Applicability of Loop-free Convergence

<draft-bryant-shand-lf-applicability-01.txt >
IETF-65 – March 2006

Stewart Bryant (Cisco Systems)
Mike Shand (Cisco Systems)

Micro-loops Are Bad

- When a network re-converges micro-loops may form.
- Micro-loops result in collateral damage to traffic not affected by the change, as well as causing the affected traffic to be lost.
- Micro-loop damage has always been accepted as a necessary evil of the routing convergence process.

Addressing the Problem

- Work is being undertaken in the RTGWG on the development of IP Fast Reroute.
- IPFRR provides a similar service to RSVP-TE Fast Reroute, but is applicable to non-TE paths.
- Early in this work it was observed that once re-convergence started, the repair would be starved, and micro-loops would form.

- The IPFRR developers have proposed a number of methods of preventing or reducing the number and impact of micro-loops.

- This draft explores the wider implications of that work.

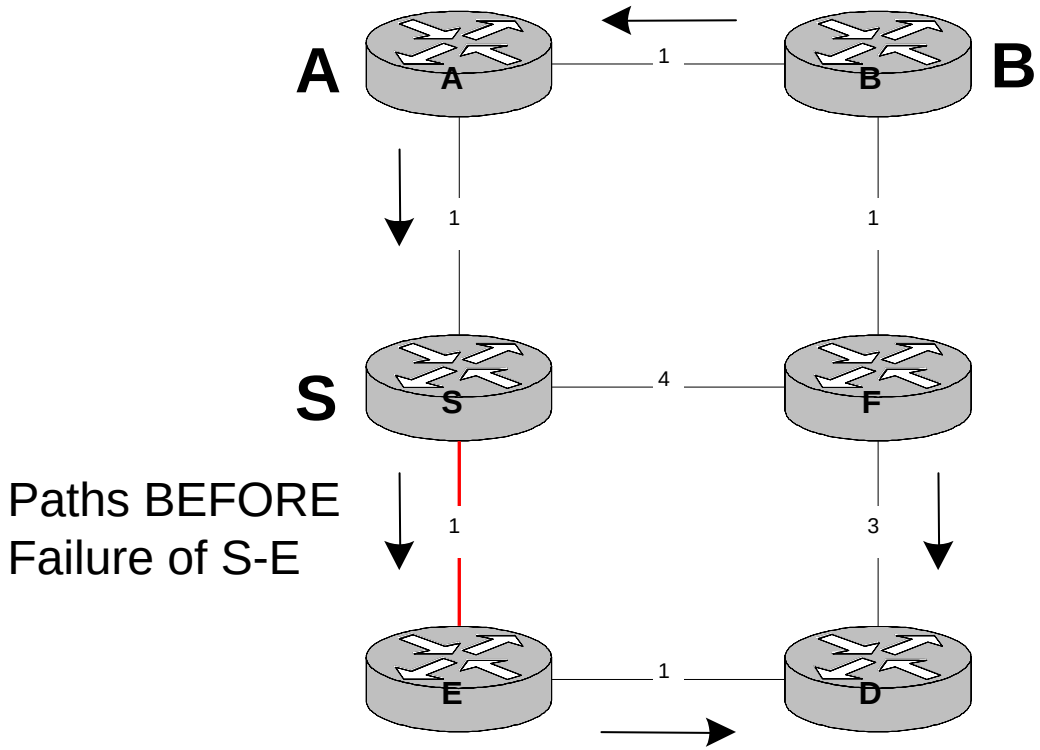
FIB Update Order

- Microloops are caused when router FIBs are updated in the “wrong” order.
- The “natural” order for failure events is the wrong order.
- Implementation specific factors will affect the exact order, but most failure events will cause loops.

“Good News” Events

- The natural order for good news events is such that loops should not occur, but implementation factors can still result in loops.
- These loops during benign events are particularly annoying!

