ECSE-6660
Availability, Survivability, Protection/Restoration, Fast Re-Route

http://www.pde.rpi.edu/
Or
http://www.ecse.rpi.edu/Homepages/shivkuma/

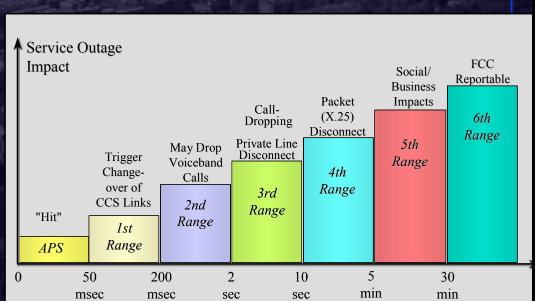
Shivkumar Kalyanaraman
Rensselaer Polytechnic Institute
shivkuma@ecse.rpi.edu

Based in part on slides of James Manchester (formerly Tellium, now RPI), and some NANOG presentations

---

- Availability: the driver…
- Survivability: protection and restoration architectures
- Fast-Reroute

Availability: Impact of Outages



Disruptions cost a lot of money!

Market Drivers for Survivability

- Customer Relations
- Competitive Advantage
- Revenue
  - Negative - Tariff Rebates
  - Positive - Premium Services
    - Business Customers
    - Medical Institutions
    - Government Agencies
- Impact on Operations
- Minimize Liability

Network Survivability: drivers

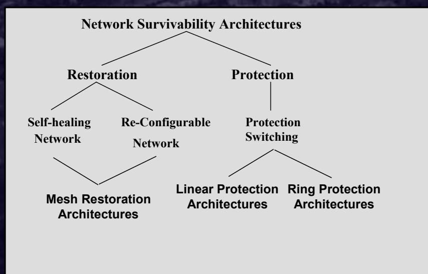
- Availability: 99.999% (5 nines) => less than 5 min downtime per year
- Since a network is made up of several components, the ONLY way to reach 5-nines is to add survivability *in the face of failures*…
  - Survivability = continued services in the presence of failures
  - Protection switching or restoration: mechanisms used to ensure survivability
    - Add redundant capacity, detect faults and automatically re-route traffic around the failure
- Restoration: related term, but slower time-scale
- Protection: fast time-scale: 10s-100s of ms…
  - implemented in a *distributed manner* to ensure fast restoration

---

Failure Types & Other Motivations

- Types of failure:
  - Components: links, nodes, channels in WDM, active components, software…
  - Human error: backhoe fiber cut
    - Fiber inside oil/gas pipelines less likely to be cut
  - Systems: Entire COs can fail due to catastrophic events
- Protection allows easy maintenance and upgrades :
  - Eg: switchover traffic when servicing a link…
- Single failure vs multiple concurrent failures…
  - Goal: mean repair time << mean time between failures…
- Protection also depends upon kind of application:
  - SONET/SDH: 60 ms (legacy drop calls threshold)
  - Do data apps really need this level of protection?
- Survivability may hence be provided at several layers

Network Survivability Architectures



Shivkumar Kalyanaraman

Network Availability & Survivability

Availability is the probability that an item will be able to perform its designed functions at the stated performance level, within the stated conditions and in the stated environment when called upon to do so.

$$\text{Availability} = \frac{\text{Reliability}}{\text{Reliability} + \text{Recovery}}$$

Shivkumar Kalyanaraman

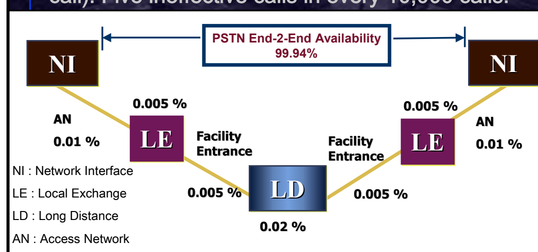
Quantification of Availability

Percent Availability	N-Nines	Downtime Time Minutes/Year
99%	2-Nines	5,000 Min/Yr
99.9%	3-Nines	500 Min/Yr
99.99%	4-Nines	50 Min/Yr
99.999%	5-Nines	5 Min/Yr
99.9999%	6-Nines	.5 Min/Yr

