DVMRP Route (In)Stability in the MBone (preliminary results)

Dan Massey <dmassey@parc.xerox.com>
Bill Fenner <fenner@parc.xerox.com>

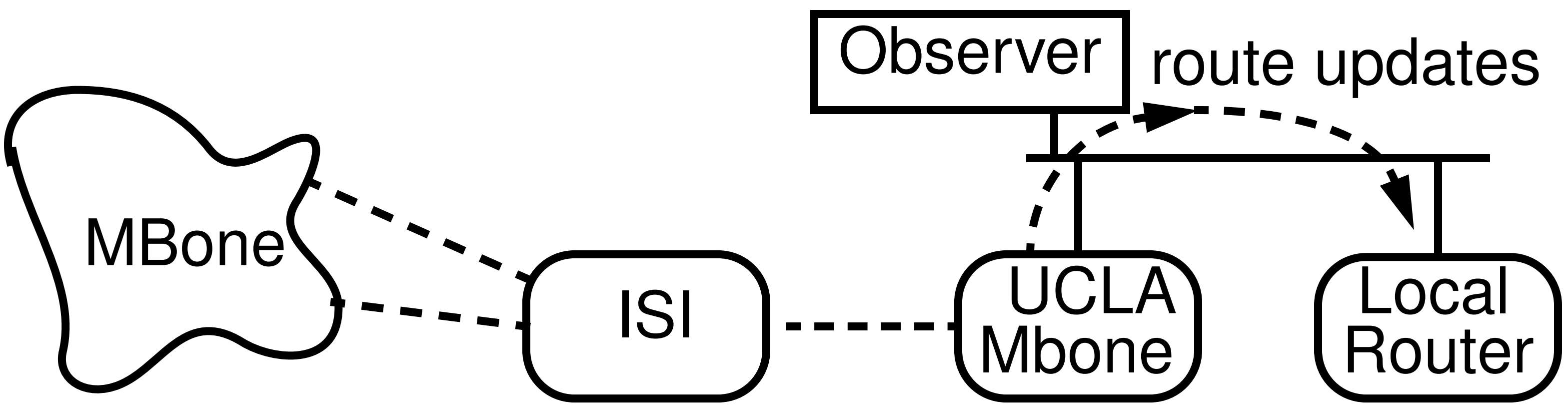
Route Stability in the MBone

- During an average hour, an MBone router hears advertisements for 5500 routes.
- Only 3665 (67%) of these routes do not change during the average one hour period.
- 366 (6%)of these routes change over 60 times an hour.
- 7950 distinct sources were seen in August, but only 5500 sources appear during any one hour period.
- Most of this instability can be explained by protocol bugs.