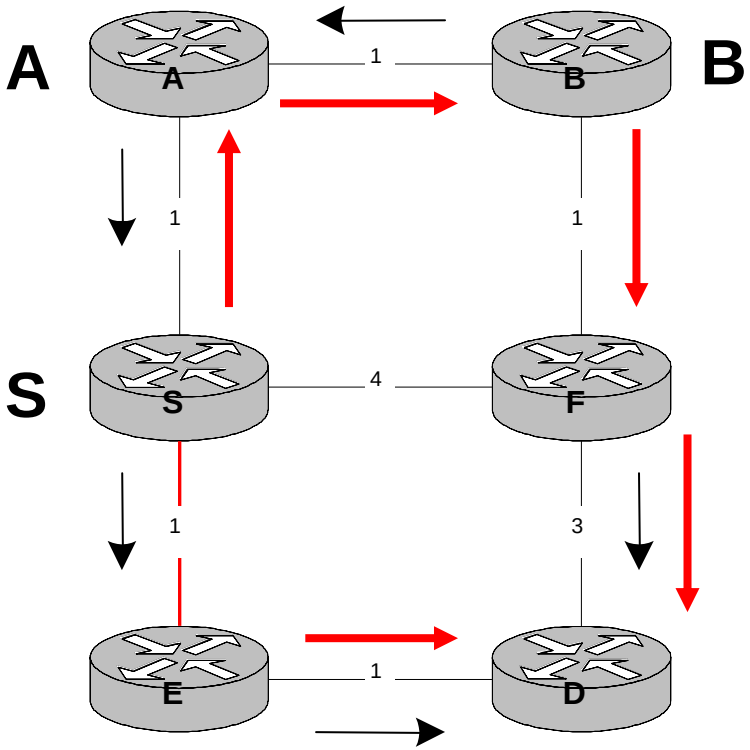
What Are Micro-loops?



What Are Micro-loops?

Paths to D AFTER Failure of S-E

Paths to D BEFORE Failure of S-E



If A changes before B, we have a loop across A-B until B changes

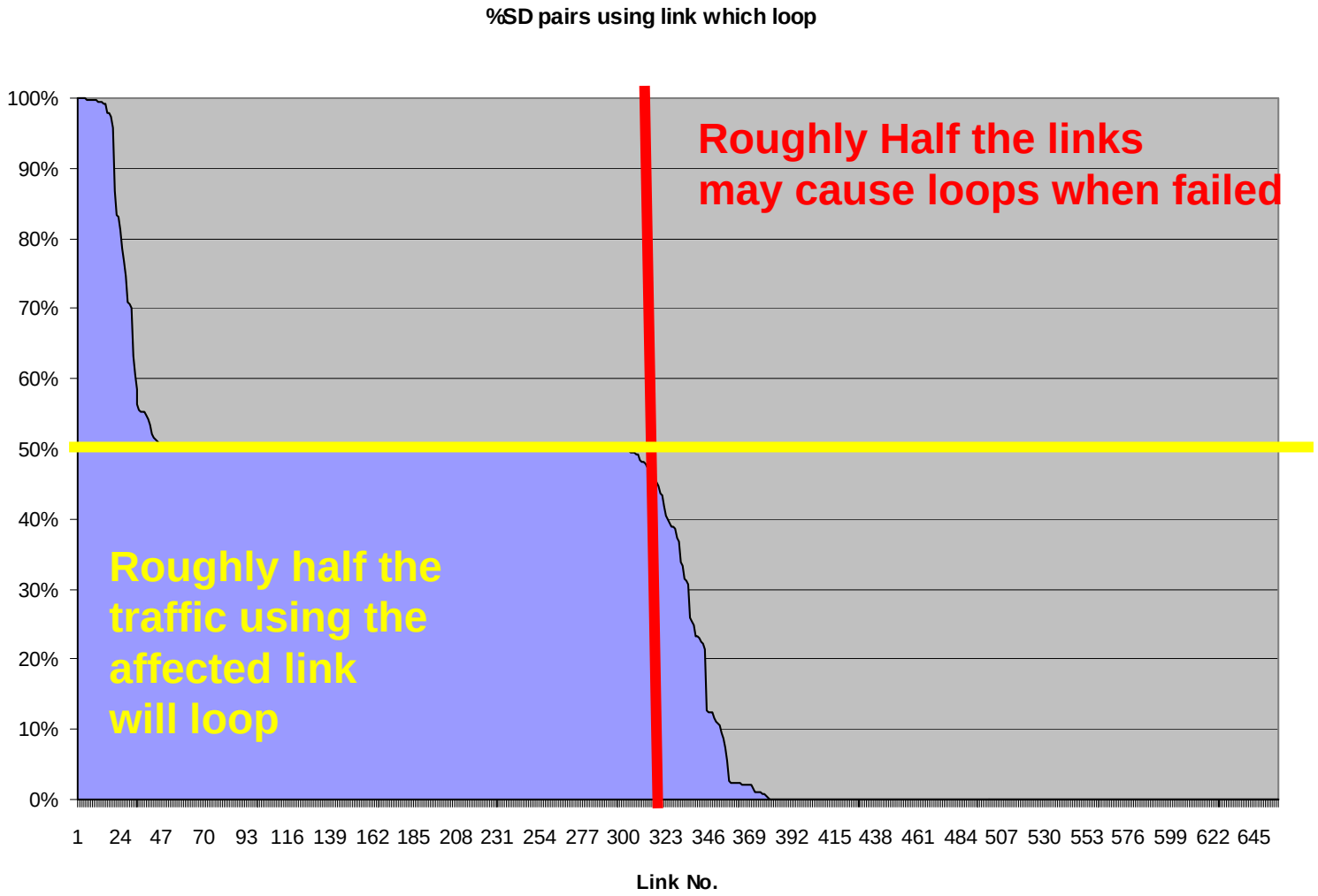
Note that we have a potential loop *anywhere* that the red and black arrows are in different directions.

