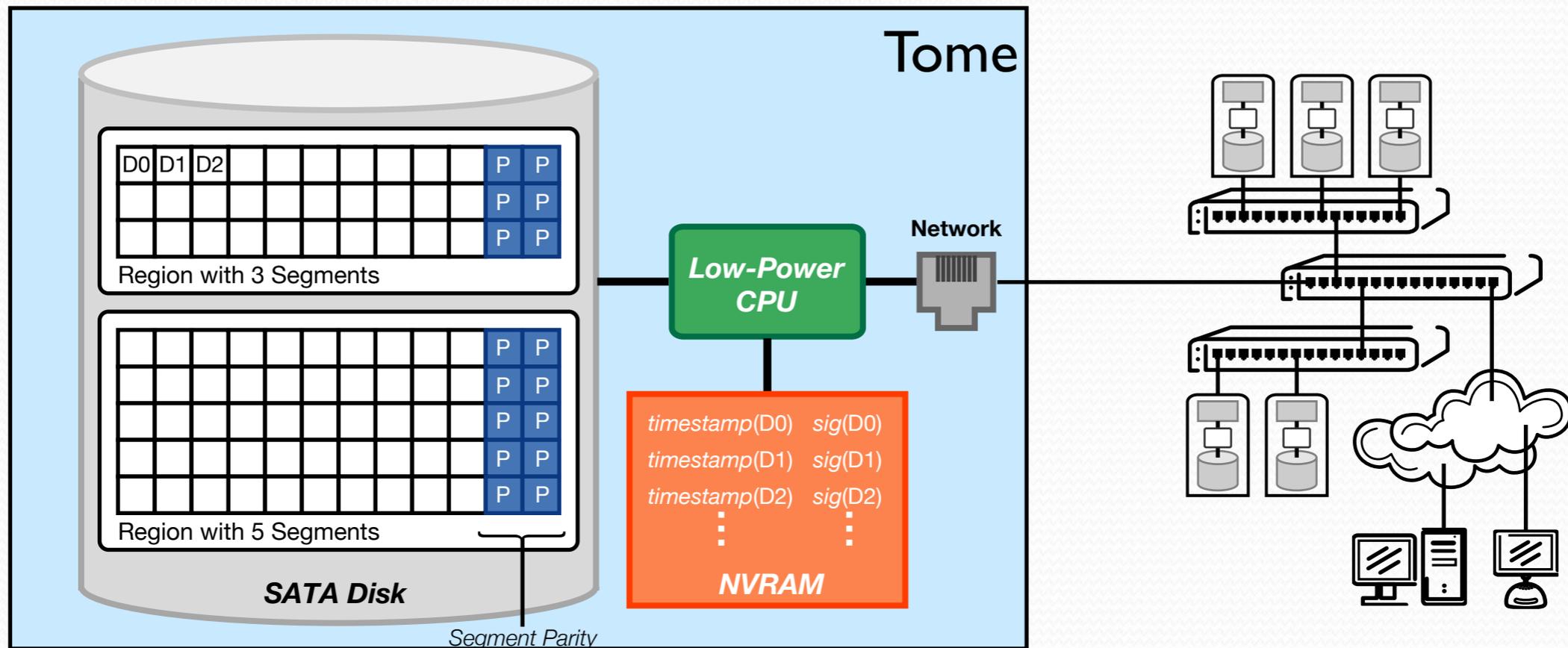# Archival Storage Architectures

## Ethan L. Miller
### University of California, Santa Cruz

# Pergamum Tome



- Low power CPU & DRAM: NAS functions, consistency checking, parity ops

- SATA hard drive: low-cost, persistent storage
  - Data stored in segments with parity appended

- NVRAM: metadata storage (data signatures, time stamps, index, etc.)

