Telcordia

the elements of success

Loop-Free IP Fast Reroute Using Local and Remote LFAPs

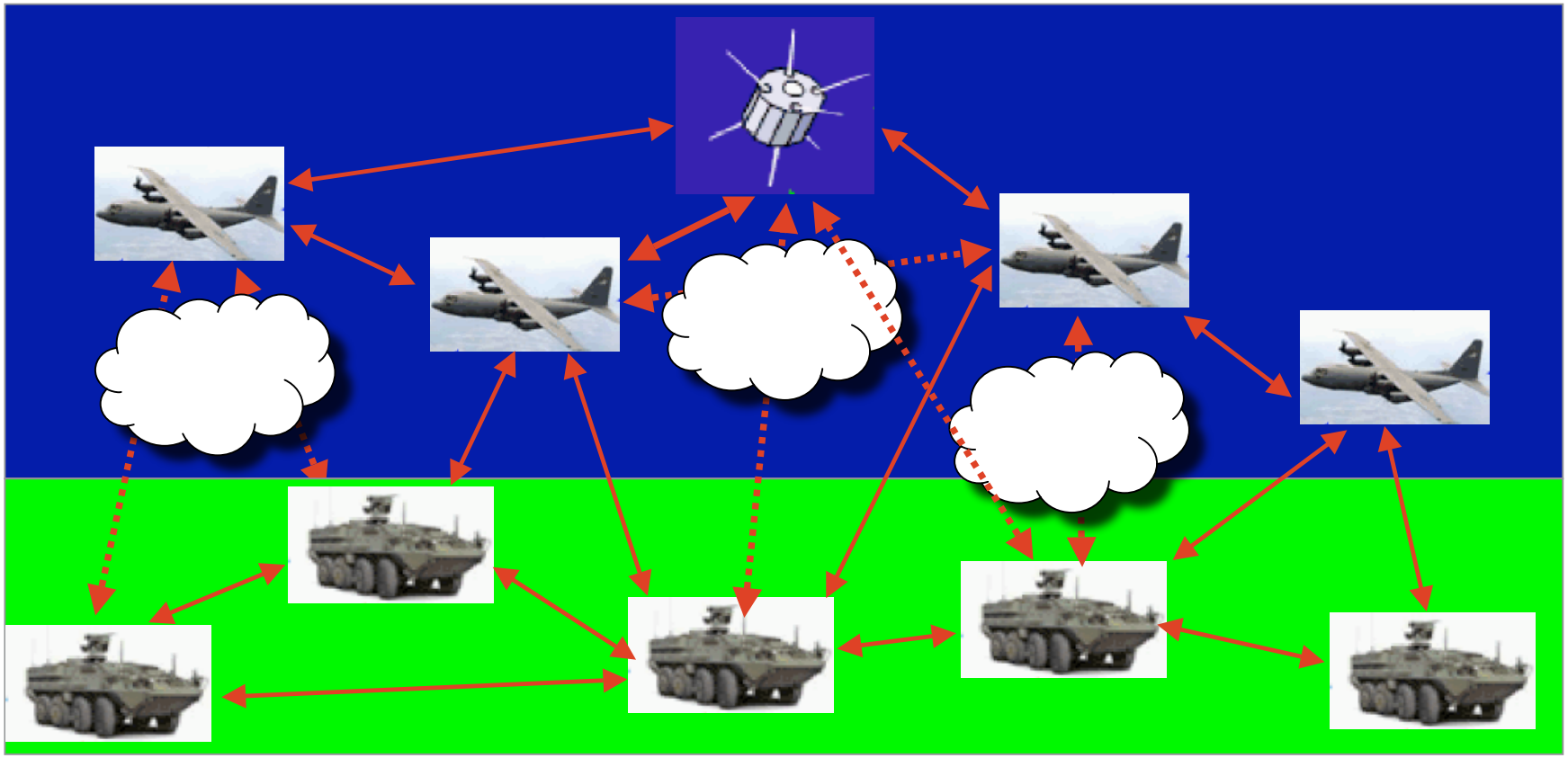
I. Hokelek, M. A. Fecko, P. Gurung, S. Samtani

Contact:
Ibrahim Hokelek
Applied Research
Telcordia Technologies Inc.
ihokelek@research.telcordia.com
732-699 3905
07/25/2007