What Initiates Micro-loops

Micro-loops may be initiated by any action that requires more than one on-path router to change its next hop.

- Component failure (link, node, shared risk link group (SRLG)).
- Component repair (as above).
- Management action to withdraw or insert a component.
- Management change of link cost (either positive or negative).
- External cost change.

SD Pairs using Link that MAY loop



What Protocols are Affected?

- Any IGP and most routing protocols that inherit their path from the IGP.
  - Link-state IGPs
  - Distance vector IGPs
  - Multicast
  - iBGP
  - LDP
- BGP?

What about MPLS?

- RSVP-TE is not affected, because the path is always complete and locked before use.
- MPLS-LDP inherits the path from the IGP and will be affected.
- MPLS-LDP networks that use RSVP-TE one-hop tunnels for path protection are protected from the time the failure is noticed until the start of convergence, BUT that protection is degraded once the convergence process starts.

Micro-loop Strategies.

Micro-loop strategies fall into three basic classes:

1. Micro-loop mitigation
   - Prevent the easy ones – put up with the rest
   - Continue with most forwarding
2. Micro-loop prevention
   - Prevent the formation of ANY micro-loops
   - Continue with all forwarding
3. Micro-loop suppression
   - Stop the co-lateral damage by dropping the affected packets
   - Forwarder recognises and suppresses micro-loop
   - Not considered a viable mechanism

Micro-loop mitigation

- Path Locking with Safe-Neighbors (PLSN)
  <draft-ietf-rtgwg-microloop-analysis-01.txt>
- In summary:
  - On a per-prefix basis, each router determines if the new next hop will loop the packet back (i.e. if a micro-loop will form between the router and its new next hop).
  - If no micro-loop next hop is changed immediately.
  - If a micro-loop will form, the router waits a bit and then changes the next hop.
- Results in a 70% to 95% reduction in micro-loops (depending on topology).
- An improvement, but we can do better if we wish.

Micro-loop Prevention

- Eight micro-loop prevention methods have been proposed
- These fall into two classes
  - Ordered Change
  - Pathlocking
- See <draft-bryant-shand-lf-conv-frmwk-02.txt> and references for more detail

Ordered Change

The router FIBs are changed in such an order that the packet either:

