IP ROUTING

BGP Collection

- Zebra BGP collectors
  - 3 PoPs (including traffic collection)
  - iBGP RRC & eBGP sessions
- Focus on iBGP updates
  - How data packets exit Sprint
  - Reflects eBGP updates, internal policy & IGP changes

Traffic Analysis Method

- Correlate iBGP & traffic
  - Find egress PoP for each data packet
    - Longest prefix match, router to PoP map
    - Ingress link to multiple egress PoP fan-out
  - Identify traffic variability due to BGP updates
    - Static BGP table + data packets
      - Shows variability due to other factors
    - Dynamic BGP table + data packets

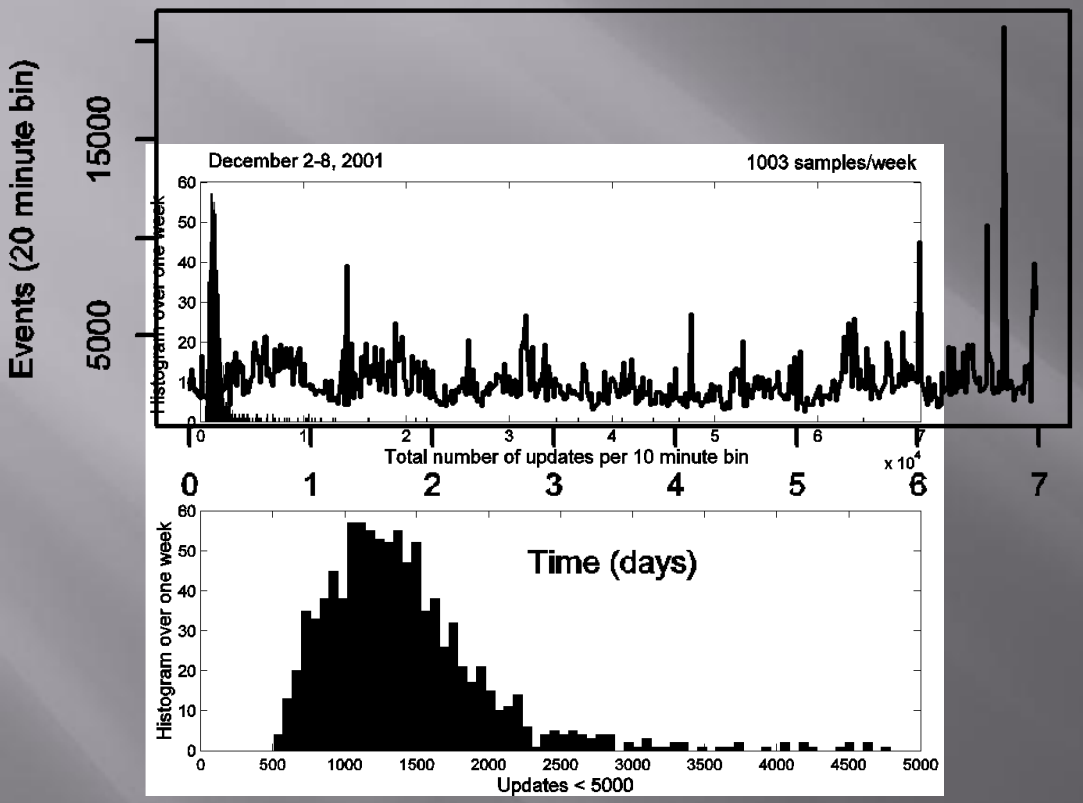
Overview

- Problem statement
- Methodology
  - Data collection & analysis
- Results
  - Likely causes
- Conclusions & implications

