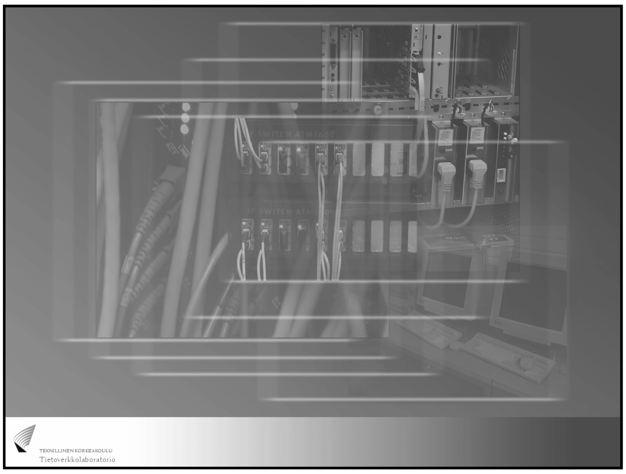




# Resiliency

**S-38.3192 Verkkopalvelujen tuotanto**
**S-38.3192 Network Service Provisioning**





## Resiliency/ Survivability

- Network Resiliency:
  - "Ability of the network to provide and maintain an acceptable level of service in the face of various faults and challenges to normal operation."
- Network Survivability:
  - "*Quantified* ability of a system, subsystem, equipment, process, or procedure to continue to function during and after a natural or man-made disturbance."



## Internet Design Goals

- Primary goal:
  - Develop an effective technique for multiplexed utilization of existing interconnected networks
- Second level goals, in the order of importance (from Dave Clark, 1988)
  1. Internet communication must continue <u>despite loss of gateways</u>
  2. …
- ➢ The concept of network resiliency has been present in the Internet from ground-up

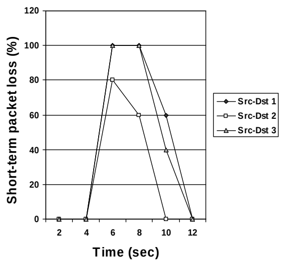
# Network Survivability

- Survivable network regains service level (SLA: availability) quickly in the presence of faults within the network
- Requires mechanisms for protection and/or restoration
  - Strength of applied mechanisms depends on the network operator's strategic goals (how important it is to regain service level quickly)
    - 2 nines -> restoration
    - 5 nines -> protection (1:1)
    - 7 nines -> protection (1+1)

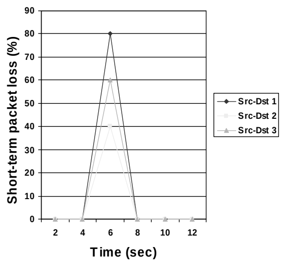


# Network Survivability illustrated

**Low survivability**




**Improved survivability**






# Protection vs Restoration

- Protection
  - Backup resource(s) are determined for primary resource(s) before a fault occurs
- Restoration
  - The (automatic) process of restoring connectivity after a fault
  - Usually slower if no protection is used



# Protection Modes

- **1+1 protection**
  - A separate secondary resource is dedicated for each primary resource
  - Traffic is sent on both resources and receiving end of resource selects one copy to be transmitted further
  - Can't carry extra traffic over secondary resource
- **1:1 protection**
  - A separate secondary resource is dedicated for each primary resource
  - If the primary resource fails, traffic switches to the secondary resource. When the failure on the primary resource is resolved, traffic automatically reverts to the primary resource
  - Extra traffic can be carried over the secondary resource but in case of fault in primary, extra traffic is pre-empted from the secondary

# Protection Modes

- **1:N protection**
  - A secondary resource is set for a group of primary resources
  - The underlying assumption is that only one of the primary resources will fail at any given time, and that the working resources are independent of each other
  - Requires less hardware than 1+1 and 1:1 schemes (however, the hardware may be more expensive – slide 9)
  - The disadvantage is that the switching between primary resources and the backup resources must occur at a higher level in the system (slower)
  - Extra traffic can be carried over the secondary resource but in case of fault in primar(y/ies), traffic is pre-empted from the secondary
    - Only a subset of extra traffic in primar(y/ies) delivered on secondary
      - Priorization of primaries



# Protection Modes

- **M:N protection** (M<<N)
  - M secondary resources are set for a group of primary resources (N)
    - Higher percentage of primary traffic is secured than with 1:N protection