Shivkumar Kalyanaraman

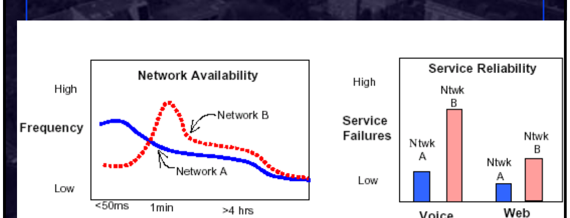
PSTN : The Yardstick ?

- Individual elements have an availability of 99.99%
- One cut off call in 8000 calls (3 min for average call). Five ineffective calls in every 10,000 calls.



Source : http://www.packetcable.com/downloads/specs/pkt-tr-voipar-v01-001128.pdf

Services Determine the Requirements on Network Availability



Source : www.t1.org

Shivkumar Kalyanaraman

IP Network Expectations

Service	Delay	Jitter	Loss	Availability
Real Time Interactive (VOIP, Cell Relay ..)	L	L	L	H
Layer 2 & Layer 3 VPN's (FR/Ethernet/AAL5)	M	L	L	H
Internet Service	H	H	M	L
Video Services	L	M	M	H

L : Low M : Medium H : High

Shivkumar Kalyanaraman

Measuring Availability: The Port Method

- Based on Port count in Network

$$\frac{\text{(Total \# of Ports X Sample Period) - (number of impacted port x outage duration)}}{\text{(Total number of Ports x sample period)}} \times 100$$

- Does not take into account the Bandwidth of ports e.g. OC-192 and 64k are both ports
- Good for dedicated Access service because ports are tied to customers.

The Port Method Example

- 10,000 active access ports Network
- An Access Router with 100 access ports fails for 30 minutes.
  - Total Available Port-Hours = 10,000*24 = 240,000
  - Total Down Port-Hours = 100*.5 = 50
  - Availability for a Single Day =

(240000-50/240,000)*100 = 99.979166 %

The Bandwidth Method

- Based on Amount of Bandwidth available in Network

$$\frac{\text{(Total amount of BW X Sample Period) - (Amount of BE impacted x outage duration)}}{\text{(Total amount of BW in network x sample period)}} \times 100$$

- Takes into account the Bandwidth of ports
- Good for Core Routers

The Bandwidth Method Example

- Total capacity of network 100 Gigabits/sec
- An Access Router with 1 Gigabits/sec BW fails for 30 minutes.
  - Total BW available in network for a day = 100*24 = 2400 Gigabits/sec
  - Total BW lost in outage = 1*.5 = 0.5
  - Availability for a Single Day =

((2400-0.5)/2,400)*100 = *99.979166 %*

Defects Per Million

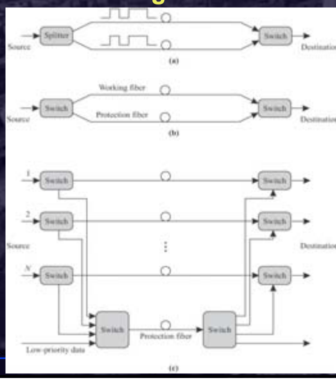
- Used in PSTN networks, defined as number of blocked calls per one million calls averaged over one year.

$$\text{DPM} = \left[\frac{\text{(number of impacted customers x outage duration)}}{\text{(total number of customers x sample period)}}\right] \times 10^{-6}$$

Defects Per Million Example

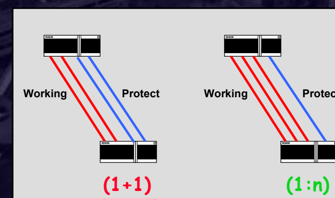
- 10,000 active access ports Network
- An Access Router with 100 access ports fails for 30 minutes.
  - Total Available Port-Hours = 10,000*24 = 240,000
  - Total Down Port-Hours = 100*.5 = 50
  - Daily DPM = (50/240,000)*1,000,000 = 208

