Adaptive Replica Management for Large-scale Object-based Storage Devices

Wei Qingsong

Network Storage Technology
Data Storage Institute, Singapore
May, 2006
1. Introduction
2. Issues of Object Replication Scheme
3. Adaptive Replica Management Model
4. Implementation
5. Evaluation
6. Conclusion
1. Introduction

Replica management is basic and challenging issue

- **Replication Advantage**
  - Increase availability
  - Balance workload
  - Reduce access latency

- **Replication Cost**
  - Consistency maintenance overhead
2. Issues of Object Replication Scheme

- **Availability**
  - Minimum replica can be maintain to ensure the given availability

- **Consistency Maintenance overhead**
  - More replicas consumes more network resource to maintain consistency
3. Adaptive Replica Management Model

- Motivation
  - How many replicas the system should keep at least to maintain certain object availability?
  - How many replicas the system can support at most to maintain object consistency under a certain network environment?

\[
\begin{align*}
[1 - p^k]^m & \geq A_{\text{except}} \\
(S_{obj} + L_{msg}) \times F_u \times k & \leq \alpha B_{\text{sys}}
\end{align*}
\]

- From the above model, the minimal replica number and maximum replica number can be calculated for any given availability under certain network environment. Each OSD runs the model and dynamically adjusts the number of object replica.
4. Implementation

**Client**
- **App1**
- **App2**
  - ... (Elipses indicating more applications)

**Object Attributes**
- Minimal Replica Number

**Replicate**
- $O_1$
- $O_2$
- ... $O_n$

**Stripping**

**Metadata Server**
- Set File Availability (0.9)
- Set File Availability (0.99)

**OSD Cluster**

**Replication**

**Master OSD**
- Request Portal
- Access Counter
- Replica Manager
- Status Monitor

**Access Requests**
- Forward Request

**Request Scheduler**

**Slave OSDs**
- OS OSDs
- Object Replication Table
- OSD Weight Tree
5. Evaluation

- Availability vs Replica Number
- Latency (ms) vs Access Frequency
- Access Frequency vs Replica Number
- Time (Minutes) vs Replica Number
6. Conclusion

- This paper builds up a dynamic model to adapt to the changes of OSD clusters and satisfy file availability in reasonable cost.
- In the future work, we will introduce more object attributes to design OSD QoS model according to object-based storage requirements such as availability, access delay, I/O speed and workload.
Thank You!