*From MSST2010*

*From MSST2011 Keynote talk this morning by Michael Ernst of BNL*

# Role of Tape in LHC Computing

**Assumption in early phase of LCG Project that there would be no Tape by the time LHC data taking starts, but …**

- Technical Evolution of Tape Technology leading to unprecedented capacity growth and reduced cost
  - Native capacity of tape cartridge surpassed capacity of biggest disk drive reducing price/GB
    - Expect ~60TB/cartridge by the end of this decade, further improving price/capacity advantage of tape
  - LTO/LTFS adds an important dimension that could help to improve access times

- Tape drives & media have steadily improved in reliability
  - Less frequent labor-intensive migrations to next gen technology
  - Lower BER and longer useful life than disk making tape better suited for long-term data-retention requirements
  - Cost-effective in terms of operating effort: at BNL ~1 FTE per 5 PB (MSS S/W + tape library and drive H/W)

# High availability tape

- Strategies for high availability tape archives
  - Harry Hulen, Consultant to IBM's HPSS project
- Tape logical block protection
  - Kevin Butt, IBM and T10/SSC Working Group
- High availability tape media
  - Todd Abrahamson, Imation Government Services
- Tape data integrity verification
  - Molly Rector, Spectra Logic
- BlueWater's archive at NCSA with RAIT
  - Michelle Butler, NCSA
- Panel discussion – YOU WILL STAY
  - Save questions for the panel

# Two main uses of tape in our community

- Disk space management
  - Free up disk space by migrating least recently used data to tape
  - Back up disk data by early writing a copy to tape
  - Data may change requiring another cycle of writing to tape
  - Often automatic – no user awareness

- Archive
  - Deliberate, user initiated
  - Object probably immutable (no changes once written)
  - Longevity requires monitoring and media upgrades

- Some sites do both

- My talk focuses on Archive

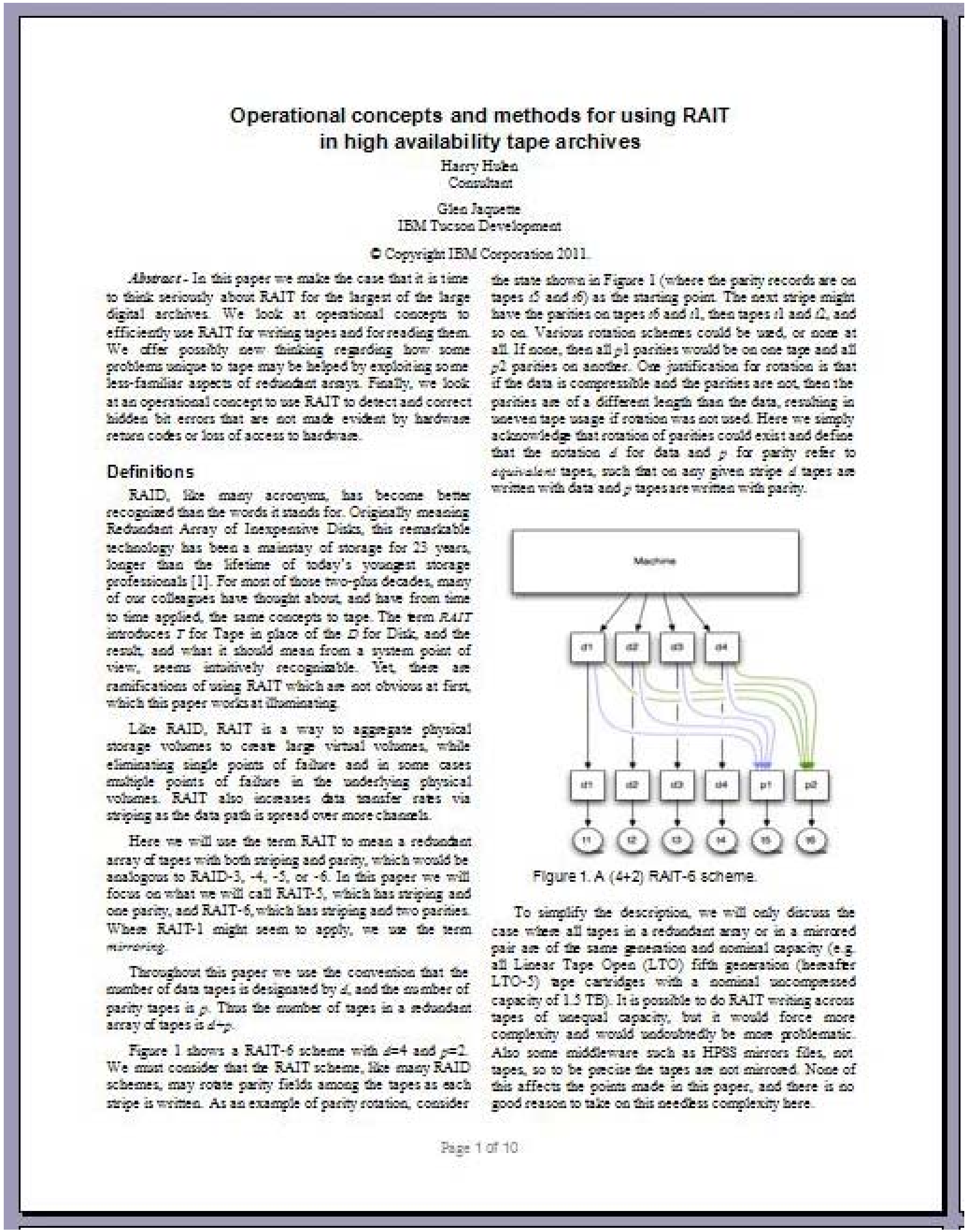
# Strategies to consider for “forever” archives