Basic Ideas: Working and Protect Fibers



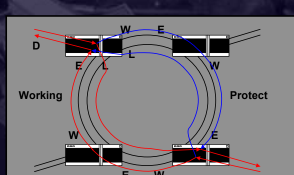
Protection Topologies - *Linear*

- Two nodes connected to each other with two or more sets of links



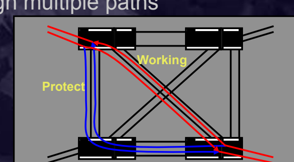
Protection Topologies - *Ring*

- Two or more nodes connected to each other with a ring of links
  - Line vs. Drop interfaces
  - East vs. West interfaces

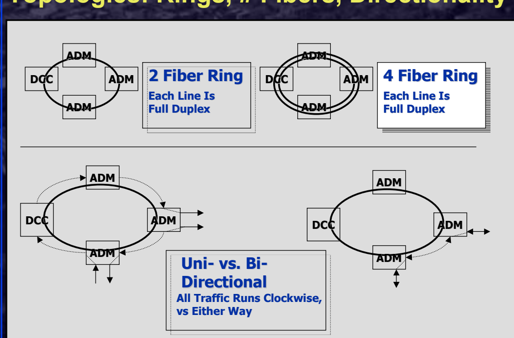


Protection Topologies - *Mesh*

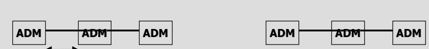
- Three or more nodes connected to each other
  - Can be sparse or complete meshes
  - Spans may be individually protected with linear protection
  - Overall edge-to-edge connectivity is protected through multiple paths



Topologies: Rings, # Fibers, Directionality



SONET: Automatic Protection Switching (APS)



Line Protection Switching
Uses TOH
Trunk Application
Backup Capacity Is Idle
Supports 1:n, where n=1-14

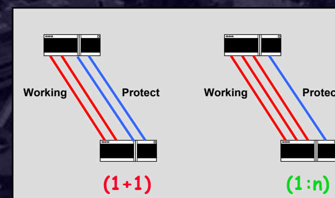
Path Protection Switching
Uses POH
Access Line Applications
Duplicate Traffic Sent On Protect
1+1

Automatic Protection Switching
- Line Or Path Based
- Revertive vs. Non-Revertive
- Restoration Times ~ 50 ms
- K1, K2 Bytes Signal Change

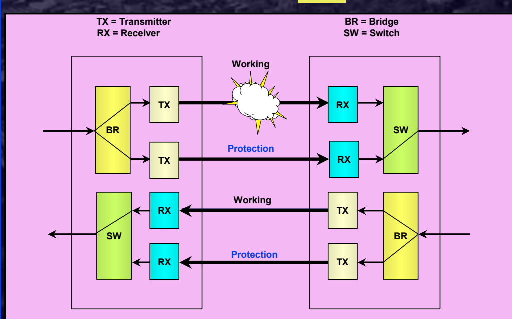
Protection Switching Terminology

- 1+1 architectures - permanent bridge at the source - select at sink
- m:n architectures - m entities provide protection for n working entities where m is less than or equal to n
  - allows unprotected extra traffic
  - most common - SONET linear 1:1 and 1:n
- Coordination Protocol - provides coordination between controllers in source and sink
  - Required for all m:n architectures
  - Not required for 1+1 architectures unless they employ bi-directional protection switching

