SAN-Engineer   Govindagouda Ranganagoudar

How EMC DSE Works   3 comments

This solution lists the best practices for implementing SRDF/A DSE.

How DSE Works

When the system Write Pending (WP) count crosses the DSE threshold (which by default is 50% of the system WP limit), DSE becomes active and starts paging SRDF/A tracks to the DSE save pool. After a cycle switch, Enginuity reads tracks from the DSE pool back into the Symmetrix system cache so that they can be transferred to the R2. The process of de-staging SRDF/A tracks from the Symmetrix system cache to the DSE pool is known as "paging out." The process of bringing tracks back into cache is known as "paging in."

Symmetrix System Resources Used by DSE

DSE paging operations require the following Symmetrix system resources:

1. DA and RA CPU cycles
2. DA bandwidth
3. Disk bandwidth.

A process known as the "DSE task" runs on all DA and RA CPUs. The process is activated when the DSE threshold is reached. There are upper limits on the CPU cycles that can be consumed by the DSE task. In Enginuity 5772.88.80, DSE can use up to 16% of DA CPU and 25% of RA CPU cycles. (The DA CPU limit can be increased up to 50% through Inlines commands.) In Enginuity 5773, the code dynamically increases the DA CPU limit to up to 50% as needed.

When DSE pages out data to the DSE pool, it always needs to write out full tracks, even if the host writes a partial track. In addition, when DSE reads in from the DSE pool, it will read in full tracks. DSE also imposes an extra burden on the disks. Every host write can result in one write to the DSE pool during page out and one read from the DSE pool during page in. If RAID-1 protection is used for DSE pool devices, each host write can result in two backend writes. These I/Os can be random in nature.

The DSE page-out and page-in rate can be impacted if any of the resources mentioned above are not available.

The following lists EMC's recommended best practices for DSE

1.   Configure DSE on both R1 and R2 Symmetrix system frames. 2.   Plan for peak workloads.

Make sure you collect data during peak workloads. If the peak workload occurs during month-end processing or during quarter-end processing, ensure that the peak workload data is used during planning for DSE configuration.

3. DSE pool configuration guidelines are as follows:

- The best practice is to spread the DSE pool devices across as many disks as possible. Do not underestimate the bandwidth and throughput demands that DSE places on the disks. When DSE is active, you could have one random write and one random read for every host write. Consider an example where DSE pool devices with RAID-1 protection are used. When DSE is active, a host write load of 1500 writes/sec will result in 3000 backend writes/second and 1500 reads/second to the DSE pool.
- Spread the DSE pool across as many disks as possible. The minimum number of disks to spread the pool across depends on the peak write rate. As an example, if the peak host write rate is 1500 writes/sec and RAID-1 pool devices are used, the DSE pool needs to be spread across about 100 disks.
- The SymmMerge user data option can model the DSE pool.
- RAID-1 is the recommended protection type for DSE pool devices. RAID-5 is another option, but it will consume more backend resources. Currently RAID-1 is the only supported configuration. An RPQ must be submitted for RAID-5.

4. Do not use dedicated disks for the DSE pool. Use only one hyper from each disk for the DSE pool. The other hypers can be used for any other purpose.

5. The size of the DSE pool depends on:

- The average write rate when DSE is active.
- The length of time DSE is active.

A simple formula for calculating DSE pool size is: Write Rate*Length of time when SRDF/A will be paged into DSEpool*64 KB

Example: If the write rate is 3000 writes/second and you want to ride over a 1 hour link outage, the size of the pool is 3000*3600*64KB = 659 GB.

This formula assumes that the I/Os are smaller than 64 KB. For I/Os larger than 64 KB, adjust accordingly. In the above example, if each write was 128 KB, the size would be 3000*3600*128 KB.

This size is before protection. If the protection is mirrored, you need to double the disk space. In the above example, you will need 659 GBx2 = 1318 GB of disk space.

This assumes that there are no rewrites, but it is better to be safe than sorry in this case. Bear in mind that the size of the pool will be an outcome of the minimum number of drives required as explained in number 3 above.

Keep the size of the DSE pool reasonable. Typically 2 to 5 times cache size is sufficient for a short outage.

6. Ensure sufficient DA and RA CPU resources are available for the DSE task.

The DSE task runs on all RA and DA CPUs. Make sure the DA and RA are not more than 50% busy without DSE being active. If they are busier, they may not be able to handle the additional load when DSE is active. A good rule of thumb to remember is that for DMX running Enginuity 5772, a DMX1500 can support approximately 1500 writes/second, a DMX2500 can support 2500 writes/sec., and a DMX4500 can support 4500 writes/sec.

7. Use as small a number of DSE pools as possible since this makes the most efficient use of resources and is easiest to manage. The number of RDF groups that can share a DSE pool is limited only by the number of RDF groups that can be created in the system. DSE imposes no limit on the

number of groups that can share a pool.

If multiple groups are assigned to a DSE pool and the need arises to perform paging operations on more than one of those groups at the same time, then there will be pool fragmentation. Sometimes pool fragmentation may impact paging performance.

8.   If there is 'spare' SRDF/A throughput, then the system will return to a normal RPO quicker after a link outage.

9.   EMC Engineering recommends lowering the SRDF/A max cache setting from the default value of 94% to 75% when adding DSE into the environment. Refer to emc142579 (http://knowledgebase.emc.com/emcice/documentDisplay.do?clusterName=DefaultCluster&preview=1&groupId=1&page=&docType=1006&resultType=5002&docProp=$solution_id&docPropValue=emc142579) for the best practice when configuring SRDF/A.

Posted August 14, 2012 by g6237118

3 responses to "*How EMC DSE Works*"

Subscribe to comments with RSS.

If suppose the DSE Pool is full, then what would be the next step(Like where the further data reside)?

Reply
Ankit

April 12, 2013 at 3:02 am
We should check why this is getting full. We should find out reasons why this is getting filled up

Reply
g6237118

June 25, 2013 at 9:53 pm
Great post.

Reply
wordpress plugins contact form

August 5, 2014 at 3:24 pm