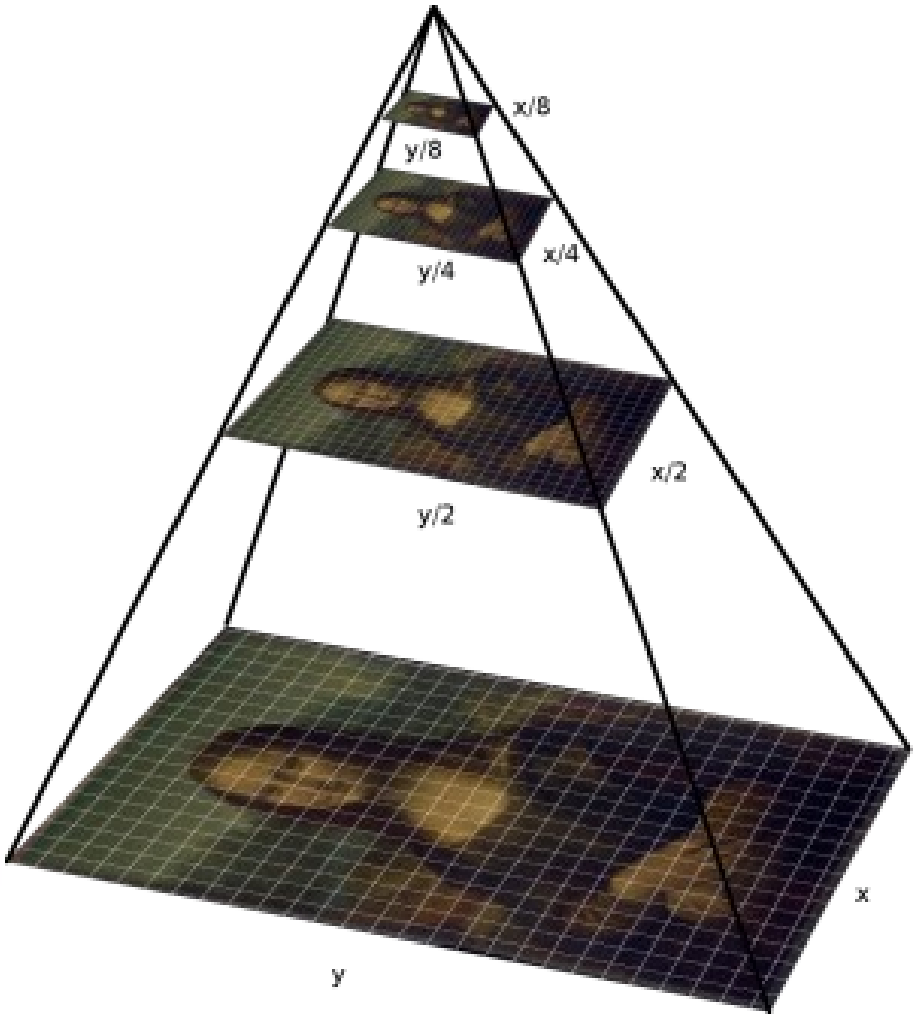
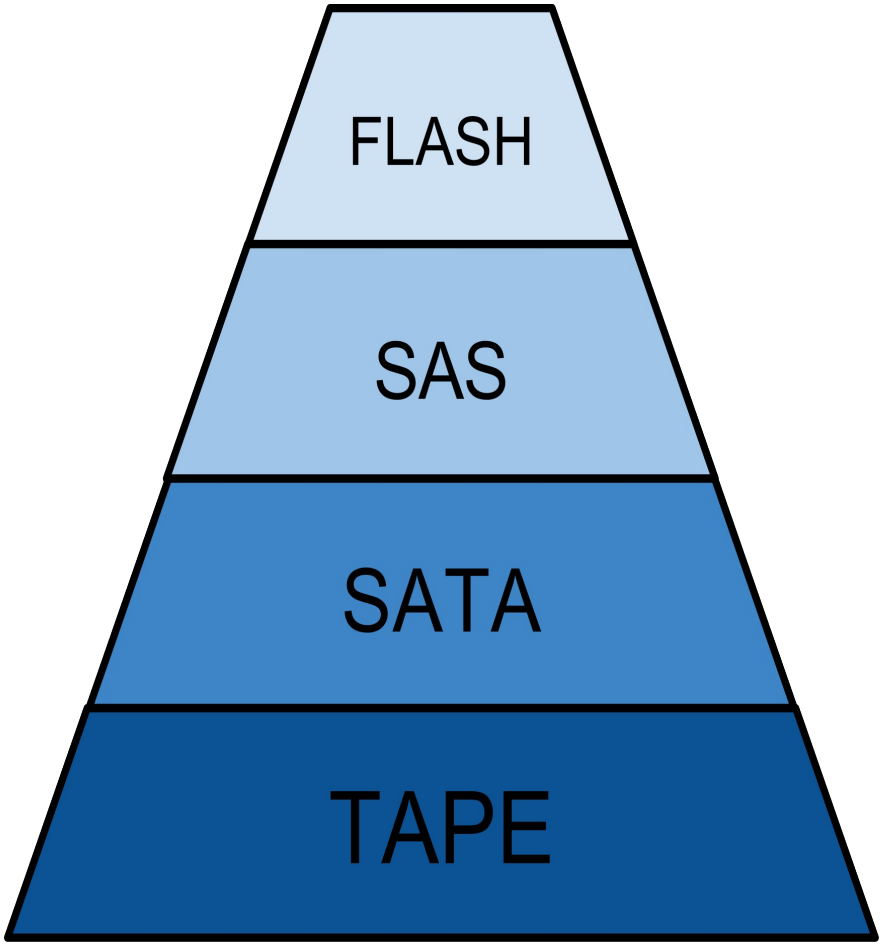
Multi-Resolution Image Store

