S-38.192 Verkkopalvelujen tuotanto

S-38.192 Network Service Provisioning

Lecture 10: Resiliency

Survivability

- Modern telecommunication network are built survivable
  - Network maintain service continuity (SLA: availability) in the presence of faults within the network
    - Requires mechanisms for protection and/or restoration
      - Level of mechanisms depend on importance of traffic
        - » 2 nines -> restoration
        - » 5 nines -> protection (1:1)
        - » 7 nines -> protection (1+1)

Protection vs Restoration

- **Protection**
  - Predetermined failure recovery
    - Protection path is precomputed and installed into the network
  - Reconfiguration
    - Switching the affected traffic from faulty entity to backup entity
- **Restoration**
  - Dynamic failure recovery
    - Recovery path is computed after the occurrence of a fault
  - Reconfiguration
    - Selection of a new path for the traffic
    - Rerouting the affected traffic

Different 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
- **1:1 protection**
  - A separate secondary resource is dedicated for each primary resource
  - Extra traffic is carried over the secondary resource but in case of fault in primary traffic is pre-empted from the secondary

Different Modes

- **1:N protection**
  - A secondary resource is set for a group of primary resources
  - Extra traffic is 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 primary traffic is delivered on secondary
      - Priorization of primaries
- **M:N protection** (M<<N)
  - M secondary resources are set for a group of primary resources
    - Higher percentage of primary traffic is secured

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 recalculates new path in case of failure in primary path
  - Precalculation of disjoint paths is possible
    - Faster switch over time

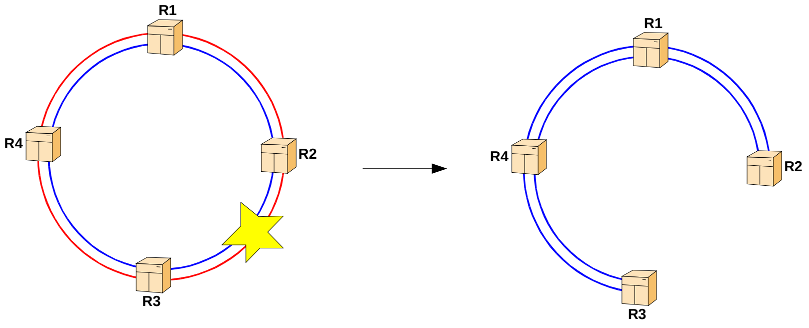
Restoration

- **Global restoration**
  - Network node that detects fault in the network informs all other nodes in the network about existence of fault
    - This depends on 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)
      - Distance vector routing: by calculating new routing vector

SDH

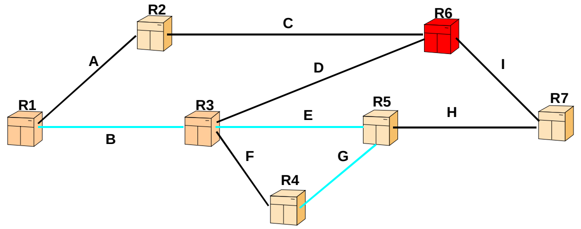
- SDH networks are famous of their fast restoration in case of fault
  - Typically less than 50ms for complete restoration
  - Based on general idea of non-arbitrary network topologies
    - Double rings which can be restored by reversing the traffic at the ends of faulty section
      - Single action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

SDH



Ethernet

- Conventional Ethernet restoration is based on spanning trees
  - Any arbitrary topology is turned into tree topology
    - Each node has weight which determines whether the root of the tree can be reached through it
      - Higher the value the more closer the root is
    - Wastes network resources by blocking loop forming interfaces



Ethernet

- Three are several versions of spanning tree protocol
  - 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 are based on same protocol operation
  - Exchange of BPDU messages to determine whether or not interface should be blocked

