EMC GDDR Solution Design and Implementation Techniques

Version 2.0

- Business Requirement Definitions
- Project Planning and Change Management Guidelines
- Technology Selection and Solution Design
- Architecture and Sizing Considerations

Dean Smith
Peter Callewaert
Chris Fallon
John Egan
Preface

Chapter 1  Defining Business Requirements

Introduction ....................................................................................... 16
Understanding Business Continuity requirements ..................... 18
   Business continuity example......................................................... 18
Disaster recovery ............................................................................. 19
Disaster restart ............................................................................. 20
Database management............................................................... 20
High Availability ........................................................................... 21
Recovery Time Objective ................................................................. 22
Recovery Point Objective................................................................. 23
EMC GDDR role................................................................................ 24

Chapter 2  Understanding GDDR

GDDR overview ................................................................................ 28
GDDR supported configurations ................................................... 30
   SRDF/S with ConGroup ........................................................... 31
   SRDF/S with AutoSwap............................................................. 34
   SRDF/A........................................................................................ 36
   SRDF/Star.................................................................................... 39
   SRDF/Star with AutoSwap....................................................... 44
GDDR requirements ......................................................................... 47
   Mainframe environment requirements .................................... 47
   Minimum hardware requirements .......................................... 48
   Minimum software requirements ........................................... 49
   Additional configuration requirements .................................... 49
GDDR fundamental concepts ......................................................... 51
# Contents

- GDDR control systems .............................................................. 51
- Workload location ..................................................................... 52
- GDDR processes ....................................................................... 54
- GDDR components ..................................................................... 56
- Parameters .................................................................................. 56
- User interface .............................................................................. 57
- Events and rules ......................................................................... 69
- Procedures ................................................................................... 70
- GDDR procedures ........................................................................ 71
- GDDR supported scripts ................................................................. 72
  - Script mapping............................................................................ 72
  - GDDR for SRDF/S with ConGroup (Two sites - DC1 and DC2)............................................................................ 76
  - GDDR for SRDF/S with AutoSwap (Two sites - DC1 and DC2)............................................................................... 78
  - GDDR for SRDF/A (Two sites - DC1 and DC3)................................................................................... 81
  - GDDR for SRDF/Star (Three sites)................................................. 83
  - GDDR for SRDF/Star with AutoSwap (Three sites)................. 89

## Chapter 3 Planning for your GDDR Implementation

- Understanding project constraints............................................... 100
- Importance of test systems............................................................ 101
- Minimizing operational risks during implementations............ 102
  - Parallel infrastructure .............................................................. 102
  - Partial production..................................................................... 103
- Importance of design................................................................. 104
  - Sizing ...................................................................................... 104
  - Storage layout ........................................................................... 104
- GDDR Solutions Support Team ................................................. 104
- Constructing timeline and milestones.......................................... 106
- Review change management procedures ................................... 107

## Chapter 4 GDDR Managed Technologies

- EMC Symmetrix VMAX with EMC Enginuity .......................... 110
  - Symmetrix VMAX hardware platform .................................. 110
  - EMC Enginuity operating environment ................................. 112
  - Symmetrix features for mainframe........................................ 112
  - EMC Mainframe Enablers....................................................... 113
- EMC ResourcePak Base for z/OS.............................................. 115
  - Tasks ...................................................................................... 115
  - Functionality ........................................................................... 116
## Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>GDDR Managed Systems Design Considerations</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>Overview</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>Ignore, Reset Clear, or IPL in place</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>Contingency systems and workload movement</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>HMC-Only systems</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>LPAR and CEC recovery</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>C-System</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Production system options</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Managed system options</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>Automated Configuration Check – DASD</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>GDDR variables influencing the expected configuration</td>
<td>155</td>
</tr>
<tr>
<td>6</td>
<td>Storage Infrastructure Design Considerations</td>
<td>158</td>
</tr>
<tr>
<td></td>
<td>Storage configuration and layout</td>
<td>158</td>
</tr>
<tr>
<td></td>
<td>Infrastructure recommendations</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>Gatekeeper device list</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>Started tasks names</td>
<td>161</td>
</tr>
<tr>
<td></td>
<td>Important dataset names</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>EMC software levels</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>LPAR configuration specifics</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td>Sysplex and coupling facility considerations</td>
<td>167</td>
</tr>
</tbody>
</table>
Chapter 7  Network Design Best Practices
Overview........................................................................................................ 172
Understanding bandwidth requirements ..................................... 173
Synchronous network design considerations............................ 176
Asynchronous network design considerations ...................... 179

Glossary
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Continuity example</td>
<td>19</td>
</tr>
<tr>
<td>RTO example</td>
<td>22</td>
</tr>
<tr>
<td>RPO example</td>
<td>23</td>
</tr>
<tr>
<td>EMC Foundation technologies</td>
<td>25</td>
</tr>
<tr>
<td>SRDF/S with ConGroup environment</td>
<td>31</td>
</tr>
<tr>
<td>GDDR complex with SRDF/S</td>
<td>33</td>
</tr>
<tr>
<td>SRDF/S with AutoSwap environment</td>
<td>34</td>
</tr>
<tr>
<td>SRDF/A environment</td>
<td>36</td>
</tr>
<tr>
<td>GDDR complex with SRDF/A</td>
<td>37</td>
</tr>
<tr>
<td>Concurrent SRDF/Star environment</td>
<td>39</td>
</tr>
<tr>
<td>Cascaded SRDF/Star environment</td>
<td>40</td>
</tr>
<tr>
<td>GDDR Complex with Star</td>
<td>42</td>
</tr>
<tr>
<td>Concurrent SRDF/Star with AutoSwap environment</td>
<td>44</td>
</tr>
<tr>
<td>Cascaded SRDF/Star with AutoSwap environment</td>
<td>45</td>
</tr>
<tr>
<td>Primary Options Menu</td>
<td>58</td>
</tr>
<tr>
<td>Change GDDR ISPF Profile Variable Values</td>
<td>60</td>
</tr>
<tr>
<td>Setup and Maintenance</td>
<td>61</td>
</tr>
<tr>
<td>View GDDR Configuration</td>
<td>62</td>
</tr>
<tr>
<td>Manage Site Roles</td>
<td>63</td>
</tr>
<tr>
<td>Perform Health Check</td>
<td>64</td>
</tr>
<tr>
<td>Select Script to Run (Cascaded SRDF/Star)</td>
<td>66</td>
</tr>
<tr>
<td>View GDDR Script Statistics screen</td>
<td>67</td>
</tr>
<tr>
<td>Actions Menu</td>
<td>68</td>
</tr>
<tr>
<td>VMAX architecture</td>
<td>111</td>
</tr>
<tr>
<td>z/OS SymmAPI architecture</td>
<td>115</td>
</tr>
<tr>
<td>TimeFinder/Mirror under Enginuity 5773 and earlier</td>
<td>120</td>
</tr>
<tr>
<td>TimeFinder/Mirror under Enginuity 5874 and later</td>
<td>121</td>
</tr>
<tr>
<td>SRDF/A delta sets</td>
<td>123</td>
</tr>
<tr>
<td>Delta set switching</td>
<td>125</td>
</tr>
<tr>
<td>SRDF/A cycles and delta sets</td>
<td>127</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>31</td>
<td>Cascaded SRDF configuration</td>
</tr>
<tr>
<td>32</td>
<td>SRDF/EDP basic configuration</td>
</tr>
<tr>
<td>33</td>
<td>Primary and secondary relationships</td>
</tr>
<tr>
<td>34</td>
<td>Failed links</td>
</tr>
<tr>
<td>35</td>
<td>Primary systems 1, 2, and 3 in a consistency group</td>
</tr>
<tr>
<td>36</td>
<td>AutoSwap example</td>
</tr>
<tr>
<td>37</td>
<td>Concurrent SRDF/Star configuration</td>
</tr>
<tr>
<td>38</td>
<td>Cascaded SRDF/Star configuration</td>
</tr>
<tr>
<td>39</td>
<td>Cascaded SRDF/Star with EDP (diskless) configuration</td>
</tr>
<tr>
<td>40</td>
<td>GDDR Actions Menu</td>
</tr>
<tr>
<td>41</td>
<td>GDDR Perform HMC LPAR Actions</td>
</tr>
<tr>
<td>42</td>
<td>Write distribution</td>
</tr>
<tr>
<td>43</td>
<td>Bandwidth requirement</td>
</tr>
<tr>
<td>44</td>
<td>DASD write activity by volser</td>
</tr>
<tr>
<td>45</td>
<td>DASD Volumes Flagged for Tuning</td>
</tr>
<tr>
<td>46</td>
<td>Recommended asynchronous network topology</td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1 Mainframe environment requirements</td>
<td>47</td>
</tr>
<tr>
<td>2 Minimum hardware requirements</td>
<td>48</td>
</tr>
<tr>
<td>3 Software requirements</td>
<td>49</td>
</tr>
<tr>
<td>4 Script names and descriptions</td>
<td>72</td>
</tr>
<tr>
<td>5 Script mapping</td>
<td>74</td>
</tr>
<tr>
<td>6 GDDR managed system recover map</td>
<td>145</td>
</tr>
<tr>
<td>7 Software levels</td>
<td>164</td>
</tr>
</tbody>
</table>
This EMC Engineering TechBook draws on EMC Geographically Dispersed Disaster Restart (GDDR) field expertise, highlighting best practices relating to both technology and project management disciplines required to achieve success during GDDR implementations. This field-inspired guide starts with the first risk area to a business continuity project: basic understanding of business continuity or high availability requirements, then proceeds through analysis, technology selection, and architecture design guidelines. Following the guidelines in this GDDR TechBook will help ensure that your business continuity or high availability project is successful.

As part of an effort to improve and enhance the performance and capabilities of its product lines, EMC periodically releases revisions of its hardware and software. Therefore, some functions described in this document may not be supported by all versions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes. If a product does not function properly or does not function as described in this document, please contact your EMC representative.

Note: This document was accurate as of the time of publication. However, as information is added, new versions of this document may be released to the EMC Powerlink website. Check the Powerlink website to ensure that you are using the latest version of this document.

Audience

This manual is intended for customers considering improvements to business continuity or high availability technology deployments.
The following is a list of related documents that may assist readers with more detailed information on topics described in this TechBook. These documents can be found on the EMC Powerlink website at http://Powerlink.EMC.com.

- ResourcePak Base for z/OS
- SRDF Host Component for z/OS
- TimeFinder/Clone Mainframe SNAP Facility
- TimeFinder/Mirror for z/OS
- Consistency Groups for z/OS
- TimeFinder Utility for z/OS
- AutoSwap
- GDDR for SRDF/Star with AutoSwap
- GDDR for SRDF/Star
- GDDR for SRDF/S with AutoSwap
- GDDR for SRDF/S with ConGroup
- GDDR for SRDF/A

Dean Smith is a Consulting Software Engineer and has recently rejoined EMC with the GDDR team after working abroad for several years. With nearly 33 years of IT experience in the Mainframe environment in both client and systems development, Dean has focused exclusively on the justification, architecture, and implementation of business continuity solutions world wide for the last 12 years.

Peter Callewaert is a Senior Solutions Architect at EMC, with an active role in the evolution of GDDR as a product, in the troubleshooting and fixing of product issues, and in providing support to the GDDR Solutions Support team. He has been with EMC for 12 years in various roles including Pre-Sales, Customer Service, and on-site Implementation. He has 20+ years of IT-experience mainly in the Mainframe environment, with an ever-increasing focus on Disaster Recovery and Business Continuity.

Chris Fallon is a Consulting Corporate System Engineer in the EMC Symmetrix Engineering and Virtualization product group with over 22 years experience in enterprise data center systems and storage expertise. Chris has been working with the Symmetrix since 1995 and
SRDF since 1997 where he has developed, architected, and deployed geographically dispersed business continuity solutions on the Symmetrix storage array.