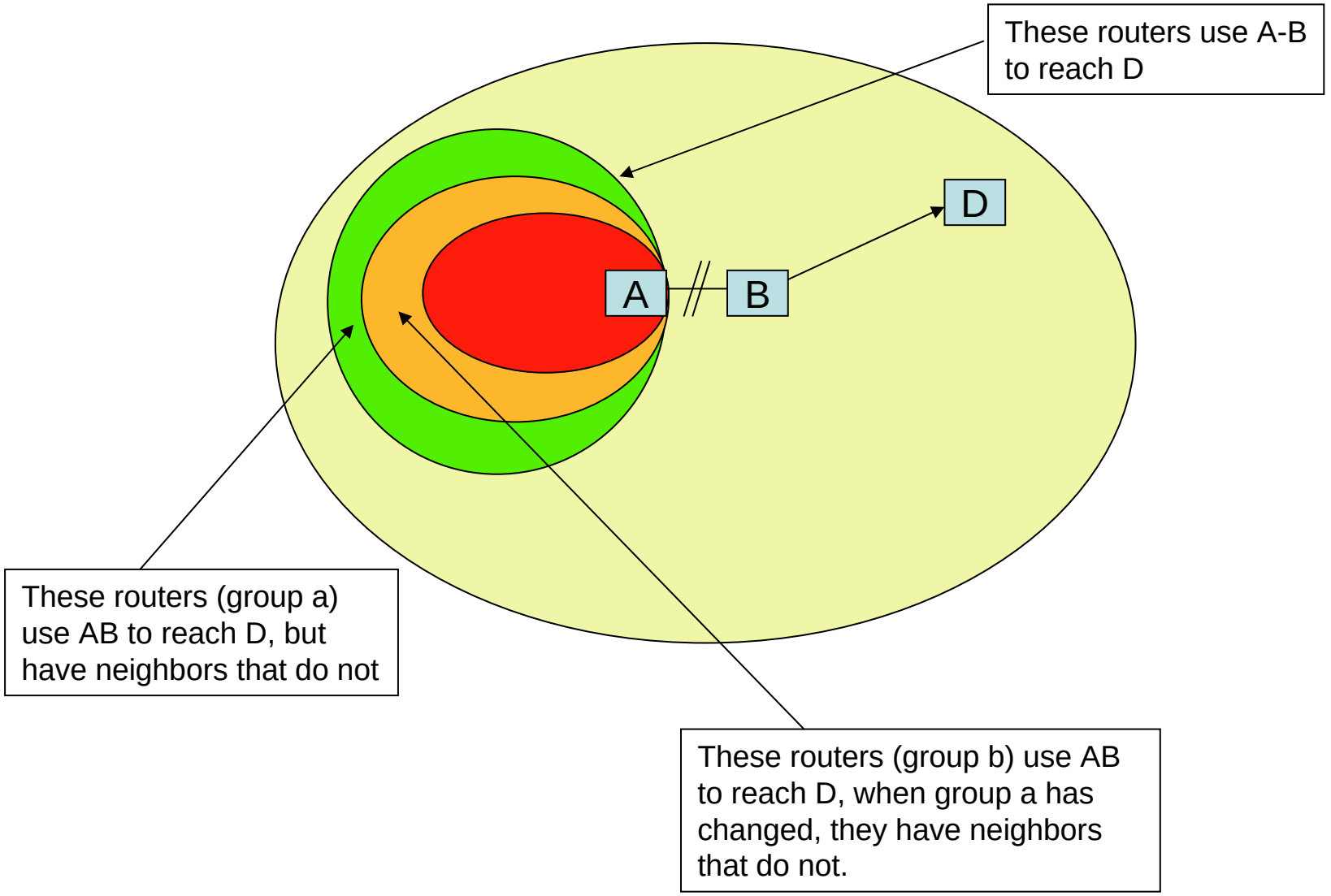
A. Continues to its destination ONLY using the old path

