An Empirical Study of Router Response to Large BGP Routing Table Load

Di-Fa Chang

Ramesh Govindan

John Heidemann

USC/Information Sciences Institute

Motivation

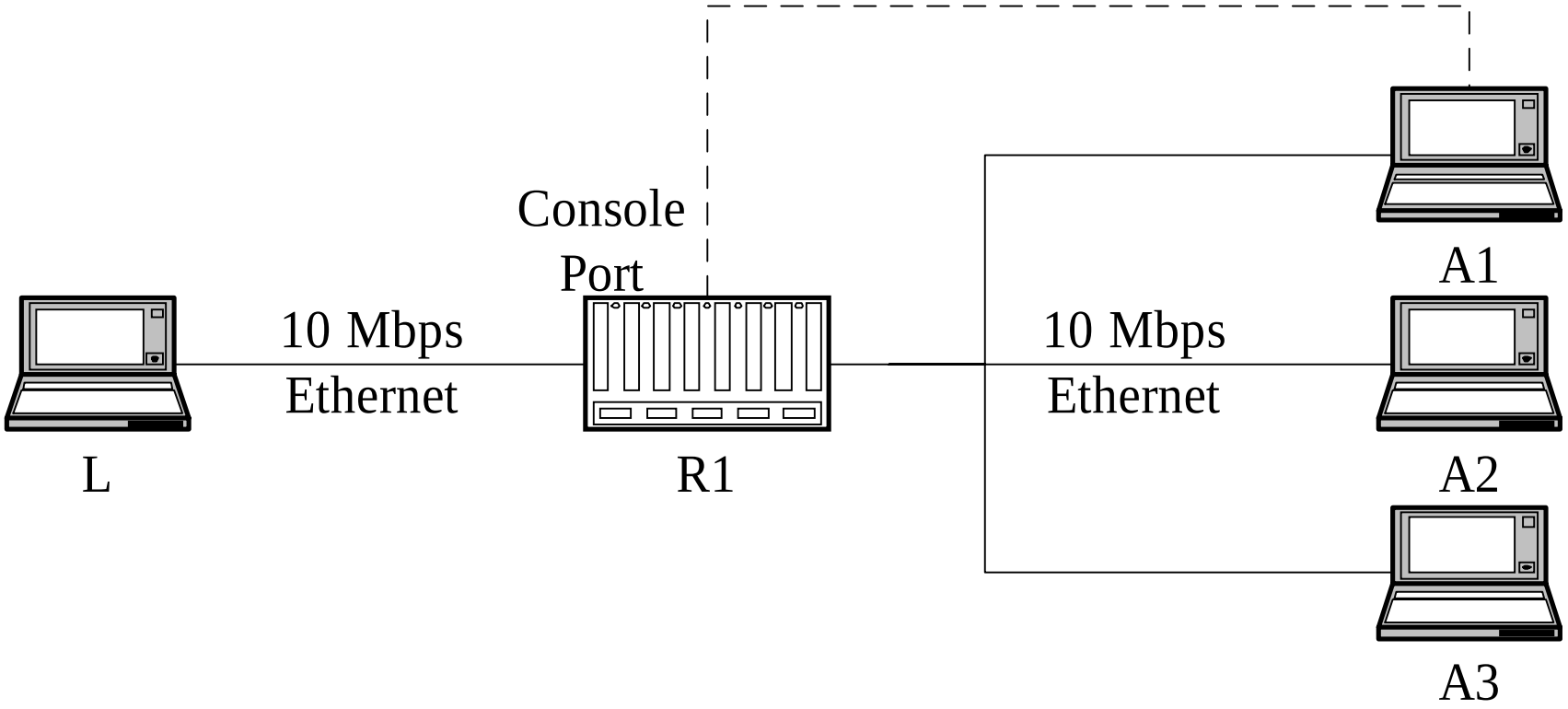
- Anecdotal evidence suggests that misconfiguration of backbone routers occasionally leads to an injection of large routing tables into the BGP routing system.
- But, current BGP protocol does not specify how to handle the routing table overloading.
- Also, no studies in literature present what is the result when routers are overloaded with a large routing table.
- Question: How do commercial routers deal with this failure?

