End-to-end Data Protection and Tapes

Kevin Butt
End-to-end data protection

• T10 standardized end-to-end data protection for blocks transferred to disk
  – This is getting adopted by an increasing number of vendors
  – An increasing number of customers are asking for this protection on tape
  – Tape solutions are losing to pure disk solutions because of a perceived data integrity advantage by disk systems

• T10 did not standardize end-to-end data protection for blocks transferred to tape
  – During the work on end-to-end for disk, the tape group was told a common solution for disk and tape was not going to happen and if tape wanted end-to-end data protection the tape group needs to do a unique solution for tapes.
End-to-end data protection is needed on tape devices

- Potential data corruption between the application and the HBA’s (same on tape as disk)
- Potential data corruption on the interfaces to the data delivery subsystem on both ends of the wire (same on tape as disk)
- Potential data corruption internal to the device (same on tape as disk)
- Potential data corruption as block is transferred between intermediate devices (e.g., protocol bridges)
What does end-to-end data protection need to cover?

• Application to HBA
  – Needs supported same as in disk solution
  – Likely will not be supported by applications for a long time

• HBA through service delivery subsystem to tape device
  – Needs supported to cover memory buffers in source and destination

• Tape drive interface through internal workings of tape drive and onto medium
  – To get true end-to-end logical block protection the protection information needs to be saved on medium with the logical block
Likelihood that protection information be saved with logical block

- Tape drives today save protection information with data blocks.
  - ECC
  - CRC
- Making the protection information fit into existing schemes in use today would greatly accelerate adoption and increase the likelihood of adoption
Existing vendor-specific methods

• IBM has been providing end-to-end logical block protection for over ten years in its enterprise tape drives.
  – Began with 3480 Tape drives
  – Continued in 3590 and 3592
  – 4-byte CRC placed on logical block and transferred with the block
    • Writes: generated at host; validated at drive and written to medium with logical block
    • Reads: read from medium and validated at the drive; validated at the host and stripped from block for application use
    • Validated at multiple points along the path
Benefits to IBM with this solution

• Prior to solution
  – Data Integrity Issues
  – Difficult to find where the problem occurred

• After solution
  – Quickly find where problems occur
  – Data Integrity issues disappeared (Integrity was assured)
  – When there is bad memory or such, it is discovered before host is told the data is on medium
Diagram of IBM Solution
(logical representation)

4-byte CRC calculated and placed as last 4-bytes of logical block – LBA size = LBA size + 4 and transferred with block (or validated and stripped if Read)

4-byte CRC validated at multiple points between source and destination (exact points not necessarily shown)
Proposal – Implement end-to-end logical block protection on tape

- There is a proven method that has been in use for more than twelve years
- Leverage this proven solution
- Make sure that options are available to meet all vendors needs
- For proposal against SSC-3 see (http://www.t10.org/ftp/t10/document.07/07-374r0.pdf)