- Ethernet controller: commodity interconnection network

# Energy Consumption

| Component | Active Power Usage | Idle Power Usage |
|---|---|---|
| SATA Hard Drive | 7.5 W | 0.3 W |
| Low-Power CPU board (w/ NIC) | 2 – 3.5 W | < 0.3 W |
| NVRAM | < 0.6 W | 0 W |

- Power usage is proportional to the square of the voltage
  - (½) Voltage ➔ (¼) power consumption
- Modern laptop processor consumes 31W
- Within the limits of Power over Ethernet
  - Could allow for very simple physical interface

# Two-Level Reliability Approach
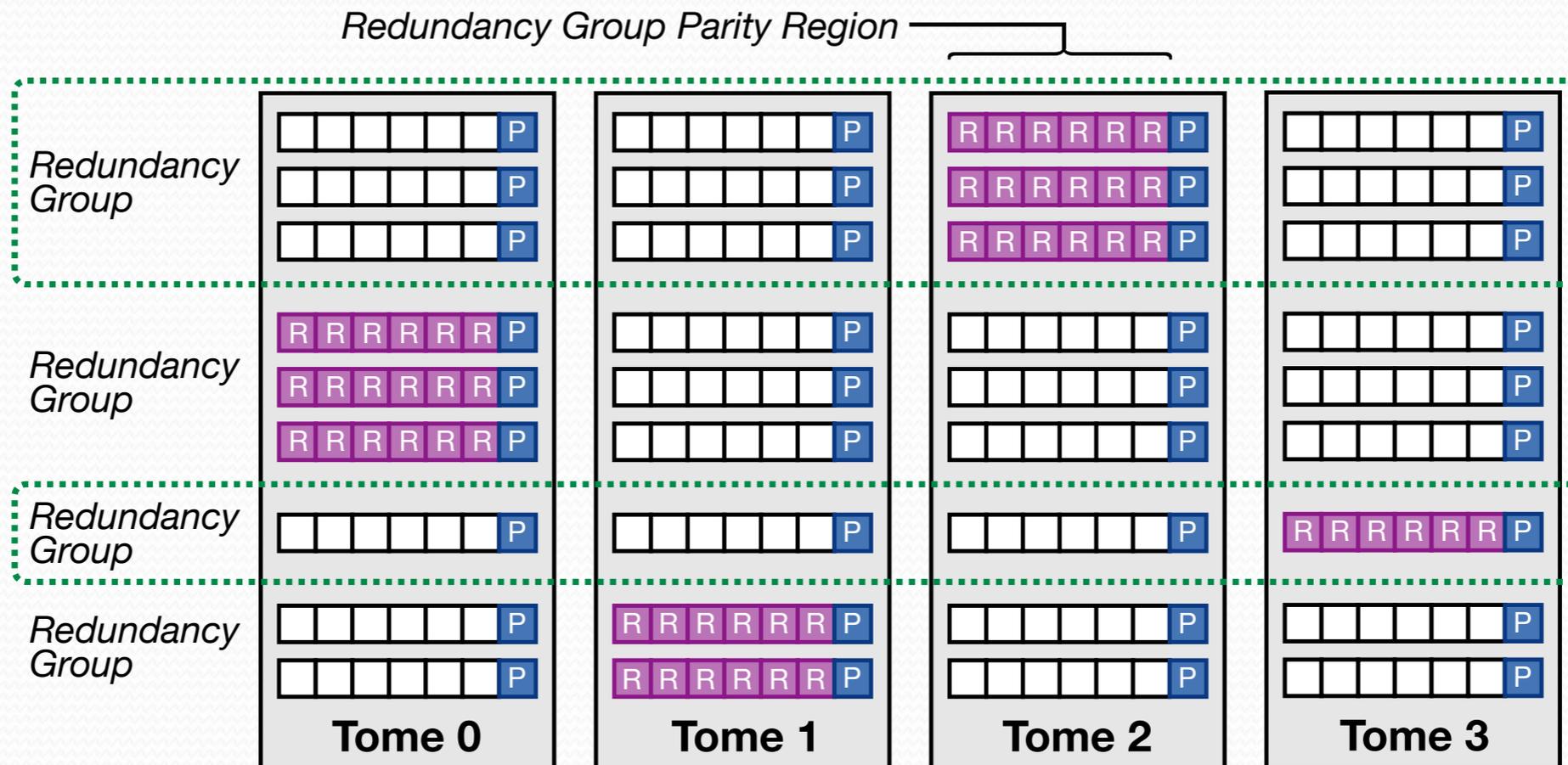


- Intra-disk - protection from latent sector errors
  - Increased reliability reduces scrubbing needs
  - Fix errors with as little spin-up as possible
- Inter-disk - protection from device loss

