STO
Strategic Technology Office

CORONET Proposers' Day - BAA 06-29

"Dynamic Multi-Terabit Core Optical Networks: Architecture, Protocols, Control and Management"

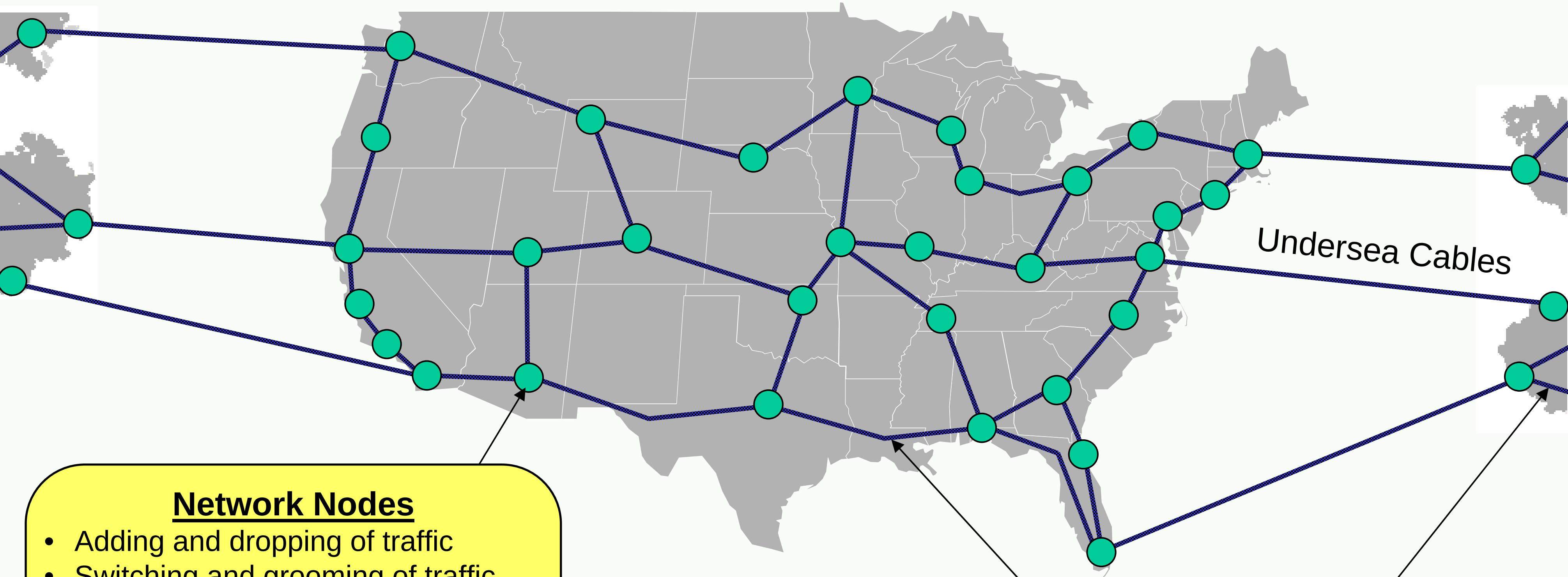
8 August 2006

Adel A. M. Saleh

DARPA / Strategic Technology Office (STO)

Adel.Saleh@darpa.mil

A Global Core Optical Network



Network Nodes

- Adding and dropping of traffic
- Switching and grooming of traffic
- Regeneration of signal as needed, e.g., every 1500 to 2000 km

Network Fiber Links

- Typically one fiber pair, for bidirectional transmission
- 10's to 100's of wavelengths per fiber
- 10's of Gb/s per wavelength
- Optical amplifiers every 60 to 100 km

In this presentation and throughout the PIP, the "***Aggregate Network Demand"*** is defined as the sum of the bit rates of all of the end-to-end demands on the network

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

STO
Strategic Technology Office

Network Requirements	Today's State-of-the-Art Networks	High-Level *CORONET* Program Goals
Maximum Fiber Capacity	Up to 1.6 Tb/s	**Up to 16 Tb/s**
Bit Rate per Wavelength	10 or 40 Gb/s	**40 or 100 Gb/s, or more**
Maximum Bit Rate per Stream	40 Gb/s	**Up to 1 Tb/s**
Aggregate Network Demand	Up to 10 Tb/s	**Up to 100 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 *Dedicated* Protection	≤ 50 msec	**≤ 50 msec**
Speed of *Shared* Restoration*	50 to 100's msec (Ring) Sec's to Min's (Mesh)	**≤ 100 msec (CONUS)** **≤ 250 msec (Global)**
** From How Many Failures:*	*Typically,* ***One*** *Failure*	*Up to* ***Three*** *Failures*
IP Services with Differentiated End-to-End QoS	Very Limited	**Basic Requirement**