Purpose of this work

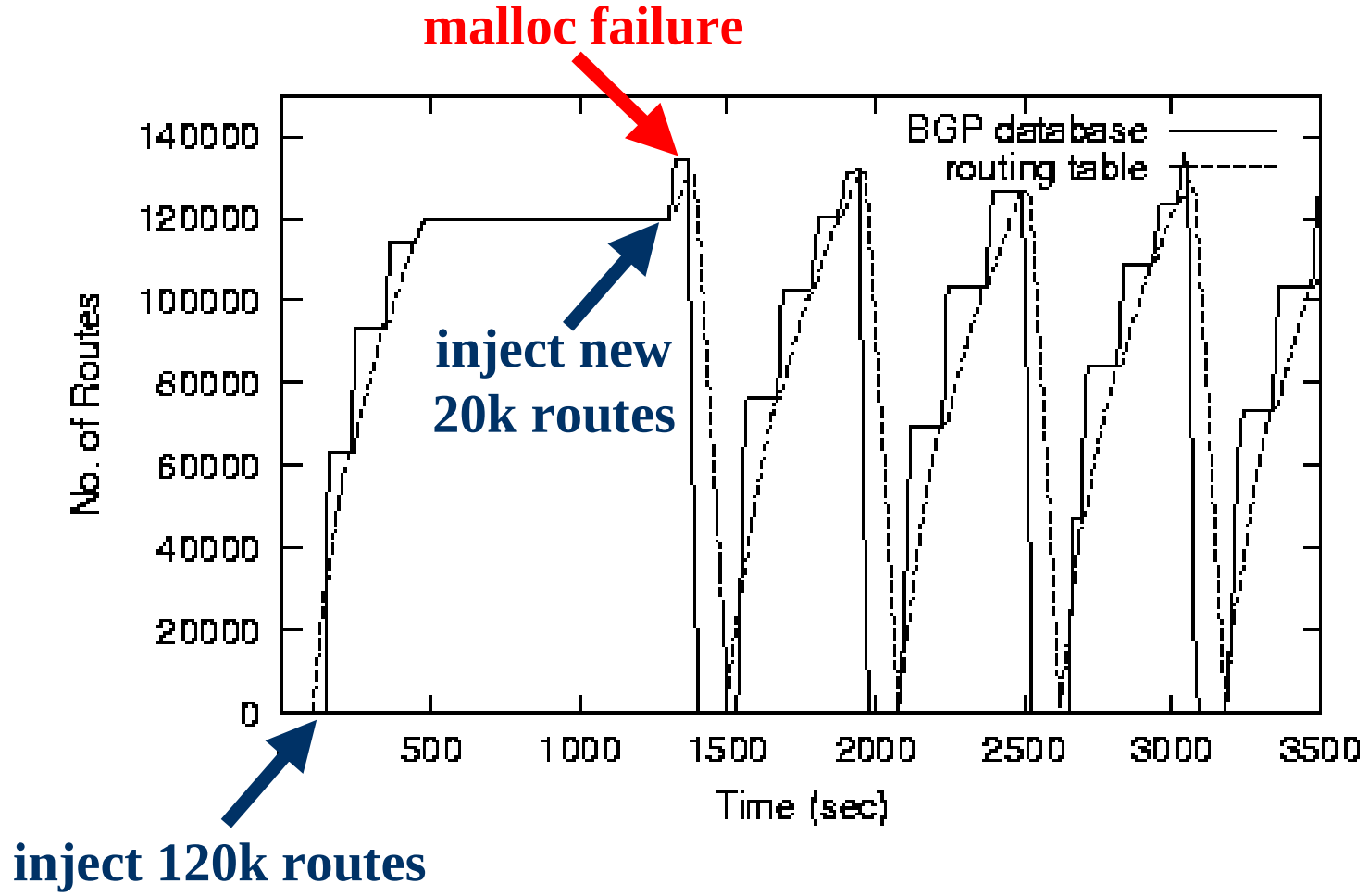
- Examine router implementations (both old and new) to understand how they respond to routing table overloading.
- Study the impact of the response on neighbors.
- Study the efficacy of existing mechanisms for controlling the response, e.g., prefix limiting and route flap damping.

Experiment I

- Test single routers without controlling mechanisms
  - To understand the default error handling behavior.



Cisco 7000 under “malloc failure”



Results

- Cisco 7000 with IOS 11.1
  - Repeated soft reset and routing table oscillation.
- Cisco GSR 12008 with IOS 12.0
  - Freeze the network interface.
- Juniper M20 with JUNOS 4.3
  - Freeze the network interface.
- Juniper M20 with JUNOS 4.4
  - Possibly degrade the forwarding performance when forwarding table is overloaded.
- Conclusion: default response is “Let the operator handle it.”