Introduction

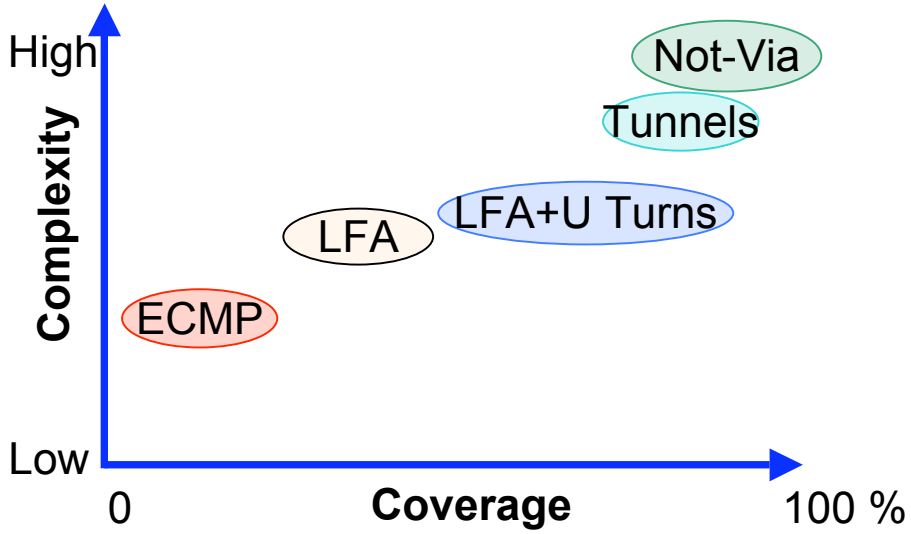
- **Goal:** Minimize disruption time for user applications in the event of node and link failures in the network
  - Low latency applications are emerging: *Voice over IP, VPN, IP TV*
  - Ubiquitous networking (anytime, anywhere, any device) focuses on access points: *WiMax (IEEE 802.16), Media Independent Handover (IEEE 802.21)*
  - Future IP backbone should be ready for ubiquitous networking paradigm
- **Why IETF draft?**
  - The existing IETF IPFRR drafts were useful for us and we would like to contribute back to IETF with our new enhancements
  - Army acquisition programs are planning to use COTS routers
  - Our goal is to engage our fast reroute protocol with the existing link state routing protocols and to support its standardization for military usage

Wireless mobile IP backbone



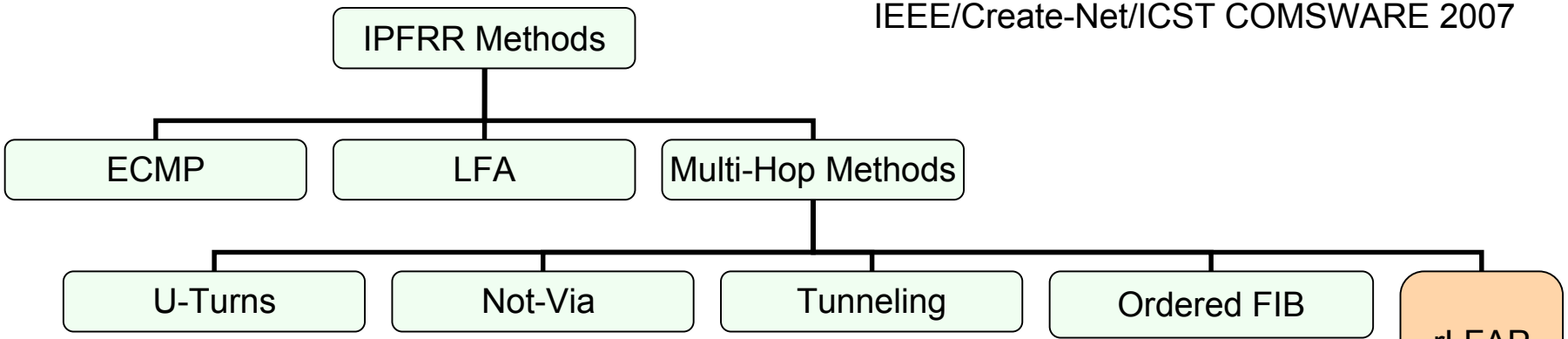
- Multiple different link types: terrestrial, air, or satellite
- Multiple simultaneous failures due to mobility, weather conditions, and peculiarities of wireless medium
- E.g., clouds may prevent free space optical link availability

