Multi-layer protection and restoration requirements

Date: **July 23rd , 2001**

Source:

Sudheer Dharanikota
Raj Jain
Nayna Networks

Yong Xue
Erik Sherk
WorldCom

ABSTRACT: **This document provides the necessary background required for multi-layer protection and restoration from the point-of-view of the faults. It then derives different possible requirements in such a topology. The goal of this document is to know the open points in the current technology and decide which should be emphasized in near future in the scope of OIF.**

For additional information contact:
The Optical Internetworking Forum, 39355 California Street,
Suite 307, Fremont, CA 94538
510-608-5990 phone ☞ info@oiforum.com

Multi-layer Protection and Restoration Requirements

Sudheer Dharanikota, Raj Jain - Nayna Networks inc.

Yong Xue, Erik Sherk – Worldcom

Goals

- To develop the requirements for protection and restoration at Optical, SONET and Packet layers.
- To determine the protocol extensions for the multi-layer protection and restoration in the scope of NNI.

Overview

- Background on P&R
- Timing analysis
- Multi-layer P&R
- Conclusions

Overview

- **Background on P&R**
  - Goal: Identify faults and find appropriate P&R mechanisms
- Timing analysis
- Multi-layer P&R
- Conclusions

Terminology

- Layers: Packet, TDM, Optical
- Servers vs Client Layers: Optical vs TDM, or TDM vs Packet
- Connection: LSP, ckt, wavelength
- Link = Group of connections = Multiple LSPs, Waveband

Domains

- A set of resources that together can survive a specified number of faults
- Examples:
  - 1+1 Protected Span
  - BLSR
  - UPSR
  - Mesh
  - Redundant set of cards or switching fabrics
- A connection may traverse a number of domains

Fault Characteristics

```
                              Fault
                    /                      \
                Scope                      Layer
               /      \                  /    |    \
      Administrative   Extent      Optical  SONET  Packet
         /    \        /   |    \
    Carrier  Customer Link Domain Multi-domain
```
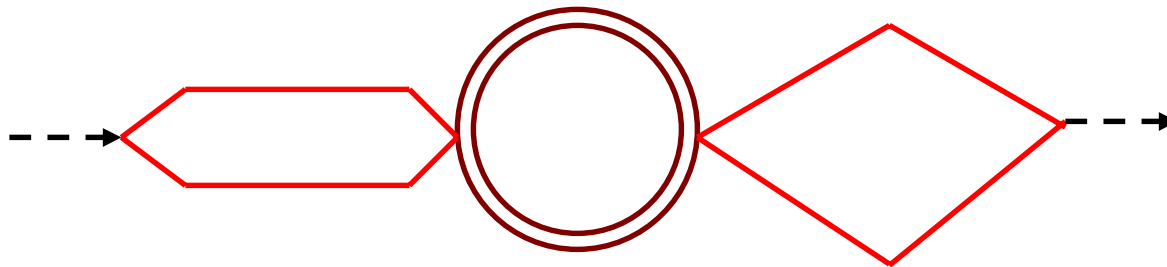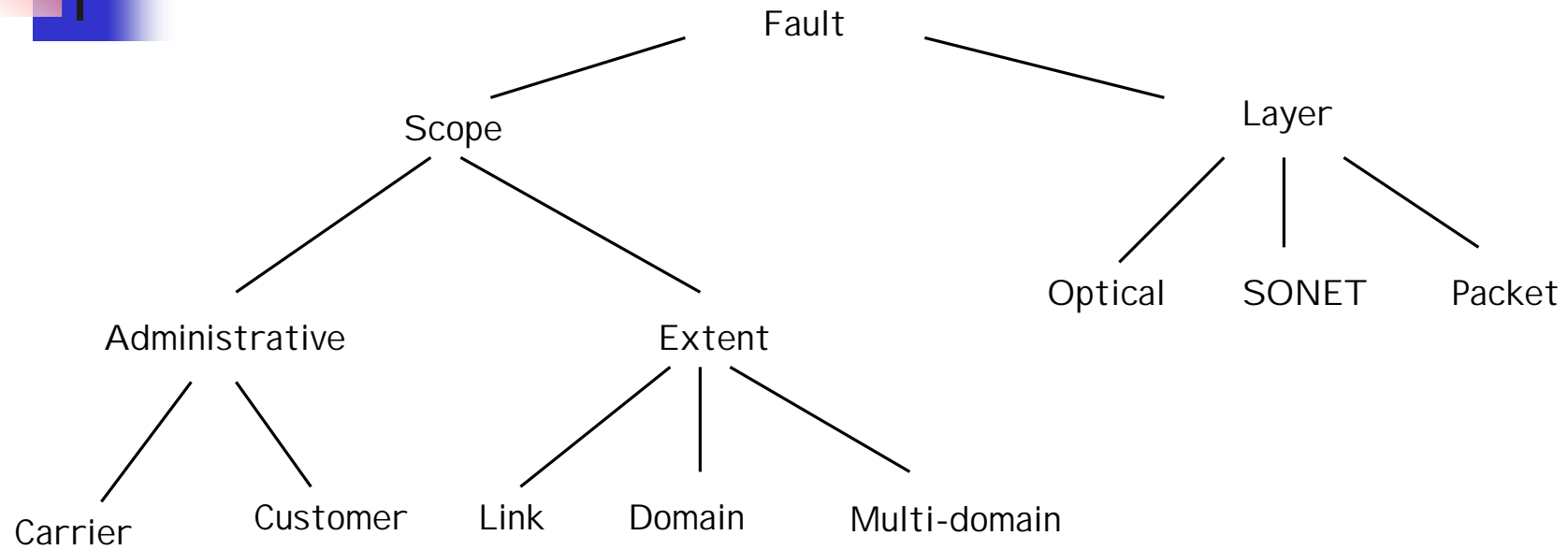
Basics of P&R - Technical