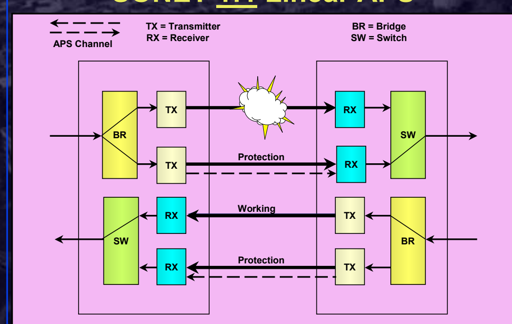
1+1 vs 1:n



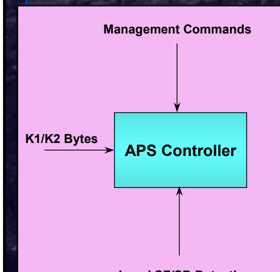
SONET Linear 1+1 APS



SONET 1:1 Linear APS



SONET Linear APS



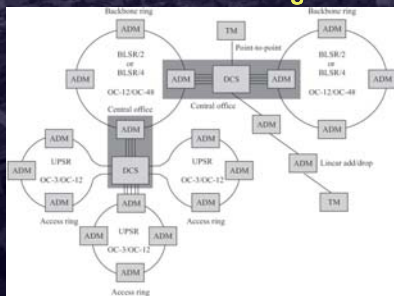
Linear APS States

K1 Byte Bits 1234	Automatically Initiated, External, or State Request
1111	Lockout of Protection
1110	Forced Switch
1101	SF - High Priority
1100	SF - Low Priority
1011	SD - High Priority
1010	SD - Low Priority
1001	Not Used
1000	Manual Switch
0111	Not Used
0110	Wait to Restore
0101	Not Used
0100	Exercise
0011	Not Used
0010	Reverse Request
0001	Do Not Revert
0000	No Request

Protection Switching: Terminology

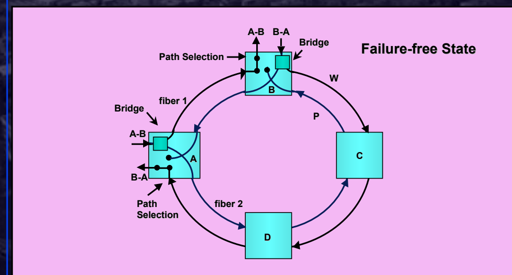
- Dedicated vs Shared: working connection assigned dedicated or shared protection bandwidth
  - 1+1 is dedicated, 1:n is shared
- Revertive vs Non-revertive: after failure is fixed, traffic is automatically or manually switched back
  - Shared protection schemes are usually revertive
- Uni-directional or bi-directional protection:
  - Uni: each direction of traffic is handled independent of the other.
  - Fiber cut => only one direction switched over to protection . Usually done with dedicated protection; no signaling required.
  - Bi-directional transmission on fiber (full duplex) => requires bi-directional switching & signaling required

Current Architectures: Ring Protection



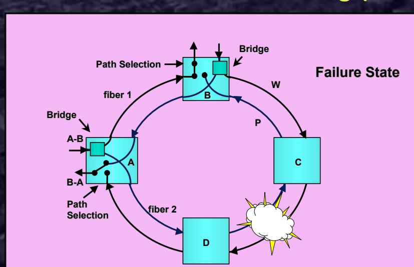
Today: multiple "stacked" rings over DWDM (different λs)

Unidirectional Path Switched Ring (UPSR)



* One fiber is "working" and the other is "protect" at all nodes…
* Traffic sent SIMULTANEOUSLY on working and protect paths…
* Protection done at path layer (like 1+1)…

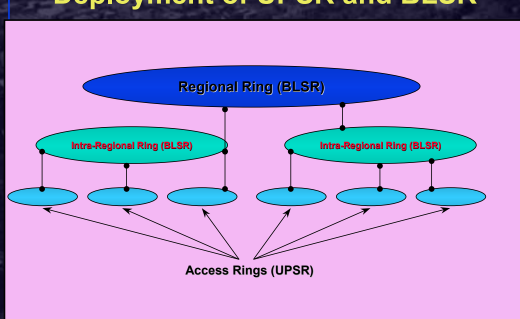
Unidirectional Path Switched Ring (UPSR)



