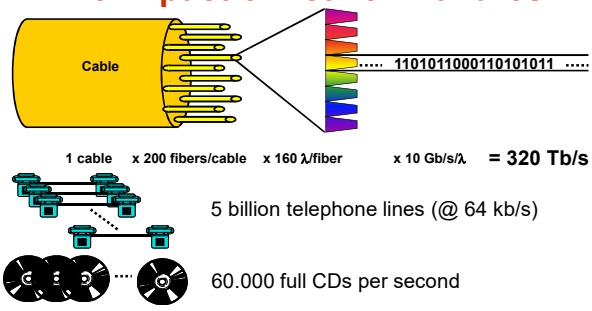
Fault management

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

Acnowledgements

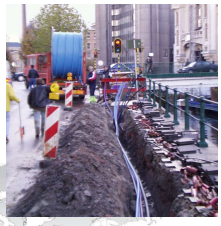
- Inspired by
  - Davide Cuda, Politecnico di Torino, now at CISCO Lausanne
  - Luca Valcarenghi, Università Pisa Sant'Anna
  - Achille Pattavina, Politecnico di Milano

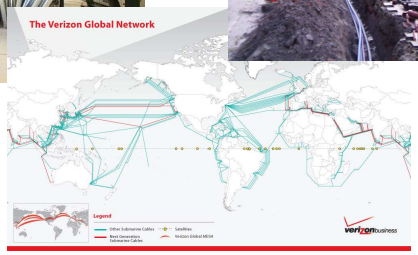
The impact of network failures



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

Fibers





Some failure rates

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

Hard Failures: service interruption

Cause	Number of failures	Percentage of failures
Damage due to thirds	19	61%
Rodents	6	29%
Malice	3	10%
Materials degradation	1	3%
Natural events	1	3%
Installation Defects	1	3%
Total	31	Source: Sirti

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

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 to the services, not only in normal (i.e., failure free) network conditions, but also when network congestion or network failure occurs

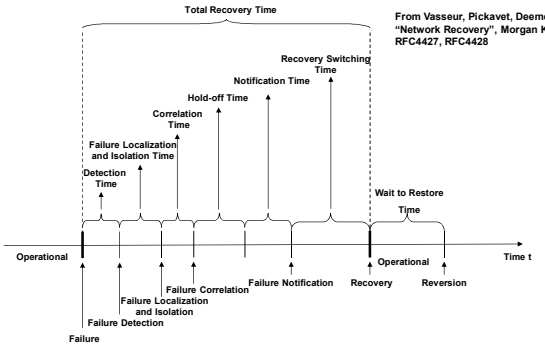
Terminology

- Survivability
  - network survivability is the fraction of a quantifiable feature x that remains after an instance of the disaster type under consideration has happened
    - x can be defined as the traffic volume, the number of connected subscribers, the network operator's revenue, the grade of service, other network characteristics that are related to the remaining "goodness" of the network
  - is a subset of integrity
    - Ability of a network to recover the traffic in the event of a failure, causing few or no consequences for the user
  - 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]

Terminology

- Reliability
  - The probability of a network element (e.g., a node or a link) to be fully operational during a certain time frame
  - R(Δt), is the probability that the system works correctly in the period of time Δt under defined environmental conditions
- Availability
  - is the instantaneous counterpart of reliability
  - Network element availability is the probability of a network element to be operational at one particular point in time
  - A(t) is the probability that the system works correctly at time point t
  - probability that an item will be able to perform its designed functions at the stated performance level, conditions and environment when called upon to do so
    - reliability_time over (reliability_time+ recovery_time)
  - Different ways to measure availability (port, bandwidth, blocking probability)

Recovery Cycle



Fault management [RFC4427,RFC4428]

- Failure Detection
  - The action of detecting the impairment (defect of performance degradation) as a defect condition and the consequential activation of Signal Fail (SF) or Signal Degrade (SD) trigger to the control plane. Thus, failure detection is the only phase that cannot be achieved by the control plane alone.
- Failure Localization (and Isolation)
  - Failure localization provides information about the location (and thus the identity) of the transport plane entity that causes the LSP(s)/span(s) failure.
- Failure Correlation
  - A single failure event (such as a span failure) can cause multiple failure (such as individual LSP failures) conditions to be reported. These can be grouped (i.e., correlated) to reduce the number of notified failures.
- Failure Notification
  - Failure notification phase is used 1) to inform intermediate nodes that LSP(s)/span(s) failure has occurred and has been detected and 2) to inform the recovery deciding entities (which can correspond to any intermediate or end-point of the failed LSP/span) that the corresponding LSP/span is not available.
- Recovery
  - Recovery of the failure through a recovery scheme
- Reversion (Normalization)
  - A revertive recovery operation refers to a recovery switching operation, where the traffic returns to (or remains on) the working LSP/span when the switch-over requests are terminated (i.e., when the working LSP/span has recovered from the failure).

