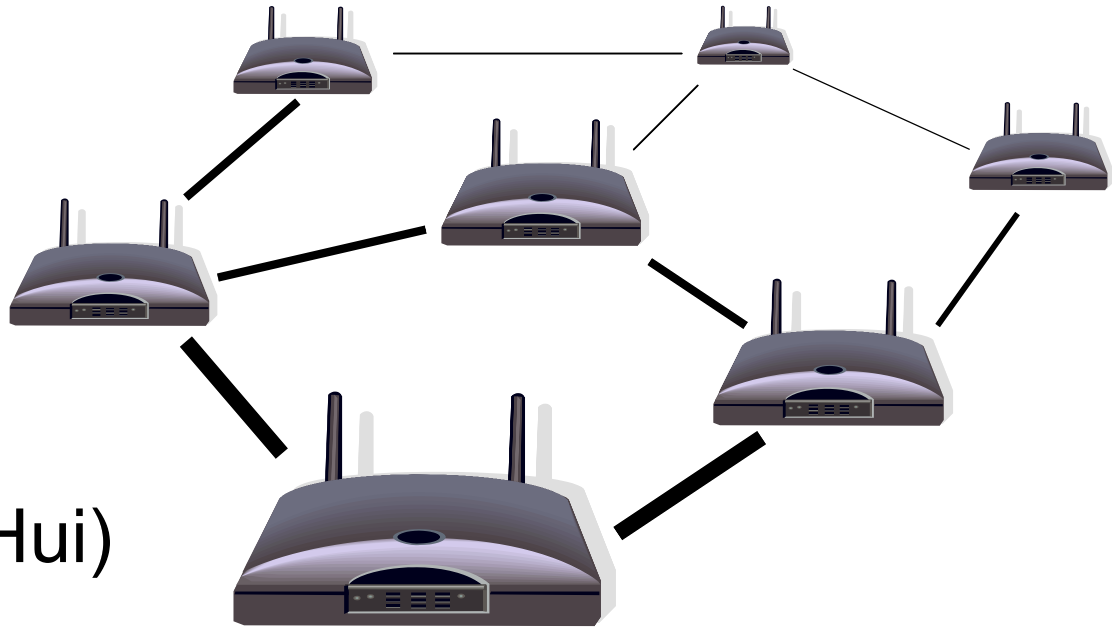
Achieving Resilient Routing in the Internet



Presented by...

Suksant Sae Lor (Hui)

Supervised by Dr Miguel Rio

NSRL, Dept. E&EE, UCL

Outline

- Introduction & Motivation
- IP Fast Re-Route Framework
- Fast Failure Detection
- Existing Repair Paths Mechanisms
- Fast Re-Route Using Alternate Next Hop Counters (ANHC)
- Performance Evaluation
- Conclusions

Introduction & Motivation

- Evidently, Internet is resilient to random failures.
- Alas, it is not tolerable for sensitive applications.
- Massive amount of packets are dropped during routing convergence.
- Several approaches have been proposed: shortening the convergence time, pre-computing backup paths, overlays, *etc*.
- Loop-free environment and routing consistency are important.

IP Fast Re-Route Framework

- Rescue packets from failures as fast as possible without waiting for the network to converge.
- Disruption time:
  - time to detect and react to failures.
  - time to implement new routes into forwarding tables.
- Two main mechanisms*:
  - Mechanisms for fast failure detection.
  - Mechanisms for repair paths.

*Internet Draft (draft-ietf-rtgwg-ipfrr-framework-11)

Fast Failure Detection Mechanisms

- In general, protocol parameters used to detect failures are:
  - Hello interval: default is ~10 seconds.
  - Dead router interval: default is ~30-40 seconds (usually multiples of Hello interval).
- Tweaking the Hello interval: ms < t < s*
- Minimum Hello interval for IS-IS, however, is 1s
- Too short interval leads to routing instabilities as the failures may be intermittent.

*Achieving Faster Failure Detection in OSPF Networks (M. Goyal, et al.)