Overview of IETF IPFRR Techniques



- Performance metrics
  - Handoff delay
  - Packet delivery ratio
  - Repair path **coverage** and quality
  - Ability to handle micro-loops
  - Additional signaling overhead
  - Backward compatibility
  - **Complexity** (processing/memory)

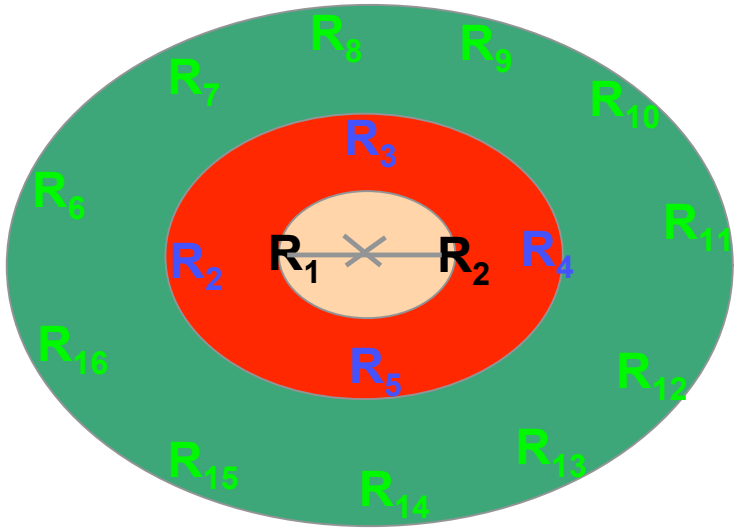
REFERENCE: M. Gjoka, V. Ram, and X. Yang, "Evaluation of IP Fast Reroute Proposals", in IEEE/Create-Net/ICST COMSWARE 2007



Telcordia

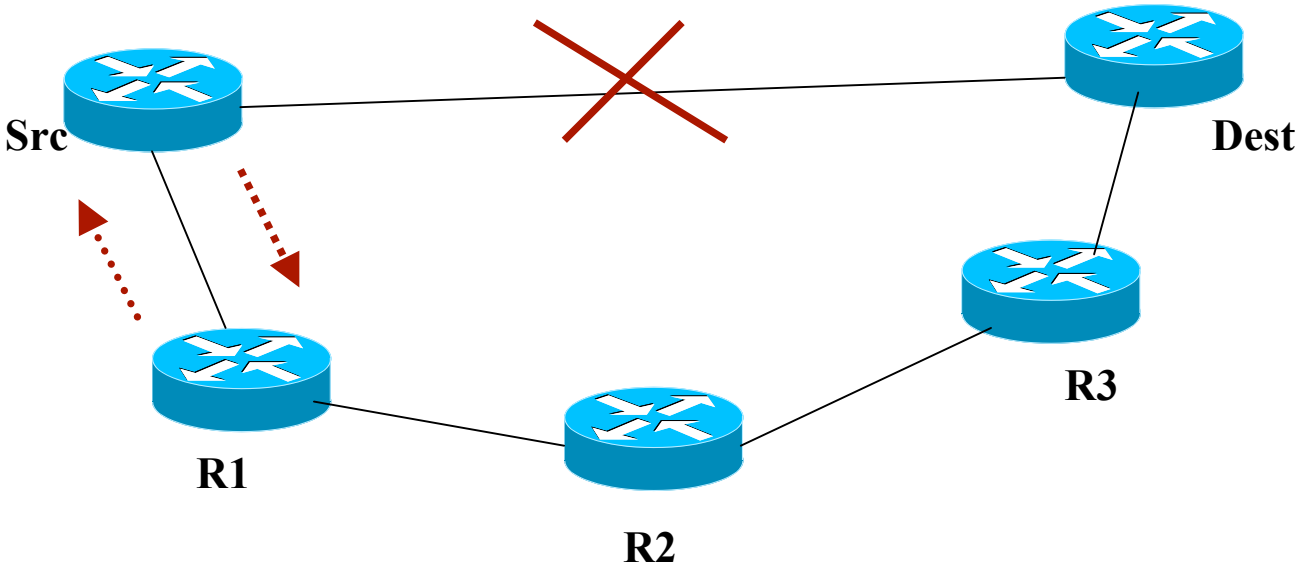
rLFAP Motivation

REFERENCE: K. Lakshminarayanan, et. al., "Achieving Convergence-Free Routing using Failure-Carrying Packets", to appear ACM SIGCOMM 2007



