Proposal for Additional Persistent Reservation Functions

SCSI Working Group

Natick, MA

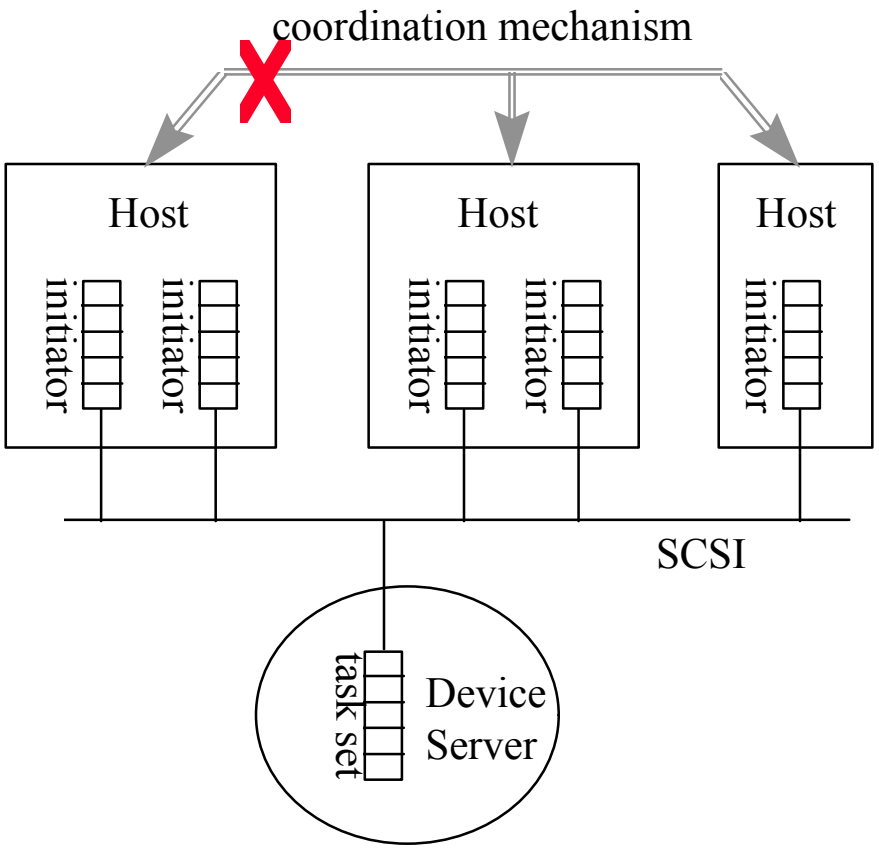
September 9-10, 1996

Tom Coughlan

Digital Equipment Corporation

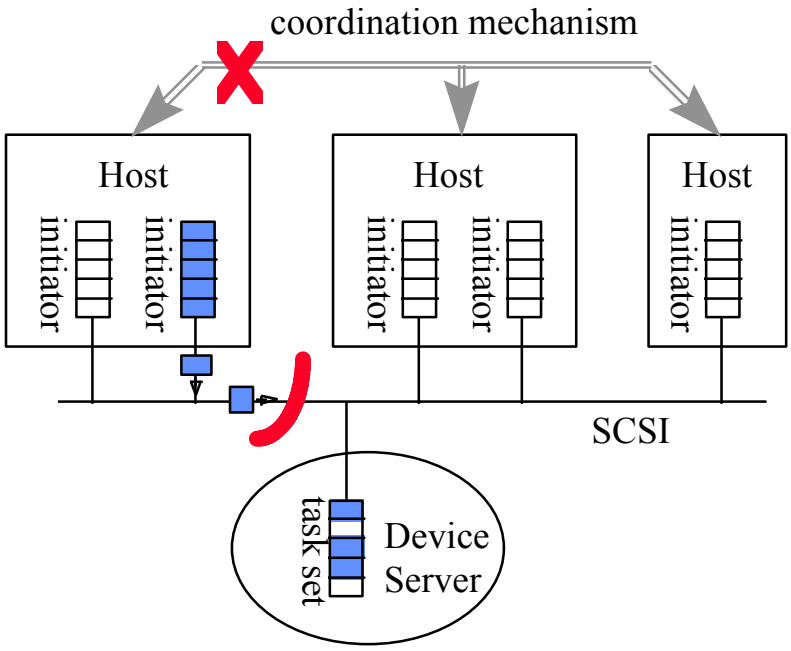
Multiple-Initiator Systems

- Multiple SCSI initiators must use some form of coordination when accessing shared storage devices.
- It is desirable for one or more initiators to take over when an initiator becomes unresponsive:



Recovery From Unresponsive Initiator(s)

- The survivors must determine that:
  - there are no tasks from the unresponsive initiator in the device server
    - and no ACA conditions
  - there are no reservations for the unresponsive initiator
  - no further tasks from the non-responsive initiator will be accepted



Constraints

- Recovery should be fast, since access to the shared resource is blocked until it completes.
- Recovery should not disrupt any tasks that the survivors have already queued to the device.

The Current Solution

- A survivor can issue a Preempt and Clear, with an Exclusive reservation.
  - this removes the appropriate tasks and reservations from the device server, clears ACA, and blocks subsequent tasks
  - unfortunately, it also blocks access by all the other initiators in the system
- A Shared Access reservation would allow multi-access, but then the unresponsive initiator's in-flight commands would not be blocked.

The Proposal

- Add a new Persistent Reservation Type, called "Shared Access, Registrants Only"
  - allows Read, Write, and Reserve, but only for registered initiators
  - does not conflict with itself
- Locking out non-registered initiators is a Good Thing
  - they are clearly not cooperating

An Example Application

- When a Host boots each initiator:
  - Registers
  - Issues a Shared Access, Registrants Only Reservation
- When an host becomes unresponsive, one or more of the survivors:
  - Issue a Persistent Reservation
    - Action = Preempt and Clear
    - Type = Shared Access, Registrants Only
    - Service Action Key = the key of the unresponsive initiator
- When a Host recovers, it:
  - Ensures all of its old commands are flushed.
  - Re-registers and re-reserves.

Summary

- One change is required to support simultaneous access multiple initiator systems.
- The initial proposal, posted to the SCSI reflector, called for an additional change, so that Preempt and Clear could be applied to all non-registered initiators.
  - the goal was to provide a means to Preempt a set of nodes quickly, and without haveing to learn each one's key
  - this has been reconsidered, and is not being proposed at this time.