Loop-Free Alternates (LFAs)

- A neighbour of a detecting node can be used as an LFA if it neither causes the traffic to traverse the failure nor creates a forwarding loop.
- LFAs are categorised by their abilities:
  - Loop-Free Condition (LFC): link protecting LFA.
  - Node-Protection Condition (NPC): node protecting LFA.
  - Downstream Condition (DSC): loop-free LFA in the presence of multiple failures.
  - Equal-Cost Alternates (ECA): equal-cost paths.
- LFAs are simple, but their repair coverage heavily depends on the underlying topologies.

Not-Via Addresses

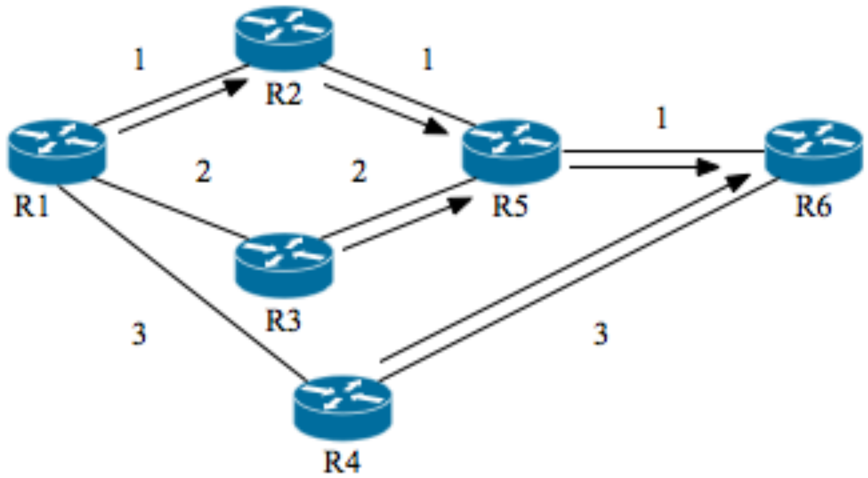
- Special addresses used to deviate the traffic around the failures.
- Requires IP-in-IP tunnelling.
- Packets are forwarded along the path avoiding the failed element.
- Guarantee 100% repair coverage for any recoverable single failures.
- However, it may degrade the performance of a router due to additional processing.

Fast Re-Route Using Alternate Next Hop Counters (ANHC)

- Guarantees 100% repair coverage for any single link failures.
- Does not employ mechanisms such as tunnels.
- Requires additional information for each existing destination in the routing table (no additional entry is required).
- Does not incur any significant overheads.
- Alternate paths are near optimal.
- Its impact on the traffic is comparable to OSPF re-route (normal convergence).

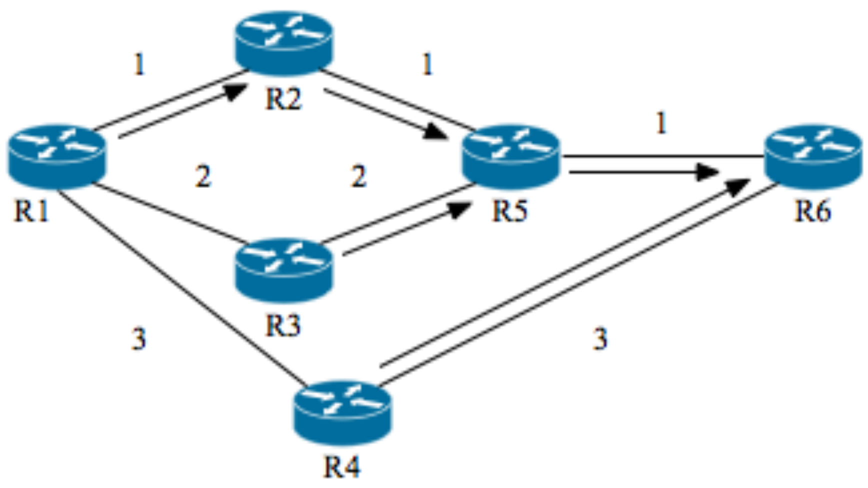
Computing the Alternate Paths (1)

- Creating some correlations between alternate paths from different origins to the same destination. The arrows form a SPT rooted at R6.