PARC

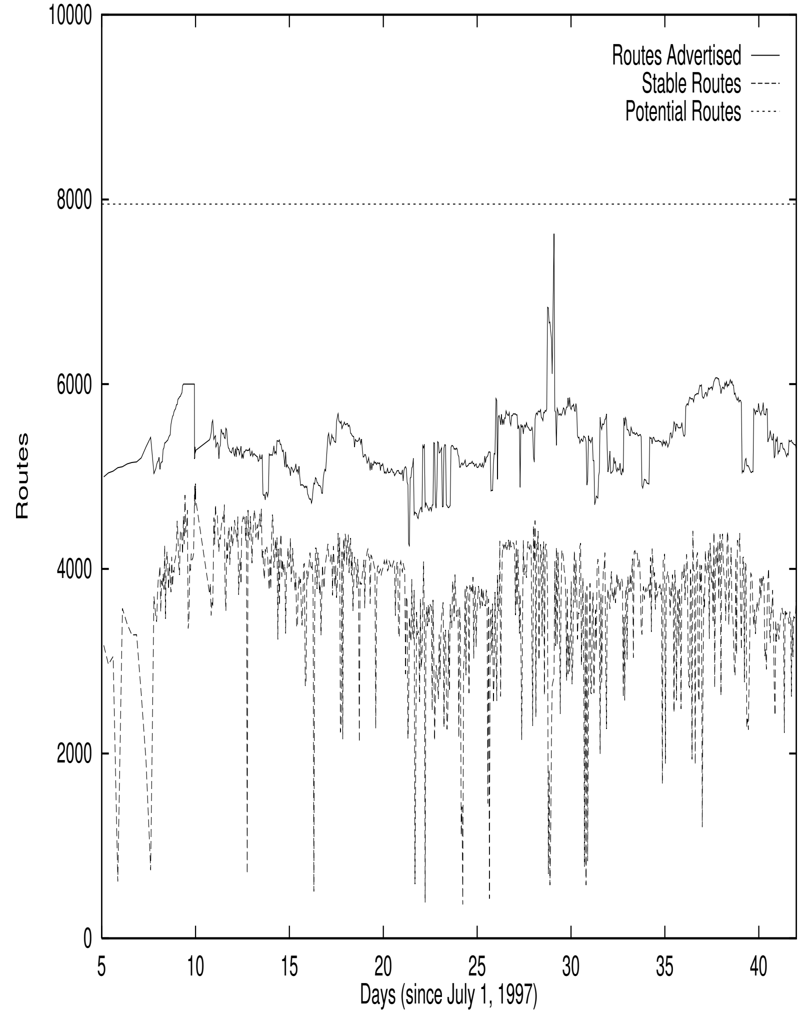
Data Collection Methodology

- Route updates sent by an MBone router are observed to deduce the router's view of the MBone topology.
- Metrics to each advertised source are recorded and changes in a metric indicate instability in the route.
- Measurements were primarily taken by snooping on updates from UCLA's main MBone router:



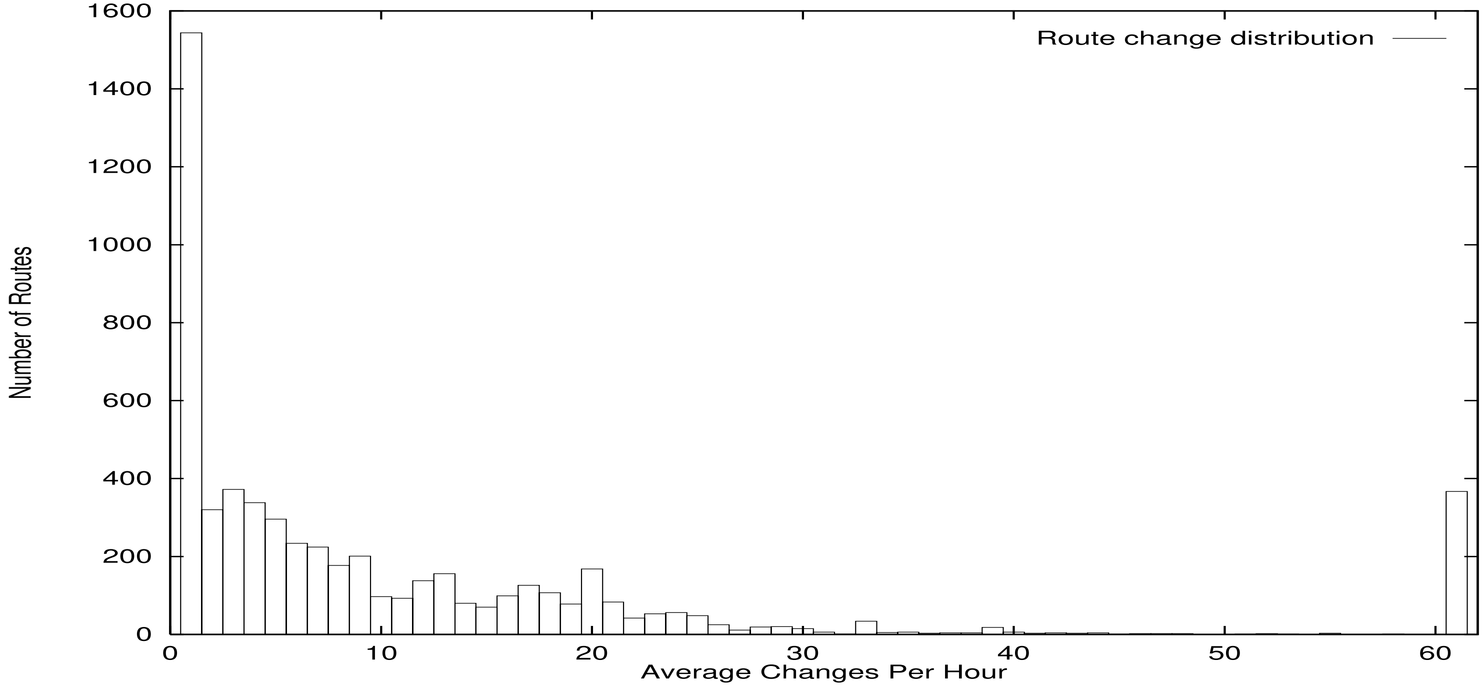
MBone Routes per Hour Viewed from UCLA

- Routes Advertised:
  # of routes seen this hour
- Stable Routes:
  # of routes whose metric did not change this hour
- Potential Routes:
  Total number of routes seen in August.



