Network QoS Assurance in the Presence of Faults*

Shrirang Gadgil, Fredrick Porter, Kirthika Parmeswaran, Ravi Vaidyanathan and Balakrishnan Dasarathy

Telcordia Technologies
One Telcordia Drive
Piscataway, NJ 08854

OMG Real-Time and Embedded Systems Workshop
Arlington, VA
July 13, 2006

•This work is supported by DARPA ARMS Program, Contract No.: NBCH-C-03-0132
Approved for Public Release, Distribution Unlimited

Telcordia

Talk Overview

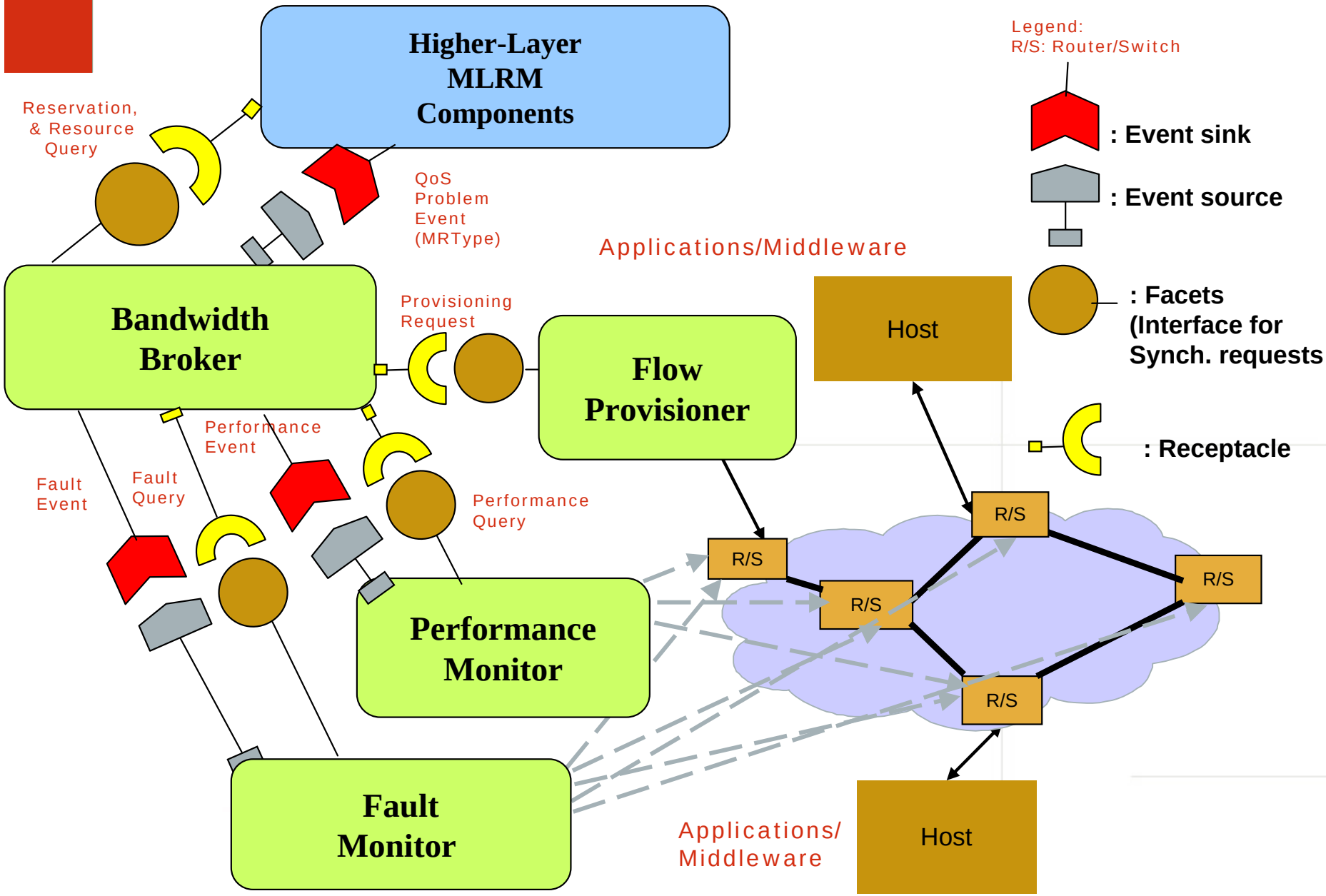
- Network QoS architecture overview
  - Bandwidth Broker overview
  - Fault Monitor
- Network fault monitoring & QoS restoration
- Software failure recovery of network QoS components

