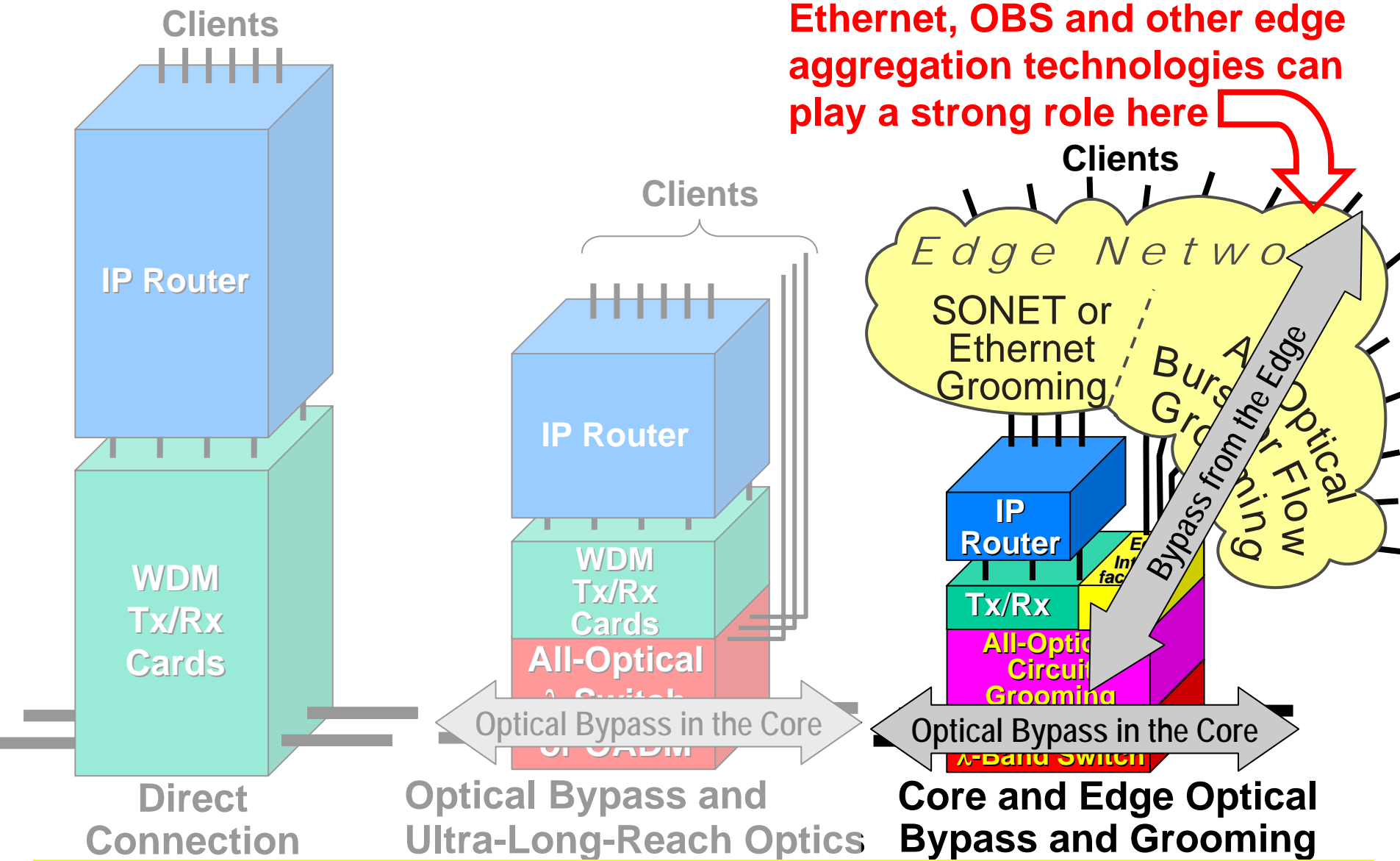
From: Saleh and Simmons, "Evolution Towards the Next-Generation Core Optical Network", *JLT*, Sept 2006



Possible Evolution of IP/WDM Network Node Architecture



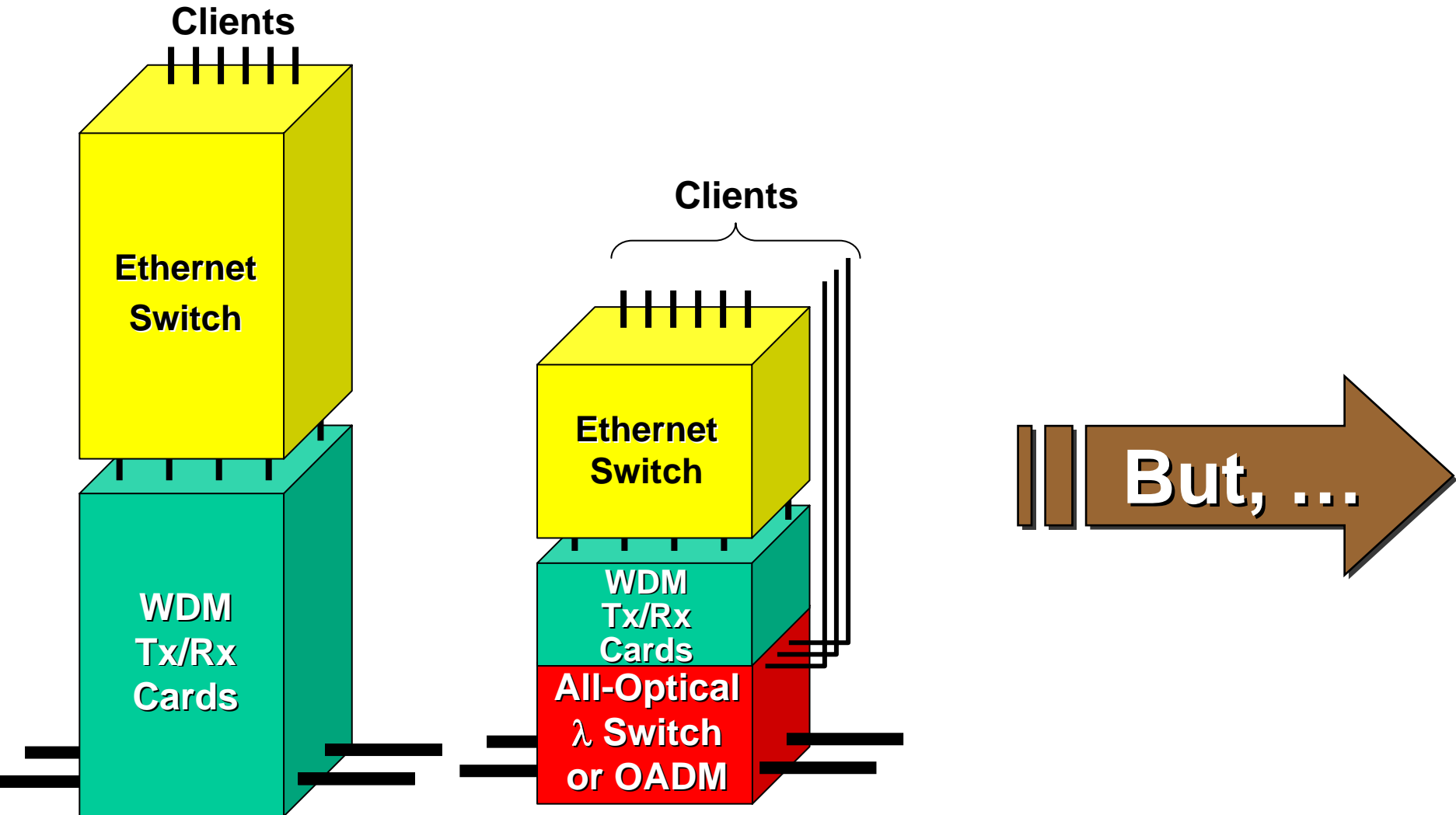
Ethernet, OBS and other edge aggregation technologies can play a strong role here



