

Module 7: Performance Considerations for Business Continuity Operations

Upon completion of this module, you should be able to:

- Discuss TimeFinder performance considerations
- Discuss SRDF performance considerations



This module focuses on performance considerations for Business Continuity operations.

Module 7: Performance Considerations for Business Continuity Operations

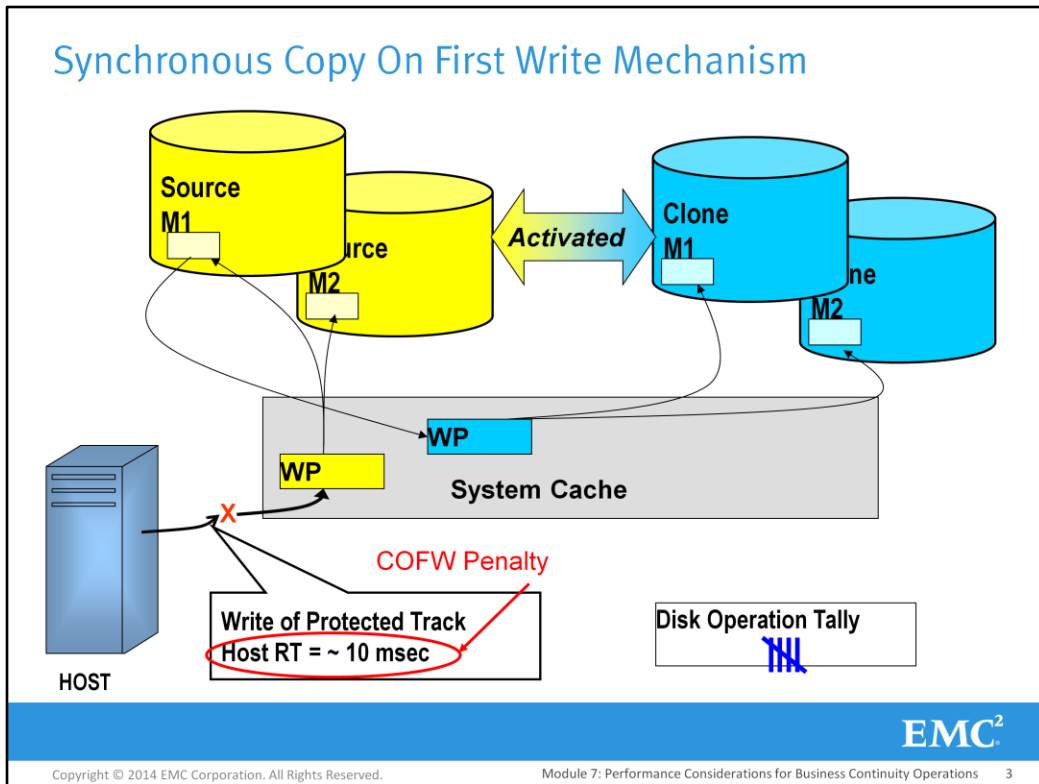
Lesson 1: TimeFinder Performance Considerations

During this lesson the following topics are covered:

- TimeFinder/Clone considerations
- TimeFinder/Snap considerations
- TimeFinder/VP Snap considerations



This lesson covers TimeFinder performance considerations.



The synchronous Copy on First Write mechanism is displayed on the slide. This was the default behavior prior to 5772.

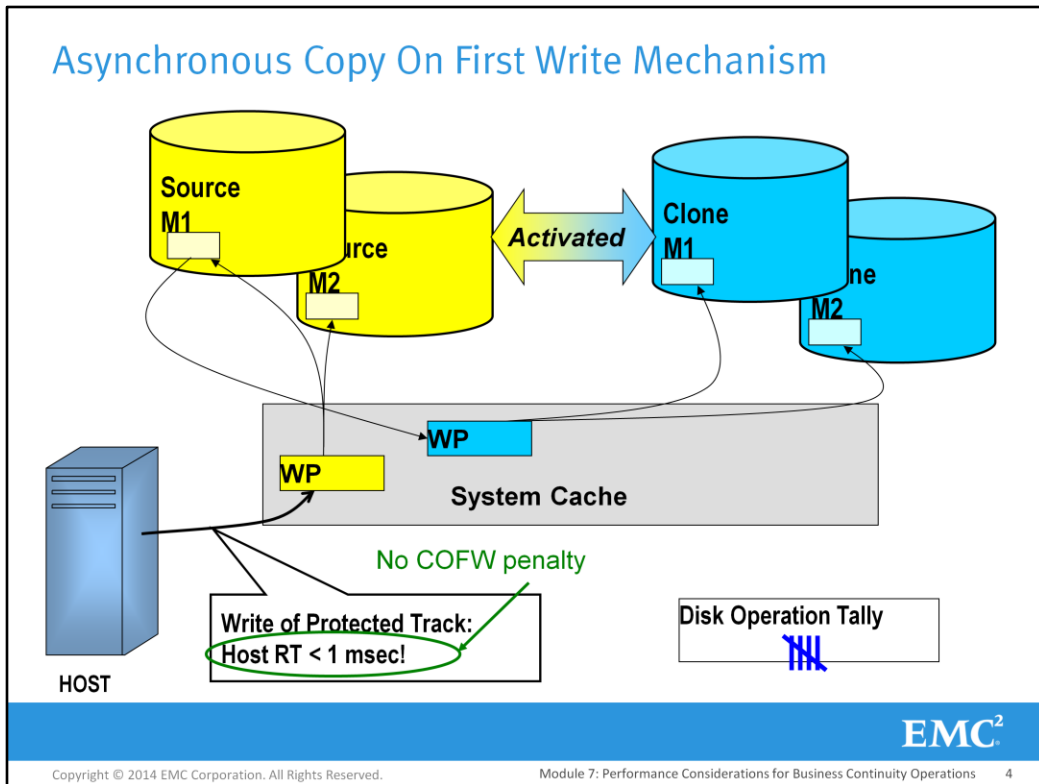
The COFW mechanism applies to both TF/Clone and TF/Snap. The illustration on the slide is for Clones. Snap would behave in the exact same manner except the tracks would be copied over to the Save pools instead of the Clone target. The example assumes RAID1 protected source and target volumes.

