Building Better Long-Term Archival Storage Systems

Ian F. Adams  Ethan L. Miller  Mark W. Storer
UC Santa Cruz  NetApp
How is archival storage used?

- Details of archival storage workload are important!
  - How often are data accessed (read & write)?
    - How dense are accesses?
  - Are there patterns in file accesses?
  - Are all accesses (users) equivalent?

- Why should we care?
  - Archival systems being designed around speculation and out of date information
  - Systems may be optimized for the wrong workload!
  - Data is either out of date, unrelated, or nonexistent
    - Last tertiary storage studies were almost 2 decades ago
    - More recent workload studies are unrelated
    - Nobody has looked at modern archival use-cases
Contributions

- Our work is bringing our knowledge of archival storage behavior up to date

- Our contributions:
  - Examine common assumptions in archival storage
    - in particular “write-once, read-maybe”
  - Examine impact on current and future architectures
  - Begin looking towards tools for future studies
Systems studied thus far

- **Los Alamos National Laboratory**
  - 55+ Million files, 1.3 PB
  - 13 months of daily FSstats histograms

- **Washington State Digital Archives**
  - 28 million web viewable records, 10+ TB
  - 3 years of record metadata and user access logs

- **California Department of Water Resources**
  - 56,000 reports on water table data, 2.5 GB
  - 3 years of access and update logs
Prevalence of mass accesses

- Frequent mass accesses
  - Google accounts for 70% of water corpus retrievals
  - Integrity checking processes (not shown) account for 99% of retrievals to historical corpus
- Future migration converts “Read Maybe” to “Read Definitely”
  - New access API?
Prevalence of mass accesses

- Frequent mass accesses
  - Google accounts for 70% of water corpus retrievals
  - Integrity checking processes (not shown) account for 99% of retrievals to historical corpus
- Future migration converts “Read Maybe” to “Read Definitely”
  - New access API?
Other findings (in brief)

• Surprisingly frequent file updates

• Strong content locality within user sessions

• Limited content popularity system wide
  • i.e. limited per-record/file popularity

• Good data is very difficult to come by
Next steps

• Analyze new data
  • National Center for Atmospheric Research
  • Additional “public use” corpora
    • We’re looking for volunteers who have access traces

• Apply findings to archival storage system design
  • Knowledge about access density / frequency
  • Batch vs. on-demand requests
  • Data grouping
Current work: DAWN

- **Durable Array of Wimpy Nodes**
  - Long life, low usage of archival storage make cost considerations paramount
    - Magnetic media dominates due to low acquisition cost
  - Consider storage class memory (flash, etc.)
    - High acquisition costs
    - Lower running costs and infrastructure needs
    - Physically robust
  - We argue SCM should be considered for archiving
    - Can be competitive with magnetic media when using a self-managing architecture
    - Self management + low infrastructure needs + long life = lower TCO