From: Saleh and Simmons, "Evolution Toward the Next-Generation Core Optical Network", *JLT*, Sept 2006



Meanwhile, everyone is saying that **Ethernet** is coming to the Core !





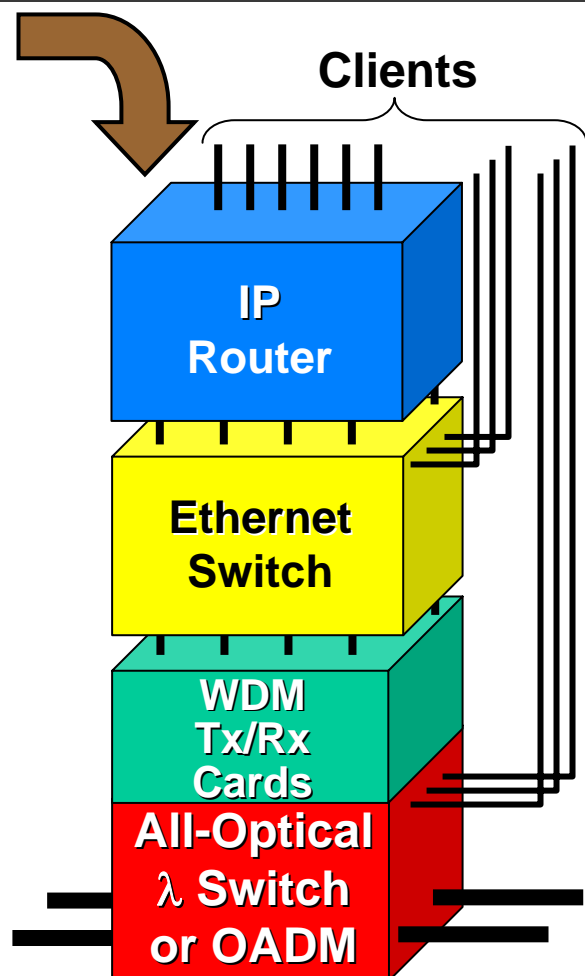
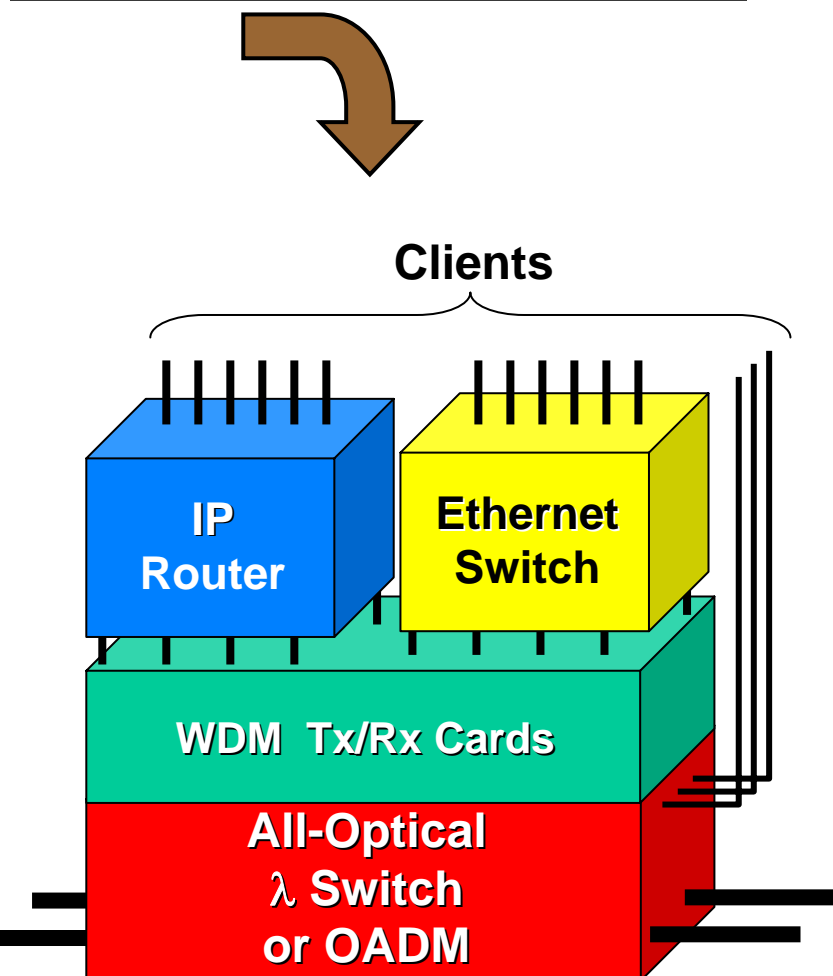
But what is the vision for the **Ethernet** architecture in the Network Core, and how does it relate to **IP** and **WDM** ?



Is it going to be **IP** beside **Ethernet**, both over **WDM** ?

OR

Is it going to be **IP** over **Ethernet** over **WDM** ?



