Fault management

Andrea Bianco
Telecommunication Network Group
firstname.lastname@polito.it
http://www.telematica.polito.it/

Acknowledgements

• Inspired by
  – Davide Cuda, Politecnico di Torino, now at France Telecom
  – Luca Valcarenghi, Università Pisa Sant’Anna
  – Achille Pattavina, Politecnico di Milano

The impact of network failures

1 cable x 200 fibers/cable x 160 λ/fiber x 10 Gb/s/λ = 320 Tb/s

5 billion telephone lines (@ 64 kb/s)

IP router failure (some data)

- route processor, line card: 70,000 - 150,000 hours MTBF (Mean Time Between Failures)
- software: 10,000 – 100,000 hours MTBF

Note: 1 year is about 10,000 hours

The impact of network failures

• A single cable cut can lead to a dramatic amount of lost traffic
• May translate to revenue losses

Some failure rates

Statistics for the year 2000 for an Optical Cable Network of 30359 km

<table>
<thead>
<tr>
<th>Cause</th>
<th>Number of failures</th>
<th>Percentage of failures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage due to thirds</td>
<td>19</td>
<td>61%</td>
</tr>
<tr>
<td>Rodents</td>
<td>6</td>
<td>29%</td>
</tr>
<tr>
<td>Malice</td>
<td>3</td>
<td>10%</td>
</tr>
<tr>
<td>Materials degradation</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Natural events</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>Hard Failures: service interruption</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

Terminology

• Many different definitions
  • Fault/Failure
    – A catastrophic event that causes the disruption of communications
    – Link failure, interface failure, node failure
  • Fault Tolerance
    – Avoiding service failures in the presence of faults
    – Network fault tolerance is a measure of the number of failures the network can sustain before a disconnection occurs
  • Resilience
    – Network resilience is the maximum number of node failures that can be sustained while the network remains connected with a probability (1-p)
    – The general aim of resilience is to make network failures transparent to users. If a failure happens to affect a circuit, it would be very desirable to reconfigure that circuit as quickly as possible with no information loss
  • Network Integrity
    – Ability of a network to provide the desired QoS in the services, not only in normal (i.e., failure-free) network conditions, but also when network congestion or network failure occurs
A single failure event (such as a span failure) can cause multiple failure (such as individual LSP failures) in the network. A network is referred to as survivable if it provides some ability to recover ongoing connections disrupted by the catastrophic failure of a network component, such as a line interruption or node failure.

Network survivability is the set of capabilities that allows a network to restore affected traffic in the event of a failure. [RFC4427]

 的 Failure Notification/Recovery [RFC4427, RFC4428]

- Failure Detection
- The action of detecting the impairment (defined as a fault degradation) as a defect condition and the consequent activation of a signal failure (SF) or signal degrade (SD), which triggers the control plane to initiate the recovery process. At this stage, the control plane may be involved in two ways:
  - Through the detection phase, it is involved in the recovery process as a fault is detected. The control plane may be involved in the recovery process as a fault is detected.
  - Through the detection phase, it may be involved in the recovery process as a fault is detected.

- Failure Localization
- The process of isolating the faulty component or network element involved in the failure.

- Failure Correlation
- The process of identifying the correlation between multiple failures or degradations.

- Failure Notification
- The process of notifying the appropriate entities about the failure.

- Failure Recovery
- The process of restoring the normal operation of the network after a failure has occurred.

- Recovery Cycle
- The sequence of operations required to restore the normal operation of the network after a failure has occurred.

- Total Recovery Time
- The sum of the duration of the detection, localization, correlation, notification, and recovery phases.

- Operational Time
- The time during which the network is operational and unaffected by failures.

- Downtime Time
- The time during which the network is not operational due to failures.

- Availability
- The probability that a network is operational under normal operating conditions.

- Reliability
- The probability that a network is operational during a specified time period.

- MTTR: Mean Time To Repair
- The average time required to repair a failed network element.

- MTBF: Mean Time Between Failures
- The average time between failures for a network element.

- MTTR: Mean Time To Repair
- The average time required to repair a failed network element.

- MTBF: Mean Time Between Failures
- The average time between failures for a network element.

- Fault types and availability

- Availability (%)
  - 99.999% 6-Nines ~1 min/year
  - 99.99% 5-Nines ~5 min/year
  - 99.9% 4-Nines ~50 min/year
  - 99.9% 3-Nines ~500 min/year

- Downtime Time
  - ~500 min/year
  - ~50 min/year
  - ~5 min/year
  - < 1 min/year

- Min:
  - Before Statistics
  - After Statistics

- Metric Bellcore-Statistics
  - Equipment MTTR
    - 2 hours
  - Cable-outs MTTR
    - 12 hours
  - Cable-cut rate
    - 1-5 cuts/year 100 Km
  - Tx failure rate
    - 10867 FIT
  - Rx failure rate
    - 4311 FIT

- Fault management [RFC4427, RFC4428]

- Failure Detection
  - The action of detecting the impairment (defined as a fault degradation) as a defect condition and the consequent activation of a signal failure (SF) or signal degrade (SD).

