



Global Distributed Storage Engineering
Citi Architecture & Technology Engineering

EMC DMX 5772 Microcode Beta Testing – Key Findings and Benefits

GDSE-07020-E

Evaluation

Version 1.0

1. Version Control

Version	Date	Author	Reviewed	Comments
1.0	23-March-07	Chris Clayton Guy Shimabuku Kevin Hobaugh		

2. Legal Notice

Information contained herein is for Internal Use and may be used only for business purposes authorized by Citigroup.

CITI, CITIBANK, CITICORP, CITIGROUP and the Umbrella Device are trademarks and service marks of Citicorp and are used and registered throughout the world. Citibank, Citicorp, Citigroup and their subsidiaries also claim rights in certain other trademarks and service marks contained in these pages.

3. Contact

Please address any comments or questions regarding the information contained in this document via email to the “*GT Global Distributed Storage Engineering” distribution list or alternatively to gdse@citigroup.com.

4. Latest Versions

It is recommended that the latest version of the document be used from the GDSE website <http://engineering.ti.citigroup.net/gdse/>. Newer versions may include certification on additional platforms, or comments on issues discovered after deployment.

5. Table Of Contents

1.	Version Control	2
2.	Legal Notice	3
3.	Contact	3
4.	Latest Versions	3
5.	Table Of Contents	4
6.	Key Findings from the Beta Testing	5
6.1.	General Comments	5
6.2.	New SRDF/A features	5
6.3.	Data Protection Schemes	6
6.4.	Solution Enabler and Microcode	7
6.5.	Cache and Device Quality of Service	7
6.6.	Security and Audit Logging	8
6.7.	“Storage tiering in a box”	8
7.	Benefits of Participation in the Beta Test Programs	9
7.1.	Advantages of Beta program participation	9
7.2.	Disadvantages of Beta program participation	10

6. Key Findings from the Beta Testing

6.1. General Comments

EMC will propel Enginuity 5772 to Generally Available (GA) status on the 29th of March. All installations of this release will be controlled through a Phased Product Release (PPR) process that will regulate which customers will have access to Enginuity 5772. During the PPR process Enginuity 5771 remains EMC's default microcode for all DMX-3 systems.

Issues with the director µcode and service processor applications plagued the majority of the beta program timeline. These issues caused communication instabilities resulting in failed commands and command files, which delayed the completion of the beta program by several weeks. Concern still exists that not all issues have been corrected.

Colour coding in this report indicate **positive features**, **minor issue** and **major issues**.

6.2. New SRDF/A features

Beta testing gave confirmation that EMC's new SRDF/A features for maintaining operational state despite network resource shortfalls/failures will be highly beneficial and that a viable operational model will be possible within Citigroup.

- 1) **Transmit Idle feature** allows SRDF/A to seamlessly ride out a transient link failure. This feature provides an extra level of protection from sessions dropping when the RDF link cannot transmit data. If the link recovers, SRDF/A can pick up operations automatically where it left off.
 - **Functionality advertised by EMC was validated.**
 - **Very robust. No failures at all during testing, when correctly set up.**
 - **Acceptable monitoring of Transmit Idle invocation monitoring (initially broken at start of Beta testing, but uncovered by Citigroup). Depends on usage of Event Daemon which will need to be brought to operational readiness as part of the Solutions Enabler 6.4 certification.**
 - **Requested EMC to change the method for enabling this feature to simplify the storage provisioning process.**

2) **Delta Set Expansion (DSE)** (a.k.a Spillover to disk or Reserve Capacity). This feature allocates a portion of disk storage as the temporary storage location, rather than storing all pending data in SRDF/A cycles in cache. This allows SRDF/A to be more resilient and flexible to handle the situations where host application throughput outpaces the RDF throughput.

- **Functionality advertised by EMC was validated i.e. clean, automatic recovery when I/O peak or temporary network outage passes. Ability to tune the threshold for when DSE is invoked on an application by application basis, if desired.**

- Robust. Few failures during testing, when correctly set up. Several reporting status errors, which EMC then fixed for us.
- Acceptable performance by end of Beta test, although concerns that an improperly configured pool will cause SRDF/A throughput performance issues and negate the benefit of DSE. We need Best Practice from EMC need on how to set up a DSE pools that will be not cause performance issues. This Best Practice is in preparation by EMC.
- DSE currently incompatible with new 5772 Cache Partitioning feature. DSE fails silently when Cache Partitioning in use. EMC will remove incompatibility in a future release.
- There will be some operational issues right-sizing the spillover pools as EMC has no method to convert existing LUNs to devices that can be used to expand these pools should the need arise. Citigroup has requested this be made possible in a future release.
- No alerting when DSE is invoked. Citigroup has requested that this be added.