Timing [RFC4427, RFC4428]

- Detection time
  - The time between the occurrence of the fault or degradation and its detection. Note that this is a rather theoretical time because, in practice, this is difficult to measure.
- Localization time
  - The time necessary to localize the failure.
- Correlation time
  - The time between the detection (localization) of the fault or degradation and the reporting of the signal fail (SF) or degrade (SD). This time is typically used in correlating related failures or degradations.
- Hold-off time
  - time between the reporting of signal fail (SF) or degrade (SD), and the initialization of the recovery switching operation with the deciding entities notification. This is useful when multiple layers of recovery are being used.
- Notification time
  - The time between the reporting of the signal fail or degrade and the reception of the indication of this event by the entities that decide on the recovery switching operation(s).
- Recovery Switching Time
  - The time between the initialization of the recovery switching operation and the moment the normal traffic is selected from the recovery LSP/span.
- Total Recovery time
  - The sum of the detection, the localization, the correlation, (the hold-off,) the notification, and the recovery switching time.
- Wait To Restore time
  - A period of time that must elapse after a recovered fault before an LSP/span can be used again to transport the normal traffic and/or to select the normal traffic from.

Fault types and availability

Availability(%)	N-Nines	Downtime Time (min/year)
99%	2-Nines	~5000 min/year
99.9%	3-Nines	~500 min/year
99.99%	4-Nines	~50 min/year
99.999%	5-Nines	~5 min/year
99.9999%	6-Nines	< 1 min/year

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

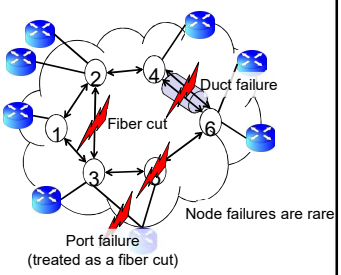
MTTR: Mean Time To Repair

FIT: Failure In Time (# failures in 10^9 hours)

1 year = 8760 hour = 500000min

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



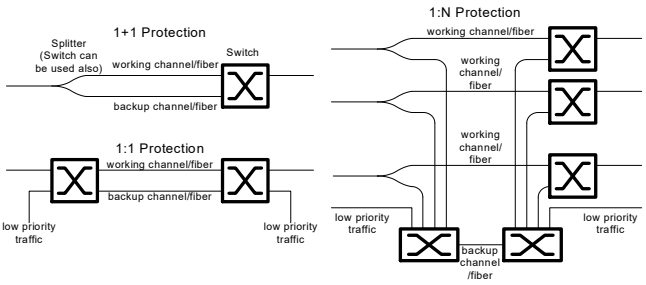
Survivability: at the physical layer?

- 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)
      - During this time packets can be routed incorrectly
    - 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

Survivability

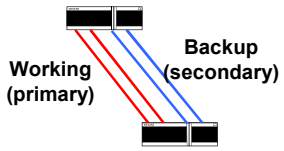
- Requires exploiting redundant capacity to deal with failures
- Often classified in
  - Provisioning (restoration)
  - Protection
- Restoration (also named dynamic resilience)
  - Redundant capacity not reserved
  - Affected traffic is re-routed by on-line processing
  - Reaction time in the order of s
- Protection (also named static resilience)
  - Redundant capacity pre-allocated
  - Automatic re-routing upon failure
  - Reaction time in the order of ms
- One single failure at the time is normally considered
  - Overall network partitioned in sub-networks
    - One failure per sub-network
  - MTTR << MTBF
- Mostly focus on links/paths
  - Node survivability is guaranteed through backup provisioning (1+1 scheme, see later)

Protection schemes



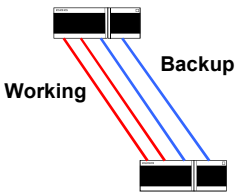
1+1 protection

- 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)

1:1 protection

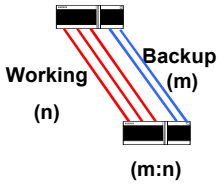


(1:1)

- 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



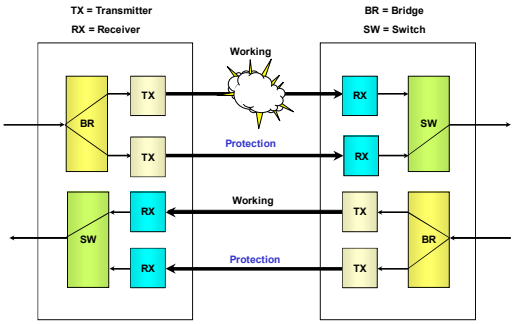
Protection techniques

- 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?