Bandwidth Broker Overview

- Bandwidth Broker provides network quality of service guarantees.
  - Both capacity (bandwidth) and delay guarantees
- Bandwidth Broker
  - Leverages DiffServ
    - Aggregate (class-based) traffic treatment
      - With the same per-hop forwarding behavior throughout the network
    - Policing at the edge routers at the level of individual flows
  - Makes use of an admission control engine based on per link book-keeping
    - This then requires knowledge of the path a flow will take on a continual basis
- Technology applicable for both layer-3 and layer-2 networks
- Bandwidth Broker provisions the ingress routers to mark and police packets appropriately using Flow Provisioner
  - QoS has to be guaranteed not only for the new flow to be admitted but also for all the flows previously admitted

Network QoS Architecture

Telcordia



Telcordia

Network QoS Feedback Mechanisms

- Performance monitoring
  - Detects and responds to performance problems
    - Active probes for measuring latency, jitter, and packet loss
- Fault monitoring & QoS restoration
  - Detects and responda to topology changes due to faults
    - One second response time is the goal.

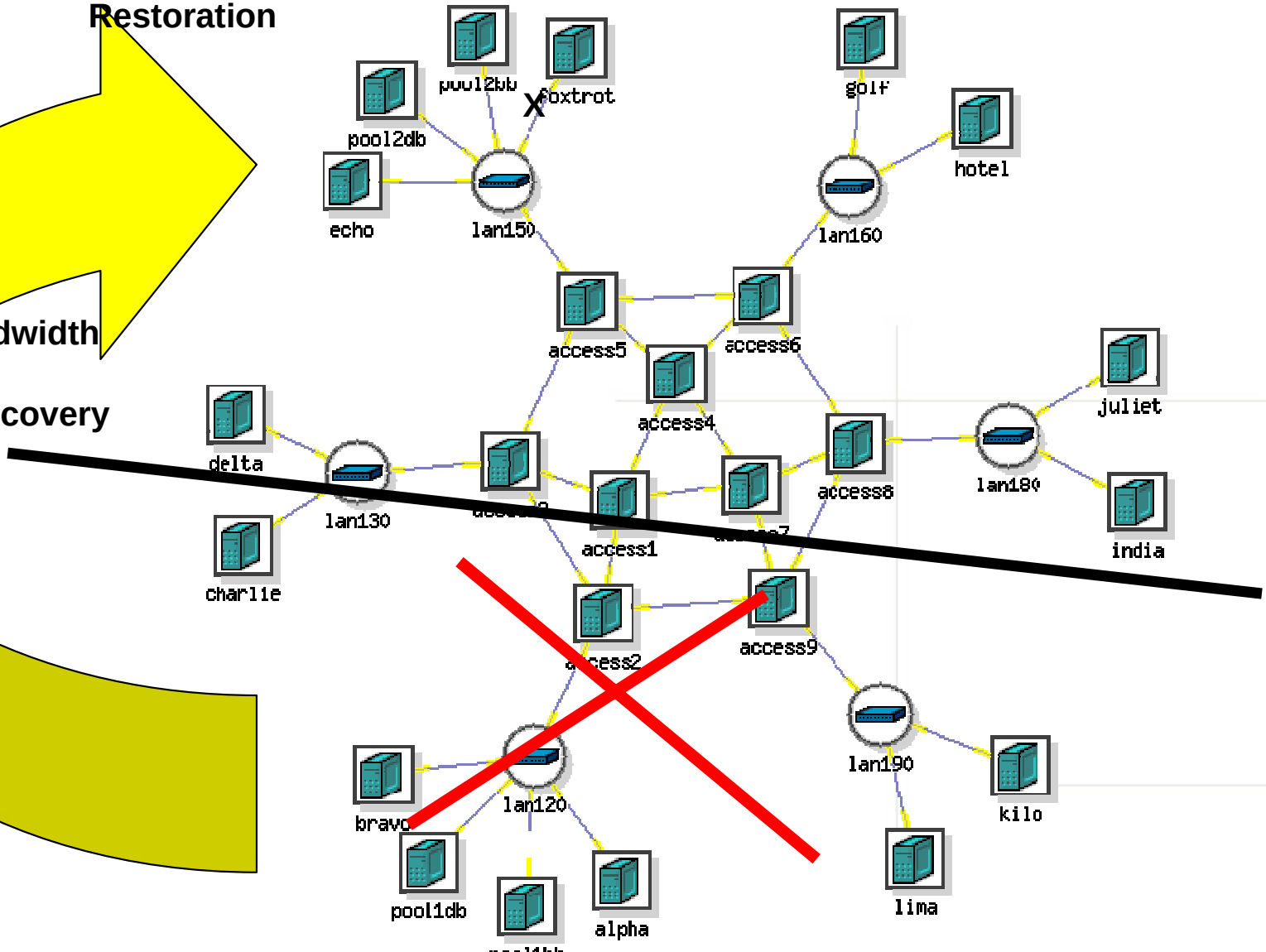
This talk is mainly on fast QoS restoration from network failures