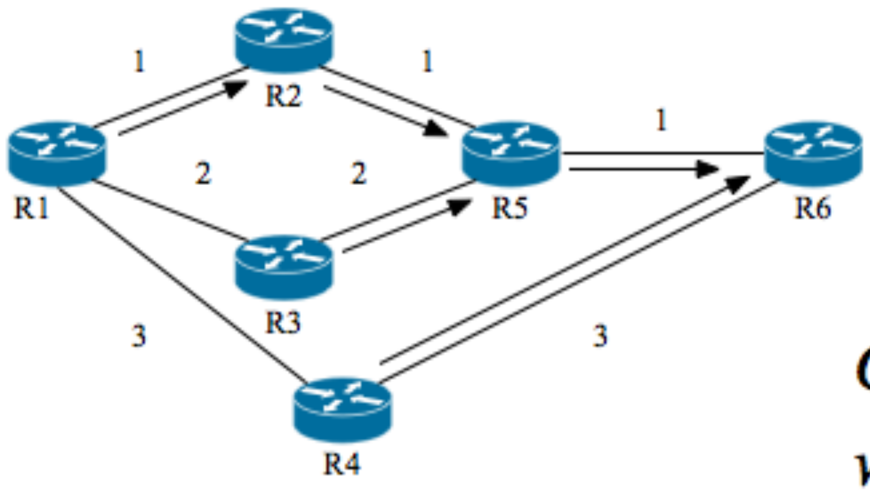
Computing the Alternate Paths (2)

- How? For all origins to the same destination, compute the alternate paths that are maximally edge disjoint from the normal paths.



Computing the Alternate Paths (3)

- How? For all origins to the same destination, compute the alternate paths that are maximally edge disjoint from the normal paths.



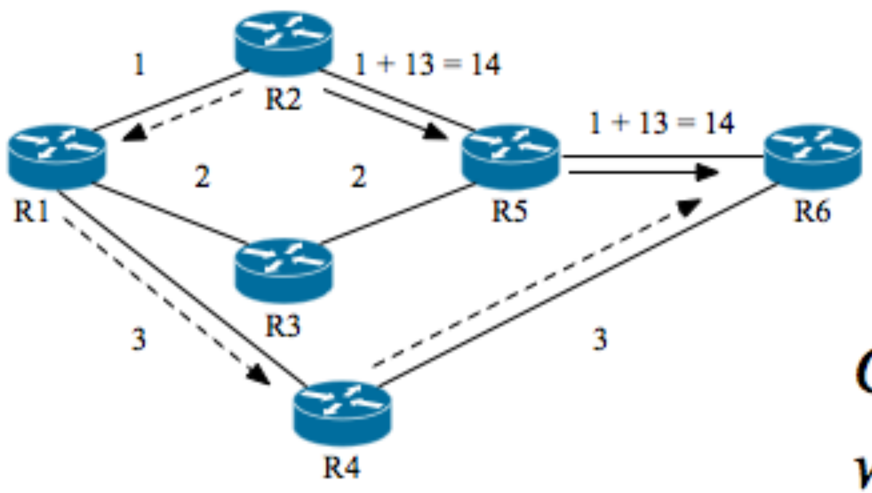
$$G = (V, E)$$

$$w(i,j) \in \mathbb{R} > 0$$

$$W_t := \sum_{(i,j)\in E} w(i,j)$$

Computing the Alternate Paths (4)

- In this topology, the total link weight is 13.
- The figure shows an example of alternate path computation of R2 to R6.



$$G = (V, E)$$

$$w(i, j) \in \mathbb{R} > 0$$

$$W_t = 13$$

Computing the ANHC values (1)

