

Standards and Digital Archives

September 23, 2013

Henry Newman

CEO/CTO

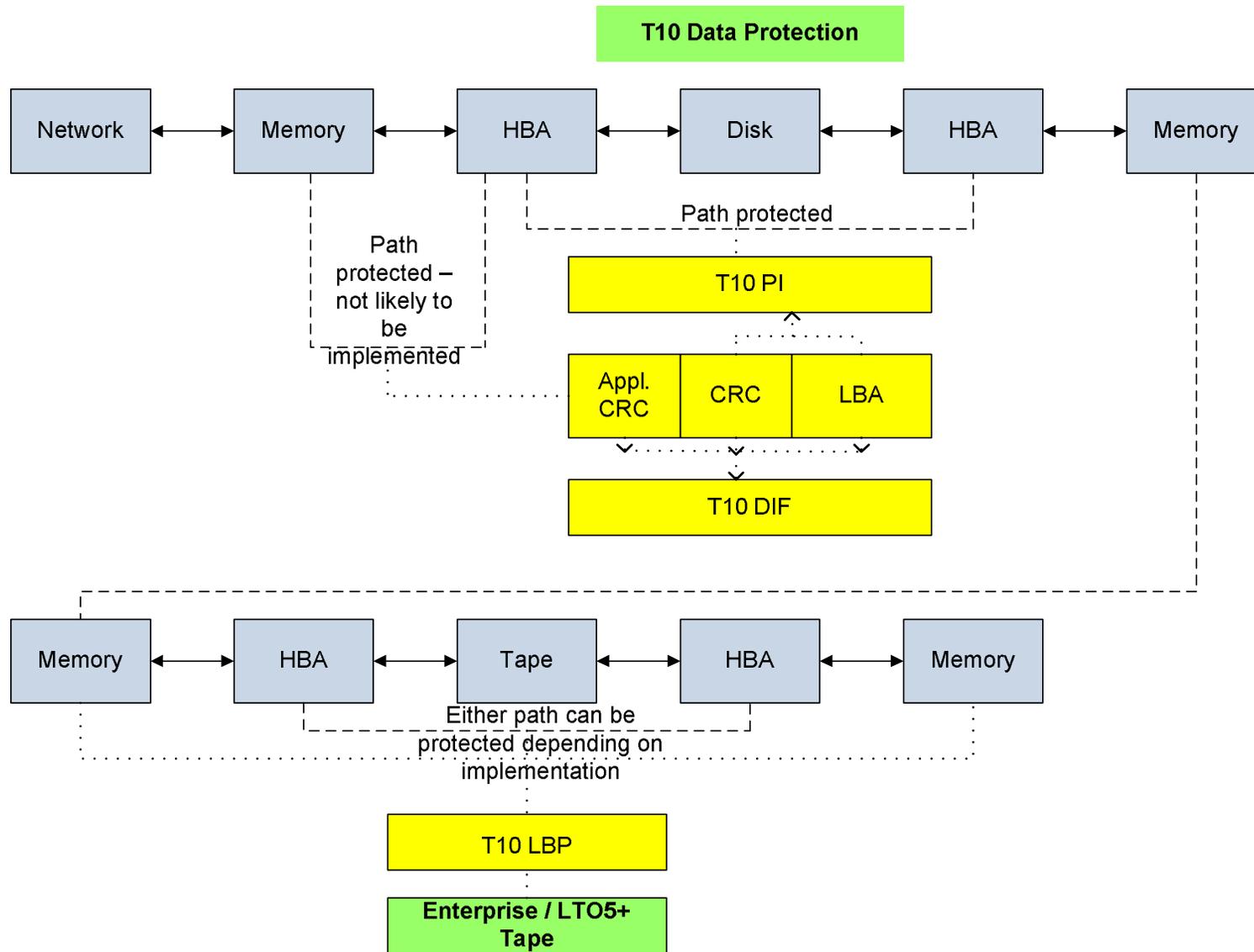
Instrumental, Inc.

hsn@instrumental.com

Background

- Problem
- History
- Challenges

What is protected and what is not



- Standards processes are long and arduous
 - Most standards take years to create - given the time it takes to come to agreement on details
 - Complex processes that require people with very detailed domain knowledge
- Vendors mostly run the standard processes
 - From low level device interface standards
 - ANSI T10 group handles the whole SCSI/SAS specification
 - To higher level standards that provide APIs
 - Openstack, openaxf and others
- Nothing about standards is easy and it is a thankless and yet required task
 - But do long-term storage managers get what they need and if not why?

- Modern history begins with TCP/IP and IETF
 - That was really the beginning of community standards for IT
 - The US government was heavily involved in the IETF along with many universities and companies
- Then came POSIX (Portable Operating System Standard)
 - Included contributors from the IETF
- Then came ANSI groups such T10 and T11
 - Far less involvement from anyone other than vendor companies in developing standards
- USL(Unix System Laboratory), XOpen, and then the Open Group for Unix standards
- By the late 1990s only vendor companies were working on these low level standards
 - Upper layer standards process was very stagnant
 - Upper layers are standards higher in the stack (e.g., ones that do not deal with devices)

Challenges - Today



HIGH PERFORMANCE INNOVATION

- Standards are technically challenging.
- It requires years of domain knowledge and detailed understanding of the technology
 - But how do we get our requirements into the standards process?
- Who is looking at things end to end?
 - Example: T10 (PI/DIF) standard supports an application checksum but it cannot be passed from the application to the file system
 - Requires a change to POSIX which is not going to happen
- Standards groups do not seem to be coordinating well from the lowest layers to the highest layers
 - Different vendors and different people involved
 - This is why it was good to have the US government previously involved to coordinate end to end requirements

- Everything seems to be cyclical in the computer industry, but the standards processes do not seem to be following that trend:
 - User requirements for standards for end to end integrity to benefit archives and long term preservation have never been adopted by anyone
 - Open Archival Information System (OAIS) was never really adopted by anyone
- There are some proprietary solutions available for archives that address end to end integrity; and
- There are some open standards, but none that address end to end integrity.
- So, there are no open solutions that meet the needs of archival community.