- Only subset of network is impacted by a failure
- Loop-free alternates are suitable choice
  - No micro-loop
  - No signalling required among routers
- *Pre-computed local LFAPs* by routers **R_1** and **R_2**
- *Pre-computed remote LFAPs* by routers in red region **R_2-R_5**
- *Instantly calculated remote LFAPs* (if needed) by routers in green region **R_6-R_{16}**

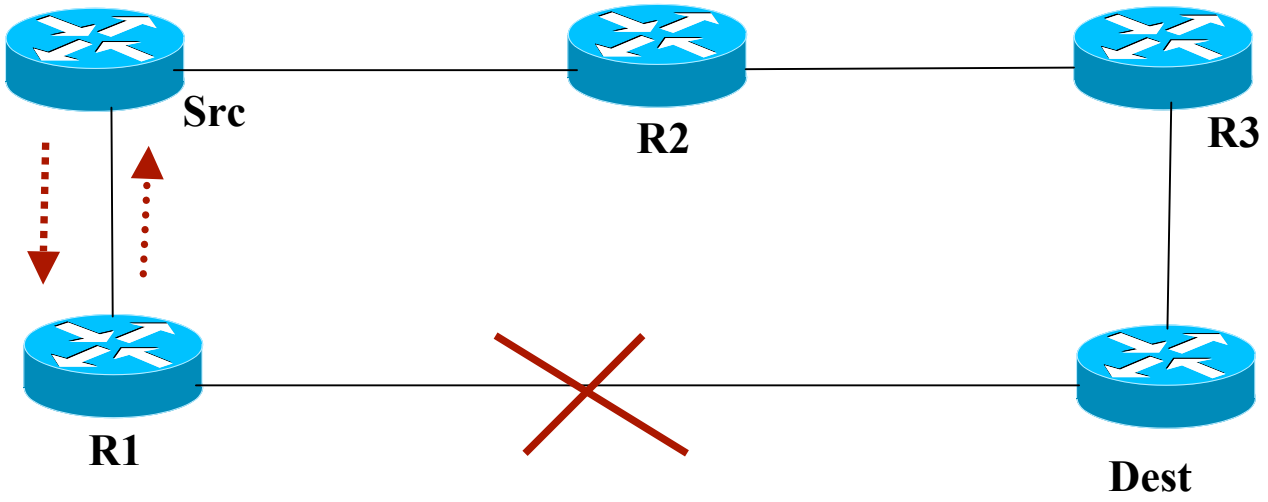
Micro-loops



- Micro-loops (i.e., routing transients) happen due to inconsistent topology view of different routers
  - Failure detection (**R1 might haven't received link failure info yet**)
  - LSP flooding and SPF triggering
  - SPF tree computation and RIB update
  - FIB computation and distribution to the Line cards (**Src is done**)
- Micro-loops can appear without IPFRR
- With IPFRR, micro-loops are more apparent