Host issues a write to a protected tracks (A write to the source volume is illustrated here).

The write is held until the protected track is read from the source volume to the cache and marked as write pending to the target, the protection bit is cleared.

The write from the host is now allowed to complete and marked as write pending in the cache. The delay in the response time can be up to 10 msec., which is what we refer to as the COFW penalty.

Eventually, the write pending tracks are de-staged to disk. There were a total of 5 disk operations in this case.



The asynchronous Copy on First Write mechanism is displayed here. This applies to writes to the source volume only.

The COFW mechanism applies to both TF/Clone and TF/Snap. The illustration on the slide is for Clones, Snap would behave in the exact same manner except the tracks would be copied over to the Save pools instead of the Clone target. The example assumes RAID1 protected source and target volumes.

Host issues a write to a protected track on the source volume.

The write is allowed to complete in cache and the track is marked as “versioned” write pending. The response time will be almost equivalent to the base response time of the application. No COFW penalty, as was found in earlier versions of the product.

The protected track is read from the source volume to the cache and marked as write pending to the target, the protection bit is cleared. The “versioned” write pending of the new track that was written is changed to write pending to the source volume.

Eventually the write pending tracks are destaged to disk. There were a total of 5 disk operations in this case, just like the case with synchronous COFW.

Asynchronous COFW

- Supported with Thick or Thin source volumes for TF/Snap and TF/Clone
- Only for writes to source devices
- Writes to target devices still use synchronous COFW and pay COFW penalty
- Number of backend operations for ACOFW is the same as SCOFW



The impact of ACOFW has been tested extensively by Performance Engineering. SPEED Gurus have access to the detailed information.

TimeFinder/Clone Considerations

- There is no impact to reads on source volumes from the existence of a clone volume (with no workload on target):
 - ▶ Workload on target could cause disk contention with source until volume is fully copied
- Write response time to protected tracks (first write to track) are higher than a normal cached write:
 - ▶ ACOFW will help to alleviate this
 - ▶ Writes to target will still have to pay COFW penalty
- First reads from target are slower than normal because of the copy on read needed for clone target device



Clone performance has no impact on reads on the source as long as there is no workload on the target. Workload could cause disk contention with the source until the clone is fully copied. First reads from the target are slower than from the STD or from a split BCV, because of the copy-on-read needed for the clone target device.

Copy Rates for TF/Clone

- Copy rate is affected by many factors:
 - ▶ Backend configuration (Disk, DA, Protection Type)
 - ▶ Contention with other activity
 - ▶ Number of changed tracks
 - ▶ RAID 5 and RAID 6 copy performance greatly improved for FBA configurations
 - ▶ Disk adapters dynamically balance copy rates with host workload



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

7

Copy rates are affected by various factors, as listed.

Benchmarks show that the Source and Clone protection types have an effect on the copy speed. Since the Clone is receiving writes, the write performance of the device is the key factor in the copy speed. RAID-1 Clones have the best performance, followed by RAID-5, and finally RAID-6.

RAID-1 Source devices also have a slight advantage over RAID-5 or RAID-6 devices, giving them a slight performance edge during a copy.

Consider this relationship if you are doing a pre-copy or background copy. The performance of the Source application may suffer during the copy process, so shortening the time would be a performance improvement.

TimeFinder/Snap Considerations

- There is no impact to reads on source from the existence of snap volume (with no workload on target):
 - ▶ Workload on target could cause disk contention with source
- Write response time to protected tracks (first write to track) are higher than a normal cached write:
 - ▶ ACOFW will help to alleviate this
 - ▶ Writes to virtual devices will still have to pay COFW penalty
- Reads from virtual devices will be slower than reads from source devices
- Save pool design is extremely important:
 - ▶ Must be made to reduce/remove physical device contention



Snap has no impact to reads on the source from the existence of a snap volume. Once there is activity, performance could be impacted and the target could cause disk contention with the source since it requires the source for its point-in-time view of production data.

Planning for Save Devices/Pools

- Save devices store the original track from the standard, copy takes place on first write
- Save devices can be RAID1 or RAID5:
 - ▶ Use RAID1 for best performance
 - ▶ RAID6 is not supported
 - ▶ Unprotected devices are by RPQ only and not recommended
- Size of each Save device should be equal to or less than the largest VDEV
- Allocate enough disks to handle change rate:
 - ▶ Spread them wide and evenly
 - ▶▶ Should be spread across DAs and disks for optimal performance
 - ▶▶ Aim for one hyper per physical disk
 - ▶ Save pool can sustain between 40 and 70 IO/s per 10 K spindle



Save devices and pools can be created and managed via Solutions Enabler and Unisphere for VMAX.

