Towards Performance-Assured Cloud Storage

Yusuke Tanimura
yusuke.tanimura@aist.go.jp

Resource Coordination

- Services on the Cloud may use ...
  - 3 basic resources: Compute, Network and Storage
    - **Compute**: Server type (CPU model, memory size, etc.)
    - **Network**: External (client-server), Internal (server-server)
    - **Storage**: Object store, Block storage, File system, etc.
      - Non-volatile, Accessible from servers

A bottleneck may exist in Storage.
- Access congestion
- Low performance
How to achieve storage QoS?

- Over-provisioning
  - Trade-off with cost efficiency

- Adaptive / reactive QoS setup
  - Runtime control only
  - Some users may be rejected at runtime in the FCFS model due to running out of resources.

- Advance reservation (Our approach)
  - Reserve performance with start and end time of the access, after negotiation.
    - Reservation interface
    - Performance-oriented resource allocation
    - Runtime I/O control

Prototype Development

Our proposed storage system built with Papio
Advance Reservation

- Expose the performance reservation interface to clients.
- We are now using GNS-WSI3.
  - A Web-Services based, resource reservation protocol
    - Handle multiple-types of resources (Compute, Network, Storage, etc.)
    - Defined in the G-lambda project (http://www.g-lambda.net)

End-to-end Performance Control

- All components involved with the reserved access on the I/O path should be controlled at the reserved time.
- We integrate the following into our system.
  - PSPacer: A network bandwidth control tool for Ethernet
  - PROBS: Software-based OSD implementation
    - Disk I/O scheduling with SSD
Access Interface

- Supported access pattern:
  - Currently we focus on **Throughput (MB/sec)** for applications which perform long and streaming-type access.
    - A single access is either read-only or write-only.
    - For write, only sequential access is supported.
- **API**
  - Provide a non-POSIX, client library
    - RID (Reservation ID) should be given at ‘open’ call.
  - Support Amazon S3 like **Bucket / Object semantics**
- **Higher-level tools**
  - FUSE?
    - We have implemented it for read-only access.
  - **Amazon S3 extension?**
    - Started development recently.

Status and Future Work

- Developing a prototype, Papio storage software
  - Demonstrated a video delivery service with our storage system and optical path networks in 2010.
    - Used PSPacer only for the storage I/O control at that time
- Plan to enhance both performance reservation and access interfaces
  - Standard protocol, requirement description, usage records, etc.

Demonstration
Thank you