# Consistency Checking: Scrubbing



Tome 0
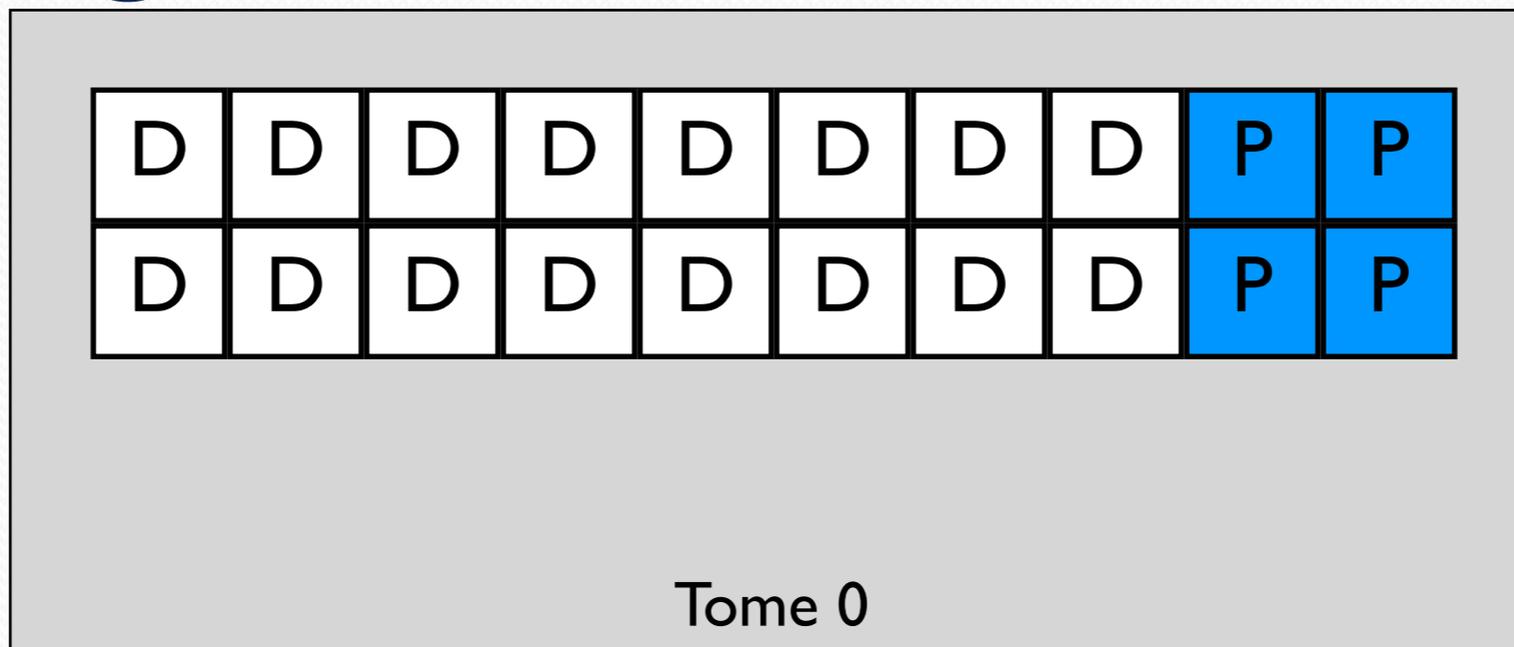
- Periodically verify data stored on a tome (by segment)
  - Read segment, including parity, and verify consistency
  - Record signatures in NVRAM

- Need to scrub entire disk ... but not all at once!
  - Scrub a gigabyte or two each time the disk is powered on
    - Checks basic disk functioning & detects dead (or nearly so) disks
  - Ensure that entire disk is scrubbed within a year
    - Intra-disk parity makes data loss highly unlikely

# Archival Storage: Usage Patterns

- Designing architectures is difficult without understanding *how* they'll be used
- Goal: study archival storage workloads to drive design of long-term storage systems
- We studied three storage systems (two public, one HEC)
  - HEC is similar to "accepted wisdom"
  - Public archives exhibit *very* different behavior:
    - Not always write-once
    - Very large batch reads from crawlers and other sources
- We'd like to study more archives: anyone have data we can analyze?