- A digital archivist should always apply a checksum to a valuable file before turning the file over to storage middleware of any kind
- Middleware and hardware should further protect data with block-level checksums, using T10 capabilities
- Drives with T10 Read-Verify greatly reduce cost of periodic tape scans
- RAIT-5 or -6 can provide better protection at less cost than mirroring can provide
  - RAIT-6 can both detect and correct a hidden error
  - RAIT is fast to write but slow to read; therefore there must be a strategy to make reads efficient
- Remote asymetric mirroring: single tape at home site, RAIT-6 at an “iron mountain” site
- Migration to new media every 5 years is an economic rule, not an archivist’s rule
  - RAIT and Read-Verify can increase the time between migrations of tape archives to new media
  - Tape drive obsolescence is often cited but is more like 15 years, not 5
  - Migrating sooner does save library slots and floor space

Look for my paper

***Operational concepts and methods for using RAIT in high availability tape archives***

in the Symposium section of the conference web pages



# Operational concepts and methods for using RAIT in high availability tape archives

Harry Hulen
Consultant

Glen Jaquette
IBM Tucson Development

*Abstract* - In this paper we make the case that it is time to think seriously about RAIT for the largest of the large digital archives. We look at operational concepts to efficiently use RAIT for writing tapes and for reading them. We offer possibly new thinking regarding how some problems unique to tape may be helped by exploiting some less-familiar aspects of redundant arrays. Finally, we look at an operational concept to use RAIT to detect and correct hidden bit errors that are not made evident by hardware return codes or loss of access to hardware.

## Definitions

RAID, like many acronyms, has become better recognized than the words it stands for. Originally meaning Redundant Array of Inexpensive Disks, this remarkable technology has been a mainstay of storage for 23 years, longer than the lifetime of today's youngest storage professionals [1]. For most of those two-plus decades, many of our colleagues have thought about, and have from time to time applied, the same concepts to tape. The term *RAIT* introduces *T* for Tape in place of the *D* for Disk, and the result, and what it should mean from a system point of view, seems intuitively recognizable. Yet, there are ramifications of using RAIT which are not obvious at first, which this paper works at illuminating.

Like RAID, RAIT is a way to aggregate physical storage volumes to create large virtual volumes, while eliminating single points of failure and in some cases multiple points of failure in the underlying physical volumes. RAIT also increases data transfer rates via striping as the data path is spread over more channels.

Here we will use the term RAIT to mean a redundant array of tapes with both striping and parity, which would be analogous to RAID-3, -4, -5, or -6. In this paper we will focus on what we will call RAIT-5, which has striping and one parity, and RAIT-6, which has striping and two parities. Where RAIT-1 might seem to apply, we use the term *mirroring*.

Throughout this paper we use the convention that the number of data tapes is designated by $d$, and the number of parity tapes is $p$. Thus the number of tapes in a redundant array of tapes is $d+p$.

Figure 1 shows a RAIT-6 scheme with $d=4$ and $p=2$. We must consider that the RAIT scheme, like many RAID schemes, may rotate parity fields among the tapes as each stripe is written. As an example of parity rotation, consider the state shown in Figure 1 (where the parity records are on tapes $t5$ and $t6$) as the starting point. The next stripe might have the parities on tapes $t6$ and $t1$, then tapes $t1$ and $t2$, and so on. Various rotation schemes could be used, or none at all. If none, then all $p1$ parities would be on one tape and all $p2$ parities on another. One justification for rotation is that if the data is compressible and the parities are not, then the parities are of a different length than the data, resulting in uneven tape usage if rotation was not used. Here we simply acknowledge that rotation of parities could exist and define that the notation $d$ for data and $p$ for parity refer to *equivalent* tapes, such that on any given stripe $d$ tapes are written with data and $p$ tapes are written with parity.

Figure 1. A (4+2) RAIT-6 scheme.

To simplify the description, we will only discuss the case where all tapes in a redundant array or in a mirrored pair are of the same generation and nominal capacity (e.g. all Linear Tape Open (LTO) fifth generation (hereafter LTO-5) tape cartridges with a nominal uncompressed capacity of 1.5 TB). It is possible to do RAIT writing across tapes of unequal capacity, but it would force more complexity and would undoubtedly be more problematic. Also some middleware such as HPSS mirrors files, not tapes, so to be precise the tapes are not mirrored. None of this affects the points made in this paper, and there is no good reason to take on this needless complexity here.