Fault Recovery Steps in a Major Catastrophe

Telcordia

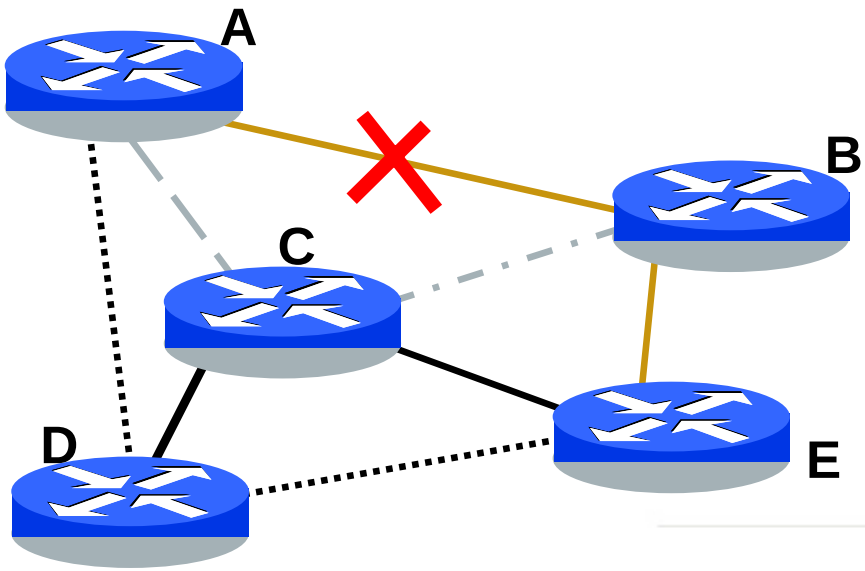
Step 2: Network Rediscovery and QoS Restoration

Step 1: Bandwidth Broker Software Recovery



Network Fault Monitoring & QoS Restoration: What is the problem?

Telcordia



- If the link between switches A and B goes down, then a flow Y between A and B may be routed through switch C (shown in dashed lines). Similarly, a flow Z between E and A that originally used the links EB and BA may now use links ED and DA (shown in dotted lines)
  - Links AC, CB, ED, and DA may now be oversubscribed causing concerns on QoS guarantees for Y and Z as well as for the flows that had been using these links prior to the occurrence of the fault

Telcordia

Network Fault Monitoring & QoS Restoration Steps

- Fault detection using SNMP traps
  - Routers allow for export of the OSPF link state database (i.e., the topology)
    - We assume OSPF as the routing protocol, as OSPF is a widely used IP routing protocol in enterprise networks.
- Network discovery
  - On a fault event, new paths between all pairs of edge routers are discovered to track the bandwidth correctly.
    - We use Floyd-Warshall all-pair shortest path algorithm to compute all pairs of routes
- Impact Analysis
  - For each admitted flow the Impact Analysis process determines whether the flow's path has changed; if so, it is a candidate for re-admission
- QoS restoration
  - Can support different algorithms with or without the use of preemption satisfying different optimality criteria (maximize the number of flows readmitted, maximize the number of highest priority flows readmitted, etc.)