OR →



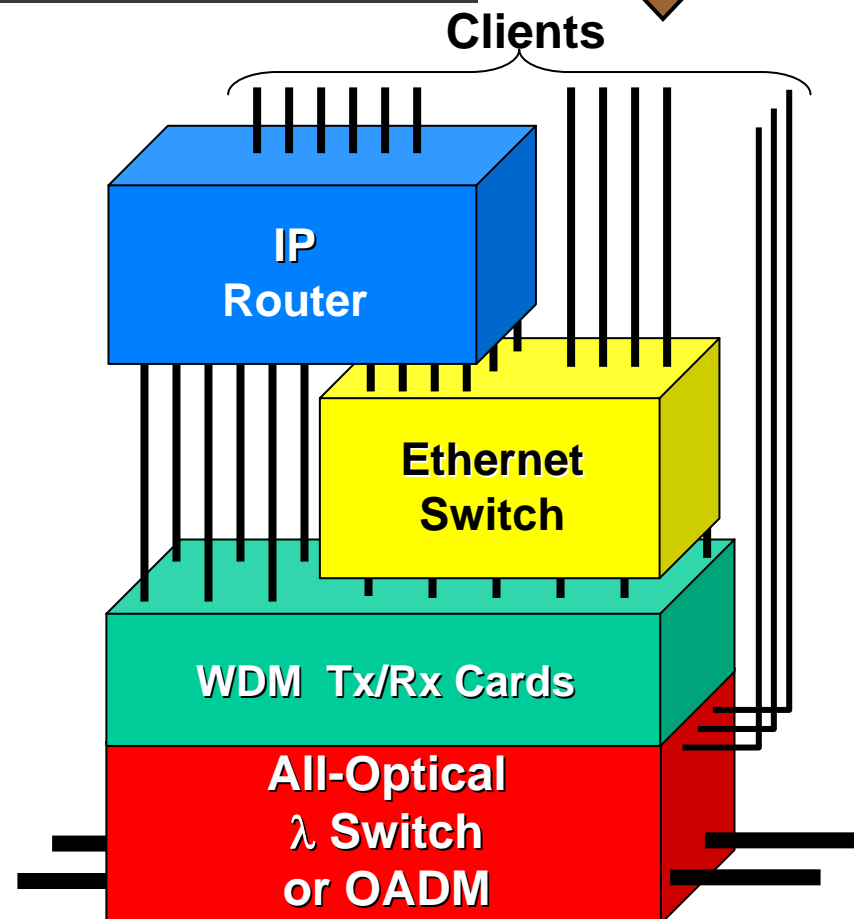
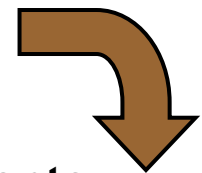
But what is the vision for the **Ethernet** architecture in the Network Core, and how does it relate to **IP** and **WDM** ?



(continued)

OR

Is it going to be a combination of the two architectures ?



These are Interesting Questions !



High-Spectral-Efficiency (SE) Fiber Transmission



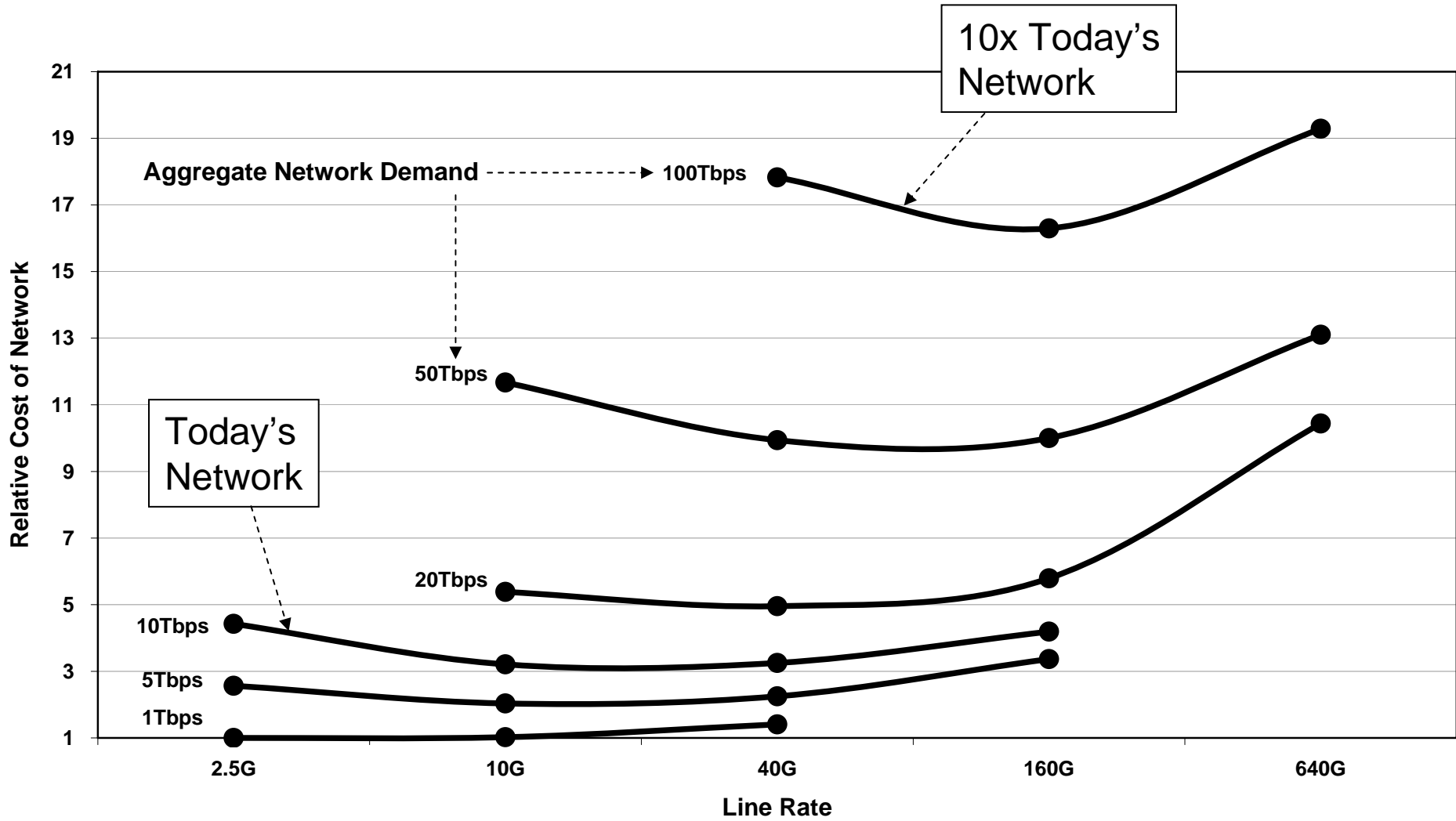
Optical Spectrum		Mod Fmt	Gb/s per λ	$\Delta\lambda$ GHz	SE b/s/Hz	Num λ 's C-Band	Tb/s in C-Band	Tb/s in C,L-Band
Progress to Today		Binary OOK	2.5	100	0.025	40	0.25	0.5
			10	100	0.1	40	0.4	0.8
			10	50	0.2	80	0.8	1.6
Today's Plans		Binary	10	25	0.4	160	1.6	3.2
		OOK	40	100	0.4	40	1.6	3.2
Near-Term Vision		Duo-Binary	40	50	0.8	80	3.2	6.4
		QPSK ?	40	25	1.6	160	6.4	12.8
Future Vision		8-PSK, QAM?	100	50	2.0	80	8.0	16.0
		16-QAM?	100	25	4.0	160	16.0	32.0



- Spectral Efficiency higher than 1.5 b/s/Hz
- 100 Gb/s (or higher?) per wavelength
- Regeneration-free optical reach greater than 1500 km
- Security “encryption” technique that is Low in Cost and in Size, Weight and Power (SWaP)
- Encryption may require a secure system to manage the encryption keys, e.g., something like quantum key distribution



Network Cost as a function of the Line Rate for various Aggregate Network Demands