John Egan is an EMC Distinguished Engineer in the EMC Symmetrix Engineering and Virtualization product group. He leads the System z Automation team, responsible for the GDDR and EzSM products. His background extends over 30 years with specialization in large scale systems integration, software engineering, project management, and design.

EMC uses the following conventions for special notices.

**Note:** A note presents information that is important, but not hazard-related.

**CAUTION**

A caution contains information essential to avoid data loss or damage to the system or equipment.

**IMPORTANT**

An important notice contains information essential to operation

**Typographical conventions**

EMC uses the following type style conventions in this document:

**Normal**

Used in running (nonprocedural) text for:

- Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus)
- Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, functions, utilities
- URLs, pathnames, filenames, directory names, computer names, filenames, links, groups, service keys, file systems, notifications

**Bold**

Used in running (nonprocedural) text for:

- Names of commands, daemons, options, programs, processes, services, applications, utilities, kernels, notifications, system calls, man pages

Used in procedures for:

- Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus)
- What user specifically selects, clicks, presses, or types
Preface

We'd like to hear from you!

Your feedback on our TechBooks is important to us! We want our books to be as helpful and relevant as possible, so please feel free to send us your comments, opinions and thoughts on this or any other TechBook:

TechBooks@emc.com
This chapter contains the following information:

- Introduction ........................................................................................ 16
- Understanding Business Continuity requirements ...................... 18
- High Availability .............................................................................. 21
- Recovery Time Objective ................................................................. 22
- Recovery Point Objective ................................................................. 23
- EMC GDDR role ............................................................................... 24
Defining Business Requirements

Introduction

This chapter introduces EMC® Geographically Dispersed Disaster Restart (GDDR) and discusses its role in High Availability (HA), Business Continuity (BC) and disaster restart (DR) strategies.

In recent years, business continuity requirements have changed for many enterprises, particularly in response to the threat of terrorism and major natural disasters. What has evolved is the notion of a regional disaster, whereby local synchronously replicated centers are in regions. However, a regional disaster, such as a major flood or prolonged and widespread power disruption, could concurrently affect both data centers. In response, many sites have re-evaluated disaster readiness capabilities.

The shift toward online/internet enabled enterprises has continued to build high availability and continuous compute architectures to support 24X7 business operations.

Consistent with these trends, some customers are building data centers and technology infrastructures to match changing business requirements. Common data center strategies to support HA requirements include:

- Customers running synchronous replication add third-site asynchronous out-of-region disaster recovery centers. This is especially true for large financial institutions.
- Customers with high (HA) or continuous availability requirements build adjacent primary and secondary data centers to gain the benefit of synchronous data replication with minimal performance impact.
- Customers modify existing data center infrastructure for redundant power connect and physical firewall separation for compute and/or storage infrastructure.

These strategies facilitate synchronous replication to operate with minimum latency overhead and also enables construction of very high-speed, high-bandwidth networks to support processor-to-processor interconnect for the IBM Parallel Sysplex Coupling Facility.
The larger financial institutions are trending toward building combined HA and Disaster Recovery (DR) solutions where both synchronous and asynchronous replication architectures are combined with data replication and management infrastructure to detect and manage potentially catastrophic disaster-in-progress events.
Understanding Business Continuity requirements

Business Continuity (BC) requirements tend to cluster around industry verticals. Examples include:

◆ Telecommunications companies typically deploy some form of two-site replication and recovery architecture.
◆ Large financial institutions are trending towards three-site architectures that support both HA and DR solutions in three data center strategies.
◆ The largest retailers tend towards geographically separated two-site data center strategies, some deploying asynchronous and others synchronous data replication strategies.

This section includes information on the following:

◆ “Business continuity example” on page 18
◆ “Disaster recovery” on page 19
◆ “Disaster restart” on page 20
◆ “Database management” on page 20

Business continuity example

Figure 1 on page 19 illustrates that as the size, sophistication, and wealth of an enterprise increase over time, business continuity through disaster recovery requirements shift from essentially no protection, to business restart, and then towards high availability.
Disaster recovery

Disaster recovery is the process of restoring a previous copy of the data as the starting point for resumption of business operations, typically from a tape image. It is then necessary to apply database logs or perform other manual processes to bring the data to the last known point of consistency. Depending on the size of the enterprise, this process can range from a few hours to days and weeks.
Disaster restart

Disaster restart is the process of restarting dependent-write consistent copies of data. This type of restart data image is created by EMC data replication technologies. The image is presented to the operating system for IPL. The data itself is in crash restart form, the exact same image that would be available following the loss of power at a primary (non-replicated) data center. Disaster restart displaces tape-based recovery and greatly reduces the Recovery Time Objective (RTO).

Database management

Database management systems (DBMS) and mainframe applications have evolved over time to restart from these data images. Any incomplete or partially complete transactions are backed out by the application layer to find a point where business resumption can occur. Any backed out or incomplete transactions are re-applied after business operations resume. The time required for restart depends on the data change rate and state of the databases at the time of disruption:

- If a database is shut down normally, achieving a point of consistency on restart requires minimal work.
- However, if the database terminates abnormally or is resumed from a restart data image, the restart process is elongated, depending on the number and size of in-flight transactions at the time of DBMS termination.

An image of a database created using EMC consistency technology while the database is in operation is in a dependent-write consistent state, also known as a DBMS restartable image. The restart of this image transforms it to a transactionally consistent data state by completing committed transactions and rolling back uncommitted transactions during the normal database initialization process.
High Availability

High Availability (HA) solutions are built on synchronous data replication to enable EMC AutoSwap™ host software to interact with storage arrays to achieve a non-disruptive swap of DASD I/O from a source array to a synchronous target array. Swap operations can occur automatically for unplanned events, such as:

- The loss of all FICON connectivity between a z/OS host and its attached EMC Symmetrix® array(s)
- Planned events through an operator command interface
- Planned events through panel dialogs supplied with GDDR

*Synchronous* data replication is required for swap operations because of the requirement for the replicated data to be an exact image of the primary disk at the precise moment of the swap event.

*Asynchronous* replication is chronologically behind the primary storage instance (by measure of seconds or minutes); therefore it is *not* possible to deploy AutoSwap into an asynchronous replication environment.

Some customers require both HA and DR solutions and deploy a combination of synchronous and asynchronous replication infrastructures. These solutions provide non-disruptive swaps to the secondary synchronous disk for localized disasters, such as severing the FICON cables between a host and its local storage, as well as DR protection in the event of a catastrophic loss of the primary and secondary data centers.
Recovery Time Objective

To gain an understanding of the type of replication and recovery infrastructure that must be in place to support an enterprise, it is typical to first classify applications by criticality.

Many customers have three or more application tiers, whereby applications are categorized by time-to-recover following a planned, or unplanned, service interruption.

Ranges can vary from high availability (RTO = zero) through to hours or days to recover; it depends on the impact to the business as a trade-off to the cost of protection against service interruption.

Figure 2 illustrates the relationship between the cost of the solution and the cost of the service interruption. Once the RTO is known, the cost of the solution can be determined.
Recovery Point Objective

Applications must also be classified according to the amount of data loss that could be sustained as result of recovery from service interruption. The measure of data loss is termed Recovery Point Objective (RPO).

Typically:
- Synchronous data replication supports a near zero RPO
- Asynchronous data replication supports RPO measured in terms of seconds, minutes or hours

For asynchronous data replication, the measure of data loss is usually proportional to the amount of Wide Area Network (WAN) bandwidth applied to the solution.

Figure 3 illustrates the relationship between the cost of the solution and the cost of the service interruption. Once the RPO is known, the cost of the solution can be determined.

![Figure 3 RPO example](image)
EMC GDDR role

EMC Geographically Dispersed Disaster Restart (GDDR) is a software product that runs on the IBM z/OS operating system to automate business recovery following both planned outages and disaster situations, including the total loss of a data center.

GDDR does not provide replication and recovery services itself, but rather monitors and automates the services provided by other EMC products, as well as third-party products, required for continuous operations or business restart.

GDDR facilitates business continuity by generating scripts that can be run on demand; for example, restart business applications following a major data center incident, or resume replication to provide ongoing data protection following unplanned link outages.

Scripts are customized at the time of invocation by an expert system that tailors the steps based on the configuration and the event that GDDR is managing. Through automatic event detection and end-to-end automation of managed technologies, GDDR removes human error from the recovery process and allows it to complete in the shortest time possible.

The GDDR expert system is also invoked to automatically generate planned procedures, such as moving compute operations from one data center to another. This is the gold standard for high availability compute operations, to be able to move from scheduled DR test weekend activities to regularly scheduled data center swaps without disrupting application workloads.
Figure 4 on page 25 shows the relationship between the software layers and the Symmetrix hardware layer.

**Figure 4  EMC Foundation technologies**

GDDR is explained in more detail in Chapter 2, “Understanding GDDR.”
Defining Business Requirements
This chapter contains the following information about EMC Geographically Dispersed Disaster Restart (GDDR):

- GDDR overview .......................................................... 28
- GDDR supported configurations .................................... 30
- GDDR requirements .................................................. 47
- GDDR fundamental concepts ...................................... 51
- GDDR components .................................................. 56
- GDDR procedures .................................................... 71
- GDDR supported scripts ........................................... 72
Understanding GDDR

GDDR overview

The EMC Symmetrix Remote Data Facility (SRDF®) family of remote replication software offers various levels of Symmetrix-based business continuance and disaster recovery solutions.

EMC Geographically Dispersed Disaster Restart (GDDR) can be implemented in a variety of configurations involving two or three-sites, SRDF/S, SRDF/A, ConGroup, AutoSwap, SRDF/EDP, and SRDF/Star. In the mainframe environment, GDDR is a requirement for a SRDF/Star configuration. GDDR can manage environments that are comprised of the following elements:

- Multiple z/OS systems
- Multiple Sysplexes
- Multiple Symmetrix controllers
- Intermix of CKD and FBA/FBAM DASD and BCVs

In each configuration, GDDR provides specific capabilities tailored to that configuration. However, the major features of GDDR are common across all topologies.

Since GDDR manages production systems following disasters, it does not reside on the same servers that it is seeking to protect. GDDR resides on separate z/OS systems from the host servers that run your application workloads.

GDDR is installed on a control LPAR at each site. These control LPARs are referred to as GDDR nodes, Control Systems, or C-Systems. Each GDDR node is aware of the other GDDR nodes through network connections between each site. This multi-site awareness allows GDDR to detect disasters and identify survivors.

- In two-site configurations, GDDR can recover business at the surviving site.
- In three-site configurations, GDDR can nominate a control LPAR to assume the leadership role for GDDR and recover business at one of the surviving sites.

To achieve the task of business restart, GDDR automation extends well beyond the disk level and into the host operating system level where sufficient controls and access to third party software and hardware products exist to enable GDDR to provide automated recovery capabilities.
GDDR can distinguish normal operational disruptions from disasters and respond accordingly. For example, GDDR is able to distinguish between network outages (SRDF link drop) and real disasters. This awareness is achieved by periodic exchange of dual-direction heartbeats between the GDDR C-Systems. GDDR constantly checks for disaster situations and ensures that other GDDR systems are healthy. This checking allows GDDR to recognize, and act on, potential disaster situations, even if only one GDDR C-system survives. Split brain problems associated with cluster technologies are avoided through operator prompts. Upon the initial recognition stage, GDDR issues messages to the operator console seeking confirmation of the event and, further, confirmation of restart actions required.

In any GDDR installation, one Control System (C-System) is designated as the Master C-System.

- In a two-site configuration, failure of the Master C-System will result in the remaining C-System becoming the Master.
- In a three-site configuration, if a local or regional disaster occurs GDDR can determine which of the surviving sites will execute the recovery.