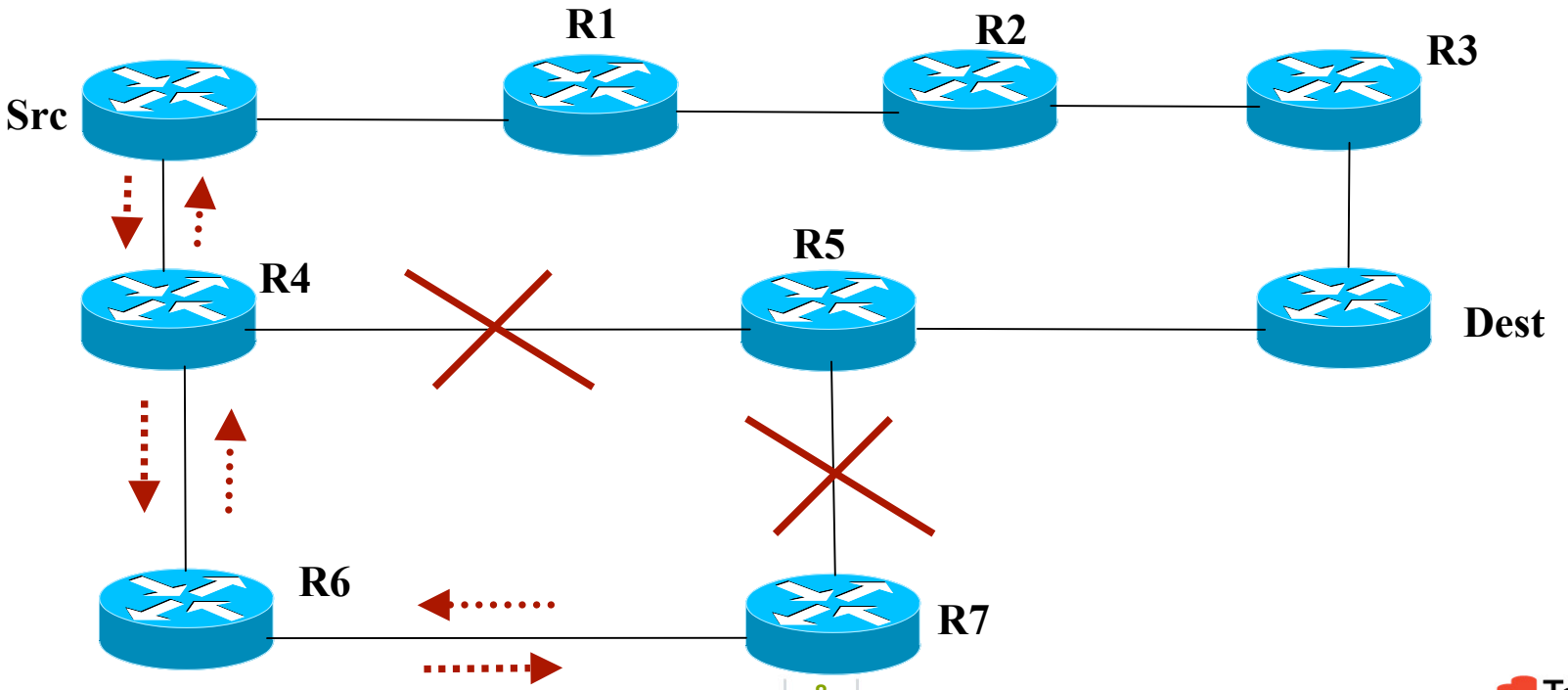
rLFAP Feature-I

- Prevention of micro-loops using remote LFAPs
- Link metrics are unity (hop-count)
- Primary path from **Src** to **Dest** is **<Src-R1-Dest>** before failure
- Instant failure propagation to multi-hop neighbors (1-hop)
- Each node maintains sets of local and remote LFAPs
- If **Src** receives the failure info, then remote LFAP **<Src-R2-R3-Dest>** can be activated



