RAID
PARITY AND STRIPING

ADAM ANDRE

DIGITAL PRESERVATION

FALL 2017

WHAT IS RAID?

- Redundant Array of Independent Disks
- - storage technology that combines multiple disk drives into what a computer's OS will recognize as one or more logical units
- Goals - Increased reliability, availability, performance, and capacity
- "RAID Levels"

BEFORE RAID

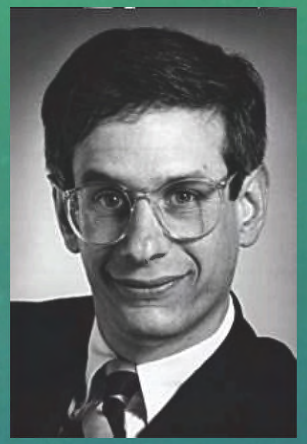
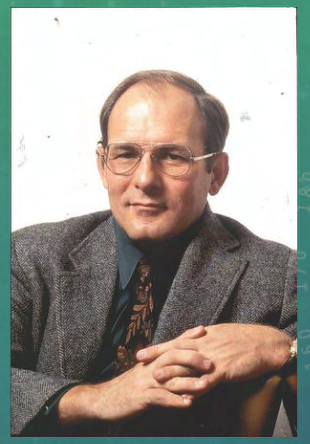
- 1956: International Business Machines Corporation (IBM) announces the RAMAC 305, the first hard disk system that holds 5MB of data
- 1961: Ampex develops helical scanning video recording which will later be adapted for high-capacity tape storage
- 1973: IBM's hermetically sealed Winchester Hard Disks become the standard design for disc drives
- 1979: Seagate Technology begins producing hard disk drives for desktop computers
- 1986: Small Computer System Interface (SCSI) is standardized by the American National Standards Institute

IBM Winchester

RESEARCH AT UNIVERSITY OF CALIFORNIA, BERKELEY

- 1983: Randy H. Katz joins Berkeley Faculty – works with Michael Stonebreaker on Ingres Database Project
  - Upper limits of performance in database systems
  - Found bottleneck in writing transaction commit log
  - Needed greater I/O rates
- XPRS Project – how to support an extensible database system on a "shared disk" multiprocessor architecture
  - Katz, Stonebreaker, and David A. Patterson
- 1986: Katz purchases an Apple MacPlus, possibly the first PC with SCSI connector
- 1987: Katz and Patterson begin research of large-scale multidisk storage systems
  - Reliability – Garth Gibson (a graduate student) became expert on performance-reliability tradeoffs