- Compare the hops of local alternate paths with the alternate next hop of intermediate nodes.
- REQUIRE:
  - Alternate path from R2 to R6
  - Alternate next hops (ANHs) from all origins to R6.
- R2s alternate path: R2→R1→R4→R6
- ANHs: R1:R4, R2:R1 , R3:R1, R4:R1, R5:R2
- ANHC(R2, R6) = **0**, R1 = R2s ANH?, YES

Computing the ANHC values (2)

- Compare the hops of local alternate paths with the alternate next hop of intermediate nodes.
- REQUIRE:
  - Alternate path from R2 to R6
  - Alternate next hops (ANHs) from all origins to R6.
- R2s alternate path: R2→R1→R4→R6
- ANHs: R1:R4, R2:R1 , R3:R1, R4:R1, R5:R2
- ANHC(R2, R6) = **1**, R4 = R1s ANH?, YES

Computing the ANHC values (3)

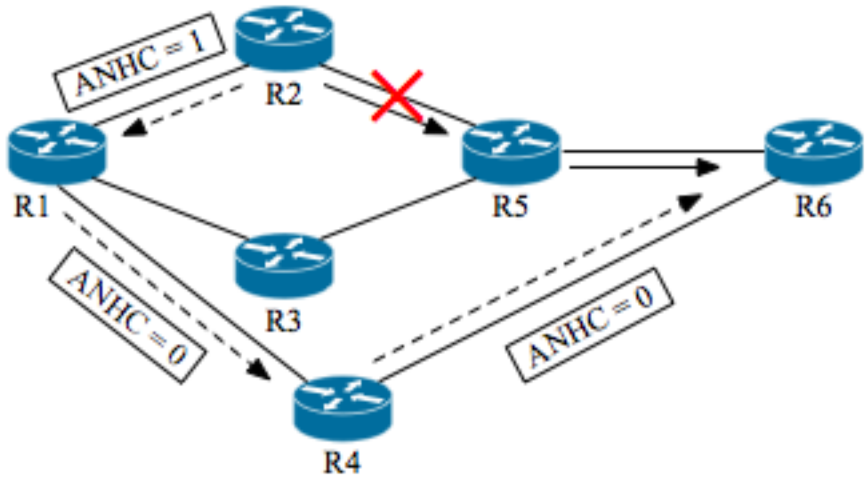
- Compare the hops of local alternate paths with the alternate next hop of intermediate nodes.
- REQUIRE:
  - Alternate path from R2 to R6
  - Alternate next hops (ANHs) from all origins to R6.
- R2s alternate path: R2→R1→R4→R6
- ANHs: R1:R4, R2:R1 , R3:R1, R4:R1, R5:R2
- ANHC(R2, R6) = **2**, R5 = R4s ANH?, NO

Alternate Next Hop Counting Mechanisms (1)

- Normal forwarding in failure-free case.
- When a failure occurs, the detecting node marks the packet with ANHC value.
- The ANHC value is decreased by 1 and forwarded to the alternate next hop.
- Each router receiving a re-routed packet determines the ANHC value.
  - ANHC > 0: decrements it and forwards the packet to its alternate next hop.
  - ANHC = 0: forwards the packet along the normal path.

Alternate Next Hop Counting Mechanisms (2)

- R2 set ANHC(R2, R6) = 2.
- R2 decreases ANHC to 1 & forwards the packet.
- R1 decreases ANHC to 0 & forwards the packet.

