Fast Fail Read

Cloud-HDD

presented by Lawrence Ying (Google LLC)
on behalf of hdd-fast-fail@ocproject.net

Jan 10, 2019

Fast Fail Read Cloud-HDD Joint-Development

Fast Fail Read spec is developed under an OCP JDA (Joint Development Agreement), using the new OCP Cloud-HDD process as defined here:

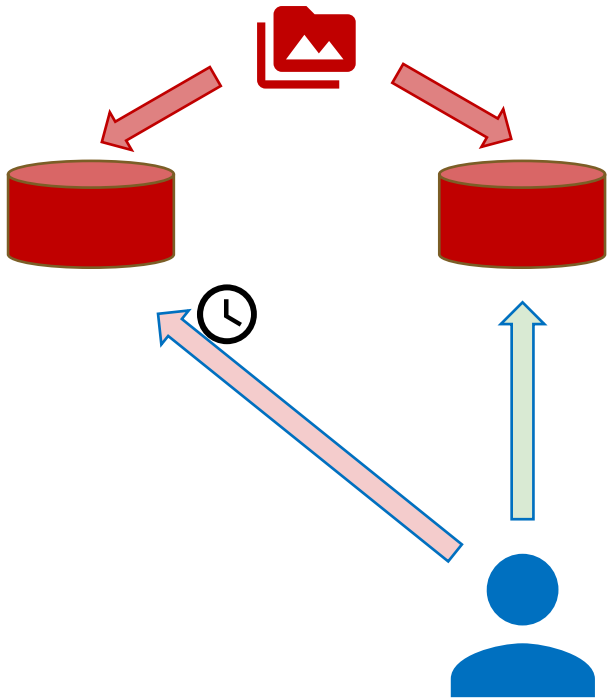
- https://www.opencompute.org/wiki/Storage/CloudHDD

Joint contributors to this **Cloud-HDD Fast Fail Read** spec (in alphabetical order):

- Avago (Broadcom), Facebook, Google, Huawei, Microsoft, Seagate, Toshiba, Western Digital

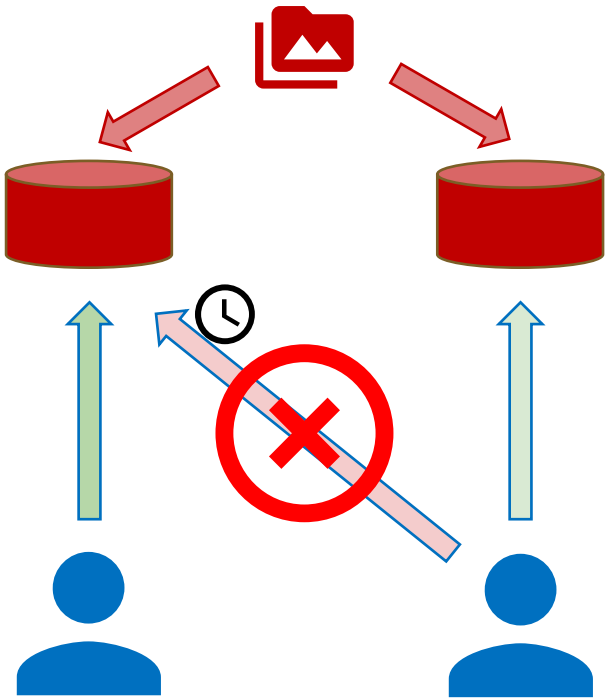
Problem Statement (Background)

- HDD can sometimes be slow to read (Ex/ 500ms read latency at 99.9%tile)
- For Cloud (large distributed file system), data is stored on >1 HDD in data center
- So when one HDD is slow to read, we can just read from another HDD instead