THE BERKELEY RAIDERS

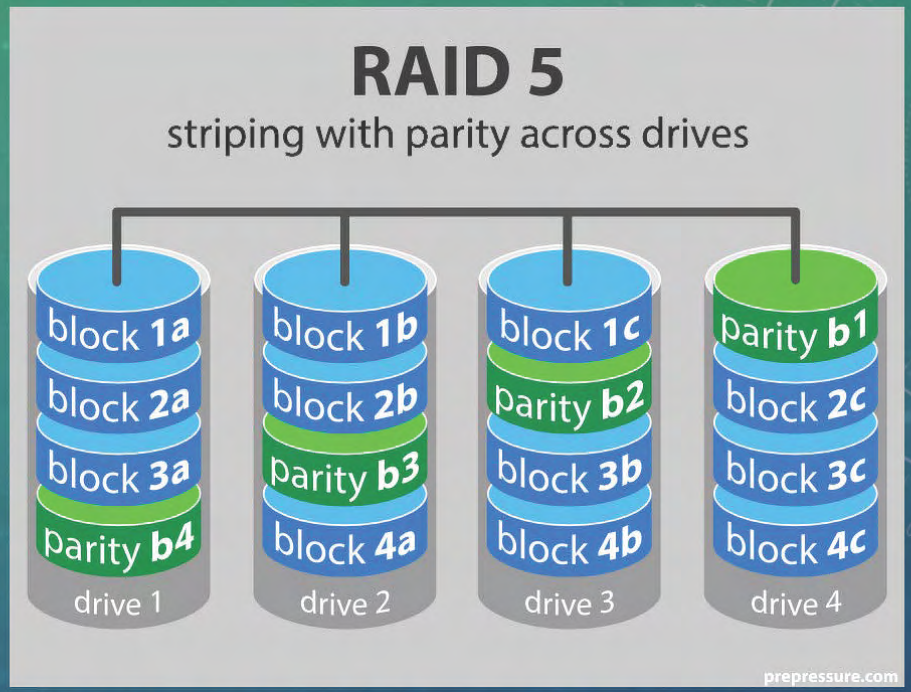




- Summer 1987: Patterson and Gibson attend disk technology course at Santa Clara University taught by Al Hoagland
  - Hoagland – California Digital Computer Project (1951) – worked with IBM to build RAMAC disk drive in 1956
- Fall 1987: Katz, Patterson, and Gibson begin developing taxonomy for RAID
  - Tandem Computers – influence for "RAID 1"
  - Thinking Machines – influence for "RAID 2"
  - Maxstor and Micropolis – influence for "RAID 3"
  - IBM – influence for "RAID 4" and "RAID 5"
- November 1987: send *A Case for Redundant Array of Inexpensive Disks (RAID)* to reviewers including Hoagland
- The report goes viral within the storage industry – is officially published in May 1988

FIRST RAID PROTOTYPE