From: Saleh and Simmons, "Evolution Toward the Next-Generation Core Optical Network", *JLT*, Sept 2006



Optical-Reach & Transponder-Cost Assumptions for the Network Study



Line Rate	Regeneration-Free Optical Reach	Relative Cost of Transponder
2.5 Gbps	4,000 km	1
10 Gbps	3,000 km	$2 \times 1 = 2$
40 Gbps	2,000 km	$2.5 \times 2 = 5$
160 Gbps	1,500 km	$2.5 \times 5 = 12.5$
640 Gbps	1,000 km	$2.5 \times 12.5 \approx 31$

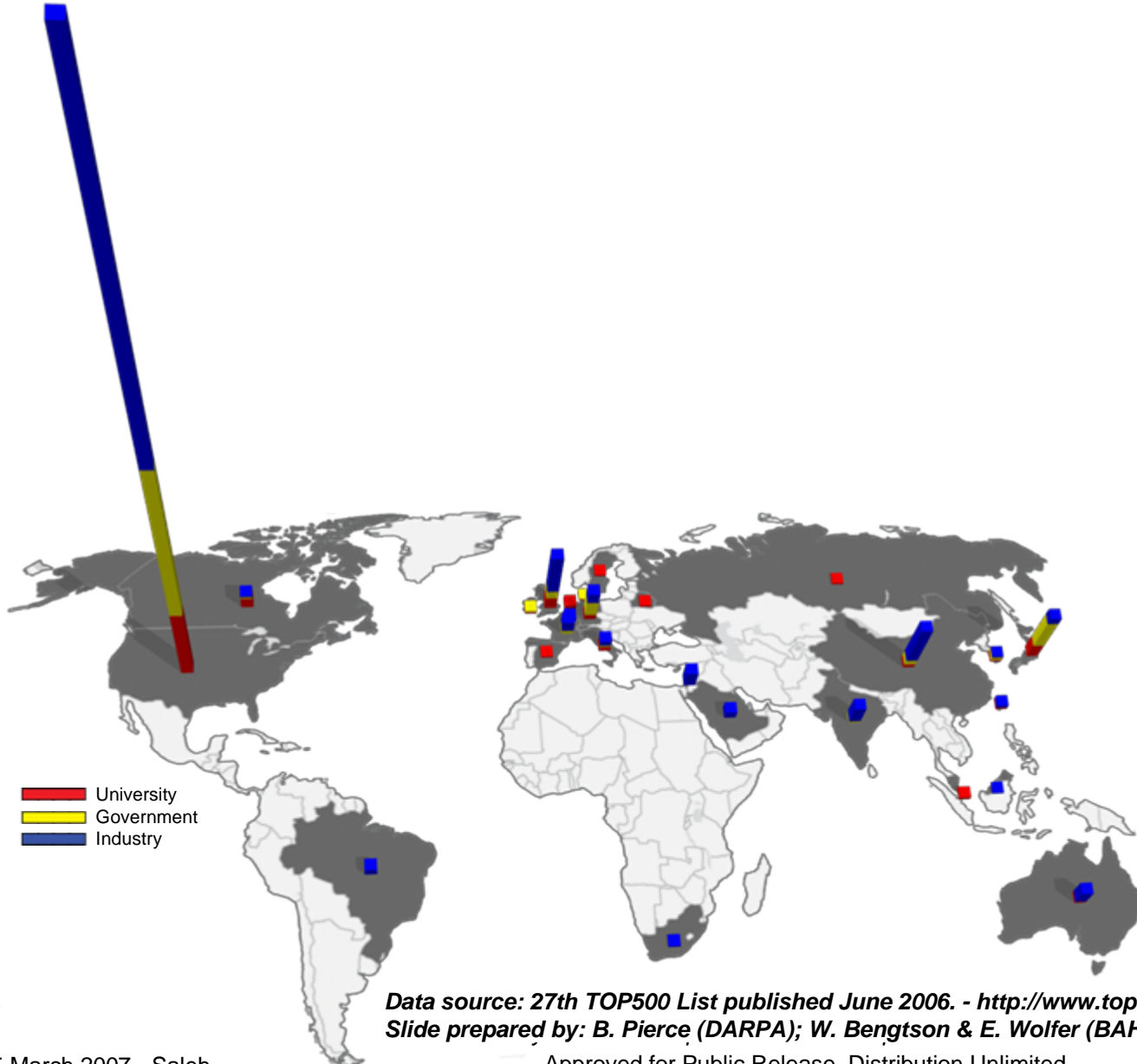
From: Saleh and Simmons, "Evolution Toward the Next-Generation Core Optical Network", *JLT*, Sept 2006



A Grid-Computing Application, and its Requirement for Very Fast Connection Set-Up



Where the top 500 supercomputers reside (June 2006)



country	supercomputers		University Government Industry		
	number	share	U	G	I
United States	299	59.8%	1	1	1
U.K.	35	7.0 %	1	1	1
Japan	29	5.8 %	1	1	1
China	28	5.6 %	1	1	1
Germany	18	3.6 %	1	1	1
India	11	2.2 %	1	1	1
Israel	9	1.8 %	1	1	1
Australia	9	1.8 %	1	1	1
France	8	1.6 %	1	1	1
Canada	8	1.6 %	1	1	1
Italy	7	1.4 %	1	1	1
Switzerland	4	0.8 %	1	1	1
Saudi Arabia	4	0.8 %	1	1	1
New Zealand	4	0.8 %	1	1	1
South Korea	4	0.8 %	1	1	1
Brazil	4	0.8 %	1	1	1
Taiwan	3	0.6 %	1	1	1
Netherlands	3	0.6 %	1	1	1
Singapore	2	0.4 %	1	1	1
Ireland	2	0.4 %	1	1	1
Belgium	2	0.4 %	1	1	1
Sweden	1	0.2 %	1	1	1
Spain	1	0.2 %	1	1	1
South Africa	1	0.2 %	1	1	1
Russia	1	0.2 %	1	1	1
Malaysia	1	0.2 %	1	1	1
Denmark	1	0.2 %	1	1	1
Belarus	1	0.2 %	1	1	1