Thrust Areas

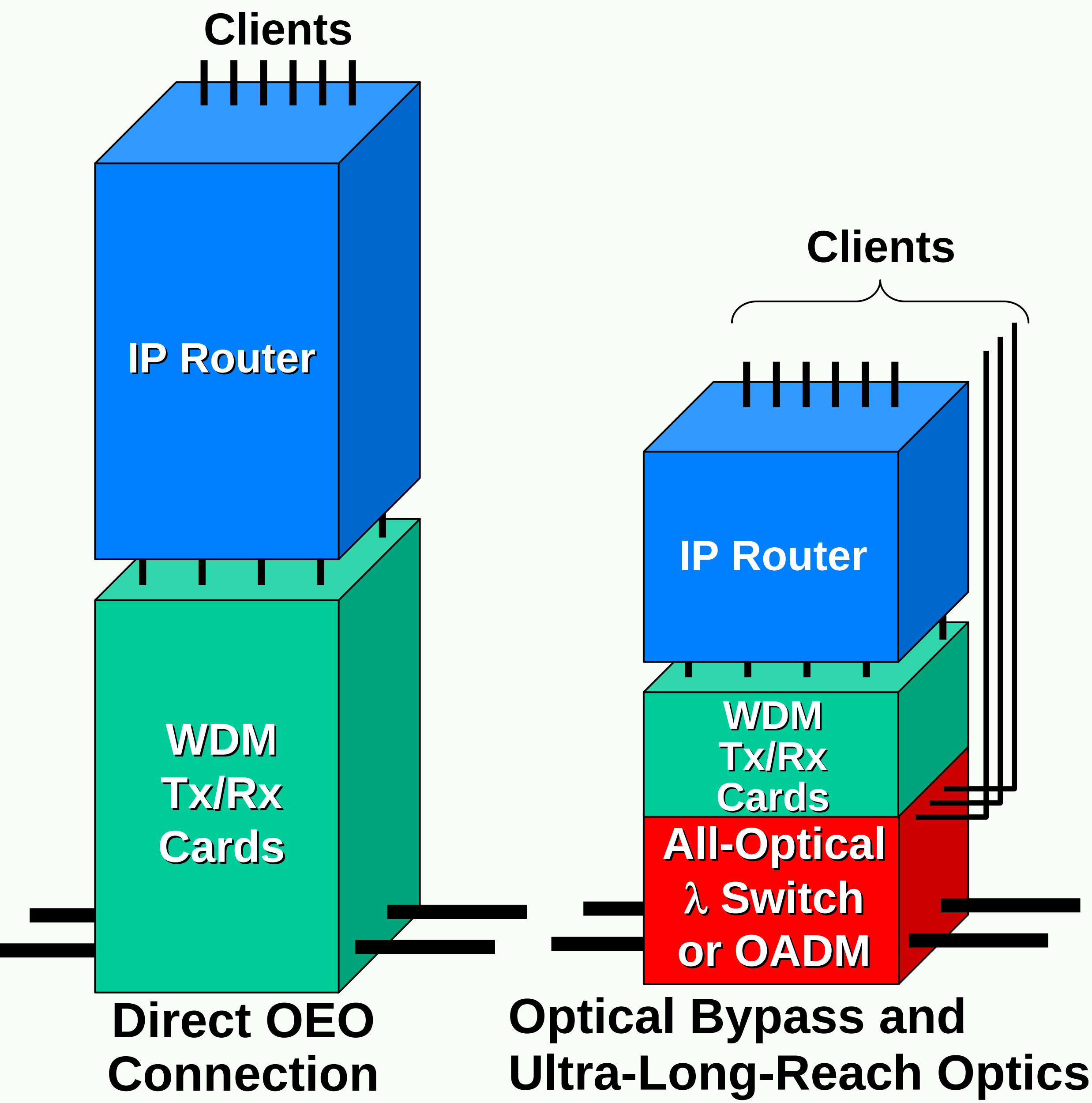
- Architecture
  - IP/WDM network architecture
  - Network node architecture
  - Network element architecture
  - Scalability in cost, size and power
  - Aggregation strategy in edge networks feeding the core
- Protocols and Algorithms
  - Fast service set-up and tear-down
  - Fast, efficient restoration from multiple network failures
  - Optimization of protection across the IP and optical layers
  - Efficient grooming of IP traffic into wavelengths
- Control and Management Planes
  - Division of functions between control and management planes
  - Interplay between the IP and optical control planes
  - End-to-end differentiated QoS for IP services
  - Scalability for large number of connections and flows
  - Hooks for security of the data, control and management planes

Hardware ?

- Hardware architecture and high-level design is <u>in scope</u> for both Phase 1 and Phase 2 of the CORONET Program
- **However, hardware building, testing, experimentation, simulation, etc., is <u>out of scope</u> for both Phases**

