Load Balancing in Ceph: Load Balancing With Pseudorandom Placement

Esteban Molina-Estolano, Carlos Maltzahn, Scott Brandt

University of California, Santa Cruz
The Load Balancing Problem

- Pseudorandom placement for distributed storage systems has several advantages
- Problem: pseudorandom placement makes load balancing harder
- Research platform: Ceph, object-based storage system developed at UCSC
Coincidental overload

clients

primary

replicas
Primary switching

clients

replicas

primaries
Read flash crowd

clients

primary

replicas
Read shedding

clients

primary

replicas
Read shedding

clients

primary

replicas
Write flash crowd

clients

primary

replicas
Write shedding
Write shedding

Serialization:
Easy with POSIX
HPC I/O extensions,
complicated otherwise

clients

primary

replicas
Preliminary Results

Throughput with overloaded OSD 0

Bandwidth (MB/sec)

Elapsed time (seconds)
Preliminary Results

Throughput with OSD 0 relieved by primary shifting

- Bandwidth (MB/sec)
- Elapsed time (seconds)