Changes to GDDR configuration information can only be made on the GDDR Master Control System. GDDR propagates these changes to the subordinate GDDR C-system(s) using inter-system communications. Restart procedures following disasters are coordinated from the GDDR Master C-System. GDDR scripts, whether for planned or unplanned scenarios, can only be executed from the Master C-System. If necessary, the Master role can be manually transferred from one C-System to another through the GDDR user interface.
GDDR supported configurations

EMC GDDR is available to monitor and manage the following configurations and topologies, each discussed further in this section:

Two-site solutions:

◆ SRDF/S with ConGroup — The two-site SRDF/S with ConGroup configuration provides disaster restart capabilities at site DC2.
  For more information, refer to “SRDF/S with ConGroup” on page 31.
◆ SRDF/S with AutoSwap — The two-site SRDF/S with AutoSwap configuration provides for near-continuous availability through device failover between DC1 and DC2.
  For more information, refer to “SRDF/S with AutoSwap” on page 34.
◆ SRDF/A — The two-site SRDF/A configuration provides disaster restart capabilities at site DC3.
  For more information, refer to “SRDF/A” on page 36.

Three-site solutions:

◆ SRDF/Star — The three-site SRDF/Star configuration provides disaster restart capabilities at either the Secondary DASD site or DC3. Concurrent and Cascaded SRDF support further minimize the DC3 recovery time objective. R22 support is available, removing the need to create device pairs during GDDR scripts.
  For more information, refer to “SRDF/Star” on page 39.
◆ SRDF/Star with AutoSwap — The three-site SRDF/Star with AutoSwap configuration provides for near-continuous availability through device failover between DC1 and DC2 as well as disaster restart capabilities at DC3. Concurrent and Cascaded SRDF support further minimize the DC3 recovery time objective. R22 support is available, removing the need to create device pairs during GDDR scripts.
  For more information, refer to “SRDF/Star with AutoSwap” on page 44.
◆ SRDF/Star with EDP.
GDDR functionality, controlled by a parameter library, can be customized to operate in any of these two or three-site configurations. During GDDR implementations, the GDDR parameter library is customized to reflect:

- The prerequisite software stack.
- The desired data center topology (two-site versus three-site, synchronous or asynchronous).

A GDDR complex consists of GDDR control systems (C-Systems), the z/OS and open systems hosts, and EMC Symmetrix storage systems which support an organization’s mission-critical workload. A single GDDR complex is able to support multiple Sysplexes.

**SRDF/S with ConGroup**

The two-site SRDF/S with ConGroup configuration provides disaster restart capabilities at site DC2. Figure 5 illustrates GDDR operation in the SRDF/S with Consistency Group environment.
As Figure 5 shows, the relationship between the DC1 and DC2 sites is maintained through SRDF/S replication of primary disk images at DC1 to DC2. Both open systems (FBA) and mainframe (CKD) disk images can be replicated. Figure 5 also shows the two GDDR C-Systems with their heartbeat communication paths, separate from the production disk and computer facilities.

The C-system at the Primary DASD site has EMC Consistency Group software installed. It is recommended that one or more of the Primary DASD site production z/OS systems also has EMC Consistency Group software installed. SRDF/S and ConGroup ensure that at the point that GDDR receives notification of an unplanned or failure event, a point of consistency is already achieved.

In this environment, GDDR can do the following:

- Manage planned site swaps
- Restart processing at the secondary site following unplanned primary site events
- Perform standard operational tasks:
  - IPL, system reset, activate, deactivate
  - Trigger stop/start of business workloads
- Actively monitor for unplanned/failure events, including:
  - Sites
  - Systems
  - Loss of SRDF/S
  - ConGroup trip
  - Inter-site communication failure

**GDDR SRDF/S complex**

Each GDDR complex manages one consistency group. A consistency group is a named group of source (R1) volumes managed by the EMC Consistency Group (ConGroup) application as a unit. The volumes can be any mix of FBA and CKD devices on multiple Symmetrix units supporting a system, a SYSPLEX, or multiple SYSPLEXs, as well as Open systems hosts (Windows and various UNIX flavors).
Figure 6 depicts a logical view of a typical GDDR complex.

![GDDR complex with SRDF/S](image)

The following are brief descriptions of the components that comprise this GDDR complex:

**BCVs** — BCVs (Business Continuance Volumes) can be supported at each of the sites. They may be established at the DC2 site and split at the DC1 site. BCVs taken at the recovery site are often referred to as **Gold Copy** devices.

**C1 and C2** — C1 and C2 are the GDDR Control LPARs (or C-Systems) at each of the sites.

**Primary site** — The primary site is the site where the production workload is located.

**Primary DASD site** — The primary DASD (direct access storage device) site is the site where the source (R1) DASD is located. The primary DASD site is the same as the primary site.

**Secondary site** — The secondary site is the site where the contingency or standby systems are located.

**Secondary DASD site** — The secondary DASD site is the site where the target (R2) DASD is located. The secondary DASD site is the same as the secondary site.

**Sites DC1 and DC2** — Sites DC1 and DC2 are the primary and secondary data centers of critical production applications and data. DC1 is the primary site, with SRDF/S data replication to the secondary site, DC2. These sites are considered fully equivalent for...
strategic production applications, connected with highly-redundant direct network links. At all times, all production data is replicated synchronously between the two sites.

**SRDF/S with AutoSwap**

The two-site SRDF/S with AutoSwap configuration provides for near-continuous availability through device failover between DC1 and DC2. **Figure 7** illustrates GDDR operation in the SRDF/S with AutoSwap environment.

As **Figure 7** shows, the relationship between the DC1 and DC2 sites is maintained through SRDF/S replication of primary disk images at DC1 to DC2. Both open systems (FBA) and mainframe (CKD) disk images can be replicated.

**Figure 7** also shows the two GDDR C-Systems with their heartbeat communication paths, separate from the production disk and computer facilities. Each of the DC1 and DC2 production z/OS LPARs as well as both C-systems has EMC AutoSwap and EMC Consistency Group (ConGroup) software installed. AutoSwap and ConGroup ensure that a point of consistency exists whenever GDDR receives notification of an unplanned or failure event.
In this environment, GDDR can do the following:

- Manage planned site swaps
- Manage recovery after unplanned site swaps
- Perform standard operational tasks:
  - IPL, system reset, activate, deactivate
  - Trigger stop/start of business workloads
- Actively monitor for unplanned/failure events, including:
  - Sites
  - Systems
  - Loss of SRDF/S
  - ConGroup trip
  - Inter-site communication failure
  - AutoSwap events
- Configure/reconfigure:
  - Couple datasets
- Manage coupling facilities:
  - Policies
Understanding GDDR

SRDF/A

The two-site SRDF/A configuration provides disaster restart capabilities at site DC3. In case DC3 was the Primary DASD site, GDDR offers disaster restart capabilities at DC1. Figures 8 illustrates GDDR operation in the SRDF/A environment.

GDDR heartbeat communication

Active Escor/Ficon channels

Standby Escor/Ficon channels

Active SRDF links

Figure 8  SRDF/A environment

As Figures 8 shows, the relationship between the DC1 and DC3 sites is maintained through SRDF/A replication of primary disk images from DC1 to DC3. Both open systems (FBA) and mainframe (CKD) disk images can be replicated. It also shows the two GDDR C-Systems with their heartbeat communication paths, separate from the production disk and computer facilities.

GDDR does not have a requirement to freeze I/O to obtain a point of consistency. Multi-Session Consistency and SRDF/A provide the mechanism to obtain a point of consistency. At the point that GDDR receives notification of an unplanned or failure event, a point of consistency is already achieved through these foundation technologies.
In this environment, GDDR can do the following:

- Manage planned site swaps
- Restart processing at the secondary site following unplanned primary site events
- Perform standard operational tasks:
  - IPL, system reset, activate, deactivate
  - Trigger stop/start of business workloads
- Actively monitor for unplanned/failure events, including:
  - Sites
  - Systems
  - Loss of SRDF/A
  - Inter-site communication failure

**GDDR SRDF/A complex**

Each GDDR complex can manage one Multi-Session Consistency (MSC) group. An MSC group is a named group, consisting of multiple RDF groups operating in SRDF/A mode, managed by the EMC MSC control software feature as a single unit. These groups can have any mix of CKD and FBA devices, on multiple Symmetrix units supporting a system, a SYSPLEX, or multiple SYSPLEXs as well as Open systems hosts (Windows and various flavours of UNIX). Figure 9 depicts a logical view of a typical GDDR complex.
The following are brief descriptions of the components that comprise this GDDR complex:

**BCVs** — Business Continuance Volumes can be supported at each of the sites. They may be established at the DC3 site and split at the DC1 site.

**C1 and C3** — C1 and C3 are the GDDR Control LPARs (or C-Systems) at each of the sites.

**Primary site** — The primary site is the site where the production workload is located.

**Primary DASD site** — The primary DASD (direct access storage device) site is the site where the source (R1) DASD is located. The primary DASD site is the same as the primary site.

**Secondary site** — The secondary site is the site where the contingency or standby systems are located.

**Secondary DASD site** — The secondary DASD site is the site where the target (R2) DASD is located. The secondary DASD site is the same as the secondary site.

**Sites DC1 and DC3** — Sites DC1 and DC3 are the primary and secondary data centers of critical production applications and data. DC1 is the primary site, with SRDF/A data replication to the secondary site, DC3. These sites are considered fully equivalent for strategic production applications, connected with highly redundant direct network links. At all times, all production data is replicated asynchronously between the two sites.
The three-site SRDF/Star configuration provides disaster restart capabilities at DC2 or DC3. Figure 10 illustrates GDDR operation in a concurrent SRDF/Star environment.

**Figure 10** Concurrent SRDF/Star environment

Figure 11 on page 40 illustrates GDDR operation in a cascaded SRDF/Star environment.
As Figure 10 on page 39 and Figure 11 show, the relationship between the DC1 and DC2 sites is maintained through SRDF/Synchronous replication of primary disk images at DC1 to DC2. Both open systems (FBA) and mainframe (CKD) disk images can be replicated. In a concurrent configuration, the asynchronous relationship is between DC1 and DC3, while in a cascaded environment, the asynchronous relationship is between DC2 and DC3.

Figure 10 and Figure 11 also show the three GDDR C-Systems with their independent heartbeat communication paths, separate from the production disk and computer facilities. The C-system at the Primary DASD site has EMC Consistency Group software installed.
It is recommended that one or more of the Primary DASD site production z/OS systems also has EMC Consistency Group software installed. During relevant GDDR scripts, GDDR will ensure the Consistency Group software is started where needed.

In this environment, GDDR can perform the following tasks:

- Manage planned site swaps
- Manage recovery after unplanned site swaps
- Manage reconfiguration of the SRDF/Star environment between concurrent and cascaded topologies
- Manage reconfiguration of the SRDF/Star environment from cascaded to concurrent with a primary processing site move
- Perform standard operational tasks:
  - IPL, system reset, activate, deactivate
  - Trigger stop/start of business workloads
- Actively monitor for unplanned/failure events, including:
  - Sites
  - Systems
  - ConGroup trip
  - Loss of SRDF/S
  - Loss of SRDF/A
  - Inter-site communication failure
Understanding GDDR

**GDDR SRDF/Star complex**

Figure 12 depicts a logical view of a typical GDDR complex in a concurrent configuration.

![GDDR Complex with Star](SYM-00220)

The following are brief descriptions of the components that comprise this GDDR complex:

**BCVs** — BCVs (Business Continuance Volumes) can be supported at each of the sites.

**Primary site** — The primary site is the site where the production workload is located.

**Primary DASD site** — The primary DASD (direct access storage device) site is the site where the source (R1) DASD is located. The primary DASD site is normally the same as the primary site.
Secondary site — The secondary site is the site where the contingency or standby systems are located.

Secondary DASD site — The secondary DASD site is the site where the target (R2) DASD is located. The secondary DASD site is normally the same as the secondary site.