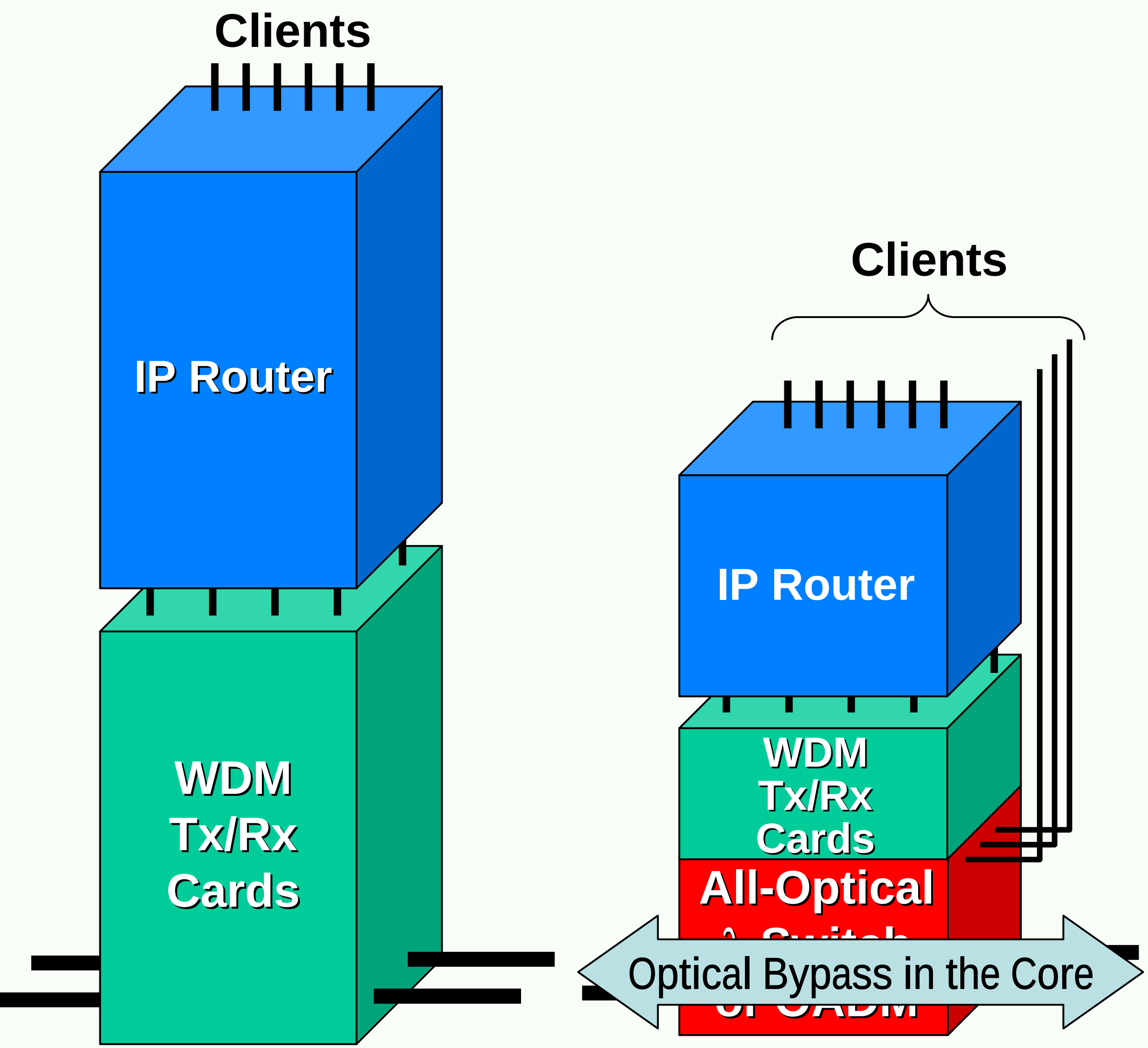
Current State of IP/WDM Network Node Architecture