# Protection Modes - example

- **Comparison of Various protection schemes in Optical WDM Mesh Networks**

| ***Protection Scheme*** | ***Cost/complexity of OXC, operation & Management systems*** | **Restoration speed** | **Wavelength resources required** | **Extra traffic that can be carried by secondary** |
|---|---|---|---|---|
| 1+1 | Lowest | Highest | Highest | None |
| 1:1 | Medium | Medium | Medium | Highest |
| 1:N, M:N | Highest | Lowest | Lowest | Medium |



# Restoration

- **Local restoration**
  - Network device that detects the error uses local capabilites to circumvent the failed part of the network
    - In case of link; possible secondary link to same destination
    - In case of node; 3rd node to circumvent failed node
  - Leads to sub-optimal network state
- **Path restoration**
  - Source of the path determines new path in case of failure in primary path
  - Pre-calculation of disjoint paths is possible
    - Faster switch over time

# Restoration

- **Global restoration**
  – Network node that detects fault in the network informs all other nodes in the network about existence of fault (by using the routing protocol)
    - Link state routing: by removing the LSA
      – Only if happens to be originator of LSA
      – Otherwise sits back and waits for timer to clean the LSDB (can be hours with basic configurations)
    - Distance vector routing: by calculating new distance vector

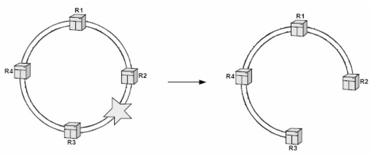


# SDH

- SDH networks are famous for their fast restoration in case of fault
  – Typically less than 50ms to complete restoration
  – Based on general idea of *non-arbitrary network topologies*
    - Double rings which can be restored by reversing the traffic at the ends of faulty section
      – Single action
      – Single failure restoration within the ring
      – 50% of network capacity reserved for restoration (1:1 protection)



# SDH



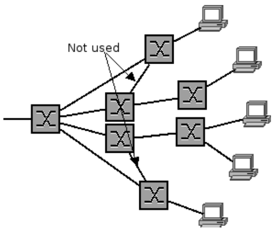



# Ethernet

- Conventional Ethernet restoration is based on *spanning tree protocol*
  – Any arbitrary topology is turned into tree topology
  – All switches in an extended LAN gather information on other switches in the network through an exchange of data messages
  – This exchange of messages results in the following:
    - The election of a unique root switch for the stable spanning-tree network topology
    - The election of a designated switch for every switched LAN segment
    - The removal of loops in the switched network by placing redundant switch ports in a backup state

## Ethernet - Spanning Tree protocol

- Loop forming interfaces are removed from the spanning tree by the switches participating in the protocol
- The tree is formed by utilizing port priorities (costs)






## Ethernet - Spanning Tree protocol

- Several version available
  - 802.1d (original spanning tree) with long convergence time (50s)
  - 802.1w (Rapid Spanning Tree) with only few seconds of convergence
  - 802.1s (Multiple Instance Spanning Tree) per VLAN operation
- All versions based on same protocol operation
  - Exchange of BPDU messages to determine whether or not interface should be blocked
- Switches usually support immediate forwarding on leaf ports
  - No additional delay in starting of communication



## Ethernet - Resilient Fast Ethernet Ring

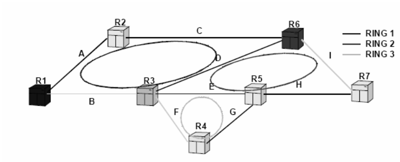
- SDH/SONET rings provide very fast restoration, but synchronous equipment is more expensive than Ethernet-based, and incurs a heavy bandwidth penalty (50%) to achieve the fast restoration
- Solution: SDH type network restoration on top of Ethernet
  - Basic idea same as in SDH:
    - Ring type network topology
    - Traffic reversion in case of error
  - Implementation 1: Ethernet Automated Protection Switching (EAPS), RFC 3619
  - Implementation 2: Metro Ring Protocol (MRP)



## Ethernet - Resilient Fast Ethernet Ring

- Each ring has a master which:
  - Blocks loop forming interface
  - In case of fault opens the loop forming interface for traffic
- Detection of fault can be based on:
  - Probes sent by the master
  - Signaling from the device that detects the fault
- Convergence time of network is dependent on time between fault and notification of master
  - Tens of milliseconds with device signalling
  - Hundreds of milliseconds with probes

## Ethernet - Resilient Fast Ethernet Ring






## MPLS

- LSP restoration processes are based on Constrained Shortest Path First routing algorithm for selecting bypass LSPs.
- Re-route options:
  - Link protection
  - Link and node protection
  - Path protection
  - Dynamic restoration