Designing Save Devices

- The number of Save devices should be proportional to the number of Virtual devices:
 - ▶ Based on percentage of change
- Spread Save devices across as many DAs and disks as possible:
 - ▶ With new installs, SymmWin does this by default
 - ▶ Current installs may require relocation of customer data unless sufficient disks are added to fulfill new capacities
- Sizing Save devices:
 - ▶ # of Save Devices = %Change * # of Virtual Devices
- If the change percentage is 30+%, TimeFinder/Clone may be a better alternative



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations 10

The number of Save Devices should be proportional to the number of Virtual Devices. To determine the number of Save Devices required by the customer, multiply the change percentage by the number of Virtual Devices.

For example, with a 20% change and 100 Virtual Devices, the number of Save Devices would be 20 (.20*100). If the change percentage is above 30%, TimeFinder/Clone is probably a more suitable solution.

TimeFinder/VP Snap

- New feature available starting with Enginuity 5876
- Provides the ability and possibility for multiple Clone Targets of Source device(s) to share track allocation within the same Thin Pool
 - ▶ All Target device(s) must be bound to the same Thin Pool
- Both Source device(s) and Clone Target device(s) must be Virtual Provisioned – TDEVs
 - ▶ Source and Target device(s) can be bound to the same Thin Pool or to different Thin Pools
- Supported with FBA and AS400 D910 iSeries devices



TimeFinder VP/Snap is a space saving local replication feature introduced with Enginuity 5876. Multiple clone targets of a source device can share track allocations within the same Thin Pool. Up to 32 sessions can be created for the same source device.

TimeFinder/VP Snap – How it works

- Sessions are created as No Copy and Non-differential
- Uses Copy On First Write mechanism (COFW) – optimized with Asynchronous Copy on First Write (ACOFW)
- On activation of the first session between the Source device and the Target device
 - ▶ Session state is set to CopyOnWrite
 - ▶ New write to a protected track on Source will trigger a copy of the original track to the Target device Thin Pool
 - ▶ Reads to protected tracks from Target device are redirected to the Source device



It works on principles similar to TimeFinder/Snap. The sessions are created as no copy and non-differential. Writes to source device will trigger a copy of the original tracks to the Thin Pool of the target devices. Reads from target devices are redirected to the source device.

TimeFinder/VP Snap – How it works (Continued)

- On activating a second session between the same Source and another Target
 - ▶ Reads to protected tracks from Target device are redirected to the Source device
 - ▶ Write to a track on the Source device
 - ▶▶ If the track has been modified after the first activation, then the second Target gets its own copy of this track
 - ▶▶ If the track has NOT been modified after the first activation, then the original track is copied ONCE to the Target Thin Pool
 - Pointers for BOTH the Targets are updated to point to this track – hence Shared Track
- Write to a shared track
 - ▶ The Target that writes will get its own copy of the track
 - ▶ Shared allocation will continue for other Targets



Potential space saving is realized when there are multiple targets for the same source. If a track is “owed” to multiple targets due to a write to the source, then the original track is copied over just once to the target device Thin Pool. All targets that need access to this track will share the allocation, and all targets will point to the same track in the Thin Pool. The conditions under which this can happen are described in the slide.

TimeFinder/VP Snap Considerations

- The `-vse` option can be specified only at the time of session creation
- Once created the session cannot be changed to any other mode
- If the Target devices are managed by a FAST VP policy, then the shared allocations cannot be relocated by FAST VP
- TimeFinder/Snap and TimeFinder/VP Snap sessions CANNOT co-exist on the same Source device
- TimeFinder/Clone and TimeFinder/VP Snap sessions CAN co-exist on the same Source device
 - ▶ The TimeFinder/Clone session CANNOT be created as a NO COPY session
- TimeFinder/VP Snap to Target device larger than the Source device is NOT allowed



Some considerations for deploying TimeFinder/VP Snap are listed on this slide.

TimeFinder/VP Snap Considerations (Continued)

- Cascaded TimeFinder/VP Snap sessions are NOT allowed
 - ▶ The Target of a TimeFinder/VP Snap session CANNOT be the Source for another TimeFinder/VP Snap session (cascading)
- TimeFinder/VP Snap session from a TimeFinder/Clone Target device is allowed
 - ▶ The TimeFinder/Clone pair state has to be Copied or Split
- TimeFinder/VP Snap from an SRDF R2 device in a Consistent and Active SRDF/A mode will require Device-Level Write Pacing on the corresponding R1 device
 - ▶ Both the source (R1) and target (R2) VMAX should be at Enginuity 5876
- Target devices should not be pre-allocated



Some considerations for deploying TimeFinder/VP Snap are listed on this slide.

Module 7: Performance Considerations for Business Continuity Operations

Lesson 1: Summary

During this lesson the following topics were covered:

- TimeFinder/Clone considerations
- TimeFinder/Snap considerations
- TimeFinder/VP Snap considerations



This lesson covered TimeFinder performance considerations.