- Recover from failures as fast as possible
- Prioritize recovered connections
- Optimize resources
- Where is P&R done?
    - Layer 1 ➜ Transmission level (Optical)
    - Layer 2 ➜ Data link level (SONET, RPR)
    - Layer 2 ½ ➜ GMPLS level
    - Layer 3 ➜ Network protocol level (IP)
    - Higher layers ➜ E.g., Transport protocol level (TCP)

- What to do when we have multiple of these P&Rs enabled?
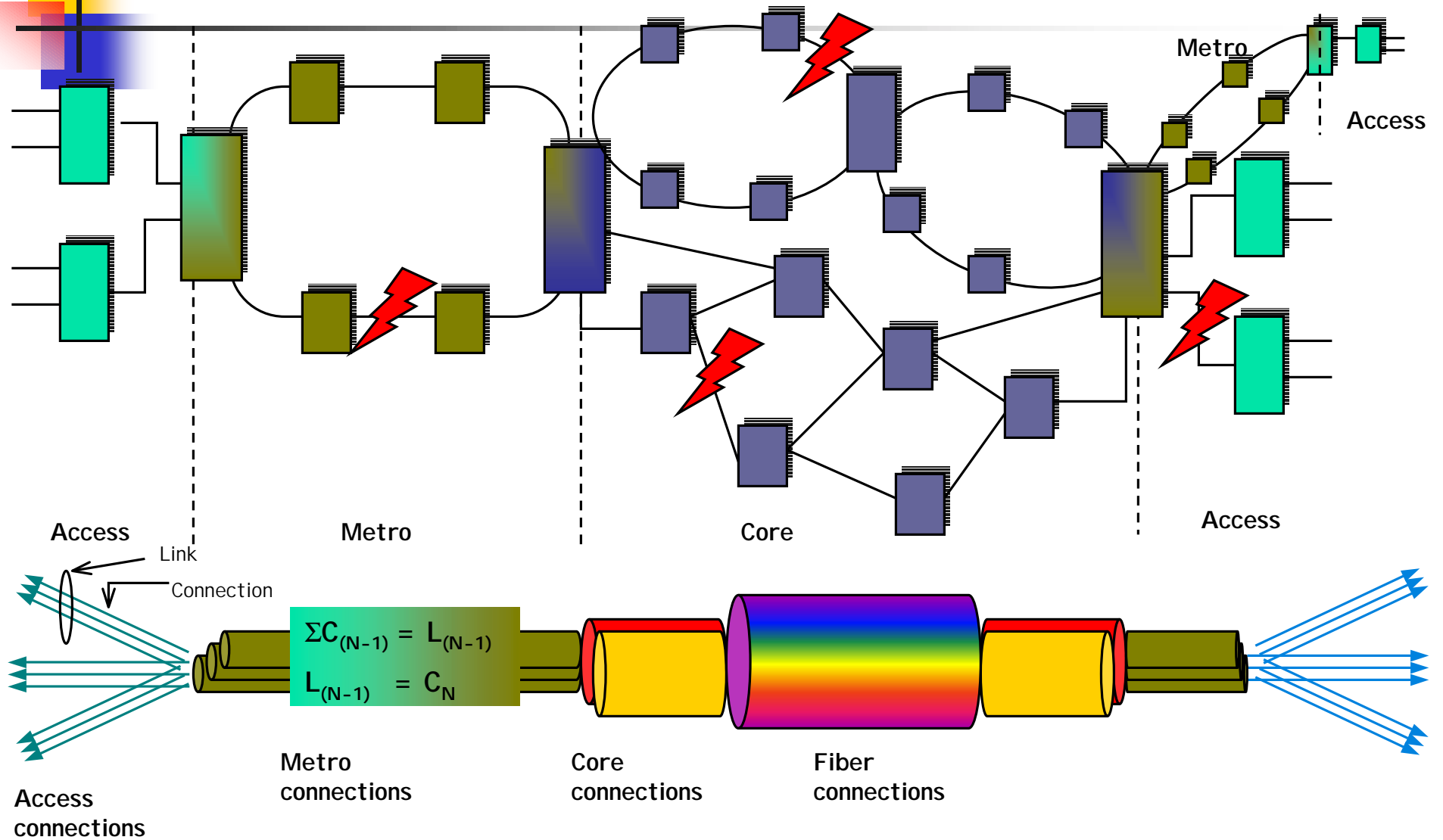
Basics of P&R - Terminology

- Protection vs restoration: Time it takes to recover from a failure.
- Survivability vs resiliency: Network design versus protocol design
- Connection vs Link : Scope of restoration
- Intra-domain vs inter-domain: Carrier, vendor, area, layer scopes
- Distributed vs centralized: Tools versus dynamic protocols
- Dedicated vs shared: Level of overbooking
- Single layer vs Multi-layer: Optical, SONET, packet or combination of them
- Server vs client layers: One who provides service versus one who use service (generally layer N-1 versus layer N)
- Level of diversity vs affordable risk: Level of paranoia – Link, Node, SRLG, SRG (COs, power plants, domains etc.)

Many dimensions of P&R

- Recovery path setup
  - Pre established, Pre qualified, Establish on demand
- Recovery resource allocation
  - Pre reserved, reserved on demand
- Scope of recovery
  - Domain (rings), Local (span, node), end-to-end (path)
- Level of recovery
  - Dedicated, shared, Preemptable
- Layers involved
  - Equipment (Cards, OXCs, ADMs, Routers), Protocols (SONET, GMPLS, IP)
- Topologies used
  - Hub and Spoke (Access), Ring (Metro, regional), Mesh (Long haul)
- Granularity of recovery
  - Fiber, Link, Connection, Partial connection

- Goal: Quickly recover from all types of failures with the least expense.