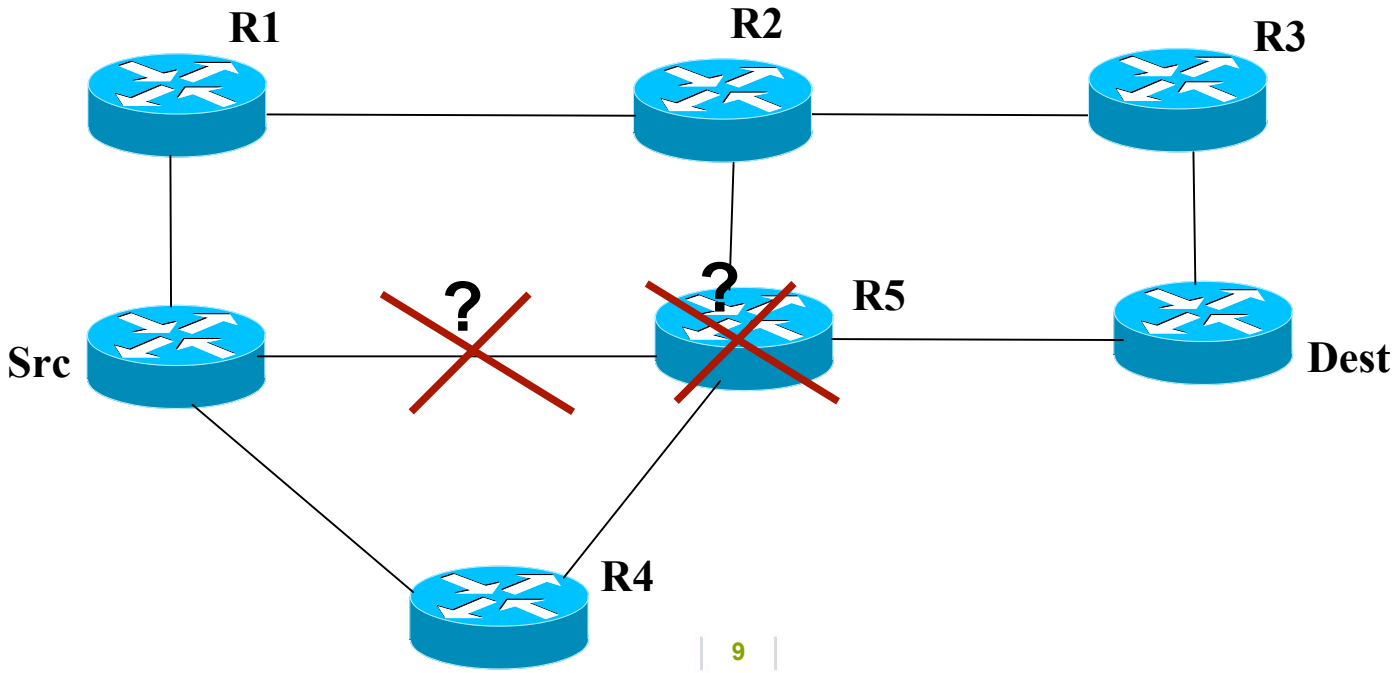
rLFAP Feature-II

- Handling multiple simultaneous failures
- Existing IETF drafts target only single failure
- Primary path from **Src** to **Dest** is **<Src-R4-R5-Dest>** before failures
- Instant failure propagation to multi-hop neighbors (3-hop)
- If **Src** receives both failure info, then remote LFAP **<Src-R1-R2-R3-Dest>** can be activated



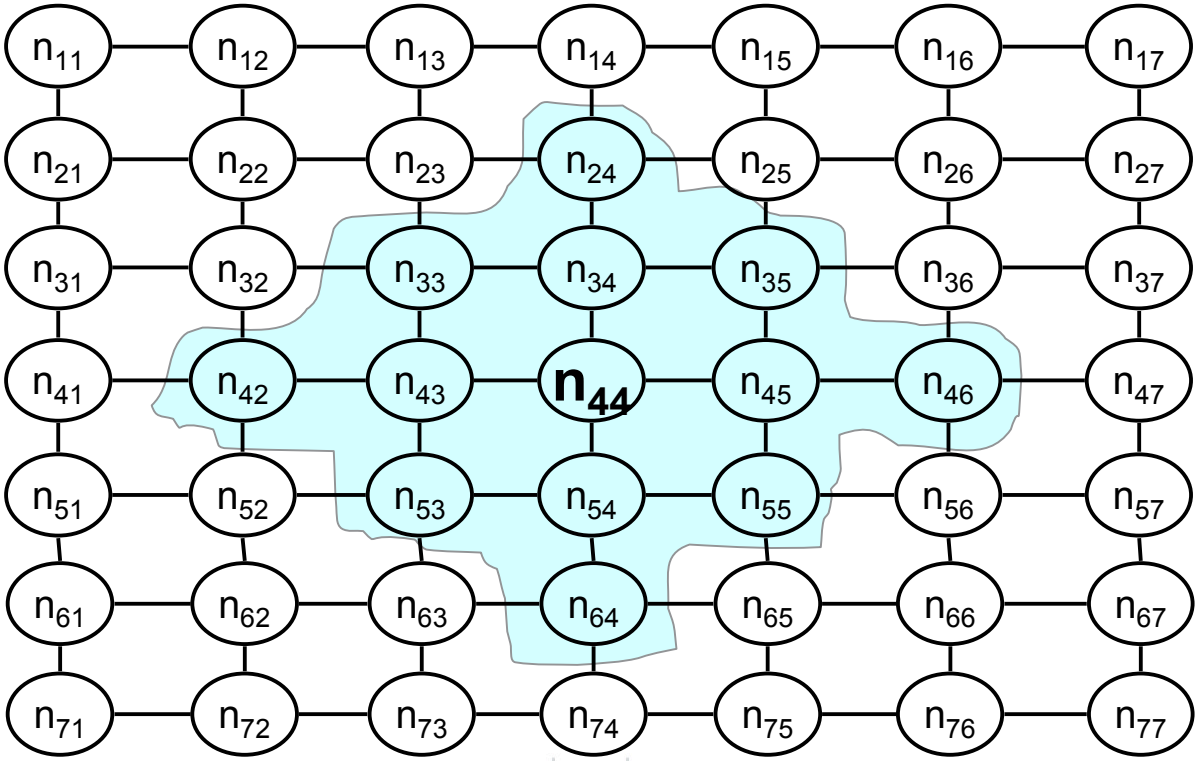
rLFAP Feature-III

- Distinguishing link and node failures
- Primary path from **Src** to **Dest** is **<Src-R5-Dest>** before failure
- **Src** cannot communicate with **R5**, Is it node or link failure?
- If link failure, local LFAP **<Src-R4-R5-Dest>** will be activated
- If node failure, local LFAP **<Src-R1-R2-R3-Dest>** will be activated
- **Src** receives failure info from other interfaces if **R5** is failed