Sites DC1, DC2, and DC3 — Site DC1, DC2, and DC3 are used together in an SRDF/Star data replication three-site topology, where DC1 is the primary site with concurrent or cascaded replication to the secondary site DC2 and the tertiary site DC3. Site DC3 is the tertiary data center for critical production applications and data. It is connected with redundant network to both DC1 and DC2. Data is replicated asynchronously from the current primary DASD site or secondary DASD site (with cascaded SRDF) with an intended recovery point objective (RPO) in a short time period.
SRDF/Star with AutoSwap

The three-site SRDF/Star with AutoSwap configuration provides for near-continuous availability through device failover between sites DC1 and DC2 as well as disaster restart capabilities at site DC3. Figure 13 illustrates GDDR operation in a Concurrent SRDF/Star with AutoSwap environment.

Figure 13 Concurrent SRDF/Star with AutoSwap environment

Figure 14 on page 45 illustrates GDDR operation in a Cascaded SRDF/Star with AutoSwap environment.
Figure 14  **Cascaded SRDF/Star with AutoSwap environment**

As Figure 13 on page 44 and Figure 14 show, the relationship between the DC1 and DC2 sites is maintained through SRDF/Synchronous replication of primary disk images at DC1 to DC2. Both open systems (FBA) and mainframe (CKD) disk images can be replicated. In a concurrent configuration, the asynchronous relationship is between DC1 and DC3, while in a cascaded environment, the asynchronous relationship is between DC2 and DC3.

These illustrations show the three GDDR C-Systems with their independent heartbeat communication paths, separate from the production disk and computer facilities. Each of the DC1 and DC2 production z/OS LPARs has EMC AutoSwap and EMC Consistency Group (ConGroup) installed.
In this environment, GDDR can perform the following tasks:

- Manage planned site swaps
- Manage recovery after unplanned site swaps
- Manage reconfiguration of the SRDF/Star environment between the concurrent and cascaded topologies
- Manage reconfiguration of the SRDF/Star environment from cascaded to concurrent with a primary processing site move
- Ensure continuous operation of SRDF/Star in the event of loss or removal of the primary MSC server when SRDF/Star High Availability is configured
- Perform standard operational tasks:
  - IPL, system reset, activate, deactivate
  - Trigger stop/start of business workloads
- Actively monitor for unplanned/failure events, including:
  - Sites
  - Systems
  - ConGroup trip
  - Loss of SRDF/S
  - Loss of SRDF/A
  - Inter-site communication failure
  - AutoSwap events
- Configure/reconfigure:
  - Couple datasets
- Manage coupling facilities:
  - Policies
GDDR requirements

This section contains the following information:

- “Mainframe environment requirements” on page 47
- “Minimum hardware requirements” on page 48
- “Minimum software requirements” on page 49
- “Additional configuration requirements” on page 49

Mainframe environment requirements

The basic infrastructure must support the specific EMC technologies involved in the configuration (for example, SRDF/S with AutoSwap or SRDF/A). In addition, GDDR has the following specific infrastructure requirements:

- One LPAR is required for each C-System, preferably using DASD isolated from the managed environment.
- There must be network connectivity between all C-Systems.
- An HMC (Hardware Management Console) must be available at each site that can be accessed from each C-System (access to these HMCs can be protected by means of a private VLAN).

GDDR mainframe environment requirements are listed in Table 1.

### Table 1 Mainframe environment requirements

<table>
<thead>
<tr>
<th>Item</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor hardware configuration</td>
<td>Any system that supports current IBM mainframe operating systems</td>
</tr>
<tr>
<td>DASD hardware configuration</td>
<td>Any supported Symmetrix DASD model at an Enginuity microcode level specified in the EMC GDDR Release Notes</td>
</tr>
<tr>
<td>Software</td>
<td>Any currently supported IBM operating system</td>
</tr>
</tbody>
</table>

GDDR supports and can manage the following combinations of DASD in a single Enterprise Consistency Group:

- Single EMC Symmetrix controllers configured with any of the following:
  - All CKD devices
Understanding GDDR

- All FBA and FBA-META devices
- Any combination of CKD, FBA and FBA-META devices
- Multiple EMC Symmetrix controllers configured with any of the following:
  - All CKD devices
  - All FBA and FBA-META devices
  - Any combination of CKD, FBA and FBA-META devices

Management and monitoring of both CKD and FBA/FBA-META devices is performed from the z/OS platform where the GDDR application resides. From the GDDR point of view, CKD and FBA/FBA-META Symmetrix devices are the same; that is, each is treated no differently than the other. They are all command targets of SRDF Host Component configuration commands using local, or remote syntax.

GDDR requires that if even only one device in an RDF group is defined to GDDR, then all devices in that group must be defined to GDDR. Most GDDR actions are directed at the RDF group level (although in some cases, GDDR will act on device ranges if that is appropriate).

GDDR has no limitations on the number of EMC Symmetrix controllers/devices that can be managed. Any limitations are subject to restrictions in EMC hardware and software.

### Minimum hardware requirements

Table 2 describes the recommended minimum processor and I/O configuration for a GDDR C-System.

<table>
<thead>
<tr>
<th>Item</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logical processors</td>
<td>1 (2 are recommended)</td>
</tr>
<tr>
<td>MSU</td>
<td>15 on IBM 2084-306 (or equivalent)</td>
</tr>
<tr>
<td>Storage</td>
<td>512 MB</td>
</tr>
<tr>
<td>Logical paths to own local DASD devices</td>
<td>4</td>
</tr>
<tr>
<td>Logical paths to managed DASD devices</td>
<td>4</td>
</tr>
</tbody>
</table>
Minimum software requirements

The minimum software requirements for GDDR for the various supported configurations are listed in Table 3. The EMC GDDR Release Notes provide the latest information regarding supported software release levels. Installation procedures for the EMC software products listed can be found in the EMC Mainframe Enablers Installation and Customization Guide.

Table 3  Software requirements

<table>
<thead>
<tr>
<th>Configuration</th>
<th>HMC API or BCPii support</th>
<th>MFE</th>
<th>Consistency Group</th>
<th>AutoSwap</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRDF/S with ConGroup</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SRDF/S with AutoSwap</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SRDF/A</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRDF/Star</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>SRDF/Star with AutoSwap</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Additional configuration requirements

The following sections list additional configuration requirements for the supported GDDR configurations.

SRDF/S with ConGroup

Refer to the EMC SRDF Host Component for z/OS Product Guide for information on configuring an SRDF/S environment.

SRDF/S with AutoSwap

Refer to the EMC SRDF Host Component for z/OS Product Guide for information on configuring an SRDF/S environment.

- SRDF/S with AutoSwap has the following additional requirements:
  - CAX protection must be added to the SRDF/S-defined ConGroups.
  - LOSTOWNERPOLICY ONSWAP=OPERATOR must be specified.

The EMC Consistency Group for z/OS Product Guide and EMC AutoSwap Product Guide provide information on these items.
GDDR Solution Design and Implementation Techniques

Understanding GDDR

**SRDF/A**

Refer to the *EMC SRDF Host Component for z/OS Product Guide* for information on configuring an SRDF/A environment. GDDR is compatible with SRDF Automated Recovery functionality.

SRDF/A MSC has the following additional gatekeeper requirement:

- There must be one or more gatekeeper devices for each MSC-controlled RDF group. These gatekeeper devices must be in OS configuration as OFFLINE at IPL as regular local devices (not BCV, SRDF, SAV, and so forth).

**SRDF/Star**

Refer to the *SRDF Host Component for z/OS Product Guide* for information on configuring an SRDF/Star environment.

**SRDF/Star with AutoSwap**

Refer to the *SRDF Host Component for z/OS Product Guide* for information on configuring an SRDF/Star environment.

SRDF/Star with AutoSwap has the following additional requirements:

- CAX protection must be added to the SRDF/Star-defined ConGroup.
- LOSTOWNERPOLICY ONSWAP=OPERATOR must be specified.

The *EMC Consistency Group for z/OS Product Guide* and *EMC AutoSwap Product Guide* provide information on these items.

In addition, there must be one or more gatekeeper devices for each MSC-controlled RDF group. These gatekeeper devices must be:

- SRDF/S-protected (you can choose any available SRDF/S RDF group)
- ConGroup and CAX-protected
- Not SRDF/A-protected
- If CKD, in OS configuration as ONLINE at IPL
- If FBA, in OS configuration as OFFLINE at IPL
GDDR fundamental concepts

This section contains the following information:

- “GDDR control systems” on page 51
- “Workload location” on page 52
- “GDDR processes” on page 54

GDDR control systems

The GDDR control systems are more commonly referred to as GDDR C-Systems. One GDDR C-System is located at each site in a separate z/OS System. Each GDDR C-System runs in monoplex mode from local DASD. GDDR C-Systems do not run any production workload. EMC recommends that GDDR is installed dedicated C-System DASD on separate controllers from the production DASD.

Deploying GDDR’s C Systems as dedicated z/OS systems insures their availability for both planned and unplanned events effecting application and/or site operations. Isolation of GDDR C-Systems from the sysplex(s) insures GDDR operational isolation and provides capabilities other automated business continuity solutions cannot offer.

The main functions of a GDDR C-System are to:

- Control the recovery after an outage
- Control a planned site swap

One of the C-Systems is designated as the Master C-System. During normal operations, the Master C-System is the central control point for all GDDR activities. The Master C-System is located at the secondary DASD site in configurations with AutoSwap and on the primary DASD site in other configurations. All GDDR C-Systems are potential candidates to take over as the Master C-System and if the Master C-System becomes unavailable for some reason, a C-System at another location/site assumes the GDDR master function ownership. This way, either the Master-C system survives a site loss, or the Master-C role is automatically transferred to a surviving site. When the original Master C-System becomes available, the master function ownership automatically transfers back to the correct location.
Some GDDR functions can only be carried out by the Master C-System, for example:

- Running planned processes
- Updating GDDR parameters

**Workload location**

In a GDDR complex without AutoSwap, the business or production workload runs at a single site; that is, one side of the Sysplex. This is the same location as the primary DASD site.

In a GDDR complex with AutoSwap, the business or production workload can run as either a single site workload, usually at the same location as the primary DASD site, or as a multi-site workload, where the production workload runs at both the primary and secondary sites.

In a GDDR complex with an SRDF/A two-site configuration the production systems and primary DASD must always be at the same site.

A contingency or standby system is a system that provides backup to a production system. The characteristics of a contingency system in a GDDR installation depend on the configuration.

Contingency or standby systems are typically located at the same location as the secondary DASD. Multiple locations containing contingency or standby systems may be used to increase availability and provide disaster restart options. Regional contingency systems are typically located in the same location as the secondary DASD, while out-of-region standby systems provide protection from geographic and infrastructure exposures that may negatively impact the primary and secondary sites.

In an SRDF/S two-site configuration, a contingency system normally provides a hot backup to the production system and is typically located at the same location as the secondary DASD. The contingency system:

- Is in the same Sysplex as its production system partner
- Is IPLed, but runs no business workload

In an SRDF/A two-site configuration, a contingency or standby system is a system that replaces production system capacity in the event of a loss of use of the primary site. It is located at an appropriate
distance from the primary systems to minimize risks from geographic and infrastructure exposures which may negatively impact primary systems availability. The contingency system:

- May be used for expendable workloads which are displaced by business workload following the loss of a primary site.
- May be cold (not powered up), or warm (powered up but not IPLed) systems reserved for business workload restart and testing of restart processes, but not in support of any meaningful day-to-day workload.

A contingency or standby system in a SRDF/Star configuration normally provides a hot backup to a production system. A contingency system:

- Is in the same Sysplex as its production system partner
- Is IPLed, but runs no business workload
- May be an out-of-region location equipped with idle processor capacity reserved for support of business workload restarts

Any production or contingency/standby system defined to GDDR is known as a GDDR managed system.

**Managed workloads**

GDDR can trigger the stop and restart of production workloads on:

- z/OS systems
- Distributed systems

**External workloads**

External workloads run on mainframe systems which do not have their DASD in the managed Symmetrix units. GDDR can coordinate stop and start of the workload on these non-managed mainframe systems with the workload stop and start actions for managed systems.