Digression: Definition of a connection

Metro

Access

Access

Metro

Core

Access

Link

Connection

$$\Sigma C_{(N-1)} = L_{(N-1)}$$

$$L_{(N-1)} = C_N$$

Access connections

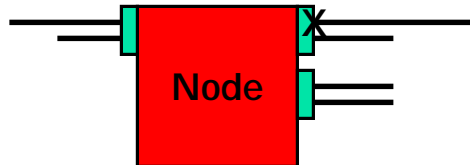Metro connections

Core connections
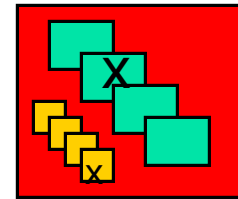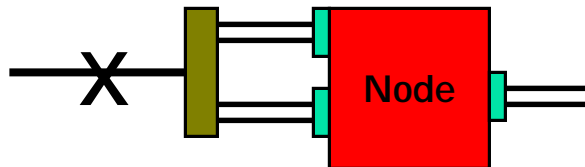
Fiber connections

Responsibility line
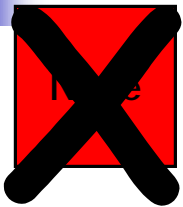
Service provider problem
Node diversity

Yes sir! I t is equipment provider
Problem... Equipment redundancy

Node

Service provider problem
Node/Link diversity

Service provider problem
External Node (SRG) diversity
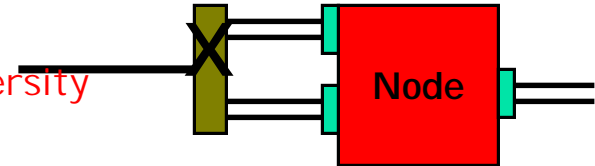
Node

Node

Service provider problem
Fiber (SRLG) diversity

Types of failures - Physical

- Link failure
    - Multiple connections are failed
    - Switch all of them together (Span or ring) or individually (Mesh)
- Equipment failure
    - Node ➔ Multiple line failures
    - Single or multiple link related failure
    - Single connection related equipment failure
- Domain failure
    - Definition: A set of resources (such as a ring of links, 1+1 Link, Redundant Cards in an Equipment) which have shared responsibility to recover from a risk (failure)
    - Domain can recover some failures relatively quickly, e.g., single failures
    - Domains may not be able to recover from some failures, e.g., Multiple failures
- Software failure (Not considered in this presentation)

A service provider has to have a strategy to recover from all the failures.
- Goal: Choose a mechanism(s) to recover quickly from the failures

Types of failures - Logical

- Region
  - Natural disasters – e.g., Earthquakes
  - Forced disasters – e.g., Nuclear disasters
- Zone
  - Intra-city failures – e.g., Power problems, fire in the Baltimore Tunnel ☺
  - Domain failures – e.g., Loosing an IGP area

- A service provider has to have a strategy to recover from all the failures.
- Goal: Choose a mechanism(s) to recover quickly from the failures

Bottom line on failures

- Bottom line – Restore a intra-domain failure by intra-domain mechanisms

- Observation: Failure recovery should be restricted to a "domain"
- A domain of a connection is where connection as a whole has a significance.
  - That is in a domain a connection is not (de)aggregated

- Example: On a wavelength failure:
  - Restoration should be at the wavelength granularity.
  - Do not restore in core – Customer have to "multi-home"
  - Restore at the packet (LSP) granularity –
    Too much repetitive work = Potentially more Expensive
    - ➔ priority grouping of previous domain links/connections
    - May not be the case always (Esp. from operations point-of-view)
  - Restore at link granularity – Unnecessary work = Potentially more expExpensive mechanisms
    - ➔ mesh versus ring P&R