STO
Strategic Technology Office



Approved for Public Release, Distribution Unlimited

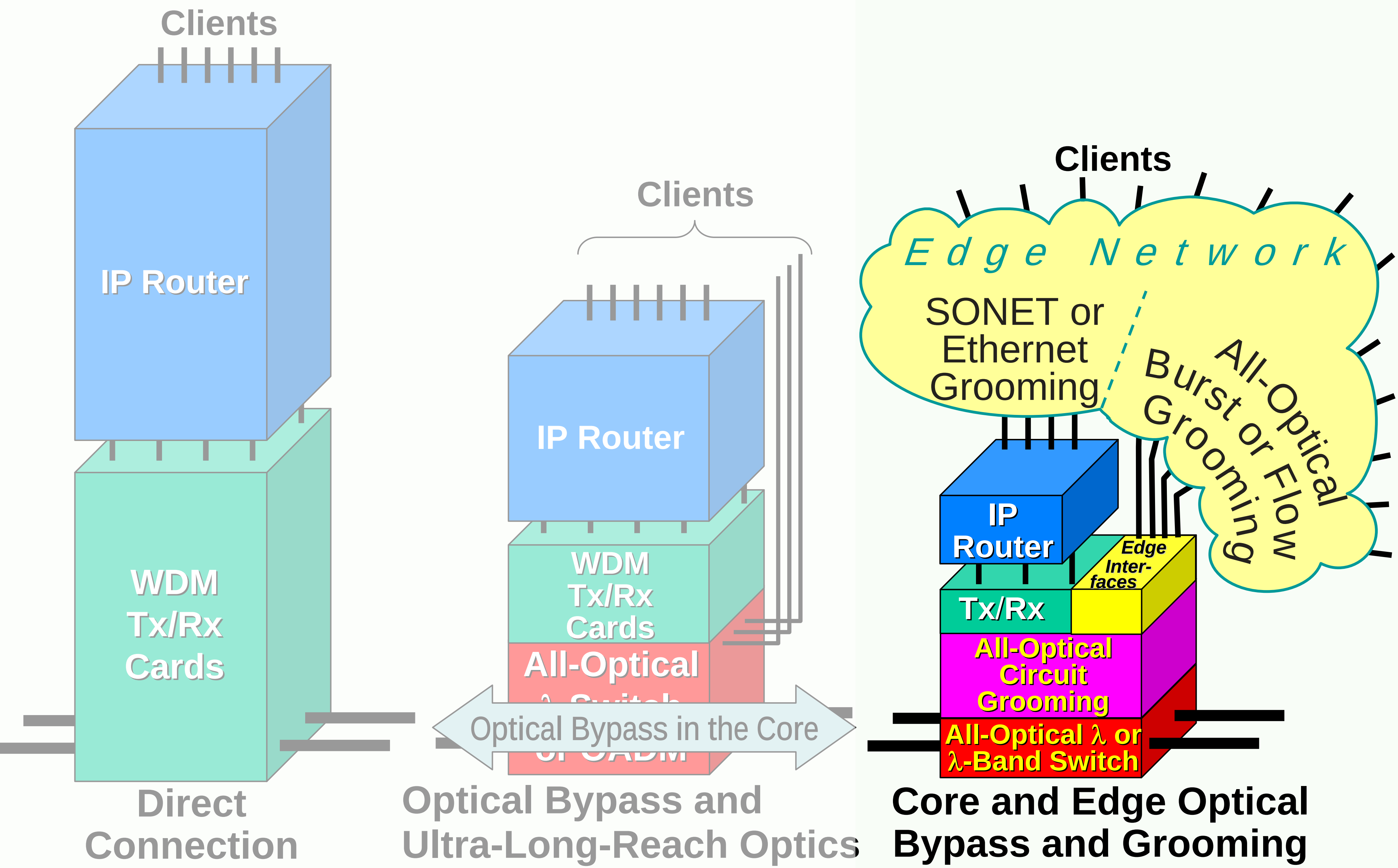
Current State of IP/WDM Network Node Architecture



Direct OEO Connection

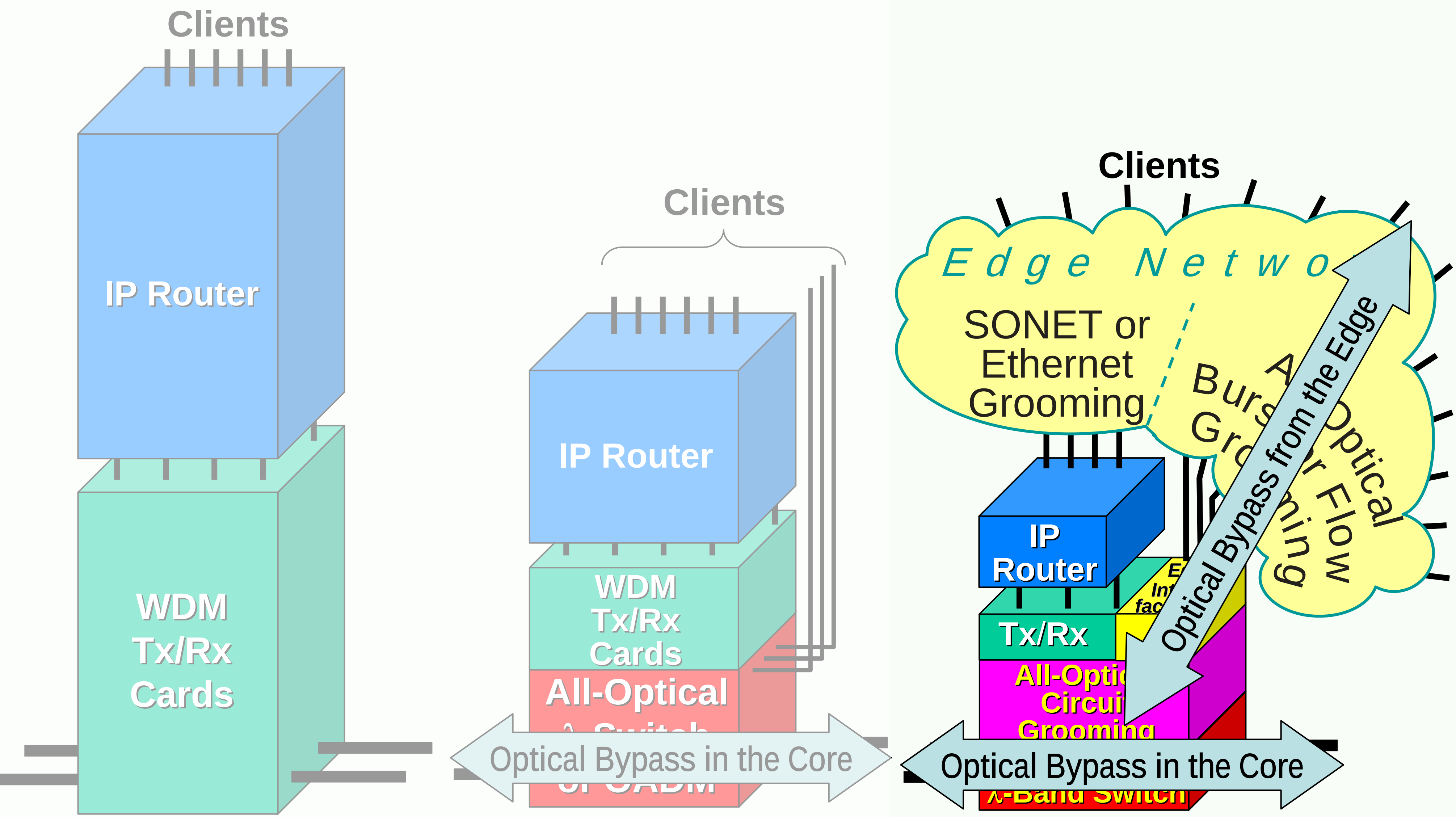
Optical Bypass and Ultra-Long-Reach Optics