- Failure Localization
  - The process of isolating the faulty component or network element involved in the failure.

- Failure Correlation
  - The process of identifying the correlation between multiple failures or degradations.

- Failure Notification
  - The process of notifying the appropriate entities about the failure.

- Failure Recovery
  - The process of restoring the normal operation of the network after a failure has occurred.

- Recovery Cycle
  - The sequence of operations required to restore the normal operation of the network after a failure has occurred.

- Total Recovery Time
  - The sum of the duration of the detection, localization, correlation, notification, and recovery phases.

- Operational Time
  - The time during which the network is operational and unaffected by failures.

- Downtime Time
  - The time during which the network is not operational due to failures.

- Availability
  - The probability that a network is operational under normal operating conditions.

- Reliability
  - The probability that a network is operational during a specified time period.

- MTTR: Mean Time To Repair
  - The average time required to repair a failed network element.

- MTBF: Mean Time Between Failures
  - The average time between failures for a network element.

- Fault types and availability

- Availability (%)
  - 99.999% 6-Nines ~1 min/year
  - 99.99% 5-Nines ~5 min/year
  - 99.9% 4-Nines ~50 min/year
  - 99.9% 3-Nines ~500 min/year

- Downtime Time
  - ~500 min/year
  - ~50 min/year
  - ~5 min/year
  - < 1 min/year

- Min:
  - Before Statistics
  - After Statistics

- Metric Bellcore-Statistics
  - Equipment MTTR
    - 2 hours
  - Cable-outs MTTR
    - 12 hours
  - Cable-cut rate
    - 1-5 cuts/year 100 Km
  - Tx failure rate
    - 10867 FIT
  - Rx failure rate
    - 4311 FIT
Fault management

*Failure types*

- Types of failure
  - Components: links, nodes, WDM channels, active components, software
  - Human errors: fiber cut
  - Fiber inside oil/gas pipelines less likely to be cut
  - Systems
  - Entire Central Offices can fail due to catastrophic events
- Most frequent fault
  - Fiber cut

*Survivability: at the physical layer?*

- Survivability can be provided at different network layers
- Example
  - Ethernet switches may rebuild the spanning tree after a link failure
  - IP routers can fight a link failure by excluding the failed route from their routing tables
- High delays (10s of seconds after the failure is detected and algorithms converge)
- Lots of signaling is required
- Handling faults at the physical is faster and protocol agnostic
- Technologies
  - SDH/SONET networks
    - Point-to-point or ring based
  - WDM networks
    - Mesh-based
- Consider a primary or working path and a secondary or protection path

*Protection schemes*

- 1+1 protection
  - Two nodes connected to each other with two or more links
  - A copy of the signal is transmitted both on the working and the protection channel
  - The receiver can select which copy to accept based e.g. on the signal quality
- 1:1 protection
  - Send data on the working channel only
  - The protection channel is reserved for future use in case of failure
  - The protection channel can be used for low priority data traffic
  - Only when a failure occurs on the working channel, the protection channel carries the traffic from the failed working channel
m:n protection

- n working links are protected using m backup links
- Working and backup path from a m:n protected group
- Backup channels can be used to carry low priority traffic

Protection techniques

- 1+1
  - the most costly
  - the fastest
- 1:1
  - higher efficiency in the protection capacity usage
  - higher restoration time
    - actions (and signalling) are needed to switch traffic from the working channel to the protection one
- m:n
  - requires less resources
  - might not be able to restore all the traffic supported before the fault
  - exploits the non-100% utilization of working links
  - how many working links may fail at the same time?

Fault management in SONET/SDH

- Dedicated vs. Shared: working connection assigned dedicated or shared protection bandwidth
  - 1+1 is dedicated, 1:n is shared
- Revertive (Non-revertive): after failure is fixed, traffic is automatically, or manually, (not) switched back on the working path
  - Shared protection schemes are usually revertive
- Unidirectional connection
  - Data transmitted both on the working (primary) and the backup (secondary) path
- Bidirectional connection
  - Data could need to be switched from the working path to the backup path (even if a fault affects only one of the primary connections)
  - Automatic Protection Switching (APS): signaling protocol to detect faults

Fault management in SONET/SDH: 1+1 APS

Fault management in SONET/SDH: 1:1 APS

Fault management in SONET/SDH: Self-healing Ring

- Much of the carrier infrastructure today uses SONET/SDH rings
  - Multiple nodes are interconnected with a single physical ring
- Rings show interesting restoration properties:
  - 2 connected topology
  - Provides 2 disjoint paths between any couple of nodes
  - Unidirectional-Line-Switched-Ring (UPLR)
  - Bidirectional-Line-Switched-Ring (BPLR)
Fault management

Current SDH architecture:
Ring protection
Multiple rings over DWDM

Unidirectional-Line-Switched-Ring:
Failure-free State
• Fiber 1: Working path
• Fiber 2: Protecting path
• Traffic is sent simultaneously on the working and the protecting path
• 1+1 protection