Categories of failure recovery mechanisms

Protection   Restoration

Pre planning of
topology

Out of band recovery

Pre planning of connections

Fast reroute

Topology optimizations

- Time   **+ Time**

Pre failure   **Failure**   **Post failure**
                         occurred
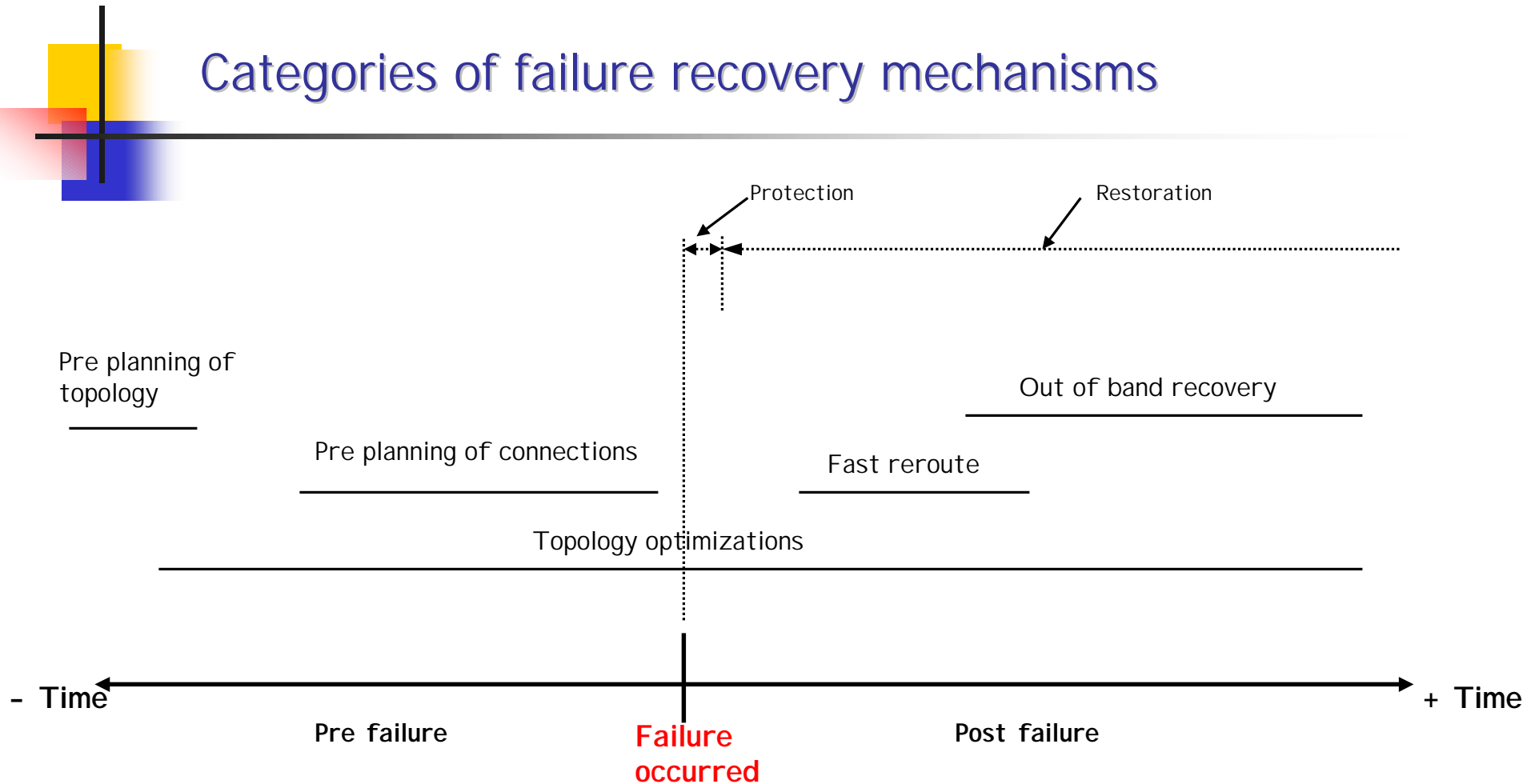
Mechanisms and rough time line of their occurrence.

Categories of failure recovery mechanisms (Contd.)

Type	Definition	Mechanisms (P – protection R – Restoration)	Faults covered
Pre planned Topology (Domain, Span) (Pre established, Pre qualified) (Dedicated, shared)	Topology is designed to be survivable	Domain (P) - Ring (BLSR)	Equipment, Link and Connection
		Span (P) – M:N span	Link and Connection
		Equipment (P) – M:N redundancy	Equipment, Link and Connection
Pre planned Connections (Domain, span, Connection) (Pre established, Pre qualified) (Dedicated, shared)`	Individual connections are designed to be survivable	Concatenation (P) per domain - Ring (UPSR) per link – 1+1 per connection	Equipment, Link, Node and Connection Equipment, Link and Connection
		End-to-end (P) Mesh – 1+1 path	Equipment, Link, Node and Connection
		End-to-end (R) Mesh (1:1 path) Mesh (1:N path)	Equipment, Link, Node and Connection Equipment, Link, Node and Connection

Categories of failure recovery mechanisms (Contd.)

Type	Definition	Mechanisms (P – protection R – Restoration)	Faults covered
Fast restoration (Pre qualify, Establish on demand) (Domain, Span, Path) (Shared, preemptable)	Mechanisms are devised To restore after failure	Domain (R) – Concatenated Mesh	Equipment, Link, Node and Connection
		Span (R) – M:N span pre qualified	Link and Connection
		Equipment (R) – M:N redundancy	Equipment, Link, Node and Connection
		Path (R) – Establish on demand	Equipment, Link, Node and Connection

Overview

- Background on P&R
- Timing analysis
    - Goal: Analyze the recovery times and missing standard pieces
- Multi-layer P&R
- Conclusions