Possible Evolution of IP/WDM Network Node Architecture



Possible Evolution of IP/WDM Network Node Architecture

STO
Strategic Technology Office



Direct Connection

Optical Bypass and Ultra-Long-Reach Optics

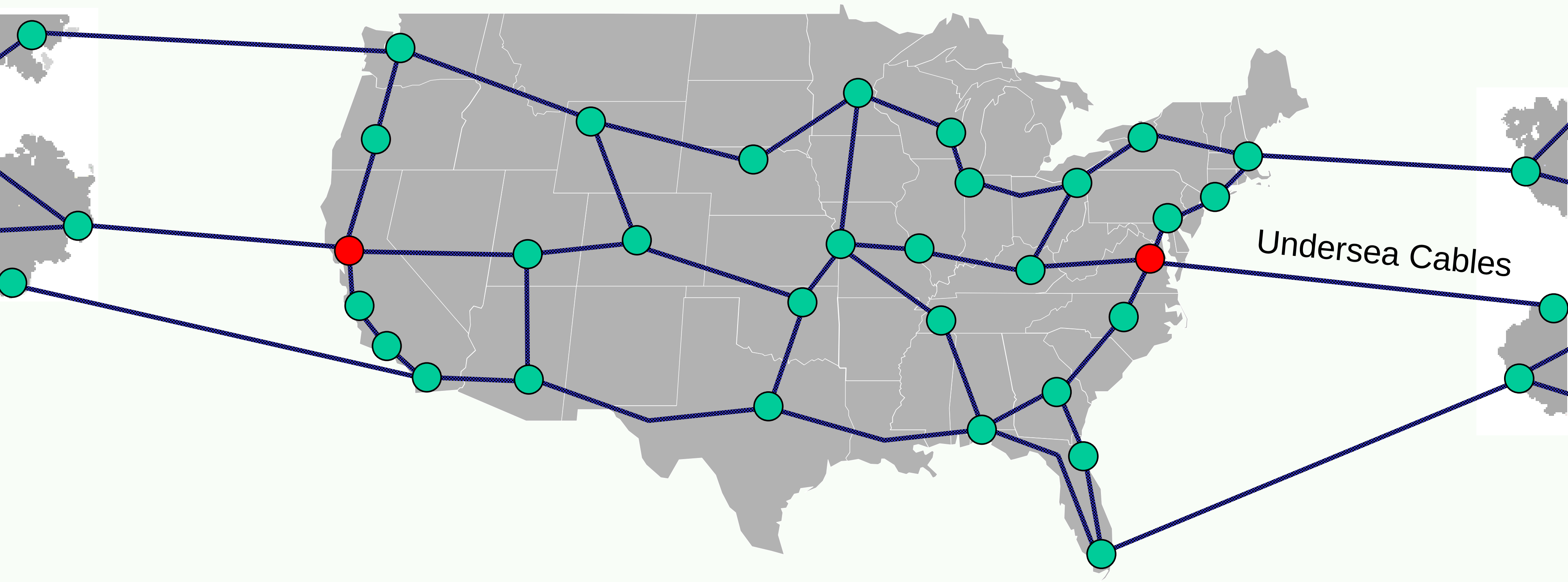
Core and Edge Optical Bypass and Grooming

Specification of the Network Topological Attributes

Table 1. General Topological Attributes of the Target Optical Network

Parameter or Item	**Assumptions**
Number of Nodes in the Core Optical Network	100 nodes in total, of which 75% are within CONUS and 25% are distributed globally
Nodal Degree Distribution	50% of the nodes are degree-2 nodes, 35% degree-3 nodes and 15% degree-4 nodes. If needed, up to 3% of the nodes can be changed to degree-5 or degree-6 nodes
Distance between Adjacent Nodes	The fiber link distance between adjacent nodes is assumed to be 20% longer than the direct 'as-the-crow flies' distance between the nodes
Fiber Topology Diversity	Three diverse east-west paths across CONUS. CONUS is assumed to be connected to the global network via stand-alone undersea systems with three diverse cables in each of the Atlantic and Pacific oceans, with sufficient capacities. Other undersea cables across the globe can be assumed as needed
Node Sizes and Traffic Distribution	A realistic, i.e., non-uniform, traffic should be assumed. As a guide, if 20% of the nodes are assumed to be "large" nodes, and the rest "small" nodes, the traffic distribution among the two types of nodes should be approximately: Large-Large: 40%, Large-Small: 40%, Small-Small: 20%

