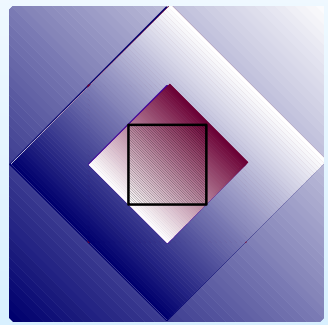
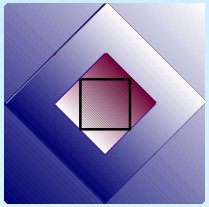
iSCSI – a SCSI over TCP mapping

London-August-2001

Julian Satran

IBM Research Lab in Haifa





Status

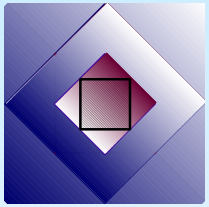
- Open “chapters”
  - Security
    - 2 “teams” working
    - SRP+keying – requires inventing a scheme
    - IKE+requires referencing a scheme
    - Encryption will probably have to be mandatory to implement
    - A separate RFC to be referenced by the main iSCSI doc
  - Framing
- Open items
  - NOP
  - Login
  - T10 ordering proposal
  - Recovery summary

NOP (1)

- Issue - NOP may close the command window
- Solution proposed – simplify NOP
  - Remove the P bit
    - Ping Data if present indicate by DataSegment Length
  - Convey the answer need through ITT/TTT
    - No valid ITT/TTT no answer needed
  - Mandate Immediate if ITT is not valid

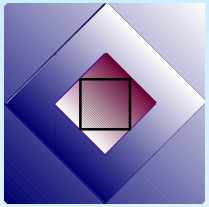
NOP (2)

- ITT valid means Initiator wants answer
- TTT valid means Target wants answer
- ITT & TTT cannot be both valid in a Nop-In (to break the loop)
- ITT & TTT can be both valid on a Nop-Out (three way handshake)



Login (1)

- Issues
  - General Structure
  - Individual Parameters
- General Structure in 07
  - 2 phases
    - Implicit
    - Optional
  - Overall concern – reduce number of handshakes and keep footprint low
  - Perceived programming complexity not a concern



Login (2)

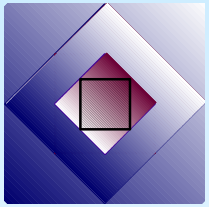
- Proposals
  - SecurityContextComplete alone – Eddy Quicksall
  - Mandatory Security – Robert Russell
  - Both Explicit & Optional
    - Through brackets SecurityPhase/OperationalPhase=<start|end>
    - Through a binary this-phase/next-phase code and reuse of the final bit

Login (3)

- SecurityPhase/OpPhase =<start|end> are the "brackets"
- Parameters for one phase only
- Legal
  - I->T Login SecurityPhase=start,…. Parameters …., SecurityPhaseEnd+F
    T->Login SecurityPhase=start,….Parameters …, SecurityPhaseEnd+F

Login (4)

- Some details about the binary-phase and final/bit proposal
  - Byte 38 in Login & Text – has 2 Nibbles Current/Next
  - Final bit means ready to move to next
  - Phases are 0-Security, 1-Op, 15-FF
  - Parameters are from one phase only
  - After the F bit Handshake they move on



Login (3)

- Miscellaneous
  - Common Header/Data CRC Negotiation (either both are on or both are off)
  - Drop Security Digest Negotiations
    - Vendors can use them as vendor specific
  - Drop Security Digests altogether
    - Nobody can use them
  - Hex/Decimal – Leave only hex?

T10 – serialization interlock

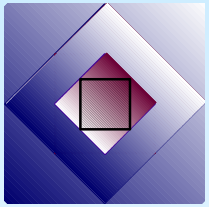
- Current proposal – Busy, Task Set Full and Reservation Conflicts become Check Condition generators under the control of bit in the LU Control Mode Page
- Issue - in single queue (per multiple initiators) devices this can cause a Denial Of Service situation
- Solutions:
  - Leave as it is – argue the case in T10
  - Use UA that with a recently proposed/adopted change can have the same serialization effect but limited to one initiator even on single queue devices
- Jim Hafner and Julian Satran will participate at the next T10 meeting attempting a closure on this issue

Interlock – Proposal Outline

- Add an Interlock Bit in the LU Control Page
- For Busy/Task Set Full/Reservation Conflict if a command form a specific initiator gets rejected the target has to "remember this event" per initiator (3 bits - cleared also by some actions like resets)
- When the the LU state changes AND the interlock bit is 1 AND the Busy/Task Set Full/Reservation Conflict reject-remembered is 1 the target enters a UA pending state for the specific initiator (the "remember" bits could be cleared here or later)
- This UA condition remains "active" until explicitly cleared by an appropriate command and prevents other commands being accepted

Interlock – Proposal Outline (cont.)

- How is it better:
  - Confined to one initiator
  - Currently executing commands are not blocked as in ACA (ACA mandates command to be blocked in order to avoid generating a second sense)
  - Successfully sent AER means (at the target getting ack!) – see SAM-2



Recovery (summary)

- SNACK is weak but useful
- The fast path price is paid
- A form of ACK might relax the need for data replay buffers at target