Signaling Convergence Proposal: Forward Path Probing

IBM Research GmbH, Zurich
June 21, 2007
Objective #1: Find a commonality in all proposals

- Forward path probing: present in all 4 proposals
  - Proposed in FECN [Dallas'06]
  - ECM has used it in early BCN ("solicit" bit)
  - E2CM and 3-point QCN have it

- Backward path probing: proposed by M. Seaman [Monterey'07]
  - Destination-initiated, moving upstream in 'reverse' path

Path probing is the one common feature found in all proposals

Why?
1. No need for CPIDs and tags
2. Provides consistent picture of the whole path, w/ local and global status
   1. Gathers local feedback from all switches along path
   2. Enables measurement of path occupancy
3. NIC-driven sampling: Big potential for optimization (adaptive sampling)
4. No constraint on feedback type or source response function
5. Can be applied to all schemes without changing their nature.
Objective #2 : Simplify signaling

• Reduce the number of signaling options

1. Forward
   + Simple: follows the traffic flow
   + Aggregates e2e feedback: conveys global status (both + and - feedback)
   + No reverse-path injection in switch
   - RX-NIC must reflect: already does it for ACKs
     - Longer lag compared to direct backward signaling
       - Worst-case lag = RTT, which implies hotspot close to source: not critical
       - Lag for critical hotspot close to destination is about the same
   - Open loop if link fails: ditto in reverse / backward case
   - Multipath ambiguation: similar to rate limiter aggregation

2. Backward
   + Less lag: fastest response
   - Local: non-coherent feedback w/ heterogeneous switches
   - Complex: down/up-stream intercoupling in switch -> reverse injection is tough

Forward is simpler, backward is faster... which one is more suitable?
Forward Path Probing: Lowest Hanging Fruit?

• Path probing is a natural match with forward signaling
  - Direct BCN alone is mutually exclusive w/ path probing

• Simpler protocol
  - Controllable overhead and complexity

• Lower entry barrier to congestion-managed networks
  - Reduces the burden on switch vendors

• OPEN
  - Enables arbitrary type of feedback and source response function
  - Allows alternate signaling options