3) Increased numbers of RDF groups per Symmetrix now available. This was a Citigroup request, and needed for SRDF/A where each application typically needs its own RDF group.

- RDF group limit per Symmetrix increased from 64 to 128 (we tested to architectural limit of 250 but EMC will not support this number at GA release) due to performance concerns.
- RDF group limit per RA increased from 16 to 32 for SRDF/S (we tested to architectural limit of 64 but EMC will not support this number at GA release). EMC is still to define new limit for SRDF/A with 5772 (currently 8 with 5x71).

6.3. Data Protection Schemes

1) RAID-6 (R6) devices provide a level of data protection between mirrored and RAID-5 devices within a DMX-3 array. R6 devices protect against data loss if one or two drives fail or unrecoverable read errors occur within the rank. EMC utilizes an evenodd encoding algorithm to create a horizontal and diagonal parities stripes within each rank, whereby allowing the dual drive protection. The 5772 Beta version provides two R6 rank configurations of 8 (6+2) and 16 (14+2) devices.

- RAID-6 devices can be created and managed in a manner consistent with mirror and RAID-5 devices.
- Performance of RAID-6 with random I/O transactions as compared to the current RAID-5 7+1 standard:
 - RAID-6 random I/O throughput test showed R6 6+2 devices with an erratic performance curve, averaging of 32% less throughput than comparable RAID-5 (R5) 7+1 devices.
 - RAID-6 random I/O throughput test showed R6 14+2 devices with performance throughput consistent with comparable R5 7+1 devices.

2) Simultaneously support of all five data protection schemes (e.g. mirrored, RAID-5 (R5) 3+1, R5 7+1, R6 6+2 and R6 14+2 devices) **within the same DMX-3 frame.**

- Utilizing a Solaris host as an I/O generator, 2 of the 5 data protection types showed extremely poor performance. Running I/O through all five protection schemes simultaneously showed mirrored, RAID-5 3+1 and RAID-6 14+2 devices averaging 1040 I/Ops, while RAID5 7+1 and RAID-6 6+2 devices averaged 3 I/Ops for sequential reads. This low level of throughput renders the RAID5 7+1 and RAID-6 6+2 device(s) non-functional.
- Utilizing a Windows host as an I/O generator, consistency throughput was displayed across all five data protection devices. With a base of 40 devices divided equally across the five data protection types, a low standard deviation of 9.5 was achieved between each device for random I/O transactions.

6.4. Solution Enabler and Microcode

1) Solutions Enabler (SE) 6.4 with a specific edit level is required to display and manage the new 5772 µcode capabilities and features (i.e. symqos, RAID-6, DSE).

- Access instability was observed with 5772.40/42/43 µcodes in communicating with the configuration server located on the service processor. This instability proved to be the most elusive and disruptive in the beta program, causing delays in configuring and running tests. EMC traced the root cause of this issue to three different sources: 1) optimization code for symapi failing to release sockets causing memory leaks, 2) ID gaps between devices being created, 3) director module sending buffer overflows.
- Updating SE fails then going from version 6.4.0.0 (edit level 54) to SE 6.4.0.0 (edit level 817) with 5772.52 µcode. The SE update, installation and removal process requires that all the stordaemons be shutdown during the task. SE edit level 54 allows the stordaemons to automatically restart causing the update procedure to terminate prior to completion.

6.5. Cache and Device Quality of Service

1) Quality of Service (QoS) provides priority service levels selection for a range of devices. This QoS feature allows the user to view or change the priority service levels of selected devices from 1 (highest throughput) through 16 (lowest throughput).

- Comparing the performance between the highest and lowest QoS priority levels showed very little throughput difference. The random I/O throughput test across a range of 2KB-64KB transaction sizes, displayed a 6% overall difference in throughput between priority 1 and 16 devices. This negligible percentage difference indicates that the QoS feature has little or no effect on throughput, concluding this feature is not functional or operating improperly.
- EMC defaults the QoS priority for each device to the highest level of 1, therefore making it difficult to configure a device with a higher priority than the default level.

6.6. Security and Audit Logging

EMC's security division, RSA, integrated its security features into 5772 microcode, providing authentication, authorization and auditing. Listed below are the highlights where this security meets and misses Citi security standards.