**Excluded systems**

Excluded systems apply to GDDR configurations with AutoSwap only. Symmetrix arrays can be shared by multiple systems and some of those systems may not require GDDR protection. GDDR can be configured to exclude certain z/OS systems from workload management, although these systems have their DASD in the
managed Symmetrix arrays. Because of this, these systems must be running AutoSwap, or else they would fail during GDDR managed site swaps.

**HMC-Bypass option**

Applies at the LPAR-level or the Site level and prevents GDDR from performing ANY action that would require HMC access for the relevant site or LPAR. Goal of this option is to allow GDDR to be deployed at sites under a third-party Facilities Management contract, where the third-party has exclusive access to the HMC, and this HMC-access by individual customers using the provided facilities is not allowed.

**HMC-only systems**

GDDR can be configured to restrict IPL and CBU (Capacity Backup Upgrade) actions for certain systems to the online interface. No other actions or automation are performed for these systems.

**GDDR processes**

A GDDR process or script is a dynamically determined sequence of function calls to achieve a predetermined result. Generally one function call corresponds to one action. A GDDR process is executed by calling GDDR provided routines, either from a batch job or as a result of specific messages being issued. There are two types of GDDR processes, discussed in this section:

- “Planned process” on page 54
- “Unplanned process or takeover process” on page 54

**Planned process**

A GDDR planned process is initiated through the GDDR interface to perform a planned task.

**Unplanned process or takeover process**

The GDDR unplanned process or takeover process can only be initiated following an error that results in a possible takeover situation. Takeover processes are initiated as a result of certain messages being issued or specific events occurring.

The messages or events that trigger an unplanned or takeover process can originate on any system, either a C-System or a production system. In response to a trigger event the current Master C-System,
following operator confirmation of any trigger event, will invoke the necessary processes based on the current GDDR complex and managed system configurations. Processes are invoked following operator confirmation of any of the following types of failure or loss:

- Sites
- DASD
- Systems
- Loss of SRDF link
- Loss of host channels

**GDDR process restart**

In the event a GDDR process does not complete properly the return codes from the function calls that make up a GDDR process are saved in GDDR global variables. For functions that issue EMC SRDF Host Component commands, the return code of the commands are also saved. If multiple commands are issued from one function, the return codes from each command are saved in GDDR global variables.

After the cause of the original failure has been identified and resolved, the GDDR process can be rerun. GDDR uses the saved return codes to establish the point of restart; that is, the point of the previous failure. This ensures that no modifications to the supplied GDDR process jobs are required in order to rerun after a failure. If the underlying problem causing a GDDR function to fail, has been resolved in a way that makes the GDDR function call unnecessary, the GDDR interface provides an easy way to skip the failed function call. If necessary, restart can be forced in a particular step by manually setting the GDDR system variables appropriately.
Understanding GDDR

GDDR components

The EMC GDDR software environment includes many components controlled by a parameter library. Parameters define the GDDR managed configuration. Some parameters are maintained in external parmlib members, but most are maintained via the GDDR user interface and loaded into GDDR global variables by the parameter activation process. The user interface is an ISPF application available only on the GDDR C-Systems. There are events in which GDDR is interested and message rules that determine the actions to be taken in response to those events. Most GDDR procedures run as started tasks and include EMC foundation products such as ResourcePak Base and SRDF Host Component.

This section contains the following information:

◆ “Parameters” on page 56
◆ “User interface” on page 57
◆ “Events and rules” on page 69
◆ “Procedures” on page 70

Parameters

This section discusses internal parameters, next.

Internal parameters

The GDDR Parameter Wizard is the control interface for managing GDDR parameters, it should only be used should only be used on the C-system which is intended to be the Master C-System for the managed configuration. During the activation process, the new parameters are loaded into GDDR global variables on the Master C-System and propagated to the remaining C-Systems using GDDR peer-to-peer communication.
The basic process of updating GDDR parameters consists of the following steps:

- Back up the existing parameters. The parameter management process uses an existing backup of parameter values as the starting point for subsequent updates.
- Select the backup as input.
- Modify parameters as required. The updates are captured in PDS members in a work dataset.
- Validate the parameter set. This is done via a batch job submitted from the user interface.
- Activate the parameter set. This is also a batch job submitted from the user interface.

**User interface**

The GDDR user interface is an ISPF application. It is used only on the GDDR C-Systems. The following are discussed in this section:

- “System components” on page 58
- “Primary Options Menu” on page 58
- “Option P: Profile - Update Personal GDDR ISPF Profile” on page 60
- “Option M: Maintenance - GDDR Setup and Maintenance” on page 60
- “Option G: GDDR Configuration - View GDDR Configuration” on page 62
- “Option R: Roles - Manage Site Roles” on page 63
- “Option C: Checkup - Perform Pre-Script Checkup” on page 64
- “Option S: Scripts - Run GDDR scripts” on page 65
- “Option T: Timing - View GDDR Script Statistics” on page 67
- “Option A: Actions - Perform GDDR Actions” on page 68
**Understanding GDDR**

**System components**

The GDDR SAMPLIB contains an exec called GDDREXC that can be customized to allocate the required GDDR and libraries. Once customized, this exec can be integrated into a site's ISPF panels in accordance with site standards. GDDR panels will display only the options the user is authorized to use.

The GDDR ISPF user interface does not currently support operating in both screens when in split screen mode.

**Primary Options Menu**

Invoking GDDR produces the Primary Options Menu, as shown in Figure 15.

```
------------------ GDDR - Primary Options Menu for GDDRPLEX ------------------
Option ===>
P rofile    Update personal GDDR ISPF Profile   This System: SYS2
M aintenance GDDR Setup and Maintenance       This Site: DC2
G DDR Config View GDDR configuration           Master-C: SYS2
G eroles   Manage Site Roles                  Primary Site: DC1
C heckup   Perform pre-script checkup           Primary DASD: DC1
S cripts   Run GDDR Scripts                    Automation: ON
T iming    View GDDR Script Statistics         Planned script: None
A ctions   Perform GDDR Actions                Unplanned script: None
E EzSM     EMC z/OS Storage Manager            EMC Geographically Dispersed Disaster Restart 04.00.00

Copyright © 2007-2011 EMC Corporation.

Select an option and press <Enter>
Press <F3> to Terminate GDDR ISPF
```

**Figure 15 Primary Options Menu**

The Primary Options Menu displays information on the current state of GDDR in the environment. It shows the name of the C-System and site, which C-System is currently Master, which site is the primary processing site, which site is the primary DASD site, whether GDDR automation is currently enabled, and the name of the planned or unplanned script, if any, that is currently in process.
From this screen, GDDR automation can be toggled on and off using the commands GDDR ON and GDDR OFF. This menu also provides access to the following functions:

- **Update Personal GDDR ISPF Profile** allows each user to change GDDR ISPF profile variable values associated with their TSO ID. This includes items like datasets required to submit scripts, and job card information. Each user must enter their profile updates before performing any setup and maintenance functions. Refer to Figure 16, “Change GDDR ISPF Profile Variable Values” on page 60 for an example of this panel.

- **GDDR Setup and Maintenance** — Provides options to manage GDDR internal parameters, system variables and system options. Refer to Figure 17, “Setup and Maintenance” on page 61 for an example of this panel.

- **View GDDR configuration** — Displays GDDR configuration details for sites, features, and C-Systems. Refer to Figure 18, “View GDDR Configuration” on page 62 for an example of this panel.

- **Manage Site Roles** — Allows manual changes to be made to site roles, including transferring the Master role to a different C-System, changing the AutoSwap owner, and moving the MSC control function. Refer to Figure 19, “Manage Site Roles” on page 63 for a complete description.

- **Perform Pre-Script Checkup** — Displays any active events, shows the status of the GDDR internal command queue, and performs and displays the result of communication checks to validate that all production systems and C-Systems are active and all inter-systems communications are enabled via the Degraded Mode YES | NO and the MPARM Consistent | Inconsistent fields. The panel also displays the status of the SRDF/Star High Availability feature. Refer to Figure 20, “Perform Health Check” on page 64 for additional details.

- **Run GDDR Scripts** — Allows the user to run GDDR scripts. Refer to Figure 21, “Select Script to Run (Cascaded SRDF/Star)” on page 66 for an example of this panel.

- **View GDDR Script Statistics** — Displays timing information for recently-executed scripts. Refer to Figure 22, “View GDDR Script Statistics screen” on page 67 for additional details.
Understanding GDDR

- Perform GDDR Actions — Provides options to perform actions such as HMC LPAR and CBU actions, HMC discovery, and managing couple datasets. Refer to Figure 23, “Actions Menu” on page 68 for additional information.

Option P: Profile - Update Personal GDDR ISPF Profile

Choosing this option from the Primary Options Menu displays the screen shown in Figure 16.

```
GDDR Subsystem Name ===> GDDR
JCL dataset ===> 
ISPF skeleton dataset ===> 
Global variable backup ===> 

Jobcards for your user: 
// 
// 
// 

Press ENTER to apply updates
Press <F3> when finished
Enter CANCEL to return without changing any profile variable values
Enter CLEAR to set all values to null and exit
Enter RESET to restore the values as they were upon entry
```

Figure 16 Change GDDR ISPF Profile Variable Values

The GDDR Subsystem name cannot be changed but all other values requested on the screen must be filled in to allow the user to perform GDDR functions like parameter maintenance and script execution.

Option M: Maintenance - GDDR Setup and Maintenance

In general, use of this screen should be avoided without guidance from GDDR Solution Support.
Figure 17 shows the GDDR Setup and Maintenance Menu.

---

**Figure 17** Setup and Maintenance

The GDDR ON and GDDR OFF commands can be used from the command line on this menu to enable and disable GDDR automation. This menu also provides access to the following functions:

- **Manage GDDR Parameters** — Provides access to a series of screens that are used to maintain the GDDR internal parameters described in 5.4.1.1 Internal parameters. These screens will be presented in detail later in this TechBook.

- **Message, Debug, and Trace options** — Provides the ability to set default debugging levels and override those values for specific programs.

- **Manage GDDR Internal Command Queue** — Provides information on the command queue used by GDDR to store SRDF commands that have not yet been executed. This screen can be used to alter GDDR processing. EMC strongly recommends that no action be taken on this screen unless instructed to do so by GDDR Solution Support.
Understanding GDDR

- Perform HMC Discovery — Discovers the HMC objects at each site. This function can be used to validate the HMC information entered during GDDR configuration as well as the GDDR HMC interface installation and setup.
- Refresh GDDR Message Table — Refreshes the GDDRMSG table managed by GDDRMAIN
- Manage GDDR System Variables — Provides the ability to view and modify all system variables used by GDDR to describe and manage the environment. Any updates done thru this panel remain local to the current C-system and are not subject to any validation.

**Option G: GDDR Configuration - View GDDR Configuration**

Selecting this option displays a screen similar to Figure 18 with the information relevant to the GDDR configuration in use. No changes can be made on this screen.

```
----------------- GDDR - View GDDR Configuration for GDDRPLEX ----------------
Command ===> 

GDDR complex name: GDDRPLEX
Sites: DC1, DC2, DC3
Features: SRDF/S, SRDF/A, STAR, AUTOSWAP, No FBA Devices
C-Systems:
   At DC1, C-System is SYS1
   At DC2, C-System is SYS2
   At DC3, C-System is SYSB

Press <F3> to return to the previous menu
```

**Figure 18** View GDDR Configuration
GDDR components

Option R: Roles - Manage Site Roles

GDDR script processing manages the location of the Master C-System, the AutoSwap Owner system, and the MSC Control site as part of normal operations. The Manage Site Roles panel is provided to address exception conditions such as preparing for a script restart. Figure 19 shows the Manage Site Roles menu.