Module 7: Performance Considerations for Business Continuity Operations

Lesson 2: SRDF Performance Considerations

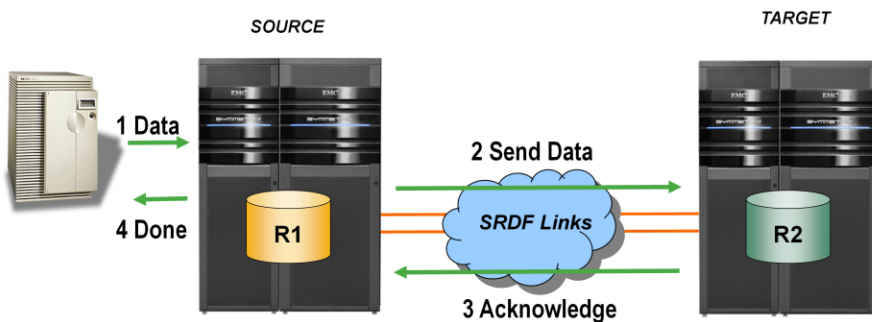
During this lesson the following topics are covered:

- SRDF/S considerations
- SRDF/A considerations



This lesson covers SRDF performance considerations.

SRDF/S Overview



- Write Data to Source Cache
- Data Sent over Link to Target Cache
- Acknowledgement received by source
- Host I/O Complete

EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

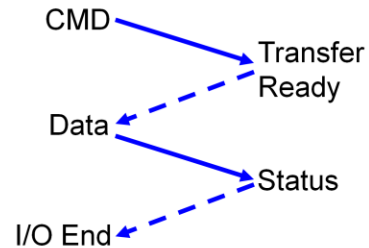
Module 7: Performance Considerations for Business Continuity Operations 18

SRDF Synchronous Mode is used primarily in SRDF campus environments. In this mode of operation, Symmetrix maintains a real-time mirror image of the data of the remotely mirrored volumes.

Data on the source (R1) volumes and the target (R2) volumes are always fully synchronized at the completion of an I/O sequence through a first-in, first-out queue (FIFO) model. All data movement is at the block level (size of the write I/O) in SRDF/S mode.

SRDF Synchronous Over Fibre Channel Remote Directors

- FC protocol requires two round trips for each write:
 - ▶ Initiator issues write command
 - ▶ Receiver issues transfer ready
 - ▶ Initiator sends data
 - ▶ Receiver sends status good
- Host to storage connectivity:
 - ▶ Not an issue
- Remote replication:
 - ▶ 1 ms latency per 200 km (Speed of light)
 - ▶ 4 ms per write for sites 200 km apart
 - ▶ Will increase host response times in Synchronous replication



EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations 19

The fibre channel protocol requires two round trips for each write. This is usually not an issue in normal host to storage operations within a SAN, because the distances that the signal has to traverse is so short that the latency due to these two round trips is negligible. The latency due to the two round trips starts to become an issue when remote replication is done over Fibre Channel. Speed of light in glass dictates that every 200 km will add 1 ms of latency. Thus if two sites are 200 km apart, two round trips would mean traversing a distance of 800 km (2 round trips * 2*200). The latency for every write would be 4 ms. The latency in the remote replication links is an issue only for Synchronous replication, because the write cannot be acknowledged till the write has been received at the remote location. For asynchronous replication, the host response times are not affected by the distance between sites.

Write Acceleration/Fast Write – FC Roundtrip Consideration

- Write Acceleration/Fast Write:

- ▶ Initiator issues write command
- ▶ Transfer ready emulated locally by switch/distance extension port
 - ▶▶ Write command sent to remote port and transfer ready from remote port will be eventually received by the local node
- ▶ Initiator begins data transfer as soon as local transfer ready is received
 - ▶▶ Does not have to wait for the transfer ready from remote node
- ▶ Receiver port sends status good back to initiator
- ▶ Latency equivalent to one round trip



EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations 20

Write Acceleration/Fast Write is a performance enhancement feature offered by some FC switch and distance extension products. The goal of Write Acceleration or Fast Write is to minimize the latency caused by the two round trips that the FC protocol requires thus improving performance during synchronous replication. The improved performance is achieved via intelligence built into the distance equipment or switches. The intelligent switch that is local to the initiator emulates a Transfer Ready when it receives a Write Command. This allows the Initiator to start the data transfer as soon as it receives the local transfer ready. In the mean while, the local intelligent switch sends the Write Command to the Receiver and eventually receives the real Transfer Ready. The latency is reduced because the Initiator did not have to wait for the real Transfer Ready from the Receiver Port. Once the data is received by the Receiver, it sends a Status Good back to the Initiator. Data integrity is ensured because the Status Good is not emulated, the Status Good is sent by the Receiver Port only if the Write was successful. The intelligent switches have additional buffers and the switches at either end have to coordinate operations. The latency due to Write Acceleration/Fast Write is equivalent to that of one round trip.

References:

1. Fibre Channel Write Acceleration – Cisco White Paper:
http://www.cisco.com/en/US/prod/collateral/ps4159/ps6409/ps5989/ps6217/prod_white_paper0900aecd8024fd2b.pdf
2. Optimizing Storage Data Transfers over FC ISLs and Long-Distance FCIP WAN Links with FastWrite – Brocade SAN Tech Note

SRDF Synchronous – Single Round Trip (SiRT)

- Symmetrix RF Directors – Single Round Trip (SiRT):
 - ▶ Latency equivalent to one round trip
 - ▶ SiRT is used only for SRDF/S
 - ▶▶ Will not be used for other RDF modes
 - ▶ Enabled by default



The Single Round Trip (SiRT) feature for Symmetrix Fibre Channel RDF (RF) directors is applicable only for the Synchronous SRDF mode. SiRT needs is enabled by default.

SiRT is enabled for all block sizes.

