

# Delayed Internet routing Convergence

Paper review by Ehsun Darody

Tuesday, 26 November, 2002

Unlike switches in public telephony network which exhibit failover on the order of milliseconds, experimental measurements shows that inter-domain routers in the packet switched Internet make take tens on minutes to stable after a failure (delayed convergence). During this period, hosts experiment intermittent loss of connectivity as well as increased packet loss and latency. It especially hurts QoS algorithms that are predicated on the existence of a stable underlying forwarding infrastructure.

Routing information shared among BGP speaking peers are announcements and withdrawals. Failover is an implicit withdrawal and replacement of a route with one having a different ASPath. To probe delayed convergence we inject many (250000) routing faults into geographically and topologically diverse peering sessions. We define four kind of routing events injected into the Internet: Tup, Tdown, Tshort, Tlong. We define latency of each injected event as the time between the injection of the fault and routing tables to reach steady state.

Experiments over tow years show that more that 20 percent of Tlong and 40 percent of Tdown events oscillated for more than three minutes. Both Tdown and Tlong events converged more slowly that Tup and Tshort. Tdown and Tlong events on average triggered more than two times the number of updates messages than both Tup and Tshort events. Results show no correlation between convergence latency and geographic or network distribution and likely primarily depend on topological factors such as the number of adjacent BGP peers and upstream provider transmit.

End-to-end measurements show that the majority of end hosts begin

responding within 30 seconds and all within one minutes after injecting fault.

Examples and analysis exist of both upper and lower computational bounds on BGP convergence. Min-RouteAdver timers in routers significantly reduce the computational and communication complexity of BGP convergence.

We can improve BGP convergence through the addition of synchronization, diffusing updates an additional state information but all of these come at the expense of more complex protocol and increased router overhead.