---

Option R: GDDR - Manage Site Roles

Option ==>  

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Current Site</th>
<th>Recommended Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>MXFR</td>
<td>Transfer Master-C system function</td>
<td></td>
<td>SYS2</td>
</tr>
<tr>
<td>MVA</td>
<td>Move AutoSwap owner</td>
<td>This Site: DC2</td>
<td></td>
</tr>
<tr>
<td>MMSC</td>
<td>Move Primary MSC Server</td>
<td>Master-C: SYS1</td>
<td></td>
</tr>
<tr>
<td>PMS</td>
<td>Planned SRDF/Star HA Takeover</td>
<td>Not Applicable</td>
<td>Primary Site: DC2</td>
</tr>
<tr>
<td>RMS</td>
<td>Restart Secondary MSC Server</td>
<td>Not Applicable</td>
<td>Primary DASD: DC2</td>
</tr>
<tr>
<td>UMT</td>
<td>Unplanned SRDF/Star HA Takeover</td>
<td>Automation: OFF</td>
<td>Planned script: None</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unplanned script: None</td>
</tr>
</tbody>
</table>

Select an option and press <Enter>

Press <F3> to return to the GDDR Primary Options menu

---

Figure 19 Manage Site Roles

The following choices are available:

- **MXFR**—Transfer Master-C system function — Displays the sites in the GDDR complex and indicates the current and recommended Master. The Master C-System function can be transferred to another site.
- **MVA**—Move AutoSwap owner — Displays the sites involved in an AutoSwap relationship and indicates the current and recommended owner. AutoSwap ownership can be transferred to another site.
- **MMSC**—Move Primary MSC Server — Displays the sites eligible to assume the MSC control function and indicates the current and recommended owner. MSC control can be transferred to another site.
Understanding GDDR

- PMTO—Planned SRDF/Star HA Takeover — In SRDF/Star HA configurations, this action performs a swap of the Primary and Secondary MSC Server roles in planned fashion.
- RSMS—Restart Secondary MSC Server — In SRDF/Star HA configurations, this action restarts a failed secondary MSC server.
- UMTO—Unplanned SRDF/Star HA Takeover — In SRDF/Star HA configurations, this action performs a re-run of an Unplanned SRDF/Star HA Takeover that was previously started by the GDDR Event Monitor but did not run to completion.

Option C: Checkup - Perform Pre-Script Checkup

When this option is selected, GDDR performs a series of health checks and returns the results. A screen similar to Figure 20 is displayed.

<table>
<thead>
<tr>
<th>Active Events:</th>
<th>MSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degraded Mode:</td>
<td>YES</td>
</tr>
<tr>
<td>MIRROR:</td>
<td>Consistent</td>
</tr>
<tr>
<td>Star-HA:</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Command</th>
<th>Queue</th>
<th>Status</th>
<th>Count</th>
<th>Active</th>
<th>Free</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

System Communication Status:

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Type</th>
<th>Status</th>
<th>Status</th>
<th>Date and Time of Status check</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC2 SY2</td>
<td>CSYS</td>
<td>Degraded</td>
<td>n/a</td>
<td>02/01/11 17:11:09</td>
</tr>
<tr>
<td>DC2 PRD1</td>
<td>PSYS</td>
<td>Active</td>
<td>Operating</td>
<td>02/01/11 17:11:09</td>
</tr>
<tr>
<td>DC3 SY3</td>
<td>CSYS</td>
<td>Degraded</td>
<td>Operating</td>
<td>02/01/11 13:52:01</td>
</tr>
<tr>
<td>DC1 PRD3</td>
<td>PSYS</td>
<td>Inactive</td>
<td>Operating</td>
<td>02/01/11 17:09:33</td>
</tr>
<tr>
<td>DC1 SY1</td>
<td>CSYS</td>
<td>Active</td>
<td>Operating</td>
<td>02/01/11 17:10:18</td>
</tr>
<tr>
<td>DC1 PRD2</td>
<td>PSYS</td>
<td>Active</td>
<td>Operating</td>
<td>02/01/11 08:52:23</td>
</tr>
<tr>
<td>DC2 PRD4</td>
<td>PSYS</td>
<td>Inactive</td>
<td>Operating</td>
<td>02/01/11 17:10:44</td>
</tr>
</tbody>
</table>

********** Bottom of data **********
In this example, the following information is provided:

- There is an MSC event.
- GDDR is running in Degraded Mode.
- The GDDMPARM member was found to be consistent.
- Star HA: 0 indicates either SRDF Star and GDDR are not configured for optional SRDF/Star High Availability support, or there is currently no Secondary MSC Server available.
- The GDDR command queue is empty.
- The DC2 C-System is operating in GDDRMAIN degraded status, and its HMC status is not available. Enter a 'D', in the Sel column for DC2 to display the System Details panel that shows additional information describing the situation.

**Option S: Scripts - Run GDDR scripts**

This option can only be chosen on the Master C-System. It displays the Select Script to Run screen, which is a dynamic display. It lists GDDR scripts that are eligible to run under the current configuration, and also specifies configuration information and scripts that are in progress, if any. Depending on the location of the primary site and primary DASD locations, the panel is customized to display valid choices. Scripts that cannot be used at the current site do not appear. The Configuration field indicates the features of the configuration that GDDR takes into account when determining which scripts are eligible to run and hence which are present in the list of scripts displayed on the panel. Figure 21 on page 66 shows a sample of the Scripts to Run screen.
### Understanding GDDR

Select Script to Run (Cascaded SRDF/Star)

<table>
<thead>
<tr>
<th>Scripts for Cascaded SRDF Reconfiguration</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconfigure to concurrent SRDF</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Reconfigure to cascaded SRDF</td>
<td>DCN CASCADE = 0</td>
</tr>
<tr>
<td>Reconfigure to concurrent SRDF with Site Move</td>
<td>DCN CASCADE = 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scripts for Planned Actions</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swap production from DC1 to DC2</td>
<td>DCN CASCADE = 0</td>
</tr>
<tr>
<td>Swap production from DC1 to DC2 - Protected</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Swap production from DC1 to DC2 - Fast</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Perform test IFL from BCVs at DC3</td>
<td>DCN CASCADE = 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scripts for Unplanned Actions</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recover after unplanned swap</td>
<td>DCN CASCADE = 0</td>
</tr>
<tr>
<td>Recover after loss of DC1 (LDR)</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Recover after loss of DC2</td>
<td>DCN CASCADE = 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scripts for Resumption</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resume SRDF/S replication after Consruct group trip</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Resume after test IFL from BCVs at DC2</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Resume SRDF/A after test IFL at DC2</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Resume SRDF/A in MSC mode to DC3</td>
<td>DCN CASCADE = 1</td>
</tr>
<tr>
<td>Resume SRDF/A (SRDF/Star AutoSwap) to DC3</td>
<td>DCN CASCADE = 1</td>
</tr>
</tbody>
</table>

---

**Figure 21** Select Script to Run (Cascaded SRDF/Star)
Option T: Timing - View GDDR Script Statistics

Choosing this option displays a list of scripts that have run on this C-System, as shown in Figure 22.

--- Script Selection for Status View ---

Option ===>

Enter S next to a script to select it for viewing
Press <F3> to return to Primary Option Menu
Last planned: Resume SRDF/A (SRDF/Star Autoswap) to DC3
Last unplanned: (none)

Script

- _GDDRPA42 - Swap processing from DC1 to DC2
- _GDDRPA21 - Swap DASD from DC1 to DC2
- _GDDRPA29 - Restart SRDF/A in MSC mode to DC3
- _GDDRPAAB - Abandon sites DC1 and DC2
- _GDDRPA29 - Resume SRDF/A (SRDF/Star Autoswap) to DC3 Last planned
- _GDDRPA23 - Resume SRDF/S replication after ConGroup trip

*********************** Bottom of data ****************************

Figure 22 View GDDR Script Statistics screen

Selecting a script displays step start and end times. If a script were currently running, it would be displayed on the "Last planned:" or "Last unplanned:" line.
### Understanding GDDR

#### Option A: Actions - Perform GDDR Actions

Figure 23 is the sub-menu displayed when Option A is chosen.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Perform HMC Discovery</td>
</tr>
<tr>
<td>L</td>
<td>Perform HMC LPAR actions</td>
</tr>
<tr>
<td>CBU</td>
<td>Perform HMC CBU actions</td>
</tr>
<tr>
<td>5</td>
<td>Manage Couple Datasets</td>
</tr>
</tbody>
</table>

From this menu, the following actions are available:

- **Perform HMC Discovery** — Discovers the HMC objects at each site. This function can be used to validate the HMC information entered during GDDR configuration as well as the GDDR HMC interface installation and setup. This function is also available from the GDDR Setup and Maintenance menu.

- **Perform HMC LPAR actions** — Provides the ability to carry out actions against the GDDR defined sites and systems including activating or deactivating an LPAR, performing a load clear, reset clear, stop or start for an LPAR, and querying an LPAR for its status. This option can only be used on the Master C-System. Refer to Figure 41, “GDDR Perform HMC LPAR Actions” on page 149 for all HMC LPAR action descriptions.
Understanding GDDR

- Perform HMC CBU actions — Presents a list of sites with processors enabled for CBU (Capacity Backup Upgrade) activation, if any. Action codes can be used to activate backup capacity (REAL mode), simulate activation of backup capacity (TEST mode), terminate actual backup capacity activation, or terminate simulated backup capacity activation. These actions can only be performed on the Master C-System.

- Manage Couple Datasets — Checks the couple dataset configurations on all systems. Currently used couple datasets are compared against those defined on the “Define Managed Couple Datasets” panel during GDDR parameter definition. If discrepancies are found, SETXCF commands are issued to bring couple dataset usage in line with GDDR parameter specifications. This function is provided to adjust couple datasets when it is not appropriate to run a site swap script, which will also align couple dataset usage with the defined parameters.

Events and rules

A GDDR event is a state change in a component part of the environment that GDDR is actively monitoring. Examples include:

- CGT — ConGroup trip has occurred / state change
- CGD — ConGroup group is disabled / state change
- SRA — SRDF/A link is down
- MHB — Missing C-System heartbeat
- LNK — Link to a C-System down
- RDF — RDF link down
- LDR — Local Disaster
- RDR — Regional Disaster

Events are stored in GDDR global variables. An event can have a state of either TRUE or FALSE. If the event has a state of TRUE, it has occurred or is currently occurring. If the event has a state of FALSE, it is no longer occurring. An event that is TRUE is considered an exception. GDDR events are used by the GDDR event monitor and GDDR processes to determine environment state. A change in state can then:

- Trigger unplanned / takeover processes
- Prevent a planned process from running
GDDR is supplied with message interception rules. These rules have two primary functions:

- To detect events that GDDR is interested in and set the appropriate GDDR event TRUE or FALSE.
- To detect events that GDDR processes have to wait for (WTOR), and reply as to the success or failure of the waited for event. This will determine if a GDDR process proceeds or terminates.

GDDR uses intersystems communication to route message traffic between production systems and C-systems. The arrival of a message at the target production system can be used to trigger an automation rule (for example in IBM Tivoli NetView or BMC Control-M). Such rules can be used to start or shut down workloads on the appropriate systems.

**Procedures**

There are several JCL procedures used in a GDDR environment, though most are not part of GDDR itself. They include the tasks for EMC foundation products, such as ResourcePak Base and SRDF Host Component. The names presented here are the vendor defaults and may not be the names used in any given environment.

In case a customer environment does not support the BCP ii HMC interface, CA-OPS is required. Refer to the GDDR product guides for more information.