Fibre Channel Considerations

- Write Acceleration/Fast Write/SiRT considerations ...
 - ▶ Synchronous SRDF
 - ▶▶ Helps
 - ▶ Adaptive Copy or SRDF/AR
 - ▶▶ Could help improve throughput by about 20%
 - ▶ Asynchronous SRDF
 - ▶▶ Typically no improvement under normal circumstances
 - ▶▶ Could help
 - SRDF/A recovery
 - Reduce resynchronization time
 - Small I/O sizes that throttle RDF throughput
 - Gaps in links due to retransmissions after link errors
- Best practice:
 - ▶ Use EMC SiRT **or** 3rd party fast write feature
 - ▶ Do not use both simultaneously
 - ▶ EMC SiRT will measure link latency and disable automatically if connected to write acceleration/fast write devices



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations 22

Write Acceleration/Fast Write/SiRT will always help in an SRDF/S environment. It can also help in an SRDF/AR or SRDF adaptive copy environment because the SRDF throughput is improved by about 20%. Typically, the two roundtrips (FC) versus one (GigE or Write Acceleration) does not impact SRDF/A performance since SRDF/A will send continuous stream of writes until the cycle is empty. If the links are kept full, there is no waiting for roundtrip delay and SRDF/A performance is not impacted by two roundtrips. Fibre Channel fast writes will improve performance for SRDF/A when there are "gaps" on the links due to retransmissions after link errors, or small writes that throttle SRDF throughput due to hitting I/Os per second limits on the links. The number of active volumes in a cycle can also affect performance on links where the links may be hitting the 32 writes/volume limit.

EMC SiRT is compatible with fast write/write acceleration switches and extenders, as it will measure link latency and disable automatically if connected to these devices. As a best practice, it is recommended that either the EMC SiRT feature or the third-party fast write feature should be used. Both should not be enabled simultaneously.

Concurrent SRDF/S Writes – Open Systems

- Multiple host writes to the same R1 device can run in parallel, as long as...
 - ▶ Each write is contained within a single Symmetrix system track
 - ▶ Writes belong to different tracks
 - ▶ Writes do not span different meta devices
 - ▶ Each write comes from a different FA path
 - ▶ Up to 8 concurrent writes
- Can improve OLTP performance



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations 23

The Concurrent J0 (SRDF/S) Writes feature, enables multiple concurrent writes to an RDF device for writes coming from different FA ports.

With this feature, writes to an RDF device are not serialized. Therefore, in most cases, the performance impact of SRDF/S mode with this feature active is significantly less than the impact when this feature is not active. If the application serializes some writes (dependent writes), then Concurrent SRDF/S Writes will not help improve application performance. If the application serializes some, but not all, writes, then you may still see a benefit, depending on the proportion of writes that are serialized. If multiple independent applications share the same Symmetrix system volume, then using the Concurrent SRDF/S Write feature will improve the write response times since the writes are not serialized. Since the serialization of the dependent writes is controlled by the application, enabling Concurrent SRDF/S Writes has no impact on data integrity (or data consistency). If the remote copy of the data was consistent with Concurrent Writes Disabled, then the remote copy will also be consistent with Concurrent Writes Enabled. This is due to the fact that any write will only be acknowledged once the data is written to the R2 Symmetrix system's cache.

Note: Mainframe has always enabled Concurrent SRDF/S Writes with PAV.

More performance details can be found on EMC SPEED Bits & Bytes article "Concurrent J0 Writes for Open Systems Devices." This article is available only to EMC SPEED Gurus.

SRDF/A Reserve Capacity – Transmit Idle

- After links fail:
 - ▶ Data transmission from source to target will stop
 - ▶ SRDF/A will remain active and the capture cycle will grow in cache
 - ▶ Session will be suspended if cache fills up
- After links are revived, cycle switching will continue
- Not supported with ESCON RAs
- Should be enabled on both Source and Target side
- Should be used if using Delta Set Extension



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations 24

SRDF/A Transmit Idle is a Reserve Capacity feature to EMC's SRDF/A feature that provides SRDF/A with the capability of dynamically and transparently extending the Capture, Transmit, and Receive phases of the SRDF/A cycle, while masking the effects of an "all SRDF links lost" event.

Without the SRDF/A Transmit Idle enhancement, an "all SRDF links lost" event would normally result in the abnormal termination of SRDF/A. The SRDF/A Transmit Idle feature was specifically designed to prevent this event from occurring.

SRDF/A Reserve Capacity – Delta Set Extension

- Allows offloading of SRDF/A delta sets from cache to specially configured device pools – Delta Set Extension Pools
- Intended to make SRDF/A resilient to temporary increases in write workloads or link loss:
 - ▶ Should be used in conjunction with Transmit Idle



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

25

SRDF/A Delta Set Extension (DSE) provides a mechanism for augmenting the cache-based Delta Set buffering mechanism of SRDF/A with a disk-based buffering ability. This extended Delta Set buffering ability may allow SRDF/A to ride through larger and/or longer SRDF/A throughput imbalances than would be possible with cache-based Delta Set buffering alone.

In the base SRDF/A implementation, an overload of writes can cause the session to drop, terminating the remote replication. This will happen if the cache consumed by the local delta sets reaches the limit set for SRDF/A (nearly all of the write pendings available on the array). With no more space to store the incoming writes, the array has no choice but to stop adding them to the sets.