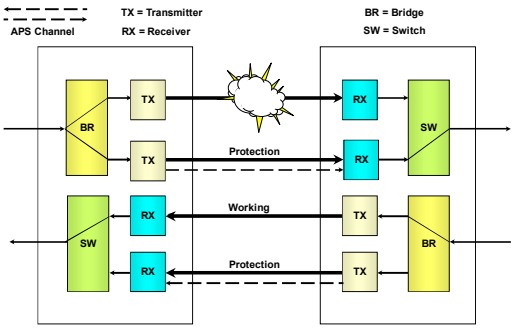
Fault management in SONET/SDH

- 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



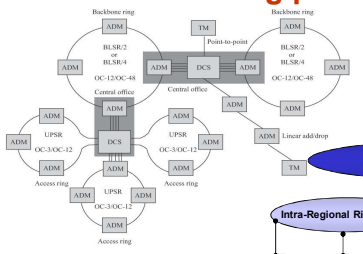
Fault management in SONET/SDH: 1:1 APS



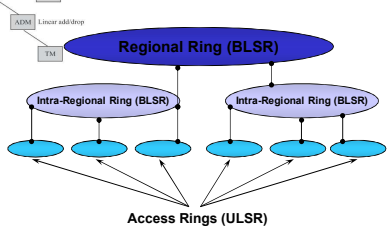
Fault management in SONET/SDH: Self-healing Ring

- 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)

Current SDH architecture: Ring protection

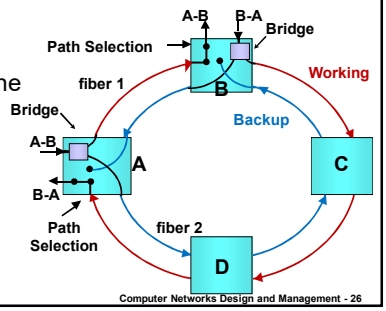


Multiple rings over DWDM



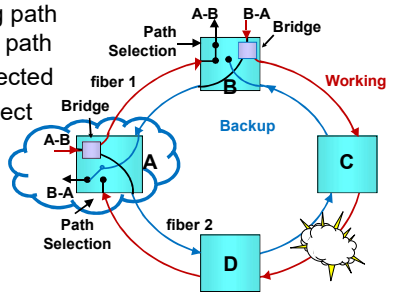
Unidirectional-Line-Switched-Ring: Failure-free State

- 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

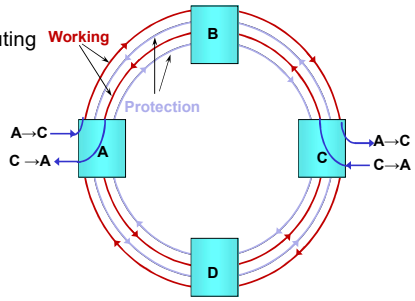


Unidirectional-Line-Switched-Ring

- 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

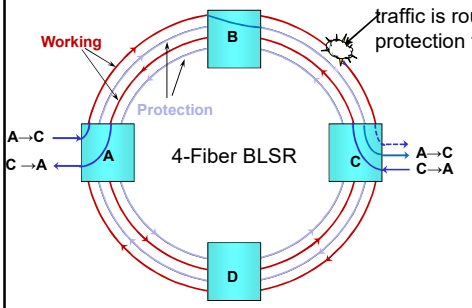
Bidirectional-Line-Switched-Ring BPLR/4

- 4-Fiber: BPLR/4
- Shortest path routing



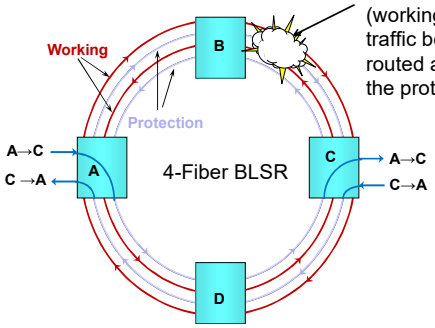
Bidirectional-Line-Switched-Ring: Span switching

- Failure of a working path: traffic is routed through the protection fiber



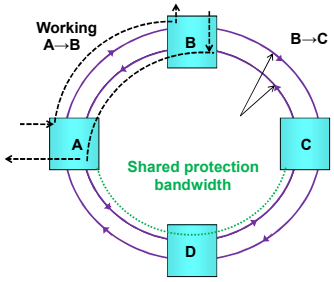
Bidirectional Line Switched Ring: Ring switching

- Node or fiber bundle failure (working and protection) : traffic between nodes is routed around the ring on the protection fiber



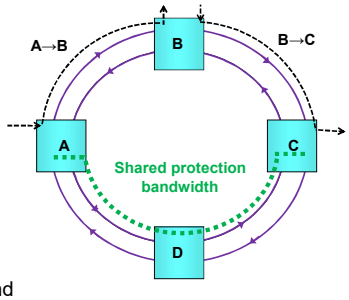
Bidirectional Line Switched Ring BLSR/2

- 2-Fiber: BLSR/2
- In each fiber half of the capacity is reserved for protection purposes
- Span switching is not possible