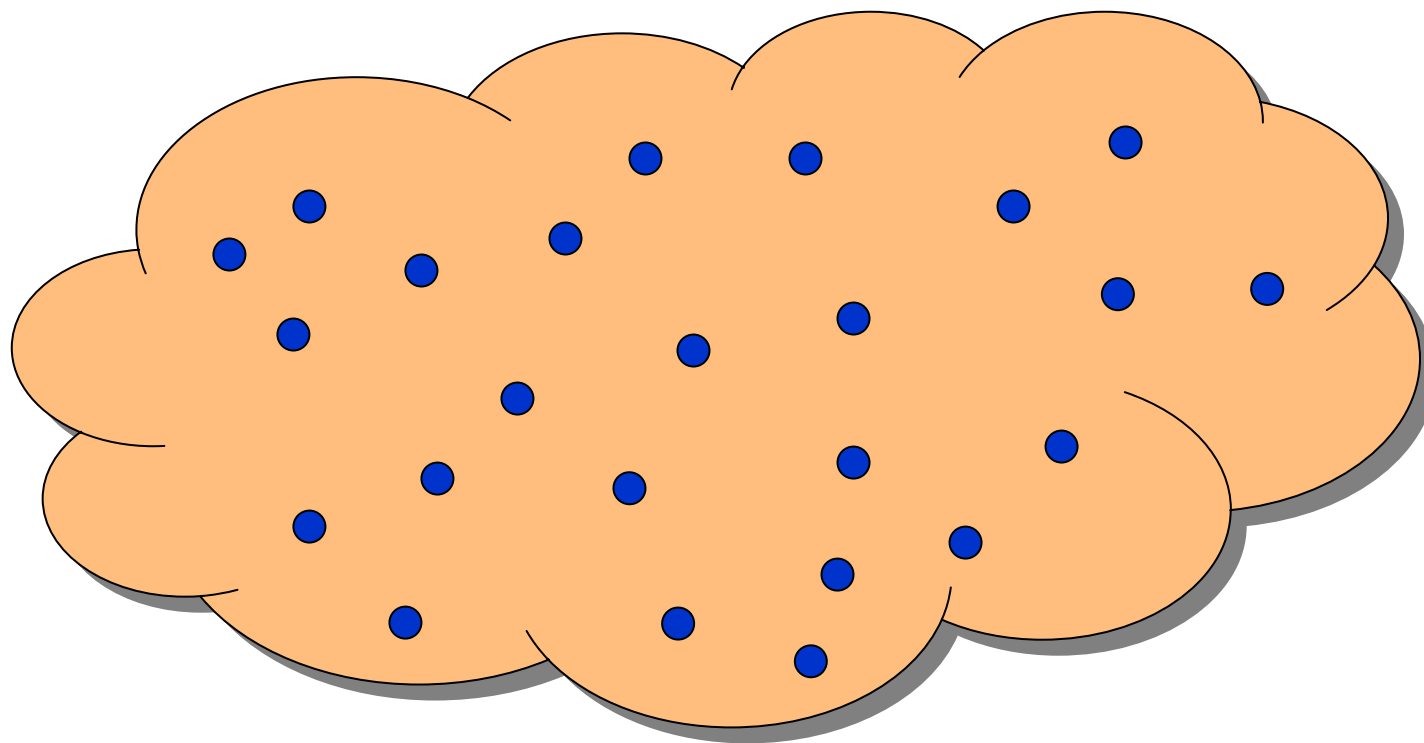
Data source: 27th TOP500 List published June 2006. - <http://www.top500.org/stats>
 Slide prepared by: B. Pierce (DARPA); W. Bengtson & E. Wolfer (BAH/DARPA Ctr.)



A Grid Computing Scenario (1)

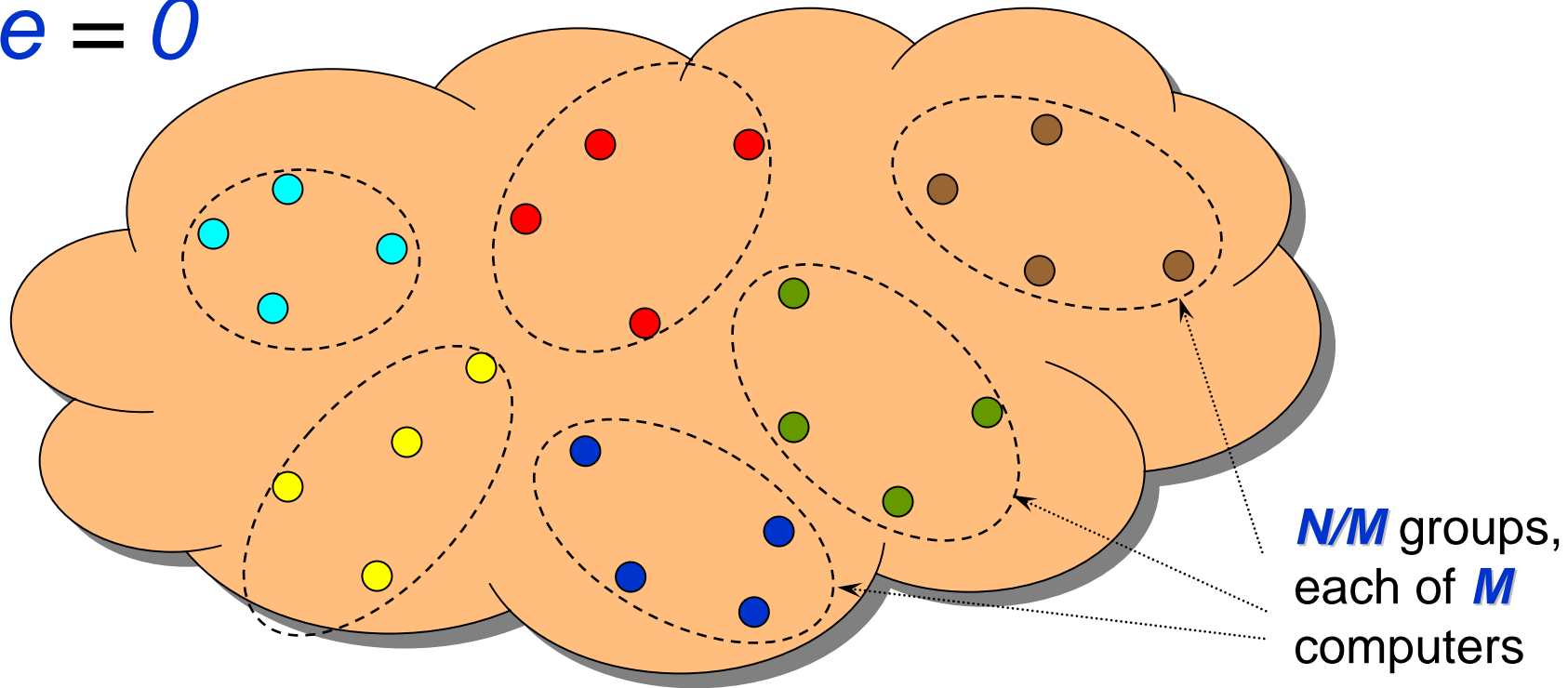


Assume N (say 100) super computers are connected to a national or global network to perform independent computing tasks by forming cooperative computational groups ...