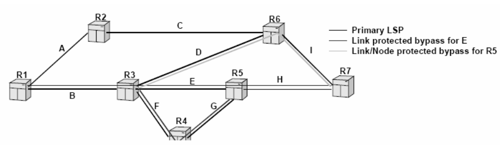
## Link Protection

- Each link on protected LSP has its own bypass for circumventing failed link
- Link protection alternatives:
  - per LSP
  - Several LSPs can be aggregated into single bypass LSP
- Requirements:
  - Separate bypass is calculated between each RSVP neighbor
  - Router tracks the interface status of egress link and reroutes the protected traffic by (two competing drafts)
    - Stacking the original label with label structure of bypass LSP
    - Changing the label for bypass label



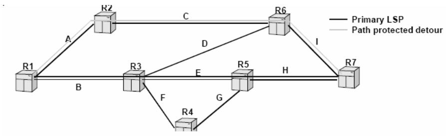
## Link/Node Protection

- Node protection is used to circumvent broken nodes (instead of links).
- Bypass LSP is established around set of next link, node and link using seprate router.
- Otherwise node protection operates like link protection



# Path protection

- Path protection is done per ingress/ egress pair and to each individual LSP
- Separate backup LSP is calculated through the network using disjoint resources (routers & links)





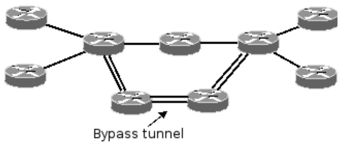

# Path protection

- In failure of primary LSP, ingress point of LSP swaps into backup
  - Question: How can ingress become aware of failure in primary
    - Upstream notification takes time to travel
    - Additional delay in restoration of network status



# One-to-one vs Facility backup

- One-to-one backup operates on the basis of a backup LSP for each protected LSP.
- Facility backup aims at using a single LSP to back up a set of protected LSPs (figure below).






# Switch Back

- *Switch back* is the process of rerouting the failed LSPs from their backups to the primary LSPs
  - For Path protected LSPs this may not be wise
    - Shifting the traffic causes always deteoriration
    - Even with make-before-brake packets usually experience sequence errors
  - Facility backups require some form of switch back
    - Into original paths ones they are up and running
    - Into new primaries if restoration of original primary is not expected to happen

## Dynamic Restoration

- If there are no other protections new LSP can also be calculated on demand
- Failure of primary triggers on-demand calculation of a new primary
- Failure is noticed from the fact that failed resources are no longer in Traffic Engineering Database (TED)
- Causes few hundred milliseconds of additional delay for restoration



## IP

- IP restoration is based on convergence of routing protocols
  - Fault is detected by using
    - Hello timers
    - (L2 indications)
    - (Bi-directional Forwarding Detection (BFD) indication)
  - Flooding of new LSAs
  - Calculation of global routing tables
  - Instantion of new forwarding table



## IP

- Detection of errors
  - Slow process if there is a L2 interconnection device between routers (link is not broken) - standard Hello based detection takes tens of seconds
  - L2 indication process works only if interconnection device fails
  - Can be sped up by using BFD
    - Probes are sent between forwarding planes of routers
    - Fault is signaled to routing process



## IP

- Convergence of IP routing depends heavily on detection time of fault
  - Hello process -> tens of seconds
  - BFD -> some hundreds of milliseconds
  - L2 indication -> few milliseconds
- Flooding process and SPF calculations take only some tens or hundreds of milliseconds
- Off-the-shelf running networks can have large deadlocks due to default timer values:
  - Hello timer of 10s -> router dead 40s
  - LS refresh time 1800s -> LSA max age 3600s

# Inter-layer Communication

- Modern telecommunications networks are layered with their structure
- Fault in lower layer affects all higher layers
- Convergence process should proceed from bottom to up
  - Unneccesary oscilation can be avoided if each layer is allowed to convergence before next layer attempts to restore the situation
  - Fast restoration in lower layer may be ignored in higher layers all together if communication partner with higher layer entity stays the same



# Summary

- Survivable network regains service level quickly in the presence of faults within the network
  - Requires mechanisms for *protection* and/or *restoration*
    - Protection = Backup resource(s) are determined for primary resource(s) before a fault occurs
    - Restoration = The (automatic) process of restoring connectivity after a fault; usually faster if protection is used
  - L 1: SDH – very fast restoration
  - L 2: Ethernet (spanning tree, resilient ring)
  - L 2,5: MPLS - LSP protection
  - L 3: IP - Fault detection bottleneck; overcome with L2 indication or BFD