**EMC foundation products**

Depending on the GDDR configuration in use, there will be two or three EMC foundation product tasks running. All environments will run EMCSFC and EMCRDF, and environments including SRDF/S will also run EMCCGRP:

- EMCSFC — ResourcePak Base or Symmetrix Control Facility
- EMCRDF — SRDF Host Component
- EMCCGRP — Consistency Group

For more information on the EMC foundation products refer to the EMC product documentation.
GDDR procedures

The following are JCL procedures needed in a GDDR environment:

- **GDDRMAIN** — Main GDDR address space.
- **GDDREVM** — GDDR Event Monitor. The event monitor is used to analyze event state changes in which GDDR is interested. On detecting the occurrence of selected events, the event monitor determines what action to take and prompts operators with the appropriate choices via WTOR.
- **GDDRHBM** — GDDR C-System heartbeat monitor. The heartbeat monitor aids the event monitor in determining the status of the GDDR managed environment. The lack of a heartbeat from a particular C-System is used to determine the state of a C-System and the site.
- **GDDRWORK** — GDDR Workers. Some pieces of work required for proper GDDR functioning will run under separate GDDRWORK started tasks. This includes all GDDR HMC functions, GDDR Command queue processing, parts of GDDR scripts shipped to a remote C-system.
- **GDDRPROC** — Used to run scripts. This is the only GDDR procedure not used for a started task that is up and running at all times.

Started task locations

The started tasks that need to run on any given system in the GDDR complex depend on whether it is a managed production or C-System, which site it is, and which configuration is in use.

- **GDDRMAIN** runs on all C-systems in the GDDR complex. In three-site configurations, it optionally does not run at DC3, if the customer selects to run DC3 as a “lights-out” data center. GDDRMAIN also runs on all fully managed production systems.
- **GDDREVM**, GDDRHBM and GDDRWORK run on all C-Systems at all sites in the GDDR complex. They are started automatically by GDDRMAIN. They are not required on the productions systems.
- **EMCSF** runs on all C-Systems and all IPLed production systems at all sites.
- **EMCRDF** runs on all C-systems.
- **EMCCGRP** runs on C-Systems and one or more production systems at the primary DASD site when SRDF/S is in use. It runs on all systems both at DC1 and DC2 in AutoSwap configurations.
GDDR supported scripts

This section contains the following information:

- “Script mapping” on page 72
- “GDDR for SRDF/S with ConGroup (Two sites - DC1 and DC2)” on page 76
- “GDDR for SRDF/S with AutoSwap (Two sites - DC1 and DC2)” on page 78
- “GDDR for SRDF/A (Two sites - DC1 and DC3)” on page 81
- “GDDR for SRDF/Star (Three sites)” on page 83
- “GDDR for SRDF/Star with AutoSwap (Three sites)” on page 89

Script mapping

Table 4 shows the mapping of script names with a short description of the script function as well as the event type that causes the script to be generated.

Table 4 Script names and descriptions (page 1 of 2)

<table>
<thead>
<tr>
<th>Script name</th>
<th>Description</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDDRPA51</td>
<td>Reconfigure to Concurrent</td>
<td>RECONFIGURE</td>
</tr>
<tr>
<td>GDDRPA52</td>
<td>Reconfigure to Cascaded</td>
<td>RECONFIGURE</td>
</tr>
<tr>
<td>GDDRPA53</td>
<td>Reconfigure to concurrent SRDF with Site Move</td>
<td>RECONFIGURE</td>
</tr>
<tr>
<td>GDDRPA42</td>
<td>Swap production from &lt;P&gt; to &lt;S&gt; (Protected)</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDDRPA45</td>
<td>Swap production from &lt;P&gt; to &lt;S&gt; (FAST)</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDDRPA21</td>
<td>Swap DASD from &lt;P&gt; to &lt;S&gt; (Protected)</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDDRPA25</td>
<td>Swap DASD from &lt;P&gt; to &lt;S&gt; - FAST</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDD2P17A</td>
<td>Abandon Site &lt;P&gt; –(site swap)</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDD2P18A</td>
<td>Restart production at &lt;S&gt; after site swap</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDD2P01A</td>
<td>Perform test IPL from BCVs at &lt;tgt site&gt;</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDDRPA27</td>
<td>Perform test IPL from R2s at DC3</td>
<td>PLANNED</td>
</tr>
<tr>
<td>GDD2P03A</td>
<td>Perform test IPL from R2s at &lt;S&gt;</td>
<td>PLANNED</td>
</tr>
</tbody>
</table>
Understanding GDDR

<table>
<thead>
<tr>
<th>Script name</th>
<th>Description</th>
<th>Event Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDDRUP31</td>
<td>Recover after unplanned swap</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDD2U09A</td>
<td>Recover after unplanned swap or LDR</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDD2U13A</td>
<td>Recover after loss of &lt;P&gt; (LDR)</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDD2U10A</td>
<td>Recover after loss of &lt;P&gt; (LDR)</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDD2U12A</td>
<td>Recover after loss of &lt;P&gt; (LDR)</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRUP41</td>
<td>Recover after loss of &lt;S&gt;</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRPA0A(1)</td>
<td>Resume replication after unplanned swap</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRPA0A(2)</td>
<td>Resume replication after loss of &lt;P&gt;</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRPA0A(3)</td>
<td>Resume replication after loss of &lt;S&gt;</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRPA0A(4)</td>
<td>Resume replication after unplanned swap</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRPA0A(5)</td>
<td>Resume replication after loss of &lt;P&gt;</td>
<td>UNPLANNED</td>
</tr>
<tr>
<td>GDDRPA23</td>
<td>Resume SRDF/S replication after ConGroup trip</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDD2P02A</td>
<td>Resume after test IPL from BCVs at &lt;tgt site&gt;</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDDRPA28</td>
<td>Resume SRDF/A after test IPL at DC3</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDDRPM29</td>
<td>Resume SRDF/A in MSC mode</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDDRPF29</td>
<td>Resume SRDF/A in STAR mode</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDD2P14A</td>
<td>Resume replication after link failure</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDD2P16A</td>
<td>Resume after test IPL from R2s at &lt;S&gt;</td>
<td>RESUME</td>
</tr>
<tr>
<td>GDDRPAAB</td>
<td>Abandon Sites DC1 and DC2</td>
<td>DC3 ONLY</td>
</tr>
<tr>
<td>GDDRPA05</td>
<td>Recover at DC3 after RDR at DC1 and DC2</td>
<td>DC3 ONLY</td>
</tr>
<tr>
<td>GDDRPA06</td>
<td>Restart Production LPARs at DC3 SRDFA to &lt;tgt site&gt;</td>
<td>DC3 ONLY</td>
</tr>
<tr>
<td>GDDRPA07</td>
<td>Recover at DC3 after LDR at DC1 SRDFA to DC2</td>
<td>DC3 ONLY</td>
</tr>
</tbody>
</table>

Where:

- **PLANNED** — Result from operator interaction with GDDR panels
- **UNPLANNED** — Result from unplanned events detected by the GDDR Event Monitor (EVM)
- **RESUME** — Resume replication after some type of service interruption (planned or unplanned)
- **DC3 Only** — These scripts only execute at DC3 for planned test events or for unplanned disaster event management.
Table 5 shows scripts mapped to the configuration type where the script executes.

Table 5  Script mapping (page 1 of 2)

<table>
<thead>
<tr>
<th>Script Name</th>
<th>Concurrent Star with AutoSwap</th>
<th>Concurrent Star</th>
<th>Concurrent Star 2 Site</th>
<th>Cascaded Star with EDP</th>
<th>Cascaded Star</th>
<th>SRDF/A with MSC</th>
<th>SRDFS with AutoSwap</th>
<th>SRDFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDDRPA51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA52</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA53</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA42</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDDRPA45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA21</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDD2P17A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GDD2P18A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GDD2P01A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GDDRPA27</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2P03A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRUP31</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2U09A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDD2U13A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2U10A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
### Table 5  Script mapping (page 2 of 2)

<table>
<thead>
<tr>
<th>Script Name</th>
<th>Concurrent Star with AutoSwap</th>
<th>Concurrent Star</th>
<th>Concurrent Star 2 Site</th>
<th>Concurrent Star with AutoSwap</th>
<th>Cascaded Star with AutoSwap</th>
<th>Cascaded Star with EDP</th>
<th>Cascaded Star with MSC</th>
<th>SRDF/S with AutoSwap</th>
<th>SRDF/S</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDD2U12A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRUP41</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA0A(1)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA0A(2)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPA0A(3)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2PA0A(4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDD2PA0A(5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDD2PA23</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDD2P02A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GDDRPA28</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPM29</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPF29</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2P14A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2P16A</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDDRPAAB</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2PA05</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2PA06</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDD2PA07</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Understanding GDDR

GDDR for SRDF/S with ConGroup (Two sites - DC1 and DC2)

This section contains the following information:

◆ “Planned script operations” on page 76
◆ “Unplanned script operations” on page 77
◆ “Resumption operations” on page 78

Planned script operations

Planned script operations include:

◆ GDD2P17A: Abandon Site DC1 (site swap) — Shuts down the single site workload at the primary site in preparation for the restart of processing at the secondary site:
  • Stops the business workload at the primary DASD site
  • Waits for the stop of all business applications
  • Resets clear all production systems managed by GDDR
◆ GDD2P18A: Restarts production at DC2 after site swap — Restart the single site workload after the GDD2P17A 'Abandon Site DC1 (site swap)' script has completed successfully:
  • Attempts reset clear of all systems at the primary DASD site
  • Activates CBU (if required)
  • Activates all needed LPARs, including CFs at the secondary DASD site
  • Creates a consistency point at the secondary DASD site
  • Prepare the SRDF environment
  • IPLs all needed production systems
◆ GDD2P01A: Performs test IPL from BCVs at DC2 — IPL contingency systems at site DC2 using BCV devices:
  • Splits BCVs; make them R/W
  • Activates test LPARs
  • and IPLs test z/OS systems using BCV volumes
  • Starts test business workload, if applicable
◆ GDD2P03A: Performs test IPL from R2s at DC2 — IPL contingency systems at site DC2 using R2 devices:
  • Confirms that SRDF/S has been stopped normally via a ConGroup trip
Understanding GDDR

- Activates LPARs and IPLs test z/OS systems using R2 volumes
- Starts test business workload, if applicable

Unplanned script operations

Operations personnel can manage unplanned events in one of two ways:

- The EMC GDDR Event Monitor prompts the operator for management confirmation of trigger events which indicate a site or DASD outage. The operator replies to the prompt in the affirmative and the GDDR recovery script is started.
- The operator may start the appropriate unplanned script and respond to prompts. The script initiates and validates that the state of the current host and storage environments matches the script prerequisites before proceeding.