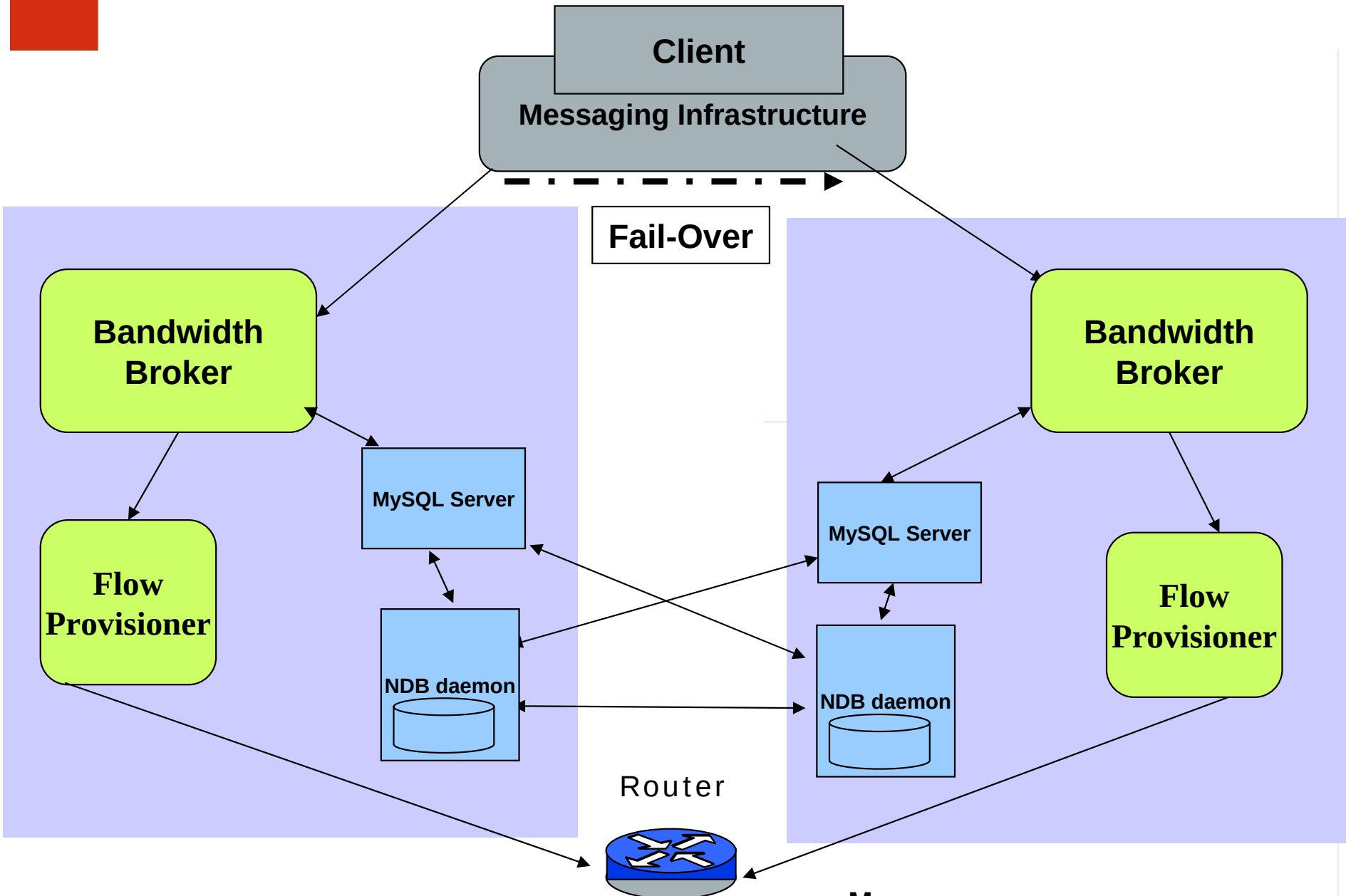
Overview of the Software Fault Tolerance Approach

Telcordia

- Based on the warm-passive process replication approach
- A particular focus is on the Bandwidth Broker's database recovery;
  - Bandwidth Broker is stateless process
    - Entire state stored in a replicated database
    - Employs a modified version of MySQL cluster technology
      - http://dev.mysql.com/doc/refman/4.1/en/ndbcluster.html
      - In-memory DB (NDB) (not disk-less), (synchronous) distributed commit
      - Survives network partitions
- A CORBA/MEAD/Spread messaging infrastructure provides server failure transparency
- Provisioning instructions to routers are not necessarily idempotent (e.g., Linux)
  - Rollback and redo router a provisioning operation, as needed