Experiment II

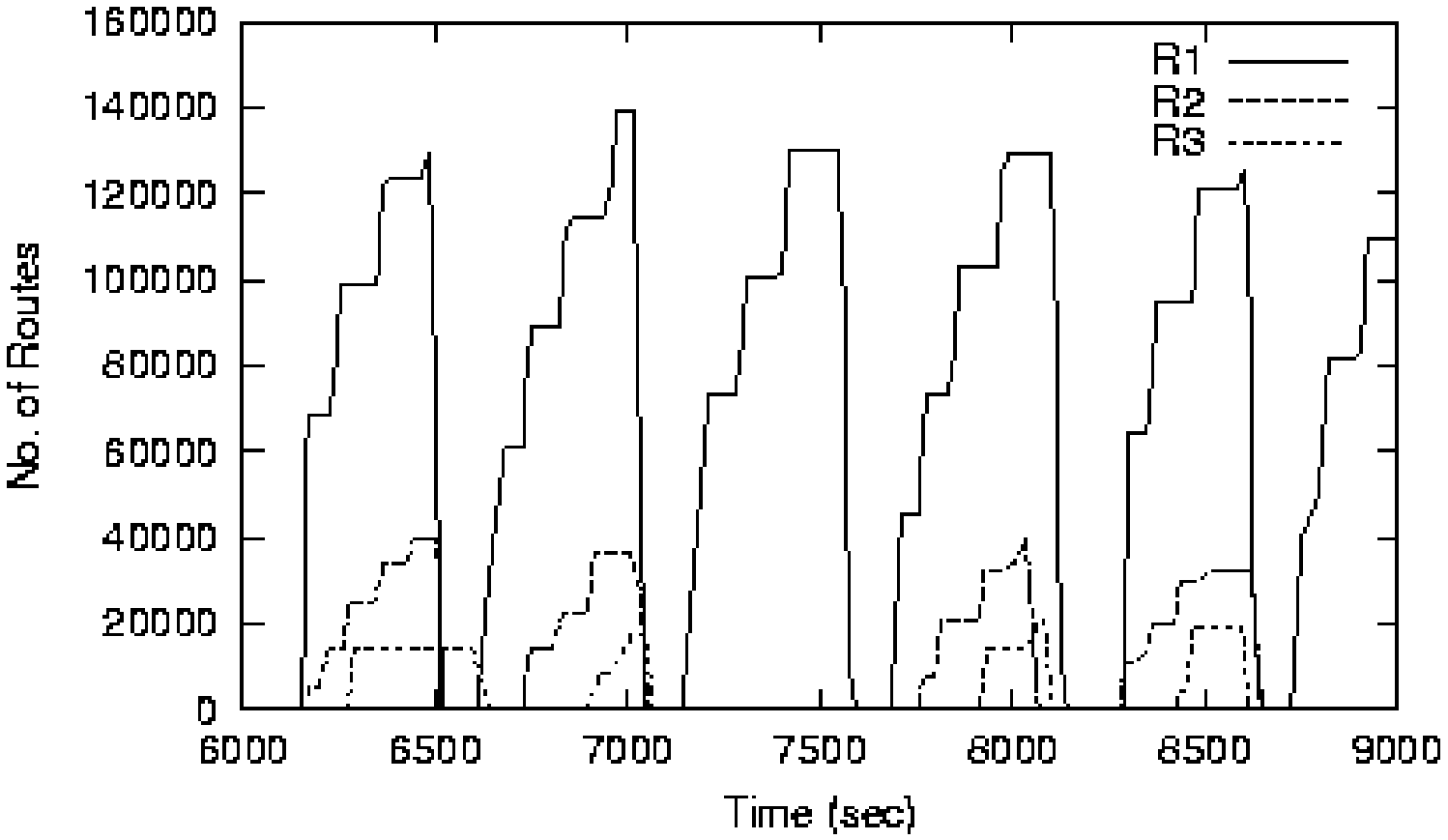
- Test a chain of routers
  - To study the impact on neighbors.

Results

- Old router OS (IOS 11.1)
  - Routing table oscillation can propagate to neighboring routers.
- New router OS (IOS 12.0, JUNOS 4.3)
  - Freezing interface has the same effect as link failure.
  - In other words, a failure caused by one routing protocol (BGP) can terminate the operation of other protocols.
- Conclusion: In some situation, the failure can cascade.

Routers with decreasing capacities

- To mimic a misconfiguration in a large ISP, and its effects on customers and smaller downstream ISPs.



Experiment III

- Test implemented controlling mechanisms:
  - Prefix limiting
  - Route flap damping

Results of Prefix Limiting Experiment

- Prefix Limiting: A mechanism that places a configurable limit on the number of prefixes that a router will accept from a given peer.
- When the limit is reached
  - Juniper M20: resets the peering session, hence results in routing table oscillation.
  - Cisco GSR: denies the peering permanently and requires manual reset.

Results of Route Flap Damping Experiment

- Flap Damping: A mechanism that limits the rate of propagation of unstable routing information.
- This can alleviate the impact of repeated resets on neighbors.
- But, due to timing of peering resets, sometimes leaked routes are not damped.

New Mechanisms Not Tested

- Graceful Restart
  - `draft-ietf-idr-restart-03.txt`, April 2002
- Peer Restart Backoff
  - `draft-hares-bgp-backoff-00.txt`, June 2002
- Both can alleviate the impact of repeated resets to neighbors.
- But, they don’t protect the router from failure. The overloaded router still can not forward traffic.

Conclusion

- Routers in other ASes can crash my router.
- My router has no ways to *completely* prevent it.
- Recovery needs operator intervention and inter-ISP negotiation.
- In all failure modes, the overloaded router loses connectivity to most prefixes.

Future Work

- Design a degradation mode, e.g.
  - Dynamically filter less significant routes.
- Degradation means:
  - Maintain connectivity to important destinations, e.g., DNS servers, popular websites.
  - Maintain connectivity to as many destinations as possible, i.e., prefixes with shorter length.
  - Lose connectivity to other destinations.