OR

B. Continues to its destination ONLY using the new path

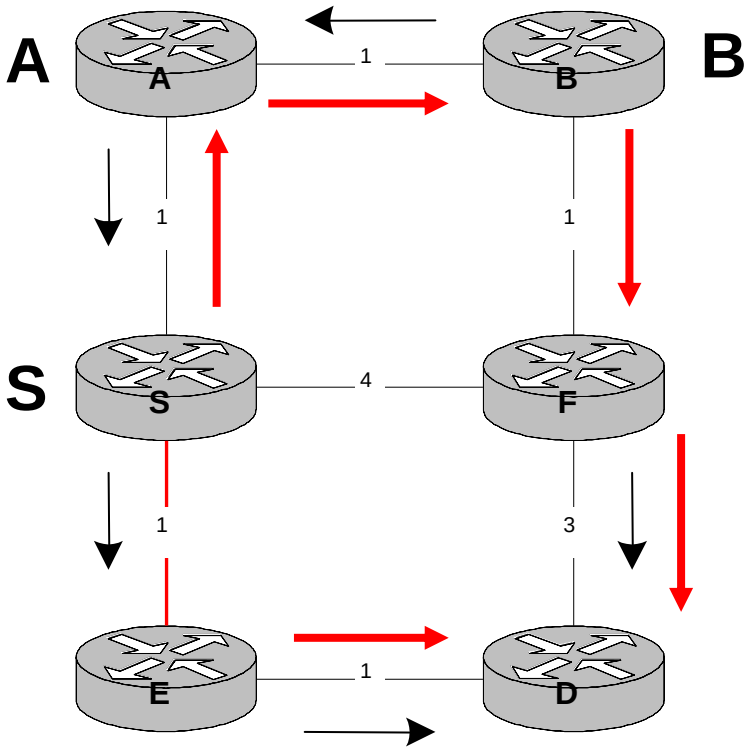
This is illustrated in the following example