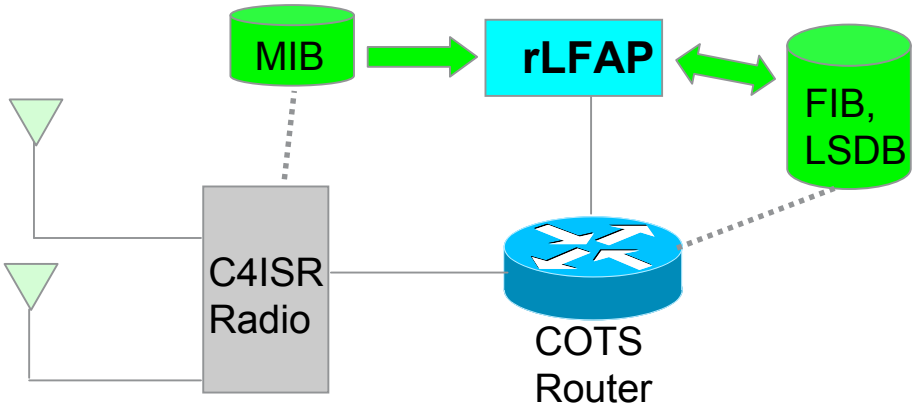
Multi-hop Neighborhood

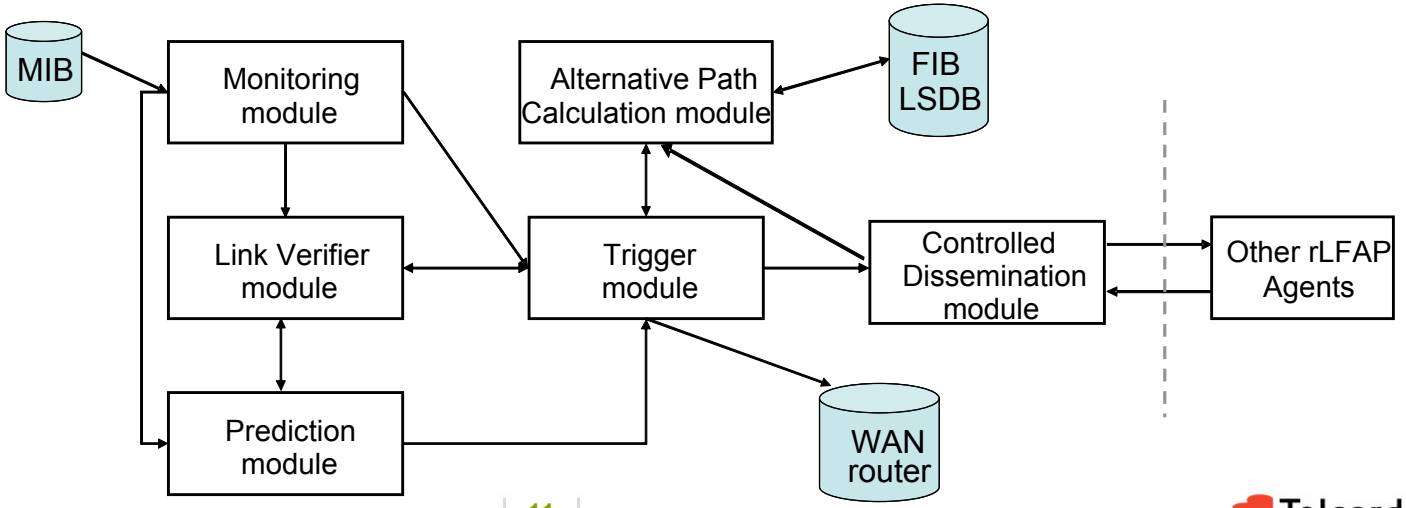
- **Node i**'s 2-hop (**X**-hop) neighborhood (NBH) includes 12 nodes and 36 links
- **Node i** pre-computes local LFAPs against its 4 local interfaces and remote LFAPs against 12 remote node and 32 remote link failures
- **NBH** defines only a local scope for each node so no additional mechanism is needed to maintain it
- **KEYLINKS** for scalability



rLFAP Architecture

- **rFAP** is running on general purpose processor (GPP)
- Link failure prediction and detection by Trigger, Monitoring, Prediction, and Link Verifier modules
- Alternative Path Calculation module pre-computes LFAPs and provides switchover mechanism between OSPF paths and LFAPs
- Controlled Dissemination module collects and disseminates failure information among multi-hop neighbors