- User ID and login policies are provided by EMC Gina developed by RSA
 - o Authentication via Active Directory
 - o Password management (no default passwords, password minimum length, alpha numeric)
 - o Passwords are encrypted (in flight and at rest)
 - o User ID disabled after 6 failed attempts
 - o Disallow display of last user name
 - o Role Based Access
 - o Entitlement Review is missing – unable to report user id and privileges
 - o Maker / Checker is missing in Solutions Enabler
- Audit Logging
 - o Failed login attempts are logged in the Audit Log
 - o Audit Log cannot be altered or deleted – tamperproof
 - o Able to ship Audit Logs to syslog server using Solutions Enabler 6.4
 - o Date/Time, User ID, Host and activity are logged in the Audit Log
 - o Unable to ship Audit Logs from SP directly to a syslog server – future enhancement
 - o All management actions aren't logged in the Audit Logs. EMC commands issued on the SP aren't logged. Suppose to be fixed in later release of microcode (3Q07)
 - o EMC ID in Audit Log is a two digit number. Unable to tie it back to an individual
- EMC Service credentials – uniquely identifies the user requesting access, their access role, and plnaUser ID,

User ID is a number. Need way to correlate ID# to an individual. In logs

6.7. “Storage tiering in a box”

Confirmation that EMC's new capability of transparent migration of volumes within a frame to a new tier is not yet useable in its current form. An engagement has begun with EMC to address this issue.

- Not strictly part of EMC's Beta test plan, but unable to get this to work during Beta period. Effectively no documentation yet.
- Unable to devise any form of Operational Model via Solutions Enabler CLI to wrap around this new feature in its current form. Citigroup will need to next investigate control via SMC (Symmetrix Management Console) to see if this is any better.

7. Benefits of Participation in the Beta Test Programs

7.1. Advantages of Beta program participation

- Early access to new features to determine if they meet Citigroup requirements. This provides an opportunity to collaborate on the direction of the product, whereby increasing the likelihood it meets our expectations. Note, similar testing would need to be done post GA as part of certification, but the opportunity for of collaboration would not be available. Therefore, reducing the chance of having enhancements or changes made on the product in the near timeframe.
 - **Example:** Storage-tiering-in-a-box. It was found that it was not possible to wrap an operational model around this new feature in its current form. We then had the opportunity to give EMC plenty of feedback on this.
 - Opportunity to significantly influence change to the product:
 - Get changes made to the product before GA release. Failure to participate can result in a GA product not up to Citigroup standards. A number of defects that would have impacted viability were fixed by EMC following feedback from Citigroup. **For example:**
 - Various significant stability issues
 - Various SRDF performance scalability issues
 - SRDF/A Reserve Capacity performance
 - Quality of service compatibility issues
 - Data protection interoperability
 - Resource utilization reporting errors
 - Event reporting failures
 - A number of enhancement requests have been submitted BEFORE product is even GA and entered into EMC's queue. **For example:**
 - Improvements to SRDF/A Reserve Capacity invocation alerting
 - Requirement for Best Practice to ensure SRDF/A Reserve Capacity is not a performance bottleneck
 - Changes required to allow better dynamic management of SRDF/A Reserve Capacity feature
 - Quality of service level definitions
 - Drive configuration for RAID protection
 - Audit log fields and format
- Example from a previous Beta:** Citigroup asked for 18 enhancement requests to SRDF/Star following the Beta. EMC then delivered 16 of them in a later release.
- Opportunity to understand the design concept and direction of the product through direct access to development engineering staff. Beta program creates an environment of cohesion between Citigroup and EMC, providing a much more open forum for information exchange and collaboration.
 - EMC dedicates a large amount of resources to beta programs; therefore, they are selective in the number of participants and who participates in the program. Beta programs help develop

close on-going relationships with the engineering and management staff, whereby providing Citigroup with a greater likelihood to participate in future beta evaluations, and elevates our product requests as it is from trusted individuals.

- o **Example:** EMC only use two external customers for SRDF Beta testing; Citigroup and BP. Hence, with such limited input, our comments carry enormous weight, and at a crucial time in the product development lifecycle.
- EMC have loaned us additional equipment, due to the valuable input we provide to them in Beta tests.
- The knowledge gained from early and detail evaluations of new features/products, provides a definite advantage when evaluating similar products from other vendors.

7.2. Disadvantages of Beta program participation

- Beta code is inherently less stable and testing can take longer than with GA code. Some repeat testing is sometime needed as new code drops can arrive on a weekly basis.
- Requires Citigroup to put hardware into the Beta test environment, and commit engineering resources at a specific time.