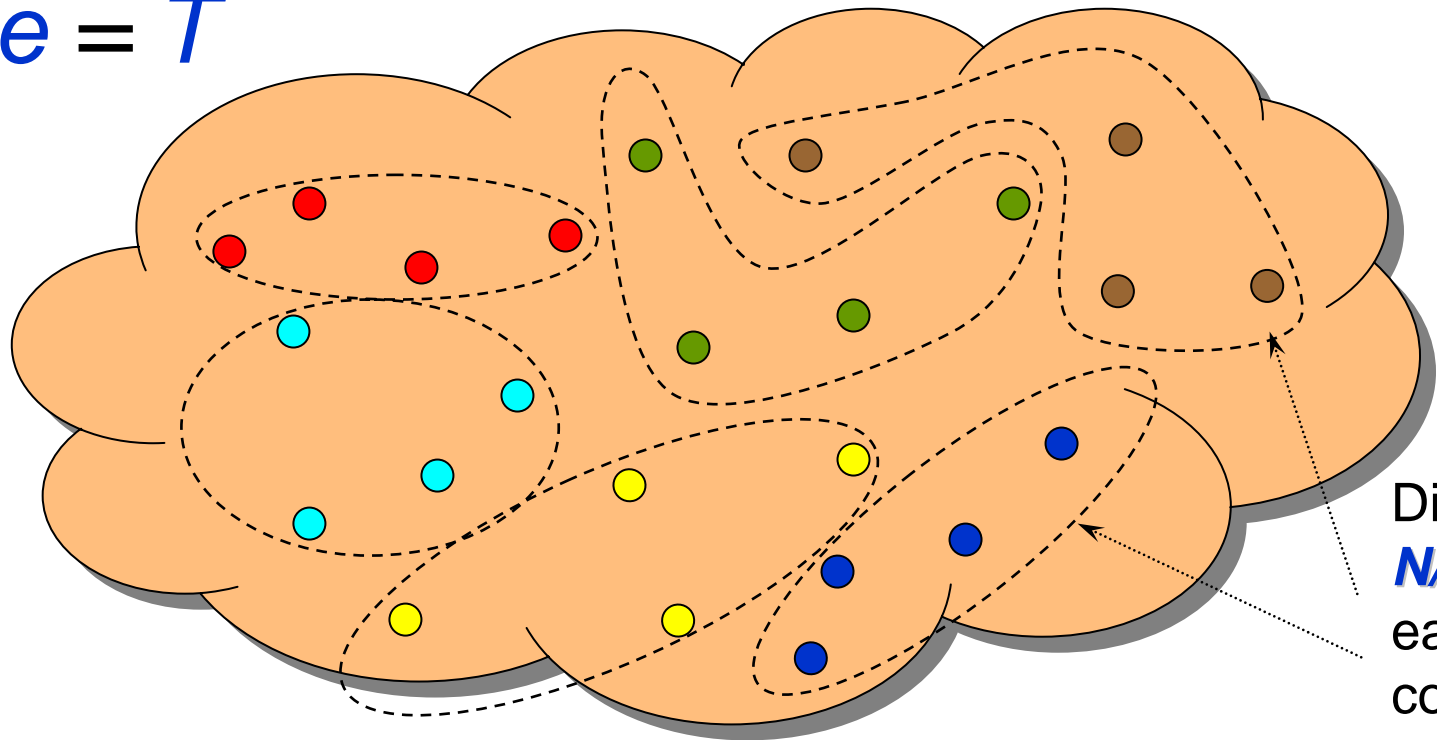
Let M (say 10, which is $\ll N$) be the average number of computers in a computation group

time = 0



The computational groups change after each computation duration, T (say minutes or more)

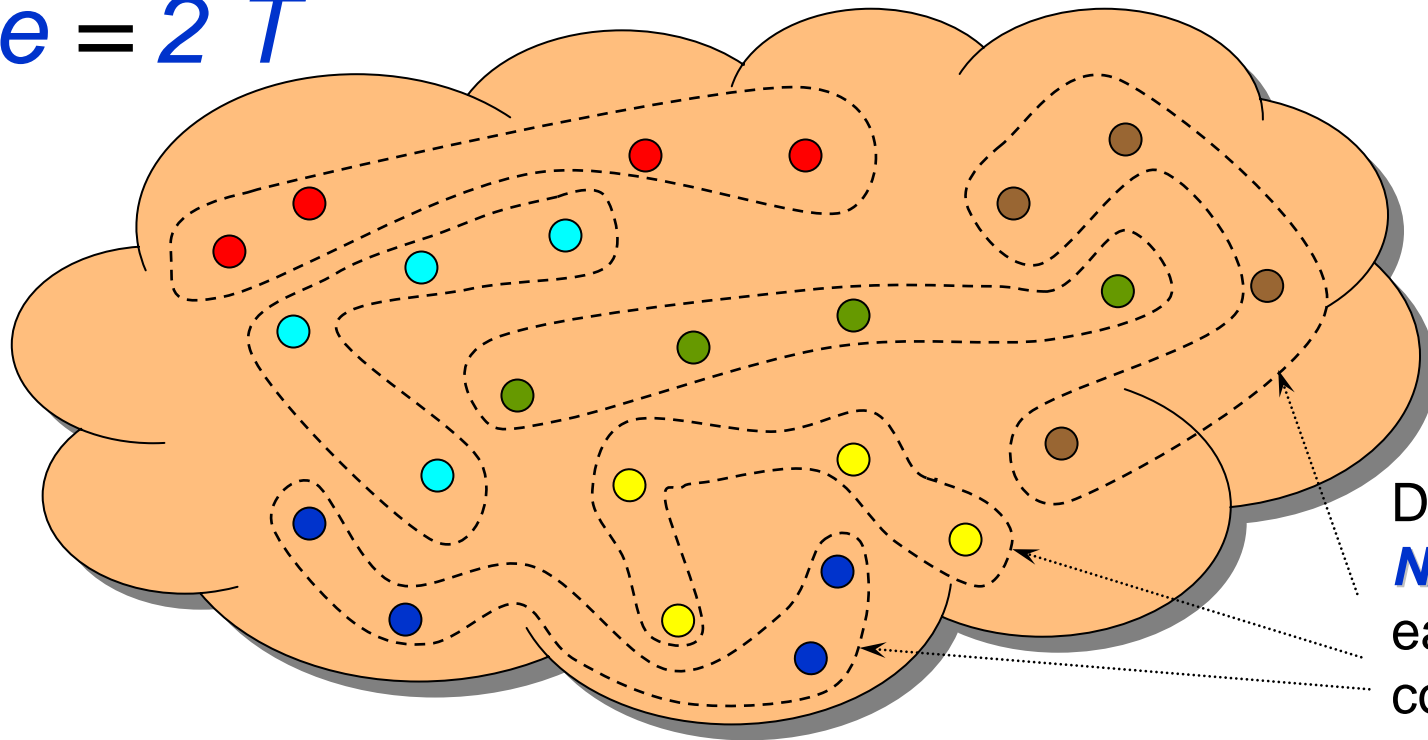
time = T



Different N/M groups, each of M computers

Let the bitrate per connection be R , the average burst duration be τ (milliseconds to seconds), the average number of bursts per computation be ν , and the corresponding average burst communications duty ratio be $\rho = \nu \times \tau / T$ ($\rho \ll 1$)

time = $2 T$



Different N/M groups, each of M computers



A Grid Computing Scenario (Summary of Definitions)