Results Presented

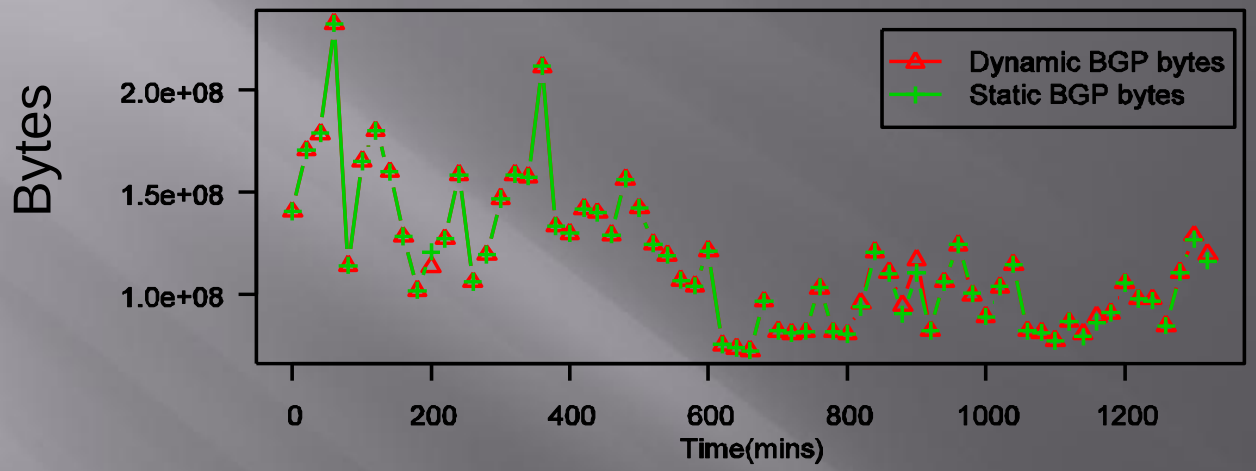
- 22 hour packet trace from Aug 6 2002
  - 112 Mbps average link utilization
- Traffic fan-out approximations
  - 2,649,315,251 packets
  - BGP table calculated every 20 minutes
  - Addresses carrying 99% of traffic
    - ~30,000 destination addrs of 200,000

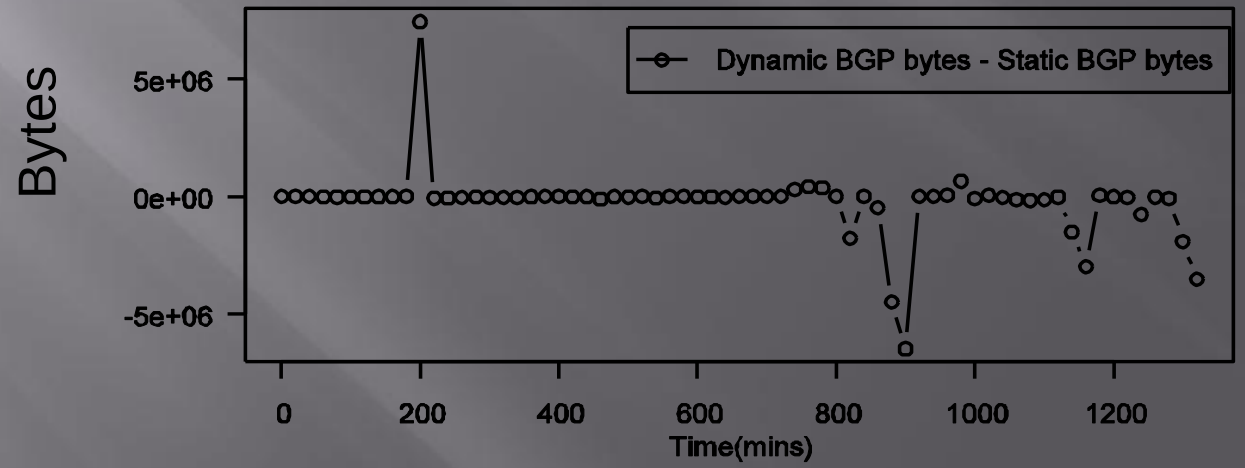
Volume of BGP Updates



- iBGP
  - Mean 1330/10min
  - Max 93320/10min
- Spikes
  - Maintenance?