Ordered FIB Concept



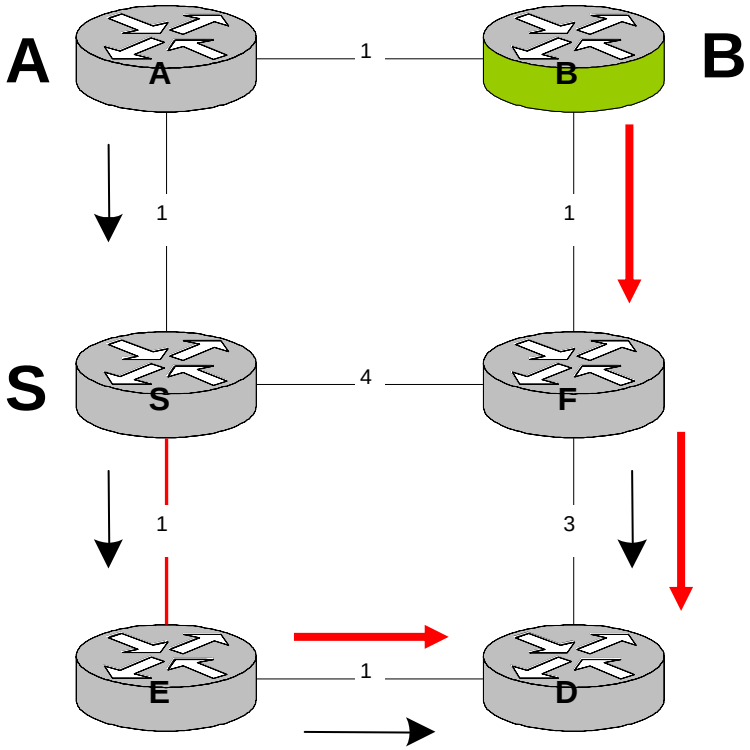
Ordered Change Example

- Ensure the changes are in the order B,A,S



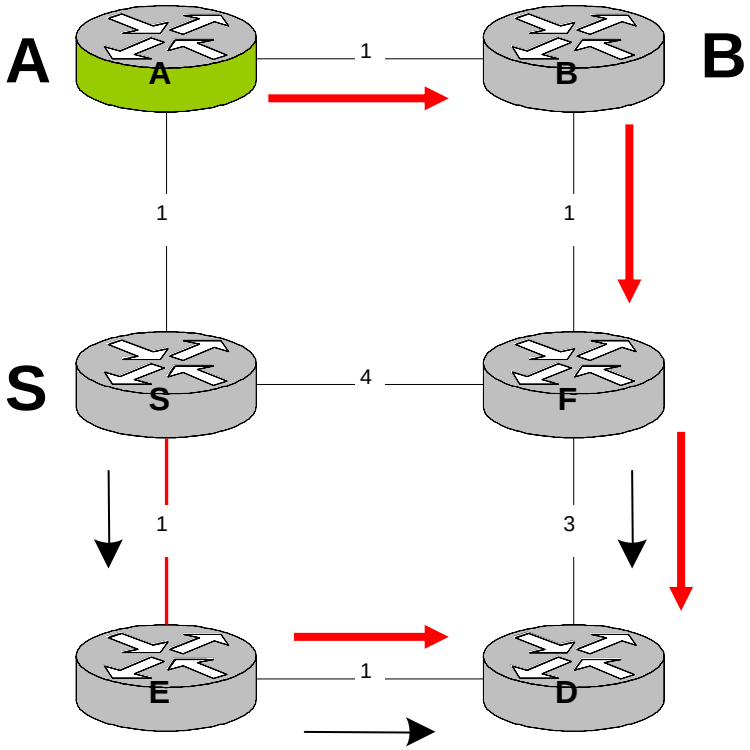
Ordered Change Example

- Ensure the changes are in the order B,A,S



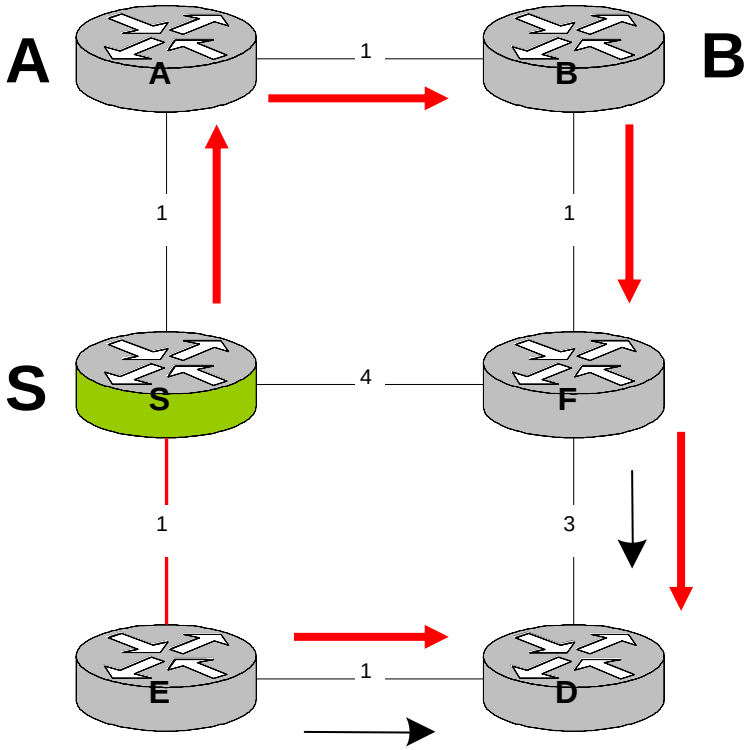
Ordered Change Example

- Ensure the changes are in the order B,A,S



Ordered Change Example

- Ensure the changes are in the order B,A,S



Pathlocking

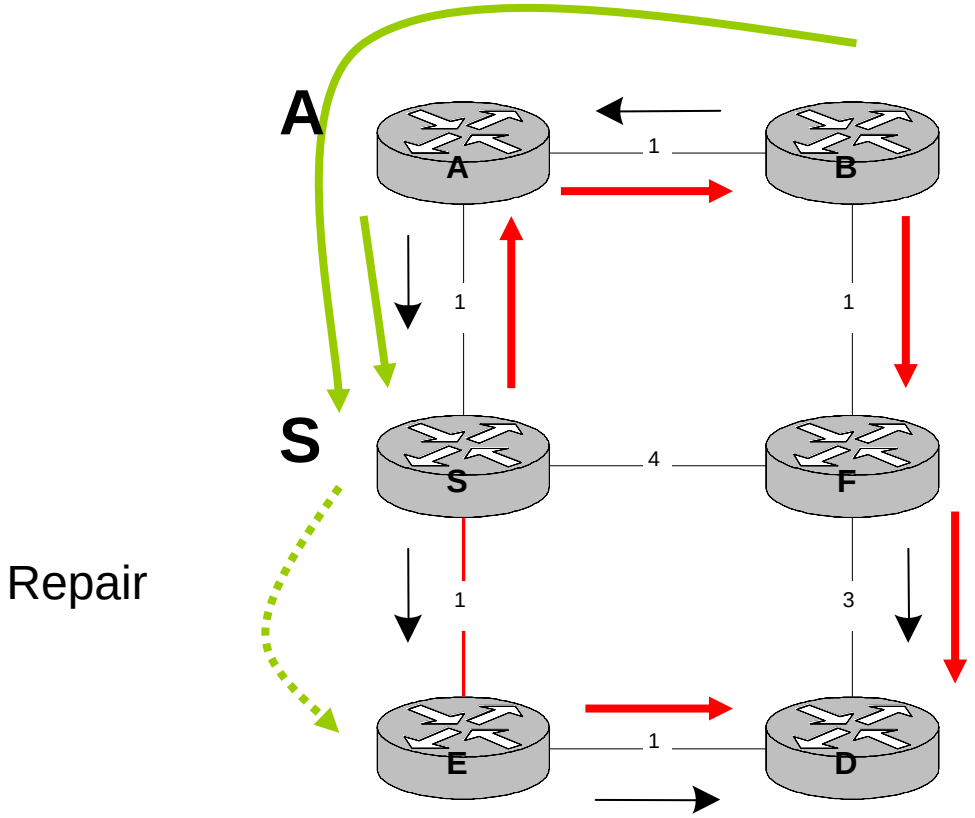
The following *approximately concurrent* changes are made to the router FIBs

A. Destinations affected by the change are identified
B. A virtual topology is created (by tunnels, packet marking or issuing new labels). This virtual topology is unaffected by the change.
C. Packets that might micro-loop are forwarded using the virtual topology
D. When all affected packets are being forwarded via the virtual topology, the real topology is modified
E. When the real topology has been completely modified packets are once again forwarded using the real topology
F. The virtual topology is removed

Note that is never necessary for any individual router to be synchronised with any other router. It is sufficient that *all* routers complete each step before *any* router starts the next step.

Path Locking Example

- e.g. locking to old topology



Packets from B to S unaffected by A's FIB change
(old route to S is same as new route to S)

Impact on Convergence

- The determination of the routes is unaffected – the routing protocols still use the **same algorithms** etc, with the **same correctness guarantees**.
- Micro-loop prevention alters the order in which the FIBs are updated.
- All micro-loop prevention mechanism can revert to normal convergence if needed.