RAID 5 ACROSS 32 DRIVES

RAID 5 AND 6 STRIPING WITH PARITY

- Striping – dividing a body of data into blocks and spreading those blocks across multiple storage devices
- Parity - technique that checks whether data has been lost or written over when it is moved from one place in storage to another or when it is transmitted between computers
- Parity bit - sometimes referred to as a check bit, is used to identify whether the moved bits arrived successfully. Also used to rebuild data after a drive failure



STRIPING CONFIGURATION

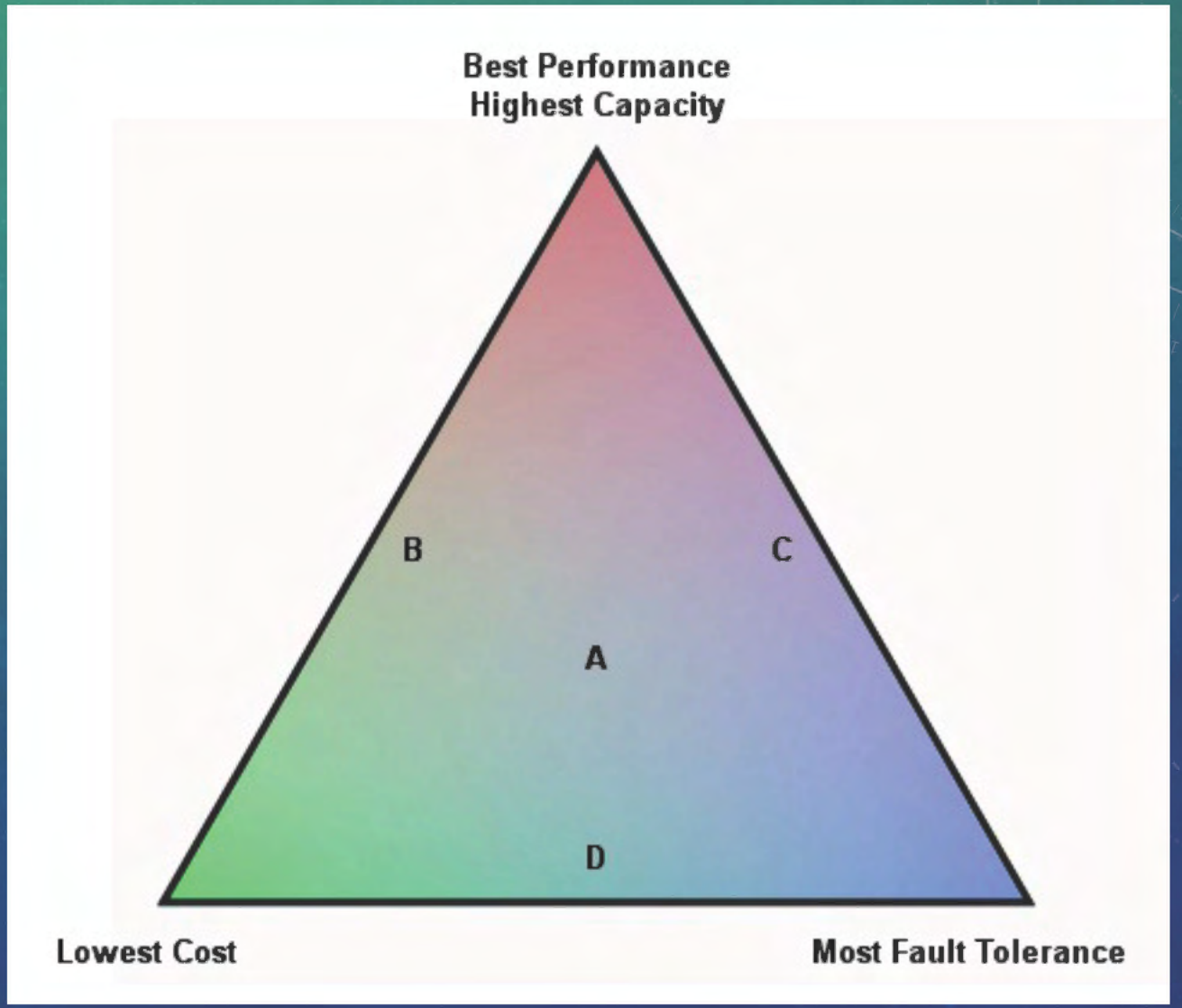
- Bits – 0 , 1
- Bytes - 01101000
- Blocks – “size on disk”
- Chunks – abstract; amount of data that an application processes from a disk at a time
- Block and Stripe size is determined by RAID Controller