Delta Set Extension is an enhancement that allows SRDF/A to continue even when there are enough writes to fill the cache. A special Delta Set Extension pool of data devices is used to offload the writes to disk before they consume too much cache. When the writes hit a threshold, they are written to the pool and removed from cache (assuming they have already been destaged locally). When needed, the writes are staged back into cache and transmitted across the SRDF links. The performance is obviously degraded, but in this way the consistent sets can be maintained.

DSE Pools

- Save Pools are designated as DSE pools at creation:
 - ▶ Contains SAVE devices of a single emulation
 - ▶▶ CKD3390, CKD3380, FBA, or AS400
- Pools can be associated (shared) with multiple SRDF/A RDF groups
- Each SRDF/A session can be associated with zero or one DSE pool of each type (e.g. FBA, 3390):
 - ▶ Must have at least one DSE pool configured with a type that matches one of the device types in the SRDF/A group in order to activate DSE
- Pools can optionally start automatically when SRDF/A is enabled



Copyright © 2014 EMC Corporation. All Rights Reserved.

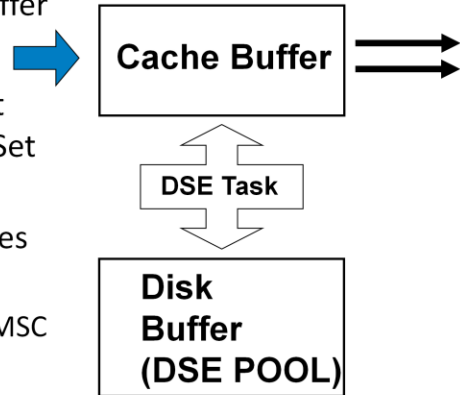
Module 7: Performance Considerations for Business Continuity Operations

26

- SRDF/A DSE Pools and Save devices are managed in the same way as TimeFinder/Snap pools
- An RDF group can have at most one pool of each emulation
- A single rdfa_dse pool can be associated with more than one RDF group, similar to snap pools shared by multiple snap sessions
- SRDF/A DSE Threshold sets the percentage of cache used for SRDF/A that will start offloading cache to disk
- DSE must be enabled on both the source and target arrays. DSE enabled on only one side of a link would lead to failure of the SRDF/A recovery with SRDF/A dropping because the R2 side would fail to have enough cache to hold the large and extended Transmit cycle.

How DSE Fits Into SRDF/A Data Flow

- Traditional SRDF/A data flows continue to interact directly with the cache buffer
- Separate DSE Task monitors Delta Set cache utilization and transfers Delta Set data between cache and disk
- Must be configured on R1 and R2 sides to be of value ...
 - ▶ And in all sessions tied together by MSC



EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

27

SRDF/A has always buffered Delta Set data in cache. However, a buffer full condition causes SRDF/A to drop. With SRDF/A DSE delta set, data is offloaded to disk buffers (DSE Pools) by the DSE task. We will look at this in more details in the next few slides.

When Can DSE Help?

- DSE is not designed to solve any permanent and persistent problems:
 - ▶ Mis-configurations:
 - ▶▶ Unbalanced
 - Cache or backend on target side is smaller/weaker than source side
 - ▶▶ Not enough cache
 - ▶▶ Not enough bandwidth
 - ▶ Host writes consistently exceeding RDF link bandwidth
 - ▶ Prolonged link outages
- SRDF/A DSE solves abnormal and temporary problems:
 - ▶ Unexpected host load
 - ▶ Link bandwidth issues
 - ▶ Temporary link loss (use with Transmit Idle)
- Increase resilience of SRDF/A



SRDF/A DSE should be used with the Transmit Idle feature. Thus, SRDF/A can ride through a temporary link loss. Once the DSE threshold is reached, data will be paged out to the DSE Pools. DSE is not designed to solve any permanent and persistent problems.

What To Monitor When Using SRDF/A

- Write % of SRDF devices
- **Cycle Time:** If a cycle has many writes, the **cycle time** may grow longer
- **Write Pendings:** If a cycle becomes too large, the **write pendings** owed to the R2 device may exceed the SRDF/A limit (often 94% of the system write pending limit)
- Delta Set Extension pool utilization



Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

29

In addition to the expected performance measures such as link I/Os and throughput, There are several other measures that should be monitored when using SRDF/A.

Watch the write percentage of the SRDF devices. A heavy write load will often lead to issues as the writes fill up cache.

Look at the cycle time of the RDF groups. As writes become a problem, the time will begin to extend. In addition to being a good indicator of increased writing, longer cycle times mean the consistent point-in-time cycle on the remote array will be further out of date.

Monitor the write pendings, especially if the tool can show the remote write pendings (owed to the remote array due to SRDF/A) separately from the local write pendings (owed to the local drives). As the remote write pendings become larger, more space is being used to store them.

Also watch the Delta Set Extension pool utilization. When no more data can be stored in the pool, SRDF/A will have to drop.

All of the SRDF/A features are designed to allow for temporary surges in writes that contain more data than can be transmitted on the links at a time. When such a surge happens, you will see the effects in these measures as writes build up. As long as they are temporary, you will eventually return to normal operations.

Measuring SRDF/A - Dashboards

- Monitor>EMC Dashboards>All RDF Directors
 - ▶ Avg IO Service Time(ms), MBs Sent/sec, %Busy Port 0, % Busy
- Monitor>EMC Dashboards>RDF Director>select RF
 - ▶ Ios/sec, MBs Sent and Received/sec, %Busy, Avg IO Service Time(ms)
- Monitor>EMC Summary Dashboards>RDF Directors>Summary
 - ▶ %Busy, %Busy Port 0, Avg IO Service Time, IOs/sec, MBs Sent/sec, MBs Received/sec



You can measure SRDF/A issues at, both, all SRDF directors or at a specific SRDF director level. The dashboards give a quick high level overview of the SRDF environment.