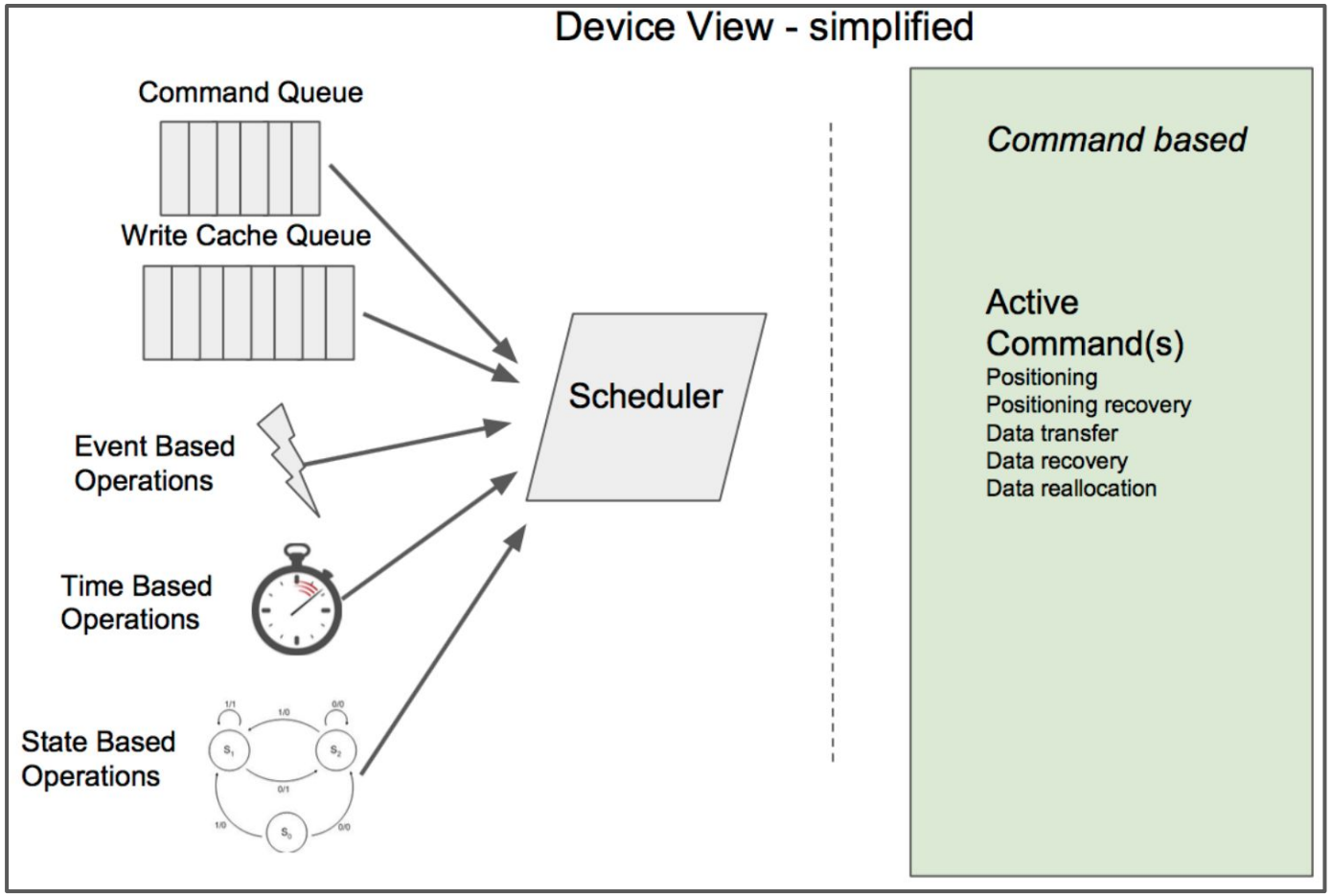


Problem Statement (continued)

- So we're now reading from another disk…
- And when this happens, we would prefer the first HDD abandon the read request (so it's "freed up" to do something else)



Cloud-HDD API Proposal: Create Two Timeouts



(left)
Inactive Command Limit

(right)
Active Command Limit

[Image courtesy of Seagate]

How to achieve “fast fail read” using the new API

By allowing host software to specify these two commands limits (aka timeouts) for an HDD, fast fail read behavior (as explained in the problem statement slides) can now be achieved.

For example, if non-queued (which is the simplest case):

- “Fast fail reads” can be achieved by setting a tight Active Command Limit.
- “Normal reads” can be achieved by having extremely relaxed or no Active Command Limit.

Partnership between OCP and T10/T13/SATA-IO

Many of the Cloud-HDD Fast Fail Read attendees are also active participants (and in some cases chairs and/or co-chairs) of the T10, T13, and SATA-IO Standards committees.

The working group consensus is to define a standard set of behavior and Host-to-HDD (Cloud software) usage specification in OCP, and to define the HDD API interface standards in T10/T13/SATA-IO as appropriate.

Expected Cloud-HDD Product Availability

Seagate, Toshiba, and Western Digital have expressed interest to enable the Fast Fail Read API in their future HDD products.

Facebook, Google, and Microsoft have also expressed interest to purchase HDDs with the Fast Fail Read API feature set enabled, and use these interfaces in the future.