Fault related actions

Post-Mortem: Analysis and Optimization

Recovery time

Reporting time

Grouping time

Detection time

T_0 T_1 T_2 T_3 T_4 T_5

Time

| Failure occurred | Failure detected | Failure correlation (Phase 1) | Failure communicated | Failure rectified | Failure correlation (Phase 2) |
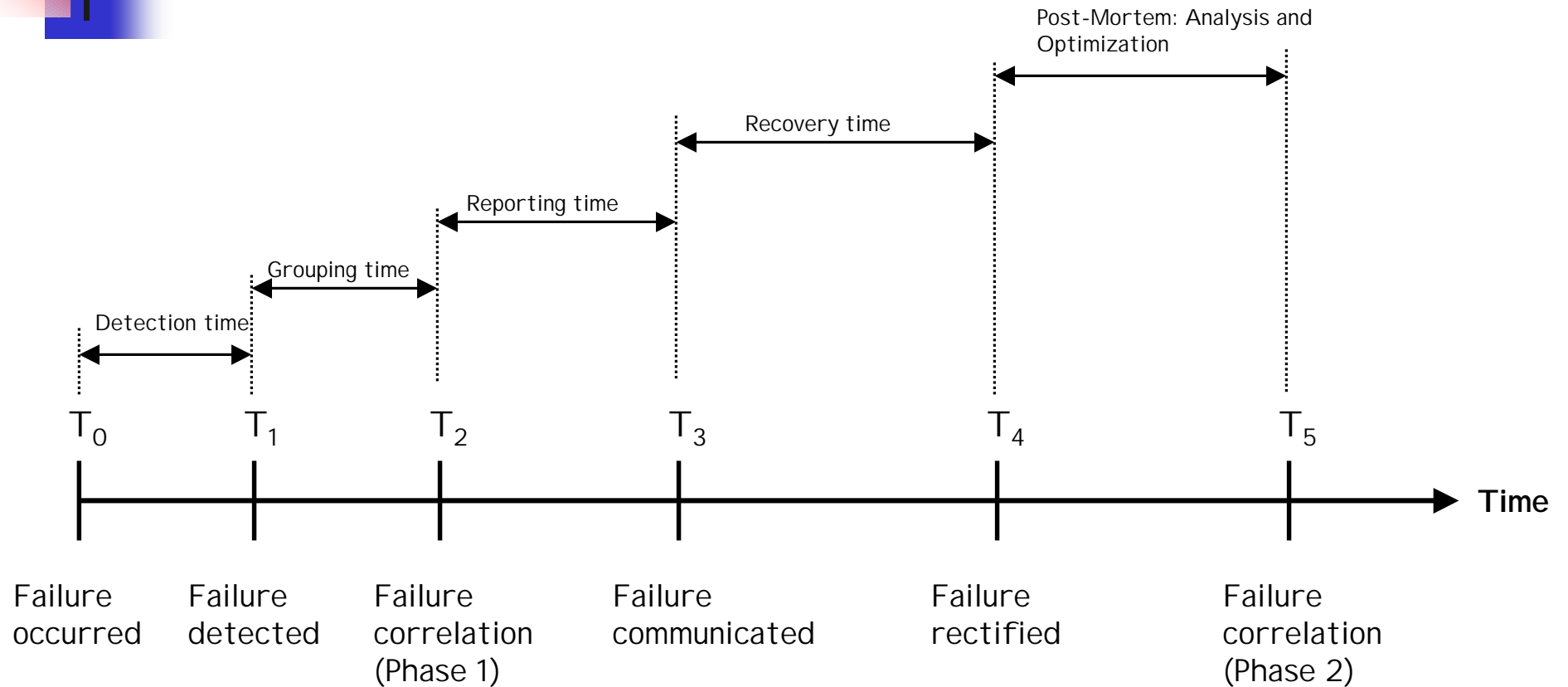
Reversion and optimization cycles are not presented.

Detection

- **Finding out the failures the neighbor(s).**
- Mechanisms
  - Automatic: A failure is implied because of an observed event.
  - In-band: A failure is informed in-band to the data channel.
  - Out-of-band: A failure is informed by OOB mechanism.
- Issues
  - Optical
    - LOL control
    - In band (Sonet header bytes)
    - OOB reporting mechanism – OLI

Grouping

- **Grouping the failures before reporting.**
- Mechanisms
    - Monitor groups
    - Wait before reporting
- Issues
    - How to monitor a link?
    - How long to wait?
    - How to group?
        - LSPs, VTs, Wavelengths, Waveband, Fiber
    - Who does grouping?
        - Detecting entity, e.g., DWDM
        - Recovering entity, e.g., crossconnect
        - Decision entity, e.g., management console

Reporting

- Reporting failures to the decision entities.
- Mechanisms
  - Automatic: Detecting Entity is also the decision entities
  - In-line: Using APS-like mechanisms
  - Out of band: Using OSC, fast reporting (OLI– like) and slow reporting (signaling, routing protocols)
- Issues
  - How far are the decision entities?
    - Neighbors (detecting entity is also decision entity) - Span, BLSR
    - Domain end points - UPSR, 1+1 and M:N
    - Connection end points - 1:1, M:N mesh
  - How granular is the reporting?
    - Per connection, link, group of links, node and domain

Recovery

- **Mechanisms**
  - Shared – Span, 2/4 wire BLSR, Mesh path (M:N)
  - Dedicated – Mesh (1+1, 1:1), UPSR
- **Issues**
  - Optical
    - Fiber/Wave bundle switching decides shared protection
  - Packet is not useful for ring
    - Possibly RPR in metro though!

Overview

- Background on P&R
- Timing analysis
- Multi-layer P&R
  - Goal: Derive requirements for multi-layer P&R
- Conclusions

Optical Layer Restoration - Motivation

4 OC12's

Service Layers

A

C

DS3, OC3, OC12
(45-622 Mb/s)

B-DCS Layer

A DCS

B DCS

C DCS

DS3, OC3, OC12
(45-622 Mb/s)

OC48

B

OC48

Digital Transmission Layer

A

B

C

D

E

F

Service Capacity
Restoration Capacity

OC48
(2.5 Gb/s)