- Consider a network with N distributed-computing nodes
- Assume that the computing nodes need to be connected to each other when and if they have to exchange bursts of data
- To speed-up computations, the bitrate, R , per connection is assumed to be equal to the optical line rate, e.g., 40 Gbit/s
- Not all N nodes need to be connected to each other to perform a single computation; thus assume that, on the average, N/M groups each with M nodes (where $M \ll N$) need to be formed, where communication is needed only *within* a group. In general, the groupings change for each computation
- The duration, T , of a computational task can be large, e.g., minutes
- The lengths of the individual exchanged data bursts, τ , can be small, e.g., milliseconds to seconds
- Let the average number of bursts per computation be ν
- Note that the average burst-communications duty ratio, $\rho = \nu \times \tau / T$, can be considerably less than unity ($\rho \ll 1$)



Networks with Different Configurability Speeds (S) with Some Average Computation Duration (T), Average Computation Burst Duration (τ), and Average Burst Activity Duty Ratio (ρ),



1) Static Network: (i.e., $S \gg T$)

In this case, all N computing nodes need to be connected to each other (a total of $N(N-1)/2$ connections), all of the time

Total Network Bandwidth Requirement:

$$BW1 = R \times N(N-1)/2$$

2) Slowly Configurable Network: (i.e., $T \gg S \gg \tau$)

In this case, one can slowly form N/M , fully connected M -computing-nodes groups on a computation-by-computation basis

Total Network Bandwidth Requirement:

$$BW2 = R \times N(M-1)/2 = (M-1)/(N-1) \times BW1 \quad (\ll BW1)$$

3) Rapidly-Configurable Network: (i.e., $S \ll \tau$)

In this case, one can quickly form any connection as needed on a burst-by-burst basis

Total Network Bandwidth Requirement:

$$BW3 = R \times \rho \times N(M-1)/2 = \rho \times BW2 \quad (\ll BW2)$$



A Grid Computing Scenario (Results for N=100 Computing Nodes)



Assume the following network switching speeds:

- **S** = Several seconds for Slow Optical Circuit Switching (Slow OCS)
- **S** = 50 milliseconds for Fast Optical Circuit Switching (Fast OCS)
- **S** = 50 microseconds for Optical Burst Switching (OBS)

Also assume the following parameters:

- Total number of computational nodes, **N** = 100
- Number of nodes in each computation group, **M** = 10
- Bitrate per connection, **R** = 40 Gbit/s
- Average computation duration, **T** = minutes or more
- Burst durations, τ , varies from 1 second down to 1 millisecond
- Burst-communications duty cycle, ρ = 10%

Approximate Total Network Bandwidth Requirement (Tbit/s)

Switching	$\tau=1$ sec	$\tau=100$ msec	$\tau=10$ msec	$\tau=1$ msec	Comments
Static	200	200	200	200	Required BW too high
Slow OCS	20	20	20	20	10% of BW of Static
Fast OCS	2	4	20	20	1% to 10% BW of Static
OBS	2	2	2	2	1% of BW of Static



The DARPA CORONET Program (BAA06-29)



- Last August, DARPA issued the CORONET BAA (#06-29), entitled “*Dynamic Multi-Terabit Core Optical Networks: Architecture, Protocols, Control and Management*”
See: <http://www.darpa.mil/sto/solicitations/CORONET/>
- This BAA is for Phase 1, which is for developing and testing (by simulations) architecture, protocols, and algorithms (18 month)
- Phase 2 will be for developing and testing a compatible network control and management software suitable for transition to Government and commercial telecommunications carriers
- Thus, the CORONET Program thrust areas are
 - Architecture (*of the Network, Network Node and Network Element*)
 - Protocols and Algorithms (*for very-fast service set-up and restoration*)
 - Network Control and Management
 - But No Hardware Development or Testing
- Proposals for Phase 1 are currently being evaluated for selection



Overview of the *CORONET* Target Network



- Global core optical network
- IP over WDM architecture
- Network services
 - Predominantly IP services (with differentiated QoS)
 - Substantial amount of λ -services
- Scalable for up to 10x increase in aggregate network demand over today's state-of-the-art networks
- Highly dynamic network with very fast service set-up and tear-down
- Resilient to multiple concurrent network failures
- Simplified network operation and increased security



High-Level CORONET Program Goals



Network Requirements	Today's State-of-the-Art Networks	High-Level CORONET Program Goals
Aggregate Network Demand	Up to 10 Tb/s	Up to 100 Tb/s
Maximum Fiber Capacity	Up to 1.6 Tb/s	Up to 20 Tb/s per Fiber
Bit Rate per Wavelength	10 or 40 Gb/s	40 or 100 Gb/s
Maximum Bit Rate per Stream	40 Gb/s	Up to 1 Tb/s
End-to-End Network Services	IP, SONET	IP (75%±) and λ-Services (25%±)
Optical-Layer Multicasting	Typically Not Possible	Basic Requirement
Performance Monitoring	Mostly in Electrical Layer	In Electrical and Optical Layers
Optical-Layer Configurability	Slow, Often Manual	Fast, Fully Automatic
Max Speed of Service Setup	Hours to Weeks	≤ 100 msec (CONUS) ≤ 250 msec (Global)
Speed of <i>Shared</i> Restoration*	50 to 100's msec (Ring) Sec's to Min's (Mesh)	≤ 100 msec (CONUS) ≤ 250 msec (Global)
* <i>From How Many Failures:</i>	Typically, One Failure	Up to Three Failures
IP Services with Differentiated End-to-End QoS	Very Limited	Basic Requirement



The Main *CORONET* Challenge



High Network Efficiency
(i.e., Low Cost, Size, Weight and Power)



