Applications of DCN: the *TeraPaths* perspective

Dimitrios Katramatos

RACF – Brookhaven National Laboratory
End-to-End Setup of Virtual Paths

- Site A: Host a1, Router, Virtual Border Router, Site Router, Border Router
- Site B: Host b1, Router, Site Router, Border Router
- Site C: Host c1, Router, Regional Provider Router, Site Host / Border Router

- WAN Domains: DCN, Pass-through, TeraPaths
L2 Issues – Need for Scheduling

- Book ahead – guaranteed time, duration, protection
- Focus on individual flows
- Need for scheduling:
  - Limitations with VLANs
    - Tag ranges
    - Tag conflicts
  - Detection of circuit failure and recovery
    - Failover to IP, re-acquire circuit to failback
    - Failover to pre-acquired (backup) circuit
  - Select route between src and dst
  - Scalability issues
    - Flow grouping / circuit consolidation
      - Flows between same src and dst sites can share circuit, policing maintains bandwidth guarantee
      - Multiple TeraPaths reservations can be associated with the same circuit reservation
Flow Grouping/Circuit Consolidation

- Easy when requirements are known in advance

- Modification of reservations required otherwise
  - Selection/optimization to minimize resource waste
  - Trade-off based on $\Delta bw$ (bandwidth difference), $\Delta t_b$, $\Delta t_a$ (time period before and after a reservation)

- Similar approach to disk buffering (read ahead / write behind)
  - Bring up ahead / teardown behind
  - Reuse existing active circuits
  - Reserve circuits with more bandwidth and longer duration depending on differences in start time, duration, bandwidth of reservations
  - Delay teardown, modify circuit duration and/or bandwidth if possible
Availability inquiries and negotiation capabilities

- **Availability**
  - `AvailablePaths(endPoint1, endPoint2, startTime, endTime, conditions...)` {returns list of paths}
  - `AvailableBandwidth(endPoint1, endPoint2, startTime, endTime, conditions...)` {returns list of lists: available bandwidth vs. time for every path}
  - Conditions include various metrics, reservations to release

- **Negotiation**
  - `TemporaryReserve(reservation parameters, timeout...)`
  - `ConditionalReserve(reservation parameters, reservations to release, timeout...)`

- **Reservation modification**
  - Beyond timewise modifications?
TeraPaths is middleware

- TeraPaths configures the network on behalf of end-users and applications
  - Configures site LANs directly, negotiates with DCN for WAN circuits
  - Reservation-oriented network scheduling
  - Flow-based end-to-end QoS guarantees
  - Fault tolerance (failover to IP, failback after re-acquiring circuit)
- Web services API
  - Direct access
    - CLI
    - Web interface
    - Remote TeraPaths instances
  - Indirect access
    - Application plug-ins
- Virtual Organization (VO) capability for grid users
  - Membership lookup (VOMS)
  - User roles