B

Optical Layer

A

C

D

E

F
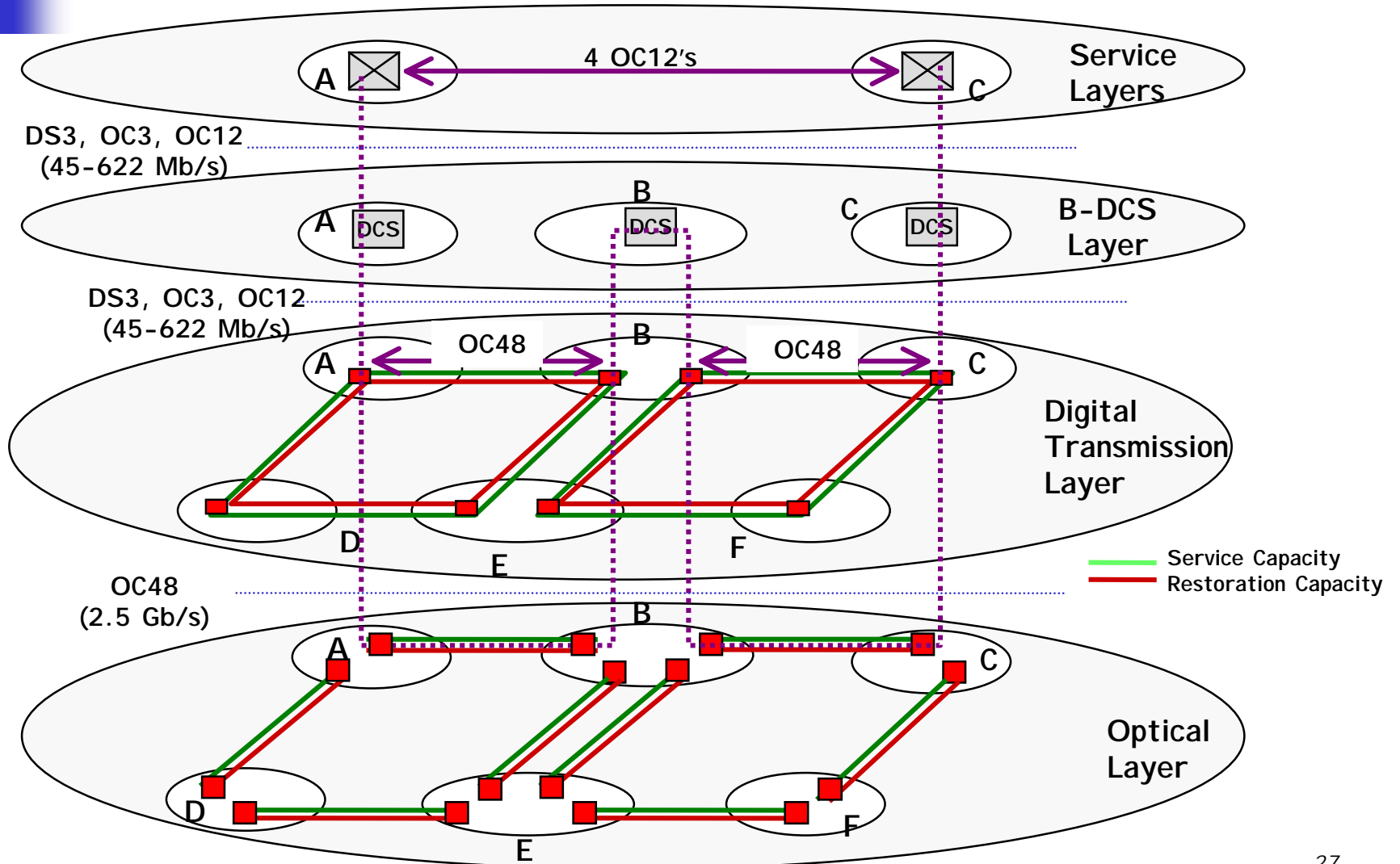
Optical Layer Restoration - Optical Layer Rings

A   4 OC12's   C   Service Layers

DS3, OC3, OC12
(45-622 Mb/s)

A   B   C   B-DCS Layer

DS3, OC3, OC12
(45-622 Mb/s)

B
OC48
A   C   Digital Transmission Layer

D   E   F

OC48
(2.5 Gb/s)

A   C   Optical Layer

D   E   F

Courtesy John Strand, AT&T

28

Optical Layer Restoration - Mesh

Service Layers

A — 4 OC12's — C

B-DCS Layer

A   B   C

Digital Transmission Layer

A   B   C
OC48

D   E   F

OC48 (2.5 Gb/s)