Telcordia

Fast Failure Notification-I

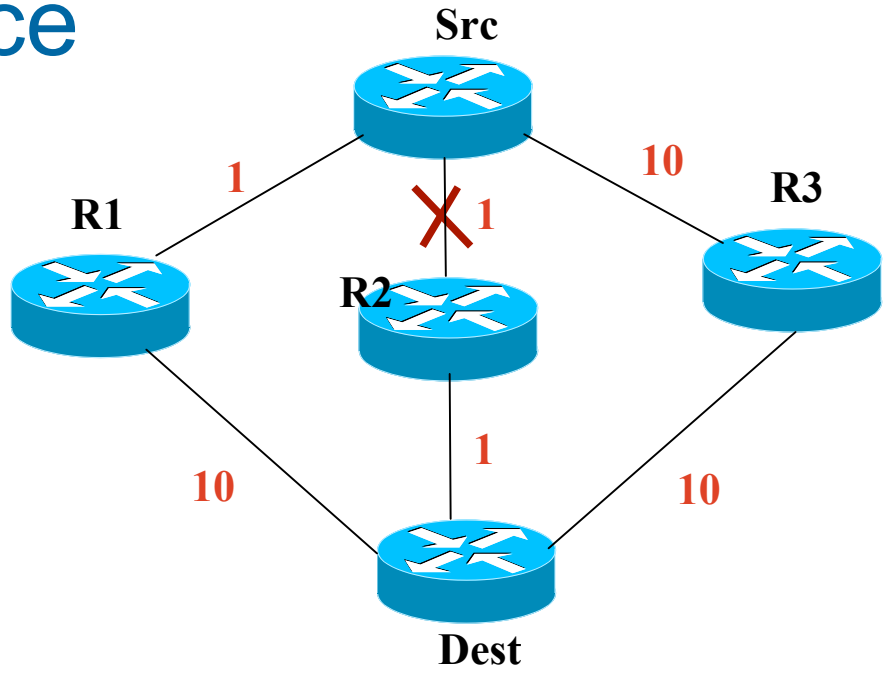
- Option 1: Configuring routing protocol's parameters
  - Link state update packets are LSAs in OSPF and LSPs in IS-IS
  - When topology changes, a new LSA is originated
  - Two instances of the same LSA may not be originated within **MinLSInterval**
  - Literature shows that **MinLSInterval** around 20-30 ms does not generate much overhead while providing fast failure propagation to multi-hop routers
  - ***Disadvantage:*** Enormous routing overhead if failures are transient (i.e., link flapping) since each LSA is flooded to the entire network

Fast Failure Notification-II

- Option 2: An Efficient Failure Flooding Mechanism for SHA
  - New link update packet (LUP) similar to LSA but two new fields:
    - **Time-To-Live (TTL) :** # of hops a new LUP will be transmitted
    - **Stop-Flooding (SF):** Indication for flooding beyond NBH
  - Initially set **TTL** to **X** and **SF** to 0 (X defines the scope of neighborhood (NBH), i.e., X-hop NBH)
  - When a new LUP is received, each router will
    - decrement **TTL** by 1
    - set **SF** to 1 only if all destinations are covered by LFAPs
    - continue flooding if TTL >0 or SF=0
  - Advantages:
    - Minimal routing overhead since flooding LUP are limited to NBH
    - Fast failure notification since the parameters of the flooding procedure (e.g., timers) are independent of routing protocol parameters

Routing Convergence

- IGP convergence from alternate paths to optimum new routes should also be loop free



- Primary path is **Src-R2-Dest** before the link between **Src** and **R2** fails
- rLFAP switches over LFAP **Src-R3 Dest** after failure detection
- New primary path is **Src-R1-Dest** after the routing convergence
- When **Src** will switch back to its new primary route?
  - With rLFAP, **Src** can immediately switch to its new route since **R1** uses LFAP **R1-Dest**
  - Without rLFAP, **Src** should wait to make sure that **R1** is also converged to its new route

Thanks for your patience!

Q&A