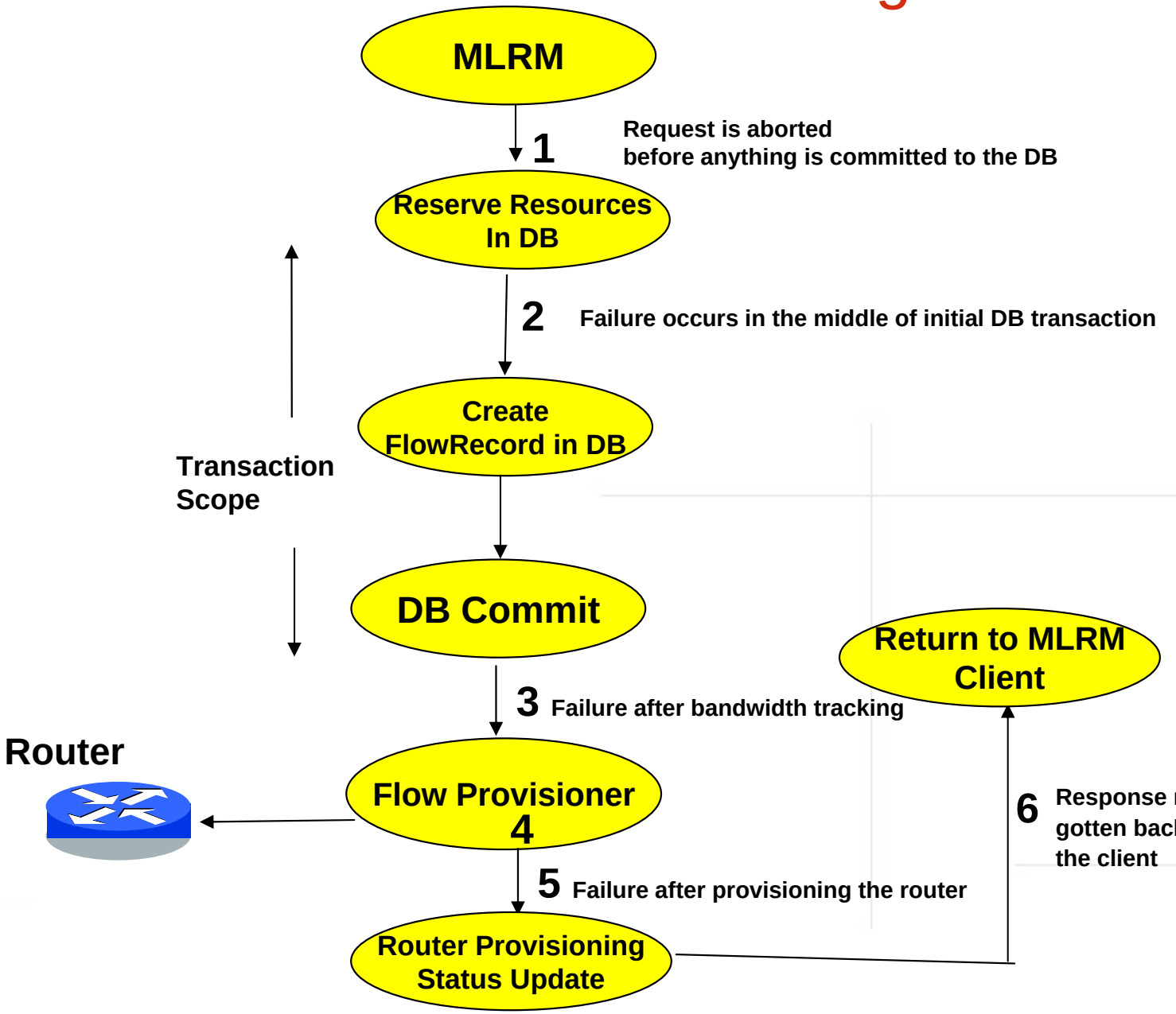
Software Failure Recovery Illustration

Telcordia



Software Failure Handling

Telcordia



Telcordia

Experiments and Measurements

- Experiments and measurements in Emulab
  - www.emulab.net
  - Using Linux routers and hosts (PCs from 800 MHz to 3000 MHz)
- Software recovery < 300 ms
  - Most of this time (about 200 ms) in failure detection so as to avoid false positives
- Network fault detection and QoS restoration
  - Path discovery using link-state information
  - Algorithms for path database updates, impact analysis and QoS restoration are currently being tuned to get timing in the second range