Service Capacity
Restoration Capacity
⊠ OLXC

Optical Layer

A   B   C

D   E   F
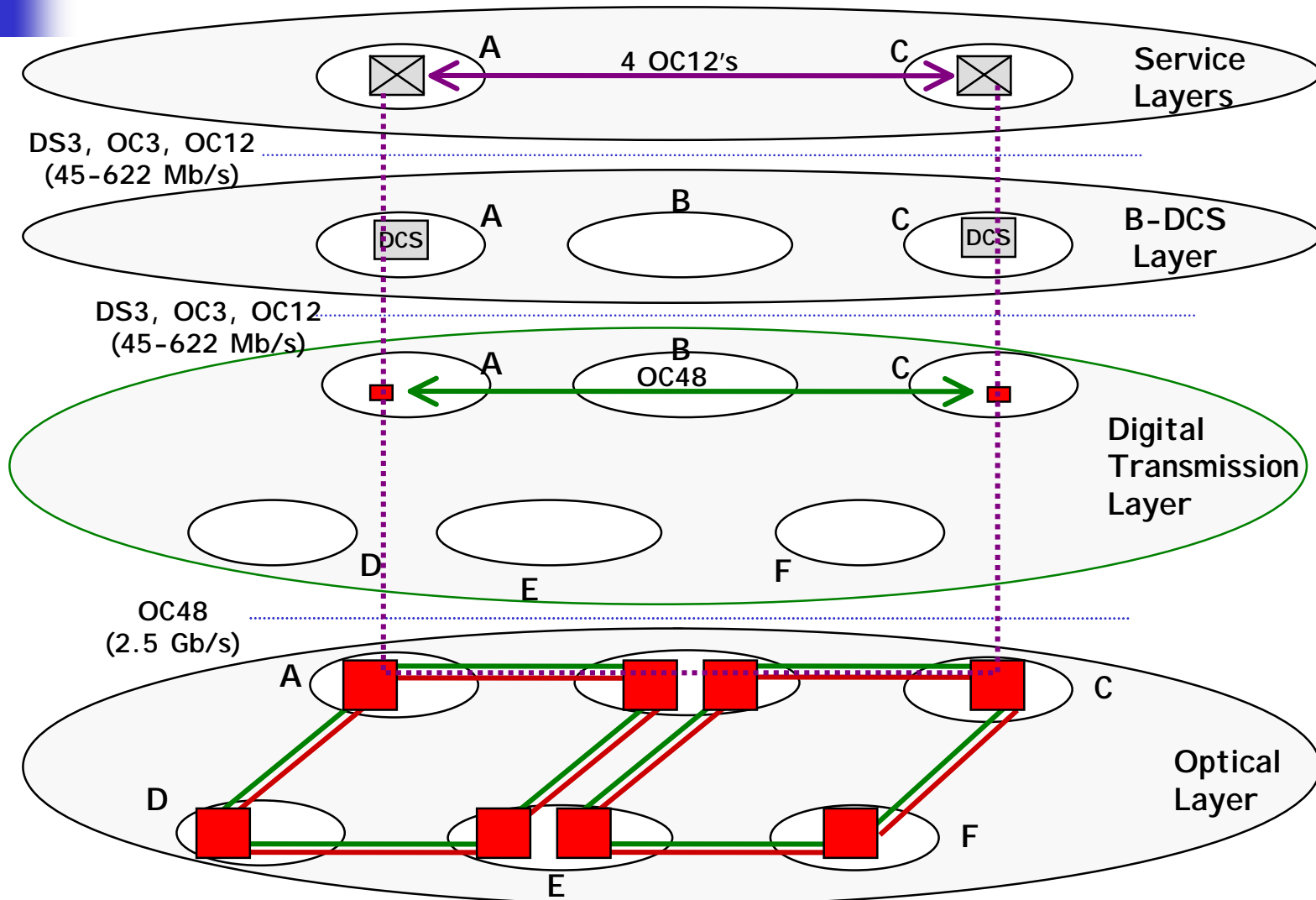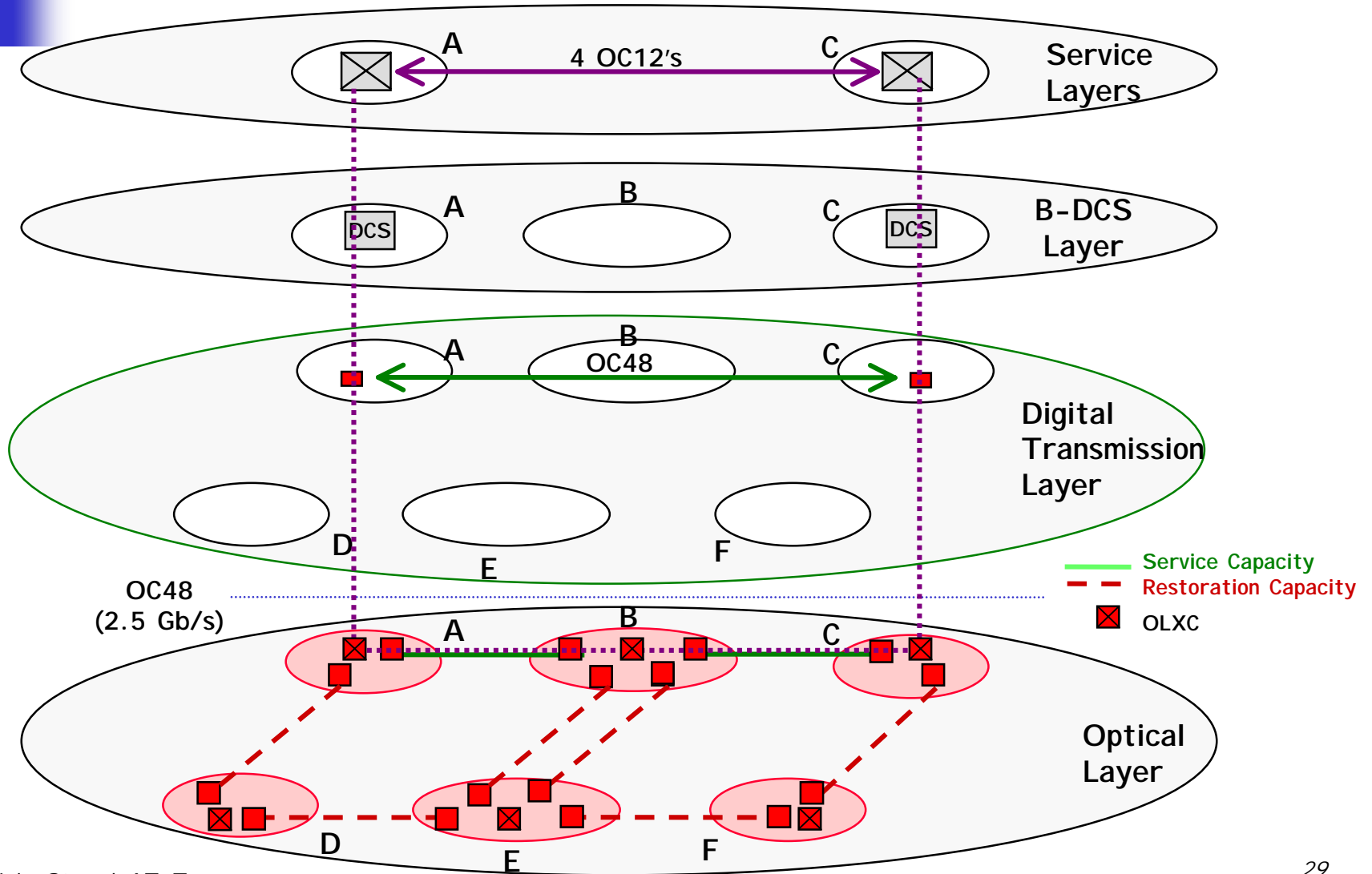
Why Multi-layer P&R?