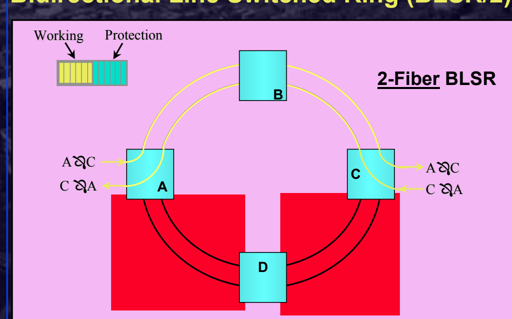
UPSR: discussion

- Easily handles failures of links, transmitters, receivers or nodes
- Simple to implement: no signaling protocol or communication needed between nodes
- Drawback: does not spatially re-use the fiber capacity because it is similar to 1+1 linear protection model
  - I.e. no sharing of protection (like m:n model)
  - BLSRs can support aggregate traffic capacities higher than transmission rate
- UPSRs popular in lower-speed local exchange and access networks (traffic is hubbed into the core)
  - No specified limit on number of nodes or ring length of UPSR, only limited by difference in delays of paths

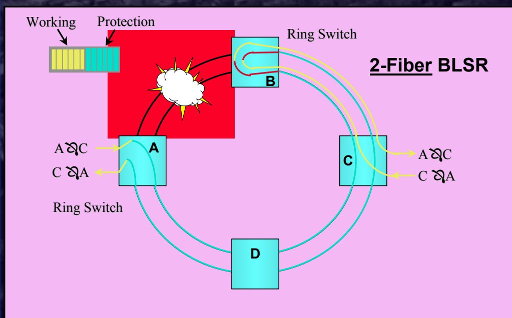
Deployment of UPSR and BLSR



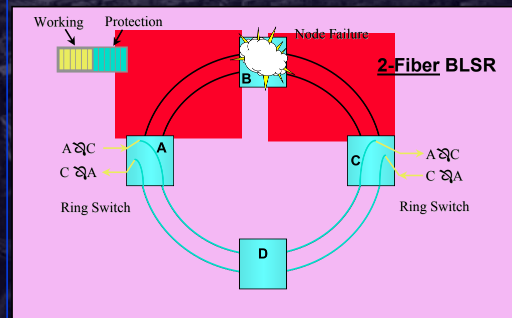
Bidirectional Line Switched Ring (BLSR/2)



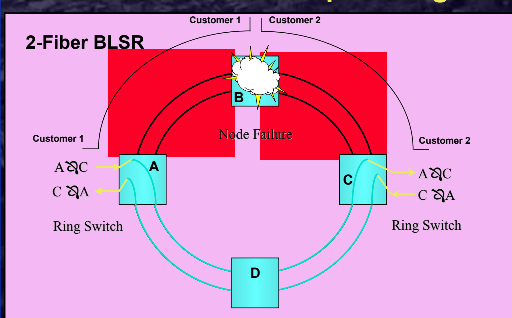
Bi-directional Line Switched Ring (BLSR/2)



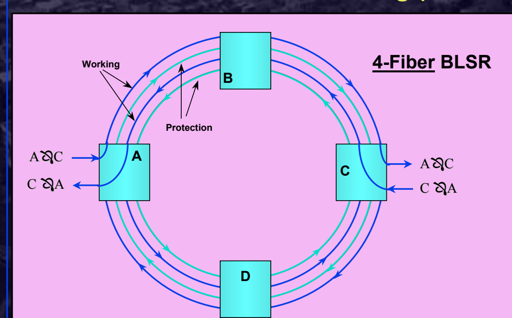
Bi-directional Line Switched Ring (BLSR/2)



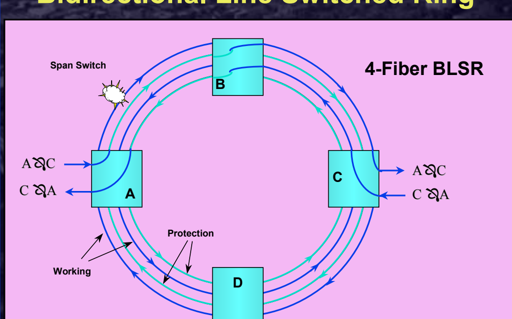
Node Failures => "Squelching"