– A Case of Size-Tiered Storage Systems

Ming Chen (mchen@cs.stonybrook.edu)

Introduction





Small Objects: Hot

- What
  - Metadata, Thumbnails
- How
  - Random: Searching, Indexing, Preview
- Where
  - Top tiers: RAM, FLASH
- Why
  - Throughput: op/sec
  - Inherent fit: Performance++;
  - Intuitive: peek before embark; size proportional

Large Objects: Not So Hot

- What
  - Large Images, Multimedia, Logs, Backup
- How
  - Sequential: Streaming, Scan, Mapping
- Where
  - Bottom tiers: SATA, TAPE
- Why
  - Throughput: mb/sec
  - Inherent fit: Capacity--; Cost--;
  - Slow seeks amortized by fast sequential I/Os

Implementation

- Schema Based on KV Store
- Log-structured Merge Tree
- Compaction (Temporal Locality)
- Multi-Tier Tablet Server Storage Layer (GTSSL)
- Deal with opposite extremes

Contributions

- Size-Tiered Property of Workloads
- Metadata Management
- Fast Prototyping of Specialized Storage Systems
- Take Advantage of Standards (JPEG 2000)

Feasibility

- Hardware
  - SSD, SAS, SATA drives
  - Servers
- Software
  - Berkeley DB, Level DB, KVDB
  - GTSSL
  - Filebench
- Intelligence
  - CSE602
  - FSL

Benchmark

- Synthetic Workloads
  - Access Model Presented in Haystack
  - Filebench
- Real Workloads
  - IIPImage deployments
  - TODO

References

- GTSSL, An Efficient Multi-Tier Tablet Server Storage Architecture
- Finding a needle in Haystack: Facebook's photo storage
- IIPImage, http://iipimage.sourceforge.net/documentation/images/
- A study of irregularities in file-size distribution
- hFS: a hybrid file system prototype for improving small file and metadata performance
- A Pseudo-Infinite Multimedia Storage System, http://www.fsl.cs.sunysb.edu/project-multimediafs.html
- Unifying biological image formats with HDF5
- In-Kernel Berkeley DB Databases and Transactional Storage, http://www.fsl.cs.sunysb.edu/project-kbdb.html