- All micro-loop prevention mechanisms slow convergence – some more than others.
- However optimised solutions are likely to result in sub-second convergence times in most cases.

Slowing Convergence

- Slowing convergence is not important when the change is as a result of management.
- When the event is “good news” the new component can be used immediately.
- When the event is “failure”
  - Packets that are being repaired get better service than with traditional convergence
  - Packets that cannot be repaired get much worse service than with traditional convergence
- The duration of the convergence process and the repair coverage are important considerations for fast re-route.

Mix and Match

- Chosen method needs to be common across the routing domain.
- To change the method needs a flag day.
- May be able to use one method to enhance another BUT this needs careful study.
- Therefore the choice of method must be made with a view to future network requirements.

Conclusion

- Work on micro-loop prevention is being carried out in the RTGWG.
- The current focus is as a component of IP Fast Re-route, and the composition of the interest group reflects this.
- However micro-loop prevention has wider implications and applicability to the work of the IETF.
- Having one solution that addressed the complete set of needs would be a good thing.

- The purpose of this draft is to drawn the attention of the IETF to this work and encourage greater participation in the formation of suitable solutions.

References

- draft-bryant-shand-lf-applicability-01.txt
- draft-bryant-shand-lf-conv-frmwk-02.txt
- draft-ietf-rtgwg-microloop-analysis-01.txt
- draft-francois-ordered-fib-01.txt
- draft-bonaventure-isis-ordered-00.txt
- draft-atlas-bryant-shand-lf-timers-01.txt