Variability in Traffic to an Egress PoP

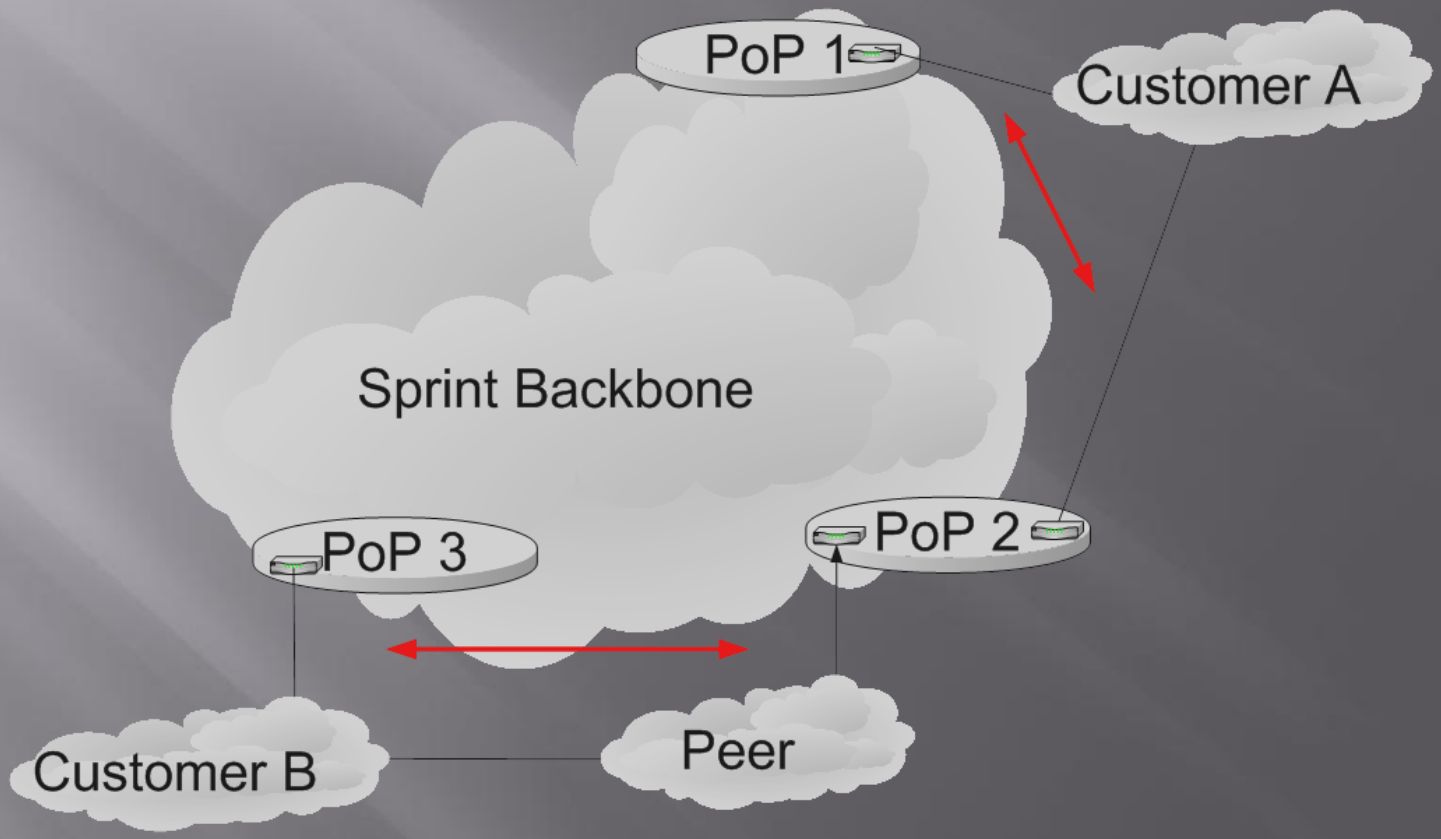




Summary of Results

- 1~3% of variability in traffic to all PoPs
  - Representative of other traces
- Specific sources of variability
  - Networks with multiple links/paths to Sprint
  - BGP updates cause shift between inter-AS paths
    - Causes shift between intra-Sprint paths

