Offloaded Data Transfer [ODX] for SPC4/SBC3 storage

By
Amit Anandram Luniya
[amit.luniya@emc.com]
Agenda

- Data Movement and Inefficiencies with traditional data movement
- ODX Insight
- Overview. How it works...
- ODX capable storage
- ODX copy sequence
- Token exchange
- ODX Read/Write Operations, Error Handling.
- Performance Tuning Parameters
- ODX and Hyper-v
- Hyper-v Operations with ODX
- ODX performance numbers
- ODX Usage models
Data Movement
Inefficiencies with traditional data movement

- Data movement consume CPU, Memory and network on the host(s).
- Data flow from and to same storage system
- Network transport bottleneck.
ODX Insight

- Microsoft has developed ODX to advance storage data movement.
- Supported with SPC4/SBC3 comply storage.
- Overcome traditional buffered copy inefficiencies.
- Minimizes latencies and maximizes array throughput.
Overview. How it works...
ODX Capable Storage

- ODX uses three new SCSI commands [XCOPY commands introduced in SPC-4]

1. POPULATE TOKEN (Also known as ‘Offload read request’)
2. RECEIVE ROD TOKEN INFORMATION [RRTI]
3. WRITE USING TOKEN [WUT] (Also knows as ‘Offload write request’)
ODX copy sequence

- PT Command to storage.
- Storage respond with status
- Read token [RRTI]
- Storage returns token [Result of PT]
- offload write command [WUT]
- RRTI command to read result of WUT.
- Returns token info, copy status.

Storage

Server

ROD Token Created

Copy manager in action.
ODX Token Exchange

- Many such commands [PT, WUT & RRTI] sent/copy.
- Separate token per file, processed in increments.
- Vendor specific 512 byte string represents data range.
- ROD token opaque, unique, and secure.

- Copy Engine [Data transfer application]
  - It must ensure both copy source LUN and copy destination LUN are ODX capable.

- Copy Manager: [Vendor specific]
  - Responsible to handle offloaded copy, maintains and validate tokens.
ODX Read/Write, Error Handling

- File data is not seen by the I/O stack.
- Must be sector aligned.
- FSCTL_OFFLOAD_READ:
  - Instructs storage to generate and return a “Token”.
- FSCTL_OFFLOAD_WRITE:
  - Performs data movement.
- Same storage link (I_T Nexus).

Error Handling:

- Fallback to traditional copy.
- Copy engine resume from first failure point.
- ODX with failover cluster:
  - Offload application must be cluster aware.
Performance Tuning Parameters

- Minimum file size requirement
  - Minimum copy offload file size set to 256kb.
  - < 256KB then legacy copy

- Storage vendor support up to “N” copy operation

- Optimal Transfer size
  - Specified by the storage target device to host.
  - Set to 64 MB, if the target storage device does not provide.
  - Set to 256 MB, if storage/target reports > 256MB.
ODX and Hyper-V

Source: hyper-v.nu
Hyper-v operations with ODX

- VM operations benefit from integration:
  - ODX useful in creating fixed-size VHD.
  - Well known token “ZERO ROD” for bulk zeroing.
  - Used in maintenance operations for VHD [e.g. snapshot delete]
  - Live migration.
Sample Performance Numbers

- **Create 100GB Fixed-Size VHDX File**
  - ODX disabled: 218.17 seconds
  - ODX Enabled: 2.12 seconds

- **16GB VHDX File Copy**
  - Seconds: ODX Disabled - 97, ODX Enabled - 40
  - MB/s: ODX Disabled - 169, ODX Enabled - 410

- **Multiple Small File Copy**
  - Seconds: ODX Disabled - 13.12, ODX Enabled - 8.22
  - MB/s: ODX Disabled - 99.9, ODX Enabled - 159.45

Source: en.community.dell.com
ODX Usage Models

- Hyper-V operations like VM Storage Migration, VHDX Creation, VM Cloning.

- Massive data migration. E.g. storage system upgrade, a new database engine.

- Host-Controlled Data Transfer within a Tiered Storage Device.
Resources

- en.community.dell.com/Dell Compellent Storage Center ODX Overview - Community
- hyper-v.nu
Thank you.

Questions?
About Author

Amit Luniya

• Working as Associate principal development engineer with EMC2 Isilon India.

• Having 9+ years of experience in Storage, Filesystem and Data Protection.