A Tier-1 ISP Perspective: Design principles & observations of routing behavior

Chen-Nee Chuah
(cchuah@sprintlabs.com, chuah@ece.ucdavis.edu)
&
Sprint ATL- IP Group (www.sprintlabs.com)

Outline

- Sprint IP Backbone: Design & Management Principles
- Routing problems can be bad!
- Analysis of ISIS & BGP routing updates

Sprint Backbone Characteristics

- Routing
  - ISIS (1 level) for intra-domain routing
  - I-BGP
    - use route reflectors to avoid full mesh
    - Redundant RRs in each POP
  - E-BGP
    - with peers and customers
    - Use community list, import/export rules
- Traffic Eng: over-provisioning, load balancing

Why Over-Provisioning?

- Keep link load less than 50% to provide
  - high quality of service
    - 0.3% loss
    - speed-of-light e2e latencies
    - 99+% high availability
  - protection against failures
    - large scale outages happen regularly (fix the train system!)
    - Effective “traffic engineering” not very easy…

Outline

- Sprint IP Backbone: Design & Management Principles
- Routing problems can be bad!
- Analysis of ISIS & BGP routing updates

Focus on routing behavior ...

- Widespread outages
  - easily identified
  - impact not fully understood
- Day-to-day operation
  - how frequent do failures happen?
  - what are the effects?
  - what needs to be improved?
- *Routing problems exist!*

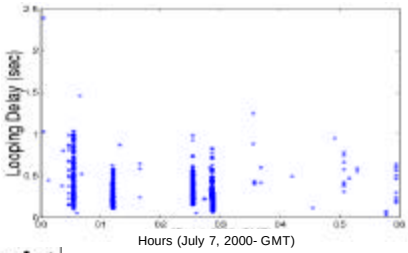
1. Fail-over creates transient link overload

Look at SNMP data on link loads

- 5-min averages
- Sept 2000 - June 2001
- Results from inter-POP backbone links show high variability in link load
  - When a link fails, traffic fail-over to backup links, causing transient overload

3. Routing loops induce delay/losses

- Measurements from 3 backbone links
  - 25% packets caught in a loop in one instance
  - 1 % lost due to expire TTL; those that escape have long delays



Outline

- Sprint IP Backbone: Design & Management Principles
- Routing problems can be bad!
- Analysis of IS-IS & BGP routing updates

The Holy Grail

- Understand routing within Sprintlink
  - ISIS & BGP dynamic behavior
    - Normal vs. Instable periods
    - Re-convergence properties after failures
    - Interactions
  - How routing instability impact traffic
  - Identify problems & find new solutions

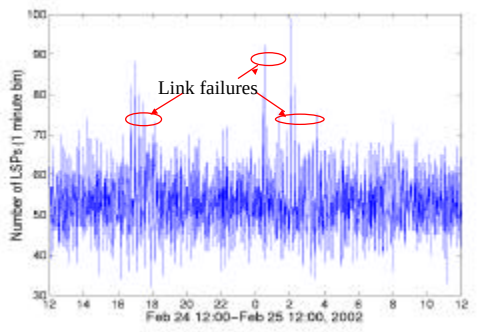
Available Tools & Data

- Python Routeing Toolkit (PyRT)
  - Software to collect ISIS/BGP messages (R. Mortier)
    - MRTD format; Microsecond timestamps
  - Deployed in SJ, NYC & PEN since November 2001
- Other tools
  - topology discovery, post-processing of routing data
- Traffic data
  - IPMON trace
  - VoIP probes between Reston & ATL

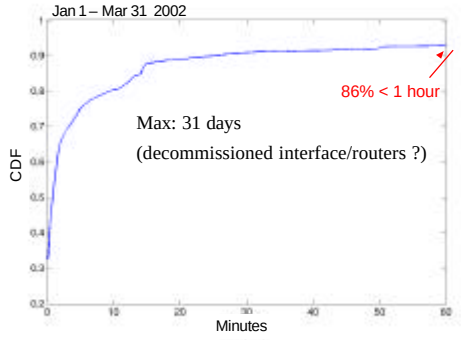
IS-IS Analysis

- IS-IS
  - Link state protocol
  - Link weights (default metric)
  - Per packet or per destination prefix load balancing
- What are we looking for?
  - ISIS churns (LSP storms)
    - Link/router failures
    - Other possible causes, e.g., administrative
    - Frequency & duration of each event
  - Anomalies
    - link weight changes; link flappings

ISIS Churns



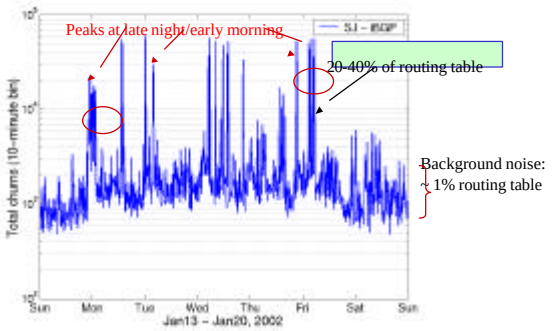
ISIS Failure Event Durations



BGP Analysis

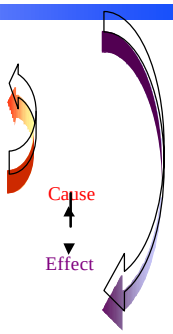
- PyRT-BGP speaker
  - iBGP session w/ a route reflector (SJ)
  - eBGP sessions (NYC & PEN)
- Study BGP Churns
  - Volume of updates/prefixes that flap
  - Classify BGP updates: which ones matter?
    - Types
      - WAdiff, WAdup, AAdiff, Aadup [C. Labovitz]
    - Which path attributes change?
      - Nexthop (affect traffic flow)
      - ASPath, Local Preference, MED, etc.

BGP Activity: Temporal Trends



Understand Cause & Effect of BGP Churns

1. Classify BGP updates
   - Which path attributes change?
   - Which updates matter?
2. Identify causes of churns
   - Who are responsible?
   - Internal vs. external policy change
3. Quantify effects of BGP churns
   - Router loads
   - Traffic delivery



Summary

- Significant no. of links “overloaded” :
  - difficult to distribute traffic evenly
  - *difficult to plan for link failures*
  - average utilization conveys incomplete picture
- Failures part of everyday operation. *Why?*
  - Maintenance, De-commissioned routers/interfaces
  - Accidental (fiber cut)
  - Others: Policy changes, bugs/misconfiguration
- Essential to dampen peaks before increasing average utilization

For more information …

Please visit Sprint ATL IP-Monitoring Project webpage

www.sprintlabs.com/Department/IP-Internetworking/Monitor/