Measuring SRDF/A – Analyze

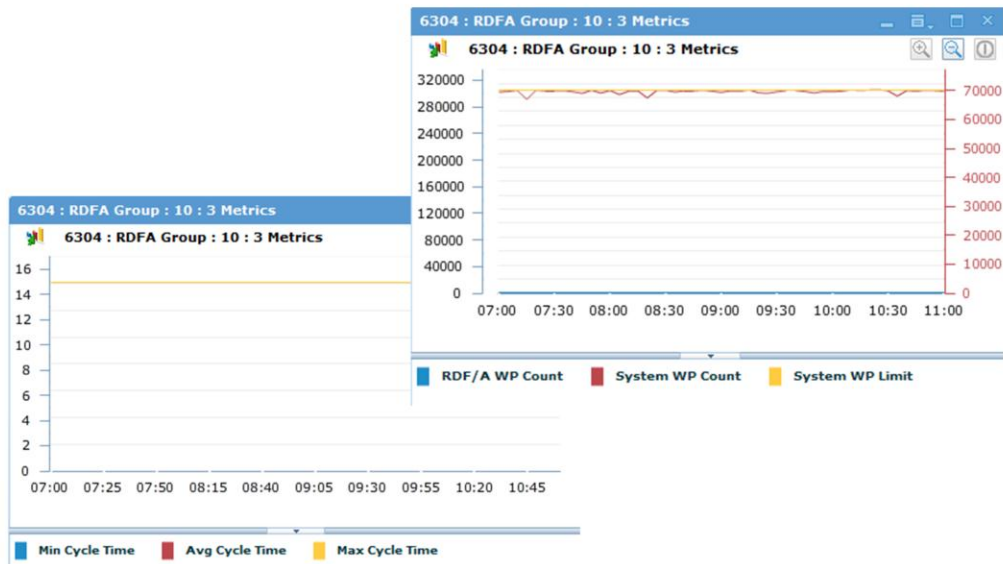
- Cache/DSE
 - ▶ Analyze>Array>DSE Pool>select pool
 - ▶▶ KPI Capacities - % Used, Total, Used, Allocated
- RDF/A Groups
 - ▶ Analyze >Array>RDF Director>RDFA Group
 - ▶▶ Avg Cycle Time, Avg Cycle Size
 - ▶▶ RDF R1 to R2 IOs/sec
 - ▶▶ % R1 Used Cache
 - ▶▶ RDF/A WP Count



The Analyze measures for SRDF/A are at the SRDF director level. You can show the write pending and Delta Set Extension pool utilization counts. You can also show the cache utilization by group, and the cycle times by group.

You can also click on an individual SRDF director (it might be transmitting data for several SRDF groups) to show write pendings, usage, and cycle times.

SRDF/A – Group Examples



EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

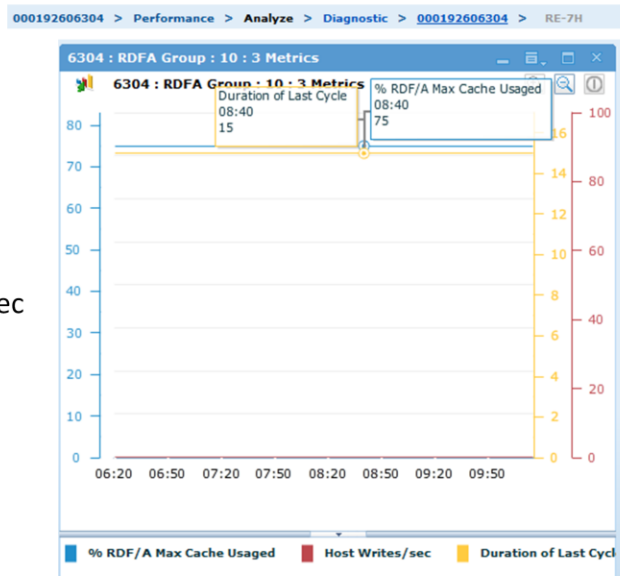
32

You can see the utility of the Performance Viewer SRDF/A measures in these examples. The graph at the upper right shows the RDF/A write pending count, System write pending count and limit. The straight yellow line marks the write pending limit. The blue line shows the write pendings used by SRDF/A. The red line is the System WP count.

The bottom left graph shows the minimum, average and maximum cycle times across all SRDF/A groups. You can see that it is in a steady state with no significant events.

SRDF/A Measures

- Cycle Time Measures
 - ▶ Average cycle size
 - ▶ Duration of last cycle
 - ▶ R1 to R2 lag time
- By RA group measures
 - ▶ Host writes per sec
 - ▶ Host MBs written per sec
 - ▶ RA group ios per sec
 - ▶ RA group MB per sec
 - ▶ Total cache slots in use



EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

33

There are many more measures specifically for SRDF/A monitoring. They appear under the RDF Director tab, RDFS and RDFA group tabs, DSE Pool tab in the Analyze view and there are also the All RDF Directors, RDF Director and Summary RDF Director Dashboards.