Note Three Diverse East-West Paths across CONUS



Network Services

Table 2. Network Services

High-Level Service Type	**IP Services**				**Wavelength Services [3]**			
High-Level Traffic Breakdown [1]	**$S \in [50\%, 100\%]$ The nominal value $S_o = 75\%$**				**1-S The nominal value $1\text{-}S_o = 25\%$**			
Service Sub-Type	**Best-Effort IP Services**	**Guaranteed-Bandwidth IP Services**			**Single-λ Services**	**Double-λ Services**	**Quad-λ Services**	**Octal-λ Services**
		Fine Granularity	**Medium Granularity**	**Coarse Granularity**				
Data Rate (bits/sec) [2]	**Bursty**	**< 400 Mb/s**	**0.4 to 4 Gb/s**	**4 Gb/s to 0.25 R**	**R**	**2 R**	**4 R**	**8 R**
Service Sub-Type Traffic Breakdown	**40% of S**	**15% of S**	**20% of S**	**25% of S**	**40% of (1-S)**	**20% of (1-S)**	**20% of (1-S)**	**20% of (1-S)**

(1) This row represents the fraction of the aggregate network demand that is in each high-level service category, where the aggregate network demand is provided in Table 3 for four design scenarios. This row is specified in terms of *S*, where *S* can take on a range of values, as specified.

(2) The data rates of some of the service types are specified in terms of *R*, the bit rate per wavelength, which is provided in Table 3 for four design scenarios.

(3) Single- or multi-wavelength services are digital-circuit services of the data rate specified.

Alternatives to the Pure IP/WDM Architecture ?

"While the CORONET program specifies an IP-over-WDM architecture, it is recognized that this is not the only model possible for carrying future services. In addition to analyzing the IP-over-WDM model, as specified in this PIP, performers are permitted to also propose and analyze other models if they can provide convincing data that these models offer significantly higher performance with comparable or lower cost, or comparable or higher performance with significantly lower cost."

[from BAA06-29 CORONET PIP, p. 11]

Asymmetric and Multicast Services ?

"For purposes of evaluation against the metrics of Section 2.5, all demands should be assumed point-to-point and bi-directionally symmetric. However, support for asymmetric services and multicast services must be included as part of the proposed solution."

[from BAA06-29 CORONET PIP, p. 11]

Network Design Scenarios

Table 3. Scenarios for Aggregate Network Demand, Bit Rate per Wavelength, and Fiber and Network-Link Capacities

	Scenario 1	**Scenario 2**	**Scenario 3**	**Scenario 4**
Aggregate Network Demand (1) (**Tb/s**)	20	40	50	100
Bit Rate per Wavelength (**R**) (**Gb/s**)	40	40	100	100
Capacity per Fiber (**Tb/s**)	4	4	10	10
Number of Wavelengths Per Fiber	100	100	100	100
Nominal Number of Fiber Pairs per Network Link (2)	1	2	1	2

(1) The *Aggregate Network Demand* is defined as the sum of all bi-directional end-to-end demands on the network.

(2) Ideally, the aggregate network demand is carried using no more than the number of fiber pairs per link specified for each scenario. However, if necessary, one extra fiber pair can be added on heavily loaded network links as specified in Table 6. The undersea cables associated with each scenario are assumed to have sufficient capacity.

Optical Hardware Technology Assumptions

Table 4. Optical Hardware Technology Assumptions for the Target Optical Network

Parameter or Item	Assumptions
Capacity per Fiber	Specified in Table 3 for each design scenario
Number of Fiber Pairs per Network Link	Specified in Table 3 for each design scenario
Bit Rate per Wavelength	As given in Table 3, the bit rate per wavelength is either 40 Gb/s or 100 Gb/s, depending on the scenario
Regeneration-Free Optical Reach	The optical reach for the 40 Gb/s wavelengths is assumed to be 2000 km and 10 optical network elements (e.g., all-optical switches or OADMs), and for the 100 Gb/s wavelengths is assumed to be 1500 km and 7 optical network elements
Node Architecture	Each network node is assumed to include an all-optical switch or an OADM (depending on the degree of the node) to add, drop and/or switch the various wavelengths, as well as to provide optical bypass for traffic that does not need to drop at the node. Electronic equipment located at the nodes (e.g., IP routers) is to be connected to the network through the drop-side of the all-optical switch or OADM
Optical Switching Granularity	Wavelengths or wavebands as dictated by the architecture
Network Reconfigurability	Completely automated without any manual intervention
Tunability	All Tx/Rx cards and regenerators (OEO or OOO) are assumed to be tunable across all wavelengths

Flexibility of the Optical Network Elements ?