+ Vendor mechanisms may not be interoperable ☺

+ A layer may not be able to
  ▪ detect a type of failure
  ▪ report a failure
  ▪ recover from a type of failure

+ A complete solution in one layer may not be fast enough

Difference between optical, SONET, packet layers

- Topologies
  - Mostly Mesh in packet (RPR is coming soon)
  - Mostly Ring in SONET
  - Mesh and Rings in Optical (being planned)
- Granularity of recovery
  - Partial connection only in packet
  - Connection, link in all
- Scope of recovery
  - Domain only in SONET, optical (in progress)
  - End-to-end, span in all

- Recovery path setup
  - Pre established in packet, (may be) in SONET and optical
  - Pre qualified, Establish on demand in all
- Recovery resource allocation
  - Pre reserved in packet (prevalently), SONET and Optical (optional)
  - Reserved on demand in all
- Level of recovery
  - Dedicated, in all
  - Preemptable Optional in all

Possible decision making process

- Identify your service
- Identify the layers involved in the service
- Understand mechanisms present (vs required) in recovery
- Decide the [mechanism, layer(s)] <-> service
- Decide how much is restored by which mechanism
  - Which mechanism is fast?
  - Which mechanism is less expensive?
  - Which faults can be covered by the mechanism?
    - What are the strategies for covering the other faults?
    - How to invoke other mechanisms?

Digression: Escalation mechanisms

- Escalation is the interworking of different survivable networks or different survivability mechanisms in a subnetwork.
- Goal: Maximize network protection at an acceptable cost.

- Survivable networks
  - Flat – Each subnetwork has its own survivable mechanism
  - Hierarchical – Strategies are arranged hierarchically to form larger groups
- Survivability mechanisms
  - Mainly depends on the economic benefits

- Strategies
  - Parallel – Different survivability mechanisms work in parallel
    - Inter-dependent – Work on the same traffic in parallel and halt when any one finishes
    - Independent – Work on different portions of the traffic
  - Serial – Pass the responsibility from one mechanism to the other
    - Issues:
      - Where to start the restoration
      - When to hand it over (Hold off timer)

Discussion

- **No escalation**
  - The choice should be the fastest of <Optical, SONET, Packet>
- **Only Optical**
  - High priority RT traffic
  - Faster than SONET, Packet
- **Only SONET**
  - Customer restoration in case of failure ➔ May not be acceptable
- **Only Packet**
  - Signaling support

- **Collaborative (Preferably serial)**
  - Hold-off timer implementation
  - Many changes to the hardware

Discussion – Multi-layer detection

Issues:

- Who detects the fault or group of faults? (Client layer dependent or client layer independent)
- How to suppress multi-layer detection?
- Which detection mechanism takes precedence?

Requirements:

- Static component
  - Possible configuration: [Service (could map to resources), Fault] ➔
    - 1st Choice ([Detection mechanism, Layer]),
    - 2nd Choice ([Detection mechanism, Layer]), …
- Dynamic component
  - Learning to correlate (Client, Server) layer topology mappings
  - Propagating the suppression logic

Discussion – Multi-layer grouping

Issues:

- Which layer is capable of grouping?
- Which mechanisms are required for grouping?
- Which equipment is capable of a given type of grouping?
- Which equipment is responsible to provide group reporting and hence what are the required protocol extensions?

Requirements:

- Static component
  - Possible configuration: [Service, Group fault] ➔
    - 1st choice ([Grouping mechanism, Layer]),
    - 2nd Choice ([Grouping mechanism, Layer]), …
- Dynamic component
  - Group reporting protocol constructs for different layers

Discussion – Multi-layer reporting

Issues:

- Which is the fastest reporting mechanism for a given fault?
- How to localize the reporting?
- Where to interpret the reporting?
- How to avoid or coordinate multi-layer reporting?
- How to enhance the reporting mechanisms?

Requirements:

- Static component
  - Possible configuration: [Service, fault] ➔
    - 1st choice ([Reporting mechanism, Layer]),
    - 2nd choice ([Reporting mechanism, Layer]) …
- Dynamic component
  - Group reporting delay coordination as part of orderly multi-layer reporting

Discussion – Multi-layer recovery

Issues:

- Which layer to do the recovery for a given fault?
- When to give up recovery in a given layer?
- What escalation mechanisms to use for recovery?

Requirements:

- Static component
  - Possible configuration: [Service, Fault] ➔
    - 1st choice ([Recovery mechanism, Layer]),
    - 2nd choice ([Recovery mechanism, Layer]), …
  - Transition timers (or hold off timers)
- Dynamic component
  - Coordination extensions
  - Increase in the recovery domain

Overview

- Background on P&R
- Timing analysis
- Case studies
- Multi-layer P&R
- Conclusions

Conclusions - General

- **On Risk**
  - Know the failures to be protected from
  - Determine the mechanism of recovery
  - Determine the layer which is responsible for recovery

- **On diversity**
  - Use diversity provided by server layer topology too
  - Group similar service requests into a diversity request

- **On recovery**
  - Recover a "connection" in a "domain"
  - Recover per service
  - Monitor a group of connections

Conclusions - General

- ## On customers considerations
  - Metro: GE/DWDM or SONET/DWDM or IP/DWDM or ATM/SONET/DWDM and others
  - Core: GE/Optical or SONET/Optical or IP/Optical

- ## On standards limitations
  - Optical
    - O-APS (linear and ring), No standard Control Channel, OLI
  - SONET
    - No wait and report Control Channel
  - Packet
    - No rings support, No out-of-band APS-like signaling

Conclusions – Multi-layer

- On detection
  - Have an alternative mechanism of detection
  - Learning to correlate (Client, Server) layer topology mappings
  - Propagating the suppression logic
- On grouping
  - Group reporting protocol constructs for different layers
- On reporting
  - Group reporting delay coordination as part of orderly multi-layer reporting
- On recovery
  - Coordination extensions
  - Increase in the recovery domain