Case Study



Few ASes Involved

- Single next-hop AS in 47% of all traffic shifts
- Single last-hop AS in 46% of all traffic shifts

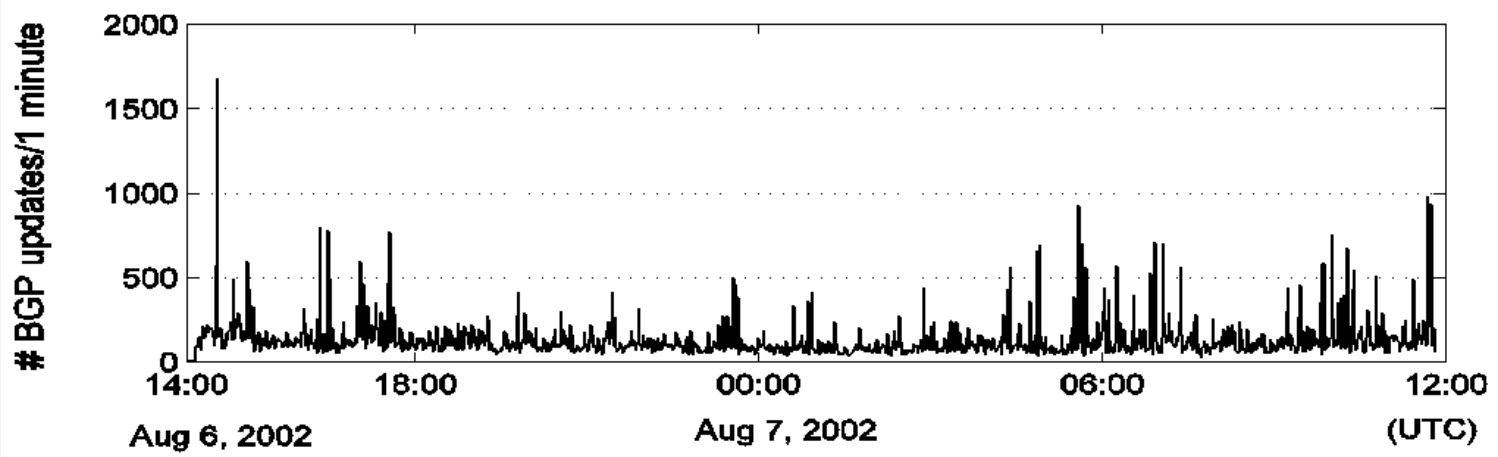
- All egress PoP shifts happen once or twice for a destination