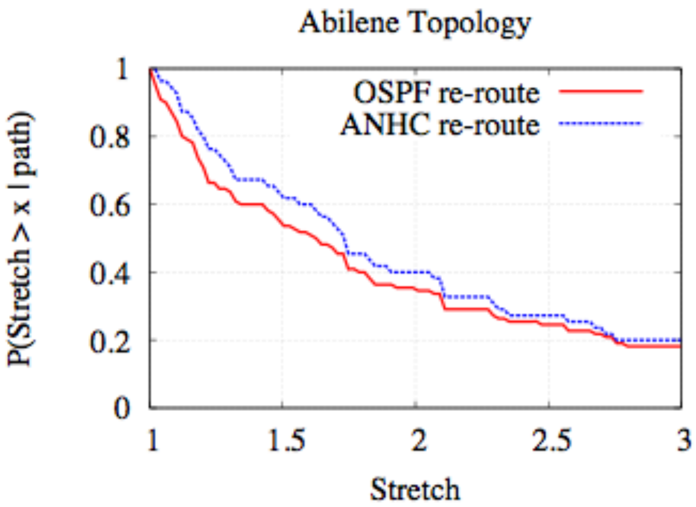


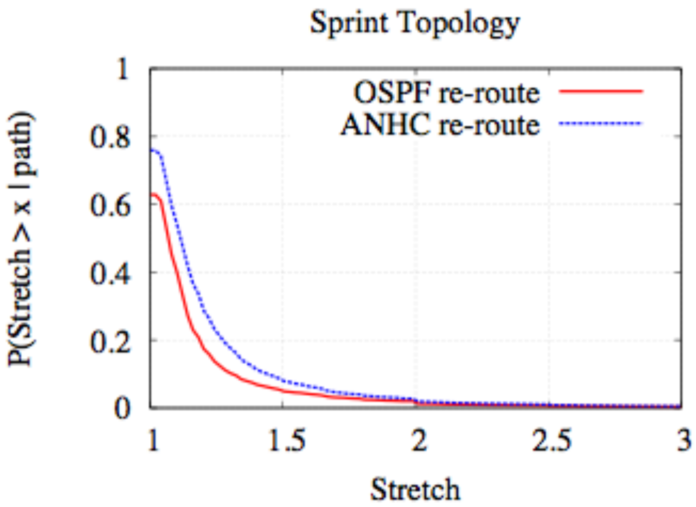
Preventing Loops Under Multiple Failures

- ANHC requires few bits in the packet header.
- Simulation results of practical topologies show that the optimal number of bits required is 3.
- In the presence of multiple failures, forwarding loops are possible.
- Employ an extra bit to indicate a re-routed packet. Thus, if a marked packet encounters another failure, it will be dropped immediately.
- Total number of bits required is 4.
- TOS in IPv4 or Traffic Class in IPv6 may be used.

Path Length Stretch

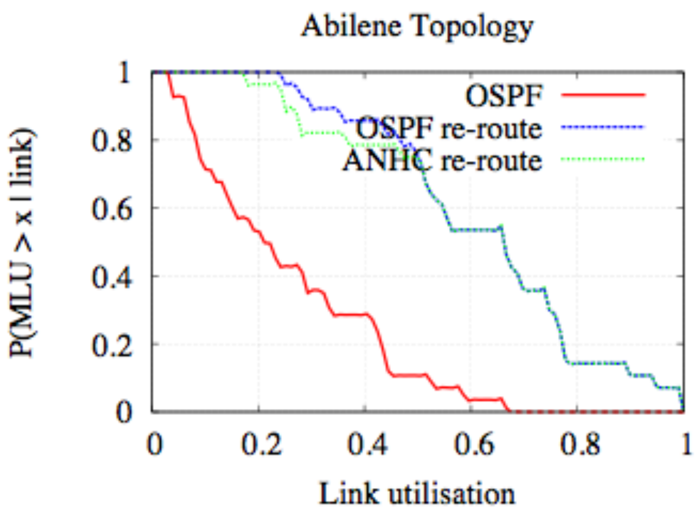
- Path length stretch :- the ratio between the alternate path cost and the optimal shortest path.