Average Number of Changes per Hour

- This is a histogram of changes vs. routes.



Not shown: 2991 routes change an average of 0 times per hour.
The rightmost bar is “more than 60.”

Explanation of distribution

- Any route with over 10 flaps per hour is clearly a pathology and not simple link failure/recovery.
- Even an average 1 flap per hour (or less) could be caused by this pathology. Routes that are sourced far away from our data collection point will be unreachable to us for much of the time that they are counting to infinity; there are many routes like this that appear and disappear with a period of over an hour.
- There are some link problems hiding in here but the flaps from these pathologies overwhelm them.

Causes of Instability

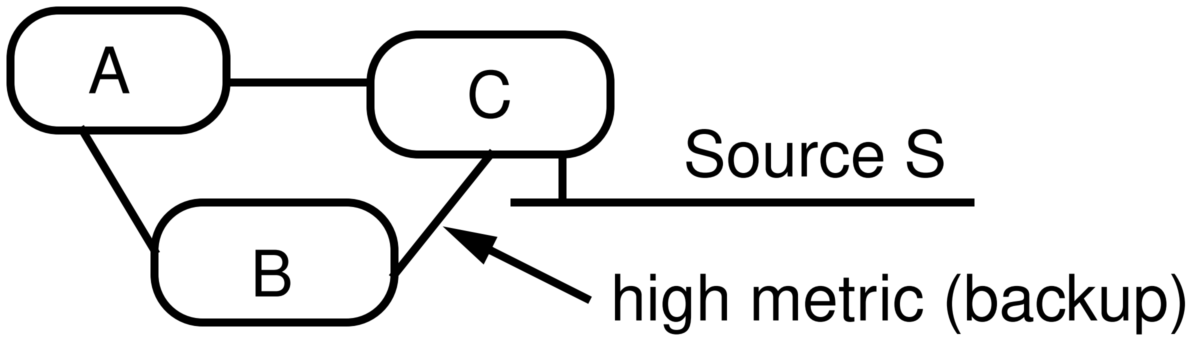
- Most instability we observed was not due to instability in links/routers and was actually caused by errors in the protocol implementation.
- Two existing bugs cause routers to advertise local sources and then fail to properly refresh the routes to these sources. The routes timeout and count to infinity. Eventually the routers re-advertise the local sources and the process repeats.
- Deploying fixes for these bugs can substantially improve route stability and reduce load on MBone routers.

PARC

Protocol Bug Details: cisco "Blaster" Bug

- cisco routers blast the entire route update at once, causing buffer overflow and dropped packets on neighboring routers. Van noticed and fixed this behavior in mrouted in April, 1994.
- Three consecutive lost updates cause the route to time out and (usually) start counting to infinity.
- The blast of updates causes socket buffer overflow and thus lost updates on neigboring mrouteds. Lost updates have also been observed on neighboring ciscos. Note ciscos send local routes last so a site's local routes are the most likely to be affected.