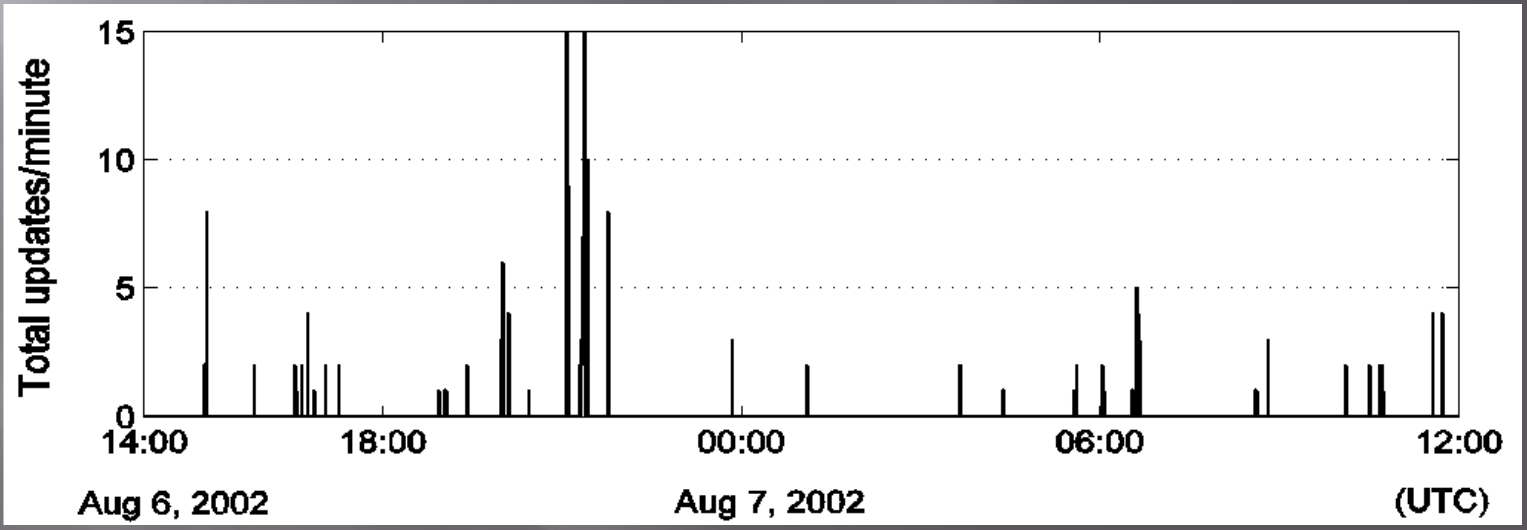
- But only 1% traffic. What of other traffic?

Backbone Traffic Characteristics

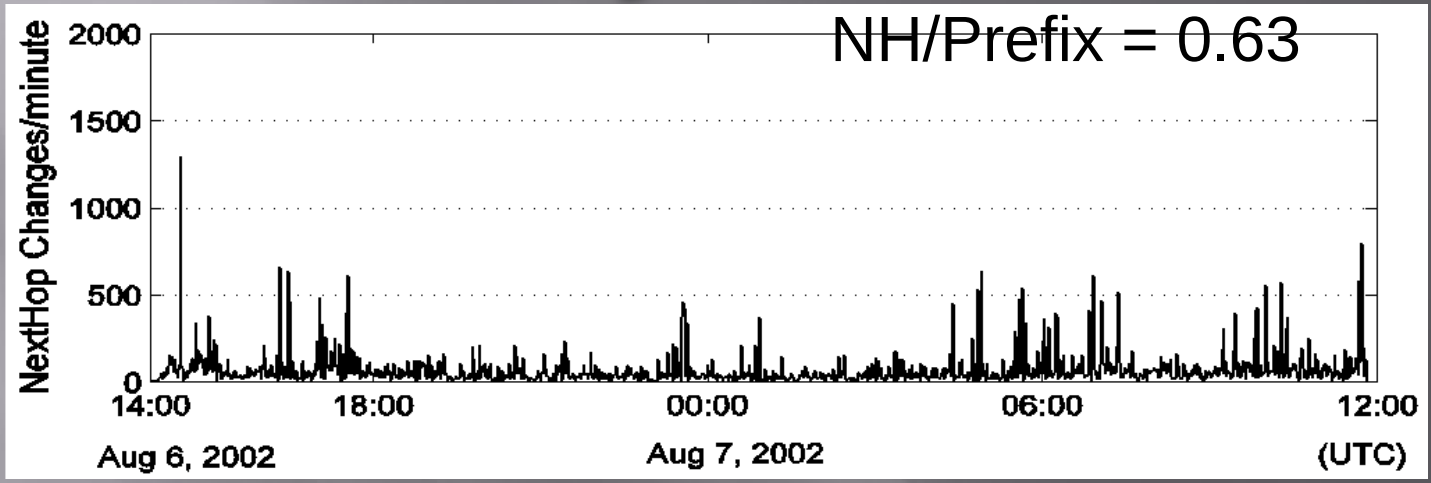
- Heavy hitters prevalent
  - ~200,000 destination addresses → 100% traffic
  - ~15 % of destination addresses → 99% traffic
  - ~1.5 % of destination addresses → 80% traffic
- Which updates affect heavy hitters?
- Which of these change egress PoP?

