Using SCTP Multihoming for Fault Tolerance & Load Balancing
Armando L. Caro Jr., Janardhan R. Iyengar
Paul D. Amer, Gerard J. Heinz, Randall R. Stewart
http://pel.cis.udel.edu

Overview

History
• Originally intended for telephony signaling
• Overcomes several TCP & UDP limitations
• Designed as a general purpose transport protocol
• Became an IETF Proposed Standard in October 2000 (RFC 2960) under the SIGTRAN working group
• Handed to Transport Area working group for continued work

Features
• Reliable data transfer
• Ordered and unordered data delivery
• Multiple streams – no head-of-line blocking
• Multihoming

Multihoming

Failover

What happens if B1 fails?

Changeover

Sender decides primary destination address for traffic
• Sender can change primary destination address during an active association
• Utilities (pending further research):
  • End-to-end mobility (with ADD/DELETE IP extension)
  • End-to-end load balancing

Related Research

Resilient Overlay Networks (RON)
• Architecture that allows a small group of Internet address to detect and recover from path outages within several seconds
• QoS enhancements for mobile computing applications
• Reliable Sockets
• Application protection from network failures common to mobile computing such as link failures and IP address changes

Migrate

End-to-end framework for Internet mobility that supports rapid re-convergence of all established TCP connections
• Fine-grained server failover mechanism of long-running connections

Rocks
• Reliable Sockets
• Application protection from network failures common to mobile computing such as link failures and IP address changes

Migrate

M-TCP
• Mechanism to migrate live TCP sessions to a redundant server upon server overload, network congestion, etc.

Current Research

SCTP’s current failover mechanism uses the Path.Max.Retrans parameter for failover

Proposed two-level (or β) threshold failover mechanism
• α – maintain primary, but failover temporarily
• β – change the primary, making failover permanent

Current Work: Changeover Issues

• Reordering causes spurious fast rtxs which cause congestion window overgrowth
• Occurrence of reordering increases due to sender introduced round trips

Future Research

• Develop an adaptive failover mechanism for SCTP
• Incorporate application requirements in the mechanism

Motivation
• End-to-end connectivity can suffer during net failures
• Internet path outage detection and recovery is slow
  (minimum 3 mins, often 15 minutes, 40% 30 mins)
• Network failures are common in mobile networks
• Network fault tolerance should cope with dynamic network conditions and varying applications needs
• Current SCTP failovers are not adaptive to application requirements and network conditions

End-to-End Load Balancing

Motivation
• Exploit all network resources visible at the transport layer
• Perform fine-grain load balancing at the transport layer
• Avoid replication of work and coarse-grain load balancing in applications

Current Research

Issues
• Scheduling of traffic on multiple paths
• Reordering
• Loss detection & recovery
• Congestion control: shared or separate?

Future Research
• Investigate changeover issues and solutions during load balancing
• Investigate dynamic shared bottleneck detection techniques for congestion control
• Investigate factors and algorithms for scheduling traffic on multiple paths

Current Work: Changeover Issues

• Reordering causes spurious fast rtxs which cause congestion window overgrowth

Future Research

• Investigate changeover issues and solutions during load balancing
• Investigate dynamic shared bottleneck detection techniques for congestion control
• Investigate factors and algorithms for scheduling traffic on multiple paths