Protocol Bug Details: cisco Multicast/Unicast Tables

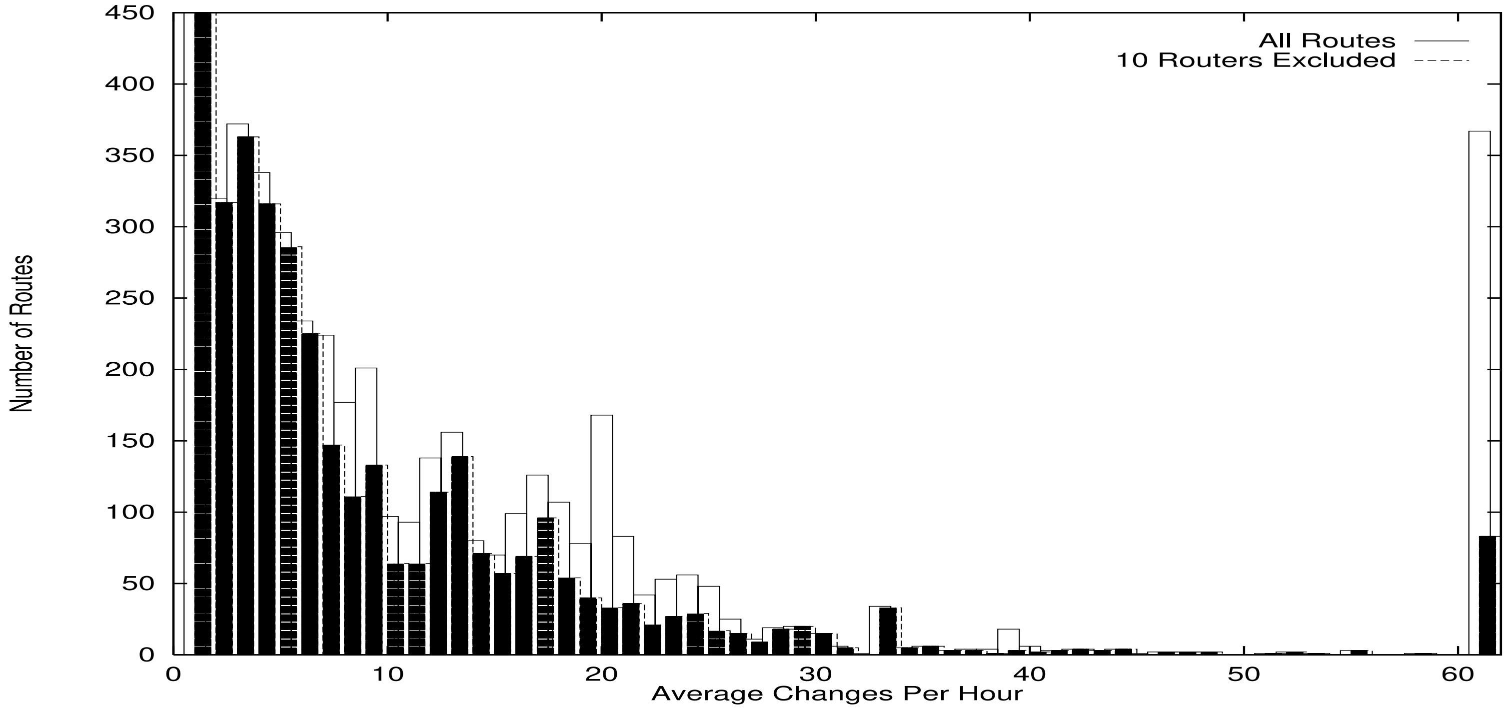


- Router C advertises a route to S from its unicast table. C learns that B has a DVMRP route to S and prefers B's DVMRP route to S over its own unicast route. This forms a loop and the route counts to infinity. Once the route is unreachable, C injects the unicast route again.

Additional Bugs That Enhance Existing Instability

- ciscos fail to trigger route updates when a route changes. This delay increases the chances that counting to infinity will occur.
- If a route has recently changed, ciscos will not accept a higher metric for the route until after a timer expires. This slows counting infinity once a loop has formed. This bug has been fixed but the fix is not widely deployed.
- ciscos also do not accept metric 32 (infinity) updates, which also slows convergence.

Distribution of Changes After Discounting the Effects of 10 Routers Suffering from these bugs.



Further Information and Data for Individual Routes: http://ganef.cs.ucla.edu/~masseyd/Route

- Information on route stability for individual sources is available on the web page. Mbone administrators are encouraged to check their addresses to see if they are contributing to the instability.
- Route changes affect a site's ability to send, not receive, so sites which are infrequent senders may not be aware that they are part of the problem.
- Information on how to obtain and deploy fixes for the bugs is also available. Route data is updated hourly.

PARC

Fixing the Instability

- mrouted 3.9-beta uses bigger socket buffers if the OS makes them available. 3.9-beta3 has a special "blaster" keyword that enables a ***partial*** workaround.
- cisco has implemented fixes for the "Blaster" and metric-32 bugs (CSCdj31841), and the unicast/multicast table problem can be fixed by configuring route distances. The bug-fix is likely to be integrated with 11.2 and a special 11.1 branch used by ISPs. If you want to see it in other versions, contact cisco.

PARC