The Main *CORONET* Challenge

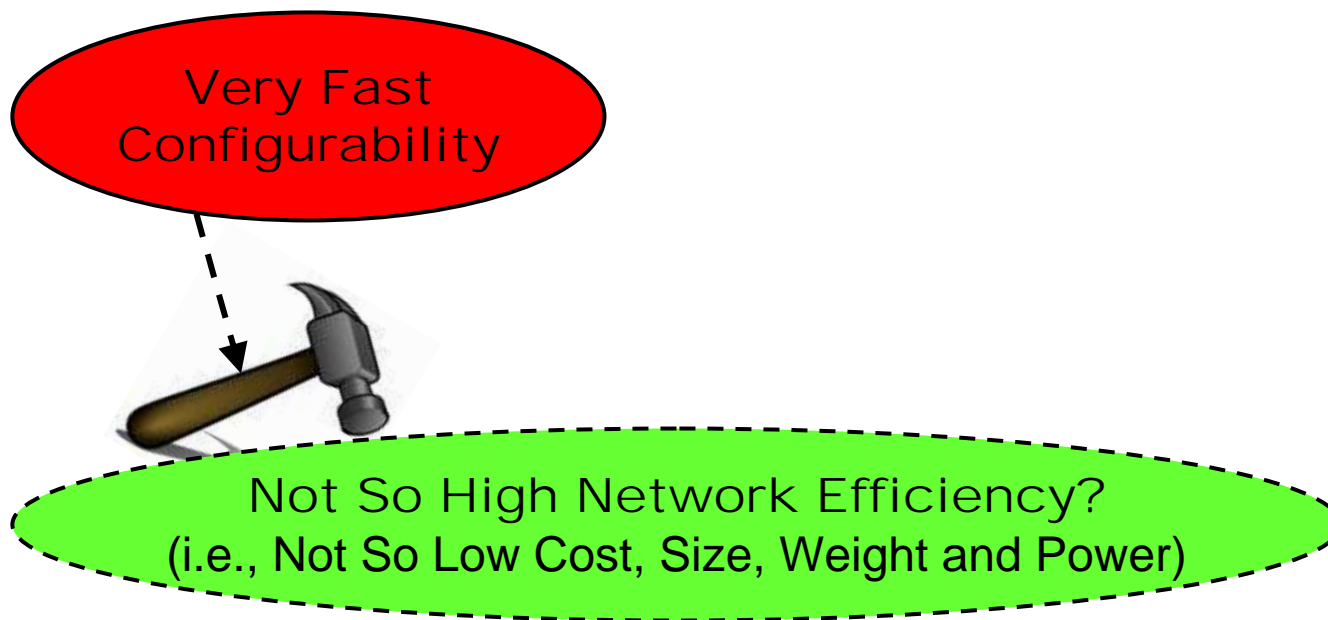


Very Fast
Configurability

High Network Efficiency
(i.e., Low Cost, Size, Weight and Power)

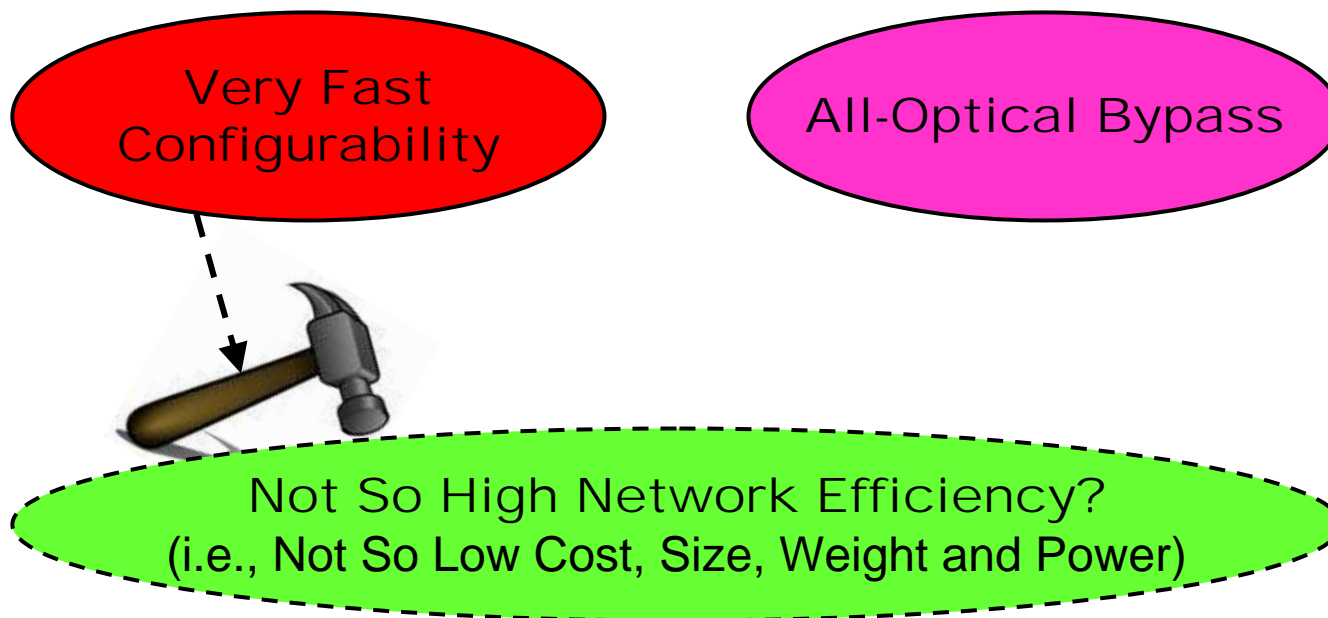


The Main *CORONET* Challenge



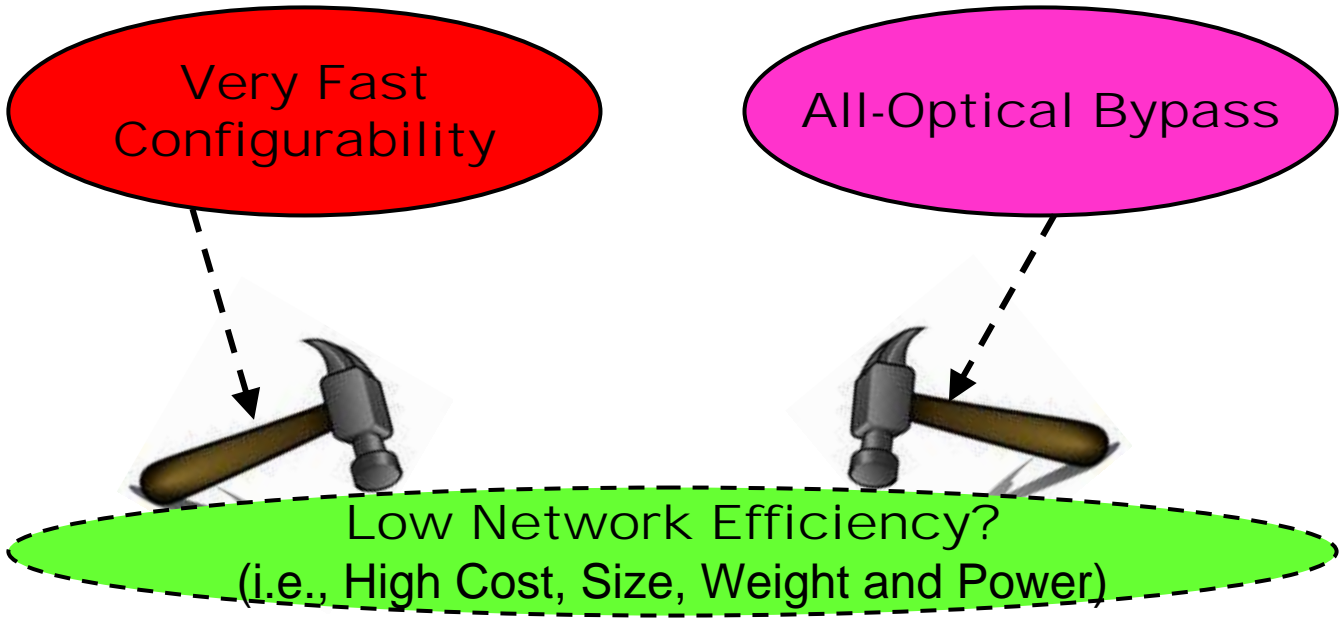


The Main *CORONET* Challenge





The Main CORONET Challenge





The Main CORONET Challenge