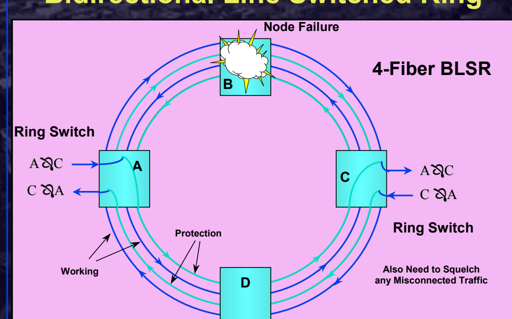
Bi-directional Line Switched Ring (BLSR/4)



Bidirectional Line Switched Ring



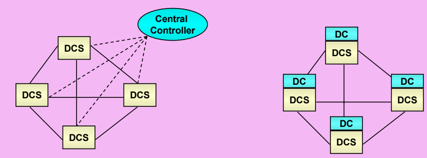
Bidirectional Line Switched Ring



BLSR: Discussion

- BLSR/2 can be thought of as BLSR/4 with protection fibers embedded in the same fiber
  - I.e. ½ the capacity is used for protection purposes in each fiber
- Span switching and ring switching is possible only in BLSR, not in UPSR
- 1:n and m:n capabilities possible in BLSR
- More efficient in protecting distributed traffic patterns due to the sharing idea
- Ring management more complex in BLSR/4
- K1/K2 bytes of SONET overhead is used to accomplish this

Mesh Restoration

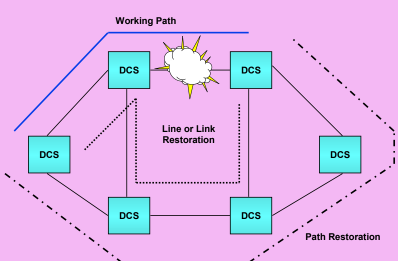


Reconfigurable (or Rerouting) Restoration Architecture

Self Healing Restoration Architecture

DC = Distributed Controller

Mesh Restoration



- Control: Centralized or Distributed
- Route Calculation: Preplanned or Dynamic
- Type of Alternate Routing: Line or Path

Mesh Restoration vs Ring/Linear Protection

Attributes	Linear APS	Ring PS	Mesh Restoration
Spare Capacity Needed	Most	Moderate	Least
Fiber Counts	Highest	Moderate	Moderate
Restoration Time	<50 ms	<50 ms	2-10 seconds
Software Complexity	Least	Moderate	Most
Protection Against Major Failures	Worst	Medium	Best
Planning/Operations Complexity	Least	Moderate/least	Most

Extracted from: T-H. Wu, Emerging Technologies for Fiber Network Survivability, See References

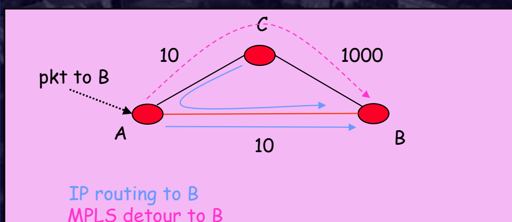
Fast Reroute

- Do the “restoration” at the MPLS (I.e. Layer 2) …
- Also possible to do fast-reroute at layer 3 (IP) with BANANAS framework.
- Issues:
  - Can MPLS re-route as fast as SONET (50ms)?
  - Can traditional IP re-route as fast as MPLS?

Fast Reroute (2)

- First question: how fast is fast?
  - Do you really need 50 ms failover?
- Second question: can you reroute really quickly while maintaining network stability?
- Third question: what are the scalability issues with fast reroute?

Fast Reroute: MPLS vs. IP



Fast Reroute vs IP Routing

IP
- All nodes must be told of failure
- Fast propagation, fast SPF trigger: how stable?
- One step to full re-convergence

MPLS (RSVP-TE)
- Only the two ends of the link need be told (no signaling)
- Local operation: explicit routing; more stable
- Two step process: detour + converge