PARITY – EXCLUSIVE/OR (XOR)

- A computer processing function that produces a parity bit based on the sequence of 1s and 0s

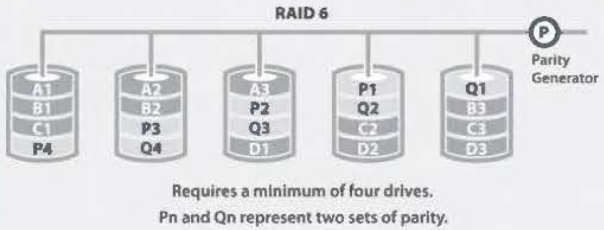
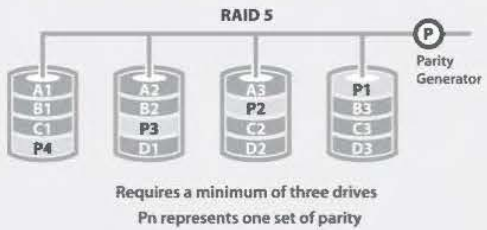
TRADEOFFS

- "Fast, cheap, good: choose two”
- Capacity
- Performance
- Cost
- Reliability (Fault Tolerance)

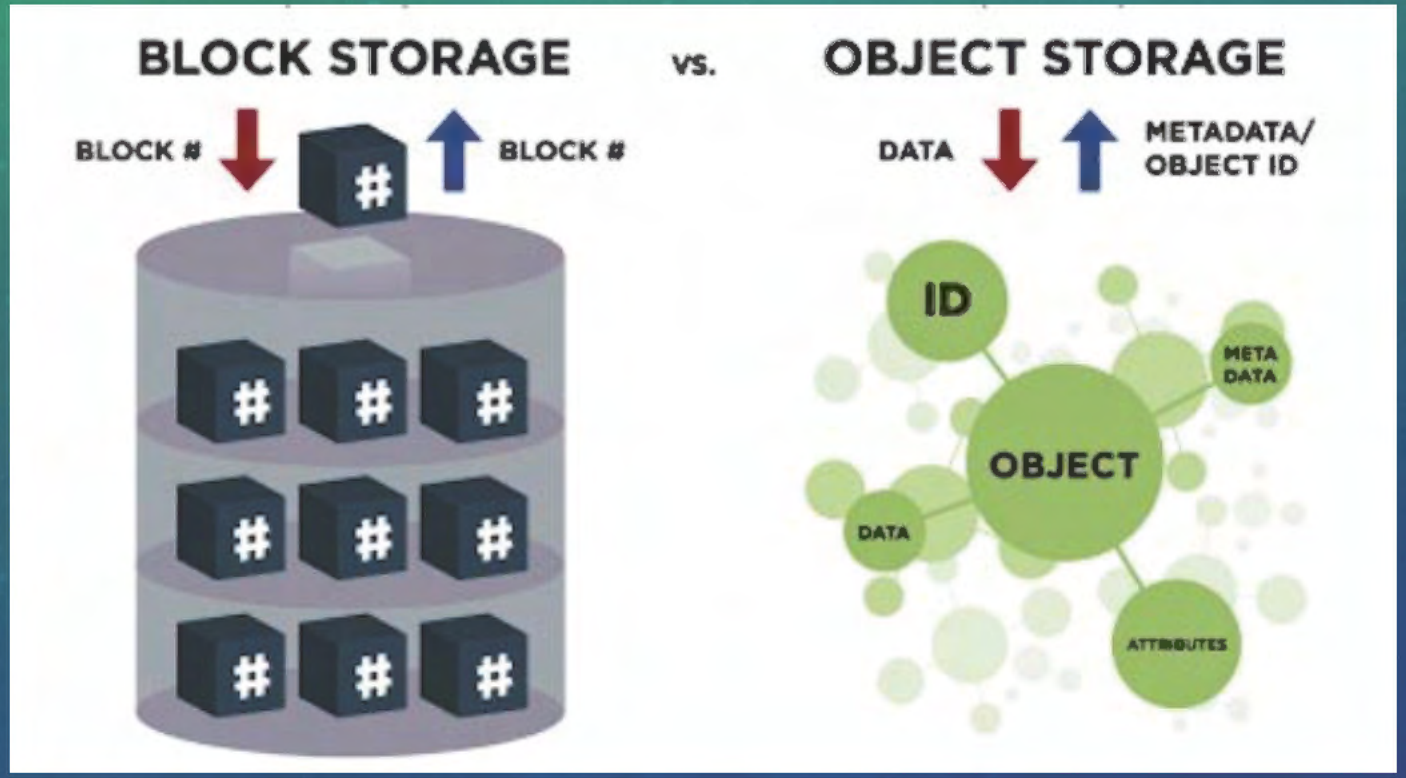


- Cinema Studies Archive
  - QNAP NAS
  - RAID 6
  - 6-Bay – 3 TB/drive
  - HGST Drives
  - Kathy thinks drives were installed in 2014
  - Software is browser based

- IndieCollect
  - Synology Server
  - RAID 6
  - 12-Bay – 10 TB/drive
  - 10 GB Ethernet connection

FEATURES	RAID 6	RAID 5
Diagram	 Requires a minimum of four drives. Pn and Qn represent two sets of parity.	 Requires a minimum of three drives Pn represents one set of parity
Minimum # of Drives	4	3
Data Protection	Two drive failure	Single drive failure
Read Performance	High	High
Write Performance	Low	Low
Read Performance (Degraded)	Low	Low
Write Performance (Degraded)	Low	Low
Capacity Utilization	50% - 88%	67% - 94%
Typical Applications	Data Archive, Backup to Disk, High Availability Solutions, Web Servers, Servers With Large Capacity Requirements	Data Warehousing, Web Serving, Archiving

RAID VS. OBJECT-BASED STORAGE



OBJECT-BASED STORAGE

- Data stored as individual object rather than in a hierarchical file system
- Each object is assigned a unique identifier
- No limit on type or amount of metadata
- Erasure Coding (Polynomial Oversampling)