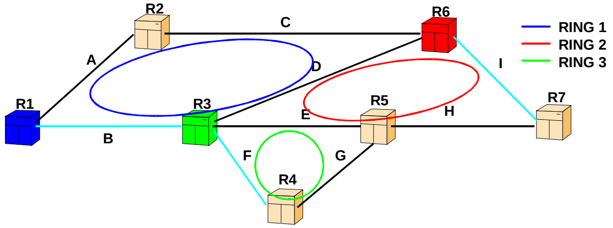
Ethernet

- SDH type network restoration on top of Ethernet
  - Two manufacturers
    - Extreme Networks: Ethernet Automated Protection Switching (EAPS) RFC 3619
    - Foundry: Metro Ring Protocol (MRP)
  - Basic idea same as in SDH
    - Ring type network topology
    - Traffic reversion in case of error

Ethernet

- 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
      - Signalling from the device that detects the fault
    - Convergence time of network is dependent on time between fault and notification of master
      - Varies between
        - » Tens of milliseconds with device signalling
        - » Hundreds of milliseconds with probes

Ethernet



MPLS

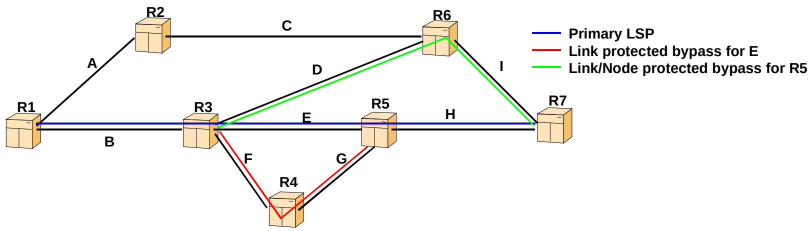
- LSP restoration processes are based on Constrained Shortest Path First routing algorithm for selecting bypass LSPs.
- Different reroute options are:
  - Link protection
  - Link and node protection
  - Path protection
  - Dynamic restoration

Link Protection

- Link protection offers per-link traffic protection
  - Each link on protected LSP has its own bypass for circumventing the failed link
  - Link protection can be made
    - per LSP
    - several LSPs can be aggregated into single bypass LSP
- Requires that
  - Separate bypass is calculted between each RSVP neighbor
  - Router tracks the interface status of egress link and reroutes the protected traffic by stacking the original label with label structure of bypass LSP

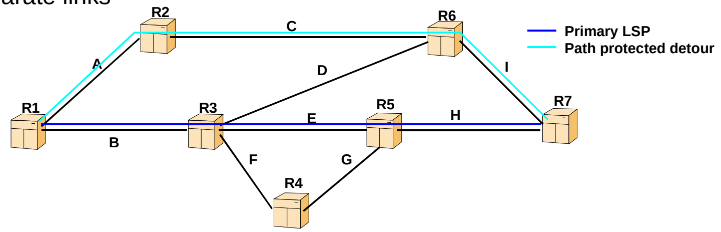
Link/Node Protection

- Node protection is used to circumvent faults which may not be due to interconnecting link rather than next node.
  - Bypass LSP is established around set of next link, node and link using seprate router.
  - Otherwise node protection operates like link protection


R2
C
R6
A
D
I
R1
R3
R5
R7
E
H
B
F
G
R4
Primary LSP
Link protected bypass for E
Link/Node protected bypass for R5

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
    - Separate routers
    - Separate links



Path protection

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

Switch Back

- Switch back is process of rerouting the failed LSPs from their backups
  - 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 circumvented by the fact that failed resources are no longer in TED
    - Causes few hundred milliseconds of additional delay for restoration

IP

- IP restoration is based on convergence of routing protocols
  - Detection of fault
    - Hello timers
    - (L2 indications)
    - (BFD indication)
  - Flood 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
    - L2 may be up even though other router is dead
    - L2 indication process works only if interconnection device fails
    - Normal Hello based detection (tens of seconds)
  - Can be speeded up with usage of bi-directional forwarding detection (BFD)
    - Probes are sent between forwarding planes of routers
    - Fault is signalled 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
- Of 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