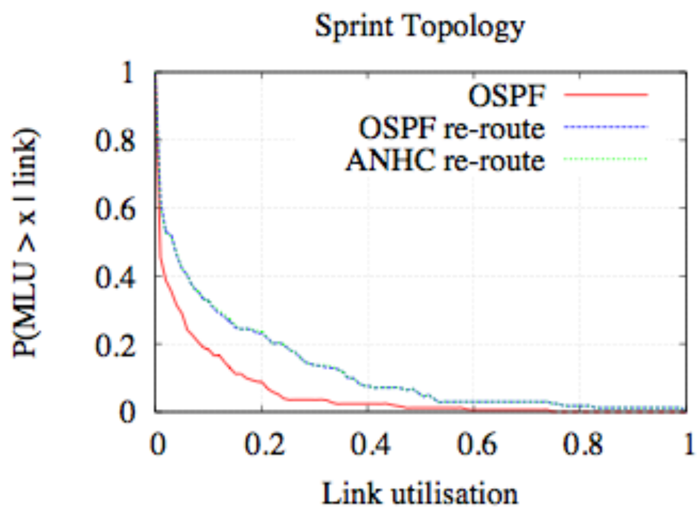




Maximum Link Utilisation (MLU)

- Abilene - real TMs*.
- Sprint - TMs* generated based on gravity model.

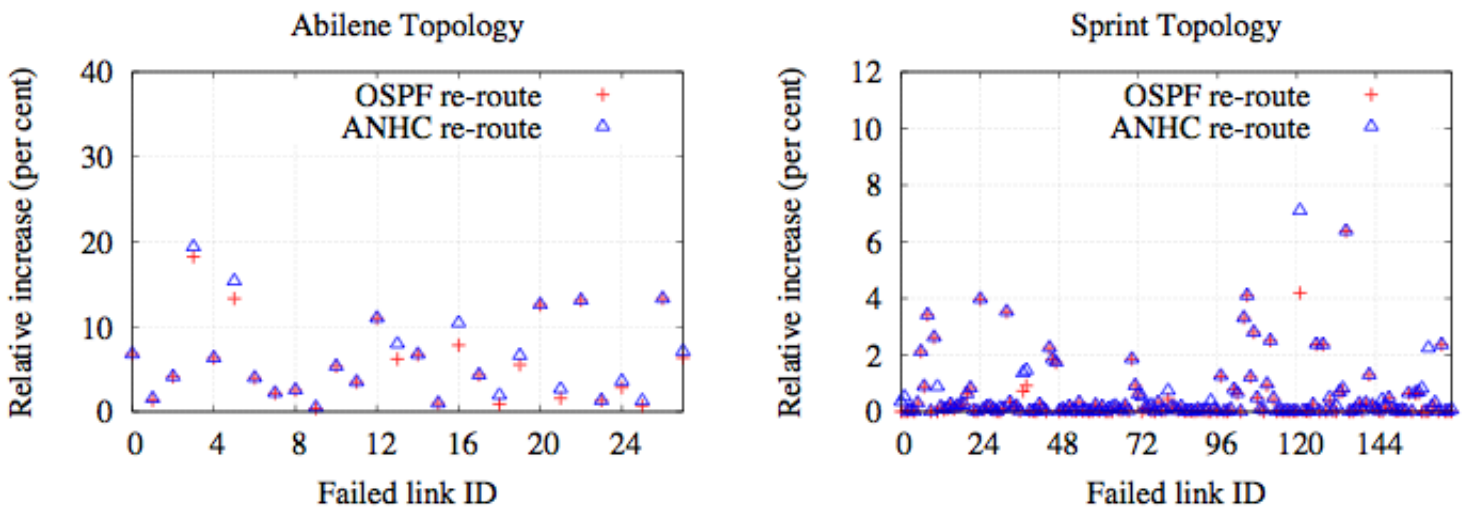




*Traffic matrices are scaled so that no MLU > 1 under normal convergence.

Total Network Throughput

- Total network throughput after different failure scenarios.



Conclusions

- Network reliability problem is very challenging due to ongoing demand for highly reliable delivery.
- Existing solutions such as LFAs, U-turn, and tunnels do not provide full repair coverage.
- Not-via addresses guarantee recovery from any single recoverable failures.
- Fast re-route using ANHC provides full protection against single link failures without using tunnels.
- Fast re-route using ANHC does not incur any significant overheads or impact on network traffic.

Thank you!

Questions & Answers