“For purposes of evaluation against the metrics of Section 2.5, assume fully flexible optical network elements; for example, an element can be remotely configured to allow any transmit/receive (Tx/Rx) card plugged into it to add/drop to/from any network port on any wavelength. However, it is also necessary to analyze how the proposed solution would perform using less flexible optical network elements that may have lower cost [or may be the only ones available].”

[from BAA06-29 CORONET PIP, p. 14]

Metrics for Service Setup

Table 5. Performance Metrics for Service Setup (applies to each design scenario given in Table 3)

		Service Setup Classes			
		Very Fast Service Setup	**Fast Service Setup**	**Scheduled Service Setup**	**Semi-Permanent**
Traffic Breakdown (1)	**Guaranteed-Bandwidth IP Services**	0%	20% of S	20% of S	20% of S
	Wavelength Services	40% of (1-S)	20% of (1-S)	20% of (1-S)	20% of (1-S)
Assumptions		These three cases represent traffic churn that must be satisfied using pre-deployed equipment			Equipment can be installed as needed
Setup Time		CONUS: ≤ **100 msec** (2) Global: ≤ **250 msec** (2)	**≤ 2 sec**	**≤ 10 sec**	N/A
Holding Time		1 sec to 1 min	10 sec to 10 hr	1 min to 1 Month	Months
Protection		Single Failures Only	See Table 6		
Blocking Probability		**10^{-3}**	**10^{-3}**	**10^{-4}**	N/A
Network-Capacity Limit		See Table 6			

(1) This represents the fraction of the aggregate network demand that is in the various setup-time classes. The aggregate network demand is provided in Table 3 for four design scenarios. *S* is defined in Table 2. Note that best-effort IP services (which comprise 40% of *S*, as specified in Table 2) are not included here.

(2) The required very fast service-setup times can be defined more generally as: 50 msec + The round-trip fiber transmission delay between the source and destination nodes.

Table 6. Performance Metrics for Fast Restoration from Network Failures (applies to each design scenario given in Table 3)

	Restoration Classes		
	Single-Failure Restoration	**Double-Failure Restoration**	**Triple-Failure Restoration**
Description	Capable of restoration from any single network failure (i.e., a link or node failure)	Capable of restoration from up to two network failures, with no more than one node failure (3), (5)	Capable of restoration from up to three network failures, with no more than one node failure (4), (5)
Traffic Breakdown (1)	(1 - 0.4 S) - 5 P	4 P	P
	Where the *Protection Parameter* P ∈ [0%, 5%]. The nominal value is P_o = 2.5%		
Restoration Time (2)	*Single-Failure Restoration Time (SFRT):* (2) CONUS: ≤ **100 msec** Global: ≤ **250 msec**	First failure: ***SFRT*** Second failure: ***SFRT* + 50 msec**	First failure: ***SFRT*** Second failure: ***SFRT* + 50 msec** Third failure: ***SFRT* + 100 msec**
Total Spare Protection Capacity Relative to Total Working Capacity	This ratio should be minimized as much as possible for each network design. For the nominal network designs (6), this ratio should **not exceed 75%** for the CONUS-only capacities (See Section 2.5.2 for details on how this ratio is to be measured)		
Network-Capacity Limit	Ideally, the fast-restoration metrics given above, as well as the service-setup metrics given in Table 5 should be simultaneously met with the link capacities specified in Table 3 for each design scenario. Extra capacity may be added if necessary, but the number of network links requiring one extra fiber pair should be minimized as much as possible. For the nominal network designs (6), **no more than 10%** of the network links should have one extra fiber pair.		

(1) This row gives the fraction of the aggregate network demand requiring the corresponding protection level. This row is specified in terms of *P*, where *P* can take on a range of values, as specified. All traffic in the Very Fast Setup class is assigned to single-failure restoration (see Table 5). All other protected demands should be arbitrarily assignable across the three protection classes, as long as the distribution of traffic falls within the indicated percentages. Best-Effort IP Services, which represent 40% of the IP traffic, are not included in any of the three restoration classes.

(2) *Single-Failure Restoration Time (SFRT)* can be defined more generally as: 50 msec + The round-trip fiber transmission delay between the source and destination nodes.

(3) Only applies to traffic between node pairs with network connectivity ≥ 3.

(4) Only applies to traffic between node pairs with network connectivity ≥ 4.

(5) In the case of multiple failures, assume there is enough time between failures to recompute backup paths.

(6) For each of the design scenarios of Table 3, the nominal network design parameters are S_o and P_o (S_o is defined in Table 2).

QoS Metrics for IP Services

Table 7. QoS Metrics for IP Services (applies to each design scenario given in Table 3)

	Best-Effort IP Services [2]	Guaranteed-Bandwidth IP Services [3]	
		Latency- and Jitter-Sensitive Services	**Packet-Loss-Sensitive Services**
Traffic Breakdown [1]	40% of S	30% of S	30% of S
One-Way Latency	CONUS: **100 msec** [4] Global: **250 msec** [4]	CONUS: **50 msec** [5] Global: **125 msec** [5]	**500 msec**
One-Way Jitter [6]	CONUS: **20 msec** Global: **50 msec**	CONUS: **10 msec** Global: **25 msec**	**100 msec**
Packet Loss	10^{-3}	10^{-3}	10^{-6}

(1) This row represents the fraction of the aggregate network demand that is in each QoS category. The aggregate network demand is provided in Table 3 for four design scenarios. *S* is defined in Table 2.