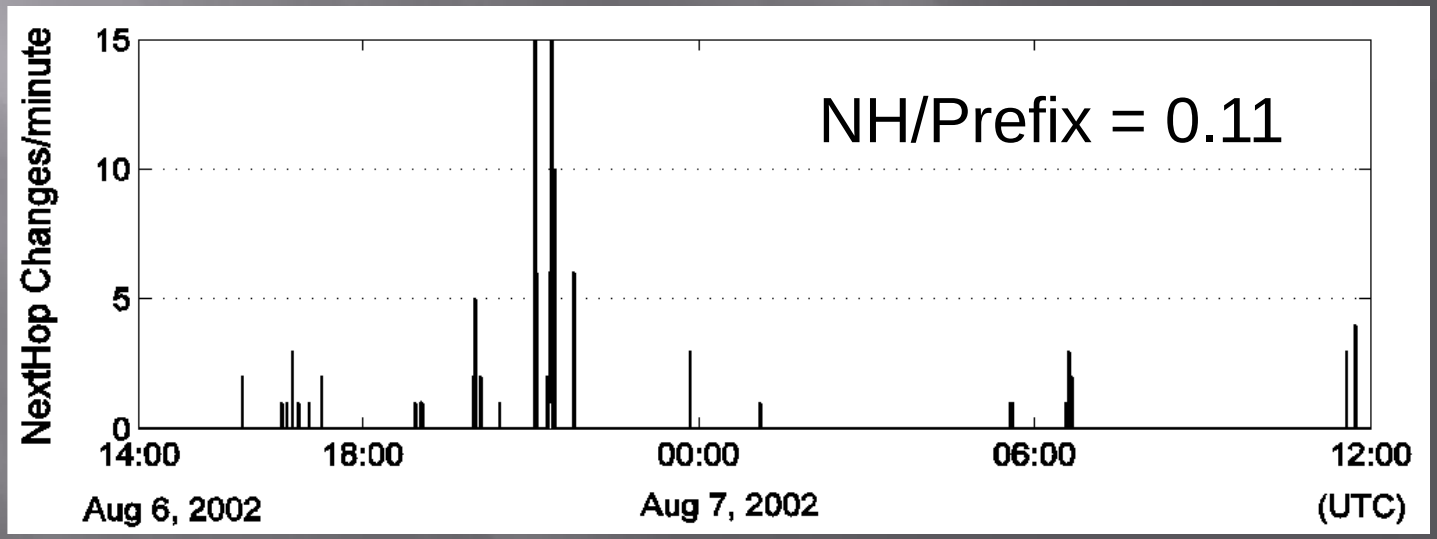
BGP Updates : Heavy Hitters





0.05% change Nexthop for Heavy Hitters





Conclusions

- BGP updates hardly affect intra-Sprint traffic fan-out
  - AT&T[Rexford02]: stable traffic → stable prefixes
  - Why?
    - Standard route filtering?
    - Stable prefixes *attract* stable traffic?
    - Layer3 protection switching and engineering?
  - Why so many other BGP updates?
  - Cause analysis : need all BGP sessions!

Implications

- BGP doesn't cause latency variation in Sprint
  - Good for applications
- BGP doesn't make link loads more dynamic
  - Provisioning / traffic engineering easier
  - Traffic matrix less variable
    - But still inherent variations in traffic

HTTP://IPMON.SPRINT.COM

Data from analysis of traces and routing

Research papers and technical reports

Contact Info