Unidirectional-Line-Switched-Ring:
Failure State
• A switches the receiver from the working path to the protecting path
• B remains unaffected
• Non trivial to detect failure in A

Unidirectional-Line-Switched-Ring
• Easy to implement
• No signaling
• Fast failure recovery
• ULSRs popular in lower-speed local exchange and access
• The delay difference between the working and the backup path affects the restoration time (max 60ms according to standards)
• Not efficient
  – 50% of capacity for protection purpose
  – no spatial reuse of wavelengths
  – no sharing of resources dedicated to protection

Bidirectional-Line-Switched-Ring
BPLR/4
• 4-Fiber: BPLR/4
• Shortest path routing

Bidirectional-Line-Switched-Ring:
Span switching
• Failure of a working path: traffic is routed through the protection fiber
Fault management in WDM mesh networks

- Working and backup path must be link disjoint
- Preplanned protection:
  - Resources dedicated to protection are allocated each time a new light path is set up
  - Simple and fast
  - Rigid resource allocation
- Provisioning:
  - The network is generically over-provisioned with respected the real traffic the network needs to support and backup connections are allocated only if primary connections fail
  - Provisioning is more complex & slower than preplanned protection
  - Provisioning can support multiple failures
Fault management

### Ring cover

- Mesh networks
  - Large scalability (physical, lower spare resources)
  - Large restoration time
- Ring
  - Easy to manage
  - Fast reconfiguration time
  - Capacity inefficient (~100% of spare capacity)
- Cover the physical mesh network with logical rings
  - Nodes are still physical nodes
  - Links are composed of one or multiple wavelength channels
  - Logical rings can behave as ULSR/BLSR or can be used for protection only

### Ring cover: stacked rings

- If rings behave as ULSR/BLSR
  - End-to-end traffic
    - Inter-Ring traffic
    - Intra-Ring traffic
      - Protected in each segment by a different ring
- A→B routed through R5 and R4

### Protection cycle (p-cycle)

- Rings are used for protection purposes only
- Find a set of directed cycles covering all links in a network
  - If a link goes down, there is a cycle to recover the traffic of the failed link
  - Given the set of all covering cycles, ILP techniques can be used to achieve different purpose (e.g., minimize reconfiguration time, minimize extra-capacity needed for protection)
  - p-cycles can protect both links on the ring and chordal (straddling) links
  - Straddling links are links having p-cycle nodes as endpoints

<table>
<thead>
<tr>
<th>Attribute</th>
<th>p-cycles</th>
<th>SONET rings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modularity</td>
<td>One spare span per link</td>
<td>OC-n modularity</td>
</tr>
<tr>
<td>Protection yield</td>
<td>Up to 2 useful restoration paths per p-cycle</td>
<td>1 restoration path per ring (or protection channel)</td>
</tr>
<tr>
<td>Protection flexibility</td>
<td>p-cycles contribute to the restoration of working links on the cycle and all straddling links</td>
<td>Rings only protect working links in spans contained within the ring</td>
</tr>
<tr>
<td>Routing and provisioning of working paths</td>
<td>Protocols without regard to structures formed in the sparing (protection) layer</td>
<td>Working path routing must be a succession of intra-ring and inter-ring traversals</td>
</tr>
<tr>
<td>Total network recoverability</td>
<td>Essentially just that of a span restorable mesh network</td>
<td>Over 100% investment in spare capacity (Can raise up to 300% depending on topology and traffic)</td>
</tr>
</tbody>
</table>
Fault management

Algorithm to compute disjoint paths: 2 steps algorithm
- 1st step: 1st path is computed using shortest path algorithm
- 2nd step: remove the edge from the original graph and compute another path using a shortest path algorithm
- Does not work

Algorithm to compute disjoint paths
- Given a graph, a source node s, a destination node d, a couple of disjoint paths is computed as follows:
  - Compute the shortest path tree rooted at node s. Let d(s,u) denote the distance shortest-path distance between s and u.
  - Transform G into an auxiliary graph G':
    - Nodes and links are kept unchanged.
    - The cost of each link (u,v) in G' is defined by c'(u,v)= c(u,v)+d(s,u)-d(s,v).
    - Reverse direction of the links along the shortest path from node s to node d.
  - Compute the shortest path from node s to node d in G'.
  - The shortest path between node s and d in G (G') is denoted as T (T'). After removing the link appearing both in T and T' (the one in opposite direction) the remaining link form a cycle. Two link-disjoint path between s and d can be found from the cycle.

Algorithm to compute disjoint paths
- Compute the shortest path tree and d(s,u).
- Transform G into an auxiliary graph G':
  - The cost of each link (u,v) in G' is defined by c'(u,v)= c(u,v)+d(s,u)-d(s,v).
- Reverse direction of links along the shortest path.
- Compute the shortest path in G'.

Algorithm to compute disjoint paths
- Transform G into an auxiliary graph G':
  - Reverse direction of links along the shortest path.
  - Compute the shortest path in G'.
- Remove links in the shortest path both in G and G'.

Working Protection