A variety of cycle time measures show when some performance issue negatively affects the SRDF/A cycles. One of the most useful of these measures is “duration of last cycle,” which measures the last transmit cycle. Under ideal circumstances, this measure would be a flat line at the chosen cycle time. But, when a performance issue causes more time to be taken to transmit the cycle, the graph will show it. A related measure shows the “R1 to R2 lag time,” which shows the most out of date data on the R2 side might be. The “active cycle size” shows the number of tracks in a cycle. These measures will typically increase when “duration of last cycle” increases.

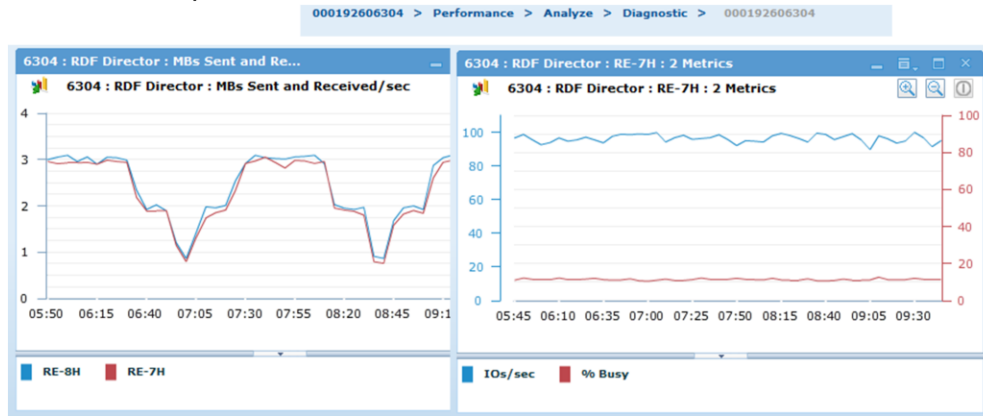
The SRDF/A category also includes standard measures by RA group (SRDF/A is enabled for entire RA Groups). You can measure the host I/Os and writes being committed to all the devices in the group, as well as the IOs and MBs transmitted by the RA group across the SRDF links. If a high locality of reference is present in the workload, you will see lower amounts of RA traffic than host traffic.

There are also a number of cache measures, including Active Cycle Size which shows the count of captured data on the R1 or R2. This is an especially useful measure of cache utilization since the normal write pending measures do not show the effects of SRDF buffering.

This is just a sample of the large number of SRDF/A measures available.

SRDF Director

- Analyze>RDF Director
 - ▶ IOs per sec
 - ▶ MBs sent & received per sec
 - ▶ % Busy



EMC²

Copyright © 2014 EMC Corporation. All Rights Reserved.

Module 7: Performance Considerations for Business Continuity Operations

34

Several metrics can be used to measure the SRDF traffic of a VMAX. The most common measures are very similar to those found on any other director type: "IOs per sec," "MBs sent per sec," "MBs received per sec," and "% Busy."

This simplistic utilization measure of the % Busy may be less useful than similar measures for other Symmetrix components since it does not factor the distance between the sites. As we shall see in the next part, the distance between the sites has a large impact on the actual utilization of the SRDF links.

Recommendations

- Use Synchronous Mode for short (<200km) distances only
- Use Adaptive Copy Modes for bulk data transfers, not online transactions
 - ▶ Write Pending mode when cache availability is not a problem, but disk access might be
 - ▶ Disk mode when cache availability is at a premium, and disk directors are not being heavily utilized
- Set limits (skew) on number of locally queued writes with Adaptive Copy Modes to reduce cache/disk traffic demands
- Consider SRDF/Asynchronous when consistent data sets are required and distances are too far for Synchronous
- Use Quality of Service to reduce copy impact of initial SRDF synchronization



Use synchronous mode only for short (<200 km) distances. Longer distances incur too much acknowledgement penalty. As technology is constantly improving the communication speeds, you can expect the definition of “short distances” to be ever changing.

Adaptive copy modes should be used only for initial data loads or bulk transfers. If cache availability is not a problem, use write pending mode; otherwise, use disk mode.

Skew limits reduce the amount of queuing in both adaptive copy modes. Setting a low skew for write pending mode reduces cache occupied by SRDF operations, while a low skew for disk mode limits the number of tracks sent to disk that must be later transferred over the SRDF lines. In both cases, a device that reaches its skew falls back into synchronous mode. This effectively halts the queuing process and forces new writes to be accepted only when the previous write has been acknowledged by the remote Symmetrix.

Initial synchronizations of R1 and R2 devices occur when mirroring has suspended and then resumed. Software commands can be issued to do this for the purpose of failover testing. Quality of Service can be used to slow down the synchronization rate in these cases, reducing the disk access demands.

Module 7: Performance Considerations for Business Continuity Operations

Lesson 2: Summary

During this lesson the following topics were covered:

- SRDF/S considerations
- SRDF/A considerations



This lesson covered SRDF performance considerations.

Module 7: Summary

Key points covered in this module:

- TimeFinder performance considerations
- SRDF performance considerations



This module covered performance considerations for Business Continuity operations.

Course Summary

Key points covered in this course:

- Relating knowledge of the VMAX architecture and I/O handling processes to their performance benefits
- Identifying the performance impact that different types of workloads have on Symmetrix architectural components
- Using key metrics to identify performance bottlenecks and component over utilization
- Making performance-oriented recommendations when allocating new storage or migrating applications



This course covered the various aspects of analyzing the performance of Symmetrix VMAX arrays.

Thank You!

