skyera

The New Era of Storage

Power Cycling Challenge

Andy Tomlin – Aug 2013

Power cycling problem

- Solid state devices robustness under power cycling is known challenge for the industry
  - https://www.usenix.org/system/files/conference/fast13/fast13-final80.pdf
  - http://www.storage-switzerland.com/Articles/Entries/2013/4/10_Power_Failure_and_Flash_Storage.html
- There are multiple levels of the challenge
  - Power protection scheme
  - Robustness of solution
  - Testability of problem
  - (http://www.flashmemorysummit.com/English/Collaterals/Proceedings/2012/20120823_S304A_Tomlin.pdf)

Flash management system challenges

- Demands of power cycling are the opposite of demands of normal operation.
  - Normal operation requires minimum write amplification
  - Less meta data written the better
  - Meta data writing lags user data writing

- For power cycling
  - Time to ready key issue
  - Complexity is the enemy
  - Synchronizing write data and meta data

Power cycling protection options

skyera

- User data
  - Command completion after write to flash (DRAM expensive + has challenges with lower page corruption)
  - Secondary, faster (more expensive) non volatile memory (eg MRAM)
  - Hold up circuit to write data in buffer (cost, component reliability)
  - Binary cache (WA impact)
  - Combination
- Meta data
  - Save via non volatile or hold up circuit – same as user data
  - Deterministic loss and limited scanning
  - Full scanning (extremely slow)

Mapping schemes

- Logical to physical mapping tables
  - Tables written to flash contains logically grouped data based on host addresses
  - Often used with secondary higher level system map to enable finding all meta data quickly.
  - Big advantage for power up as upon loading system map (and possible minimal scanning) the mapping table is coherent and host read write can start
    - For lower end client type applications that do not use external DRAM – only solution
    - For mid range SSD with external DRAM provides fairly consistent time to ready independent of capacity point
    - For high end flash appliances with thin provisioning – only solution

Mapping schemes

- Physical to Logical mapping tables
  - Tables written to flash contains physically grouped data based on physical addresses
  - WA advantage in normal operation as mapping matches write process resulting in less overhead
  - Problematic for power cycling as map only coherent after ALL mapping data read.
    - Makes time to ready capacity dependent
    - Time to ready gated by FW processing time

Conclusions

- Power cycling challenge typically drives flash management system design
- Complexity of problem and challenge of testing typically makes it the long pole for product development
- Anything that is good for normal operation is usually bad for power cycling
- Compression makes problem more difficult
- Thin provisioning also makes the problem more difficult
- Snapshots make problem more difficult
- The right architecture plus experience and strength of design team critical to success