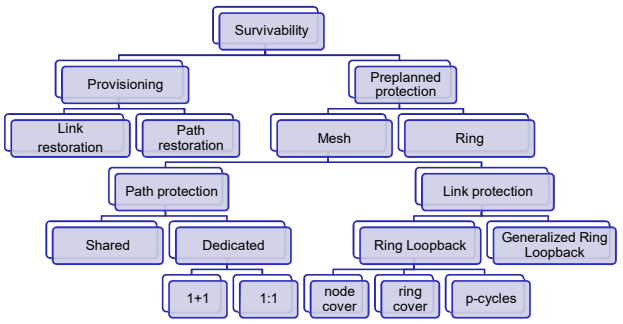


Bidirectional Line Switched Ring

- 1:n and m:n are possible in BLSR
- Protection bandwidth can be shared by spatially separated path
- BLSR are
  - More capacity efficient than ULSR
  - More complex (extensive signaling and protection mechanism)



Fault management in WDM mesh networks

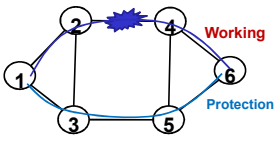


Fault management in WDM mesh networks

- 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

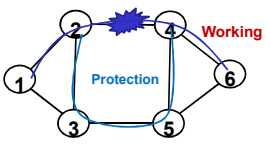
Fault management in WDM mesh networks

- Path protection
  - Traffic rerouted through a backup route
  - Primary and backup route must be link-disjoint



- Sub-Path protection: divide path in a sequence of segments and protect each segment separately

- Link protection:
  - Traffic rerouted around the failed link

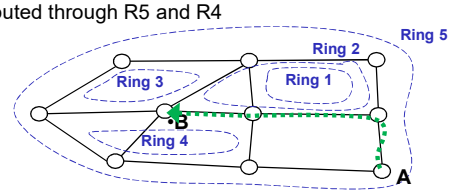


Ring cover

- 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

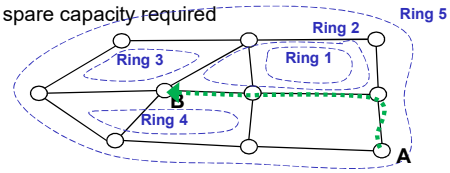
Ring cover: stacked rings

- 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



Ring cover: stacked rings

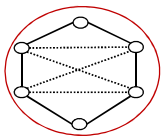
- If rings behave as ULSR/BLSR
  - Horizontal dimension: adjacent rings provide geographic coverage
  - Vertical dimension: rings stacked on top of each others providing additional capacity
  - Usually achieved by WDM
  - 100% of spare capacity required



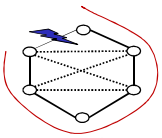
Protection cycle (p-cycle)

- 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 end-points

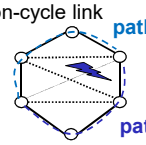
Protection cycle (p-cycle)



p-cycle



Protecting failure of a on-cycle link



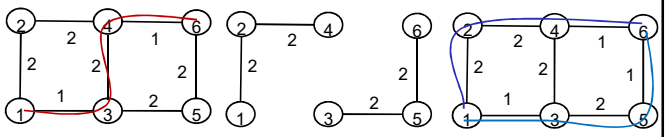
Protecting failure of a straddling link (2 paths)

Protection cycle (p-cycle)

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

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

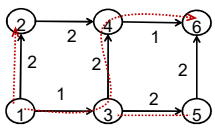


Algorithm to compute disjoint paths

- Given a graph, a source node s, a destination node d, a couple of disjoint paths is computed as it 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 define 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.

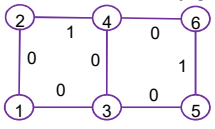
Algorithm to compute disjoint paths

- Compute the shortest path tree and d(s,u)



d(1,2) = 2
d(1,3) = 1
d(1,4) = 3
d(1,5) = 3
d(1,6) = 4

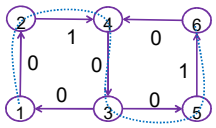
- Transform G into an auxiliary graph G'



c'(u,v)= c(u,v)+d(s,u)-d(s,v)
c'(1,2) = 2 + 0 – 2 = 0
c'(1,3) = 1 + 0 – 1 = 0
c'(2,4) = 2 + 2 – 3 = 1
c'(3,4) = 2 + 1 – 3 = 0
c'(5,6) = 2 + 3 – 4 = 1
c'(4,6) = 1 + 3 – 4 = 0
c'(3,5) = 2 + 1 – 3 = 0

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'