(2) For Best-Effort IP Services, the QoS metrics are statistical averages that are not guaranteed for any individual flow.

(3) For Guaranteed-Bandwidth IP Services, the QoS metrics are to be guaranteed for every individual flow.

(4) In general, this is equal to: 2×(25 msec + the one-way, source-to-destination fiber transmission delay).

(5) In general, this is equal to: 25 msec **+** the one-way, source-to-destination fiber transmission delay.

(6) In general, this is 20% of the one-way latency.

Teaming Recommendations

"Proposals addressing only a subset of the topics will be considered non-responsive to this BAA. As such, teaming of proposers to assemble the necessary breadth and depth of expertise and resources is recommended."

[from BAA06-29 CORONET PIP, p. 4]

Program Phases

- Phase 1 – 18 months
  - Design overall architecture and necessary protocols and algorithms
  - Develop appropriate-fidelity-models of the system
  - Deliver valid and verifiable simulation results that meet or exceed the metrics (Tables 5-7)
  - Deliver full proposal for Phase 2 (more details on this will be provided during Phase 1)
- Phase 2 – time period not yet determined
  - Develop, integrate, demonstrate and test the network control and management software for implementing the Phase 1 vision
  - An important objective of Phase 2 is to make the results of the program attractive for transition to commercial telecom carriers
  - Add (as fully costed *Options*) features to the Phase 1 vision that may help in achieving the above objective, e.g.,
    - *Support for evolution of technology over the next several years*
    - *Support for hybrid line-rate systems (e.g., 10, 40 and 100 Gb/s)*
    - *Extension of the operation across different administrative domains*
    - *Addressing necessary security concerns*
    - *…, etc.*

Programatics

- This BAA is open for three years from the date of issue, but the Government does not anticipate awarding any contracts outside two selection windows:
  - Phase 1 selection window is open for proposals received by 10/27/2006
  - Phase 2 selection window is to be announced during Phase 1 execution
- Proposers [primes] for Phases 1 and 2 must be U.S. companies
  - Contractor teams for Phase 1 may include foreign firms, foreign universities and foreign individuals *(see conditions on page 4 of CORONET PIP)*
  - It is possible that the Phase 2 effort, or portions thereof, will significantly limit foreign participation in accordance with U.S. Export Control Laws
- Primes can form new teams when bidding for Phase 2
- Bidders for Phase 2 should have verifiably demonstrated that they successfully met all Phase 1 goals and metrics

Independent Test and Evaluation in Phase 1

- MIT Lincoln Lab has been chosen as the independent test and evaluation (ITE) agent for Phase 1 of the program
- The demonstration of Phase 1 metrics is to be done via verifiable simulations using the OPNET environment
  - By verifiable, it is meant that example network and models could be run, with assistance from the performers, on an OPNET simulation environment at the ITE agent's facility
- More on this during the next presentation by MIT LL

Overview of Technical Items Required in the Proposal

- High-level description of proposed nodal architecture
- High-Level description of proposed network elements
- High-level description of proposed architecture for the control and management planes
- High-level description of proposed algorithms and protocols
- Results of the network study defined in the CORONET BAA/PIP (to be discussed in the next slide)
- High-level proposal for Phase 2 of the program

Network Study Required in the Proposal

STO
Strategic Technology Office

- The main purpose of the network study is to gauge preliminary algorithmic performance with respect to efficient restoration from multiple network failures
- The study is required only for Design Scenario 1 in Table 3 (i.e., aggregate network demand of 20 Tb/s)
- The network study is restricted to the CONUS portion of the topology; i.e., the 20 Tb/s is entirely CONUS traffic
- The study includes a mix of IP and wavelength services
- 100% fixed traffic is to be assumed, i.e., no churn
  - Assume all demands are added at one time
- Network topology, traffic demand and performance curves are required
- Restoration times, service setup times and IP QoS are <u>not</u> a part of this study
- The full report on this study is to be limited to 7 pages

Proposal Evaluation Criteria

- Innovative Technical Approach
  - Efficacy, novelty and practicality of proposed solution
  - How convincing that proposed solution will achieve goals
  - Realism in technical and economic assumptions
- Network Study
  - Practicality and efficiency of preliminary restoration algorithms
  - Level of networking expertise demonstrated
- High-Level Phase 2 Proposal
  - Appropriateness of the high-level development and test plans
  - Realism in estimating the required effort and ROM cost
  - Adherence to the Phase 1 vision
  - Credibility of the commercial transition strategy
- Work Plan, Management Approach and Past Experience
  - Appropriateness of statement of work, timeline, and milestones
  - Areas of expertise and teaming arrangements
- Potential Contribution and Relevance to DARPA's Mission
  - Identifying path for transition to commercial and government networks
- Cost Reasonableness and Realism