Unplanned script operations include:

- GDD2U10A: Recovers after loss of DC1 (LDR) — Restarts the single site workload after the GDDR Event Monitor has detected a ConGroup trip and the GDDR Heartbeat Monitor has timed out. These events occurring concurrently are viewed by GDDR as a local disaster:
  - Confirms that a ConGroup trip occurred
  - Confirms that SRDF links failed
  - Confirms that a local disaster (LDR) event occurred
  - Shuts down applications at the primary site, if applicable
  - Splits BCVs and condition R2s at secondary site for restart
  - Activates contingency systems
  - Restarts applications

- GDD2PA0A: Resumes replication after loss of DC1 — Resume the SRDF/S link to the secondary site after a primary site disaster:
  - Confirms SRDF/S links are down
  - Splits BCVs at the secondary site, if applicable
  - Issues ConGroup cleanup and SRDF/S restart commands
  - Reestablishes BCVs at the secondary site
Understanding GDDR

Resumption operations

Resumption operations include:

- GDD2P02A: Resumes after test IPL from BCVs at DC2 — Reset clears contingency systems after a test at DC2:
  - Stops test business workload, if applicable
  - Reset clears test system LPARs
  - Reestablishes the BCVs
- GDD2P16A: Resumes after test IPL from R2s at DC2 — Resets clear contingency systems after a test at DC2:
  - Stops test business workload, if applicable
  - Resets clear test LPARs
  - Restart SRDF/S to DC2
- GDD2P14A: Resumes replication after link failure — Restores the SRDF/S link after a loss of the link:
  - Confirms SRDF/S links are down
  - Stops ConGroup on all systems
  - Splits BCVs at the secondary site, if applicable
  - Issues ConGroup cleanup and restart commands
  - Reestablishes BCVs at the secondary site

GDDR for SRDF/S with AutoSwap (Two sites - DC1 and DC2)

This section contains the following information:

- “Planned script operations” on page 78
- “Unplanned script operations” on page 79
- “Resumption operations” on page 80

Planned script operations

Planned script operations include:

- GDDRPA42: Swaps production from DC1 to DC2 — Swaps the DASD and the single-site workload from site DC1 to site DC2:
  - Stops the business workload at the primary DASD site
  - Swaps the DASD to the secondary DASD site (AutoSwap followed by SRDF/S personality swap)
- Resumes SRDF/S
- Restarts the business workload
- GDDRPA21: Swaps DASD from DC1 to DC2 — Swaps the DASD only from site DC1 to site DC2.
- GDD2P01A: Performs test IPL from BCVs at DC2 — IPLs contingency systems at site DC2 using BCV devices:
  - Splits BCVs; makes them R/W
  - Activates test LPARs using BCV volumes
  - Starts test business workload, if applicable
- GDD2P03A: Performs test IPL from R2s at DC2 — IPLs contingency systems at site DC2 using R2 devices:
  - Confirms that SRDF/S has been stopped normally via a ConGroup trip
  - Activates LPARs using R2 volumes
  - Starts test business workload, if applicable

**Unplanned script operations**

Operations personnel can manage unplanned events in one of two ways:
- The EMC GDDR Event Monitor prompts the operator for management confirmation of trigger events which indicate a site or DASD outage. The operator replies to the prompt in the affirmative and the GDDR recovery script is started.
- The operator may start the appropriate unplanned script and respond to prompts. The script initiates and validates that the state of the current host and storage environments matches the script prerequisites before proceeding.

Unplanned script operations include:
- GDD2U09A: Recovers after unplanned swap or LDR — Restarts the single site workload after the GDDR Event Monitor has detected a ConGroup trip and the GDDR Heartbeat Monitor has timed out. These events occurring concurrently are viewed by GDDR as a local disaster:
  - Confirms that a local disaster (LDR) event occurred
  - Conditions R2s at the secondary site for restart
  - Shuts down applications at the primary site, if applicable
Understanding GDDR

- Splits BCVs at the secondary site
- Activates contingency systems
- Manages Couple Facility structures and datasets
- Restarts applications

- **GDD2PA0A**: Resumes replication after unplanned swap — Resumes the SRDF/S link to the secondary site after a primary site disaster:
  - Confirms SRDF/S links are down
  - Splits BCVs at the secondary site, if applicable
  - Performs ConGroup cleanup, transfer AutoSwap ownership and restarts SRDF/S
  - Manages Couple Facility structures and datasets
  - Reestablishes BCVs at the secondary site

**Resumption operations**

Resumption operations include:

- **GDD2P02A**: Resumes after test IPL from BCVs at DC2 — Reset clears contingency systems after a test at DC2:
  - Stops test business workload, if applicable
  - Reset clears test systems LPARs
  - Reestablishes the BCVs

- **GDD2P16A**: Resumes after test IPL from R2s at DC2 — Reset clears contingency systems after a test at DC2:
  - Stops test business workload, if applicable
  - Reset clears test system LPARs
  - Restarts SRDF/S with AutoSwap to DC2

- **GDD2P14A**: Resumes replication after link failure — Restores the SRDF/S link after a loss of the link:
  - Confirms SRDF/S links are down
  - Stops ConGroup on all systems
  - Splits BCVs at the secondary site, if applicable
  - Issues ConGroup cleanup and restart commands
  - Reestablishes BCVs at the secondary site
**GDDR for SRDF/A (Two sites - DC1 and DC3)**

This section contains the following information:

- “Planned script operations” on page 81
- “Unplanned script operations” on page 81
- “Resumption operations” on page 82

**Planned script operations**

Planned script operations include:

- **GDD2P17A: Abandons Site DC1 (site swap)** — Shuts down the single site workload at the primary site in preparation for the restart of processing at the secondary site:
  - Stops the business workload at the primary DASD site
  - Waits for the stop of all business applications
  - Reset clears all production systems managed by EMC GDDR

- **GDD2P18A: Restarts production at DC3 after site swap** — Restarts the single site workload after the ’Abandon Site DC1 (site swap)’ script has completed successfully:
  - Attempts reset clear of all systems at the primary DASD site
  - Activates CBU (if required)
  - Activates all needed LPARs, including CFs at the secondary DASD site
  - Creates a consistency point at the secondary DASD site
  - Prepares the SRDF environment
  - IPLs all needed production systems

- **GDD2P01A: Performs test IPL from BCVs at DC3** — IPLs contingency systems at site DC3 using BCV devices:
  - Splits BCVs; make them R/W
  - Activates and load test LPARs using BCV volumes

**Unplanned script operations**

Operations personnel can manage unplanned events in one of two ways:

- The EMC GDDR Event Monitor prompts the operator for management confirmation of trigger events which indicate a site or DASD outage. The operator replies to the prompt in the affirmative and the GDDR recovery script is started.
The operator may start the appropriate unplanned script and respond to prompts. The script initiates and validates that the state of the current host and storage environments matches the script prerequisites before proceeding.

Unplanned script operations include:

- **GDD2U12A: Recovers after loss of DC1 (RDR)** — Restarts the single site workload after the GDDR Event Monitor has detected an SRDF/A link down event and the GDDR Heartbeat Monitor has timed out. These events occurring concurrently are viewed by GDDR as a regional disaster:
  - Confirms that an MSC drop occurred
  - Confirms that SRDF links failed
  - Confirms that a regional disaster (RDR) event occurred
  - Shuts down applications at the primary site, if applicable
  - Splits BCVs and condition R2s at secondary site for restart
  - Activates contingency systems
  - Restarts applications

- **GDD2PA0A: Resumes replication after loss of DC1** — Resumes the SRDF/A link to the secondary site after a primary site disaster:
  - Confirms SRDF/A links are down
  - Splits BCVs at the secondary site, if applicable
  - Issues MSC cleanup and SRDF/A restart commands
  - Reestablishes BCVs at the secondary site

**Resumption operations**

Resumption operations include:

- **GDD2P02A: Resumes after test IPL from BCVs at DC3** — Reset clears contingency systems after a test at DC2:
  - Stops test business workload, if applicable
  - Reset clears test LPARs
  - Reestablishes the BCVs

- **GDDRPM29: Resumes SRDF/A after link loss** — Restores the SRDF/A links after a planned or unplanned stop of SRDF/A.
GDDR for SRDF/Star (Three sites)

This section contains the following information:

◆ "Reconfiguration operations" on page 83
◆ "Planned script operations" on page 83
◆ "Unplanned script operations" on page 84
◆ "Resumption operations" on page 86
◆ "DC3-specific operations" on page 87

Reconfiguration operations

Reconfiguration operations include:

◆ GDDRPA51: Reconfigures to concurrent SRDF — Reconfigures a Cascaded SRDF/Star environment to a Concurrent SRDF/Star environment. The result is that site DC1 is protected at site DC2 using SRDF/S, and DC1 is the source of the SRDF/A replication to DC3. The workload continues at DC1:
  • Stops SRDF/A and deletes the SRDF/A relationship between DC2 and DC3
  • Performs a differential CREATEPAIR from DC1 to DC3
  • Reestablishes the SRDF/Star environment (DC1 to DC3)

◆ GDDRPA52: Reconfigures to cascaded SRDF — Reconfigures a Concurrent SRDF/Star environment to a Cascaded SRDF/Star environment. The result is that site DC1 is protected at site DC2 using SRDF/S, and DC2 is the source of the SRDF/A replication to DC3. The workload continues at DC1:
  • Stops SRDF/A and deletes the SRDF/A relationship between DC1 and DC3
  • Performs a differential CREATEPAIR from DC2 to DC3
  • Reestablishes the SRDF/Star environment (DC2 to DC3)

Planned script operations

Planned script operations include:

◆ GDD2P17A: Abandons Site DC1 (site swap) — Shuts down the single site workload at the primary site in preparation for the restart of processing at the secondary site:
  • Stops the business workload at the primary DASD site
  • Waits for the stop of all business applications
Understanding GDDR

- Reset clears all production systems managed by EMC GDDR
  - GDD2P18A: Restarts production at DC2 after site swap — Restarts the single site workload after the GDD2P17A ‘Abandon Site DC1 (site swap)’ script has completed successfully:
    - Attempts reset clear of all systems at the primary DASD site
    - Activates CBU (if required)
    - Activates all needed LPARs, including CFs at the secondary DASD site
    - Creates a consistency point at the secondary DASD site
    - Prepares the SRDF environment
    - IPLs all needed production systems
  - GDD2P01A: Performs test IPL from BCVs at DC3 — IPLs contingency systems at site DC3 using BCV devices:
    - Splits BCVs; makes them R/W
    - Activates test LPARs using BCV volumes
    - Starts test business workload, if applicable
  - GDD2P03A: Performs test IPL from R2s at DC2 — IPLs contingency systems at site DC2 using R2 devices:
    - Confirms that SRDF/S has been stopped normally via a ConGroup trip
    - Activates LPARs using R2 volumes
    - Starts test business workload, if applicable
  - GDDRPA27: Performs test IPL from R2s at DC3 — IPLs contingency systems at site DC3 using R2 devices:
    - Confirms that SRDF/A has been stopped normally via an SRDF/A PENDDROP
    - Activates LPARs using R2 volumes
    - Starts test business workload, if applicable

Unplanned script operations

Operations personnel can manage unplanned events in one of two ways:
- The EMC GDDR Event Monitor prompts the operator for management confirmation of trigger events which indicate a site or DASD outage. The operator replies to the prompt in the affirmative and the GDDR recovery script is started.
The operator may start the appropriate unplanned script and respond to prompts. The script initiates and validates that the state of the current host and storage environments matches the script prerequisites before proceeding.

Unplanned script operations include:

- **GDD2U13A**: Recovers after loss of DC1 (LDR) — Restarts the single site workload after the GDDR Event Monitor has detected a local disaster (LDR) at the primary site. The following events occurring concurrently are considered a local disaster: Loss of SRDF/A, Loss of SRDF/S, or a ConGroup trip, or the primary site GDDR Heartbeat Monitor has timed out:
  
  Note: In the cascaded SRDF with ConGroup topology, the SRDF/A event is not considered for LDR.

  - Confirms that a local disaster (LDR) event occurred
  - Deactivates systems at old primary site
  - Prepares the SRDF environment at secondary site
  - Creates a consistent point on BCVs at DC3
  - Performs a differential CREATEPAIR between the new primary site and DC3
  - Starts SRDF/A in MSC mode
  - Reestablishes the SRDF/Star environment
  - Reestablishes BCVs at DC3
  - Activates production systems

- **GDDRUP41**: Recovers after loss of DC2 — In environments which support Cascaded SRDF mode, GDDRUP41 restores data protection for the primary site should the secondary site suffer an outage resulting in loss of SRDF/S and SRDF/A. This script is submitted by the GDDR Event Monitor when a ConGroup trip is detected and the GDDR Heartbeat Monitor at the secondary site has timed out. These events occurring concurrently are viewed by GDDR as a secondary site disaster. Following the restoration of the secondary site, the script GDDRPA0A: ‘Resume replication after loss of DC2’ is used to resume SRDF/S:

  - Provides the option to stop primary site and secondary site workloads since all remote protection is lost
  - Splits BCVs at DC3
Understanding GDDR

- Removes the SRDF pairing between DC3 and the secondary DASD site
- Performs a differential re-synchronization between the primary site and DC3
- Starts SRDF/A in concurrent MSC mode
- Provides the option to start primary site workloads
- Reestablishes BCVs at DC3

- GDDRP0A0A: Resumes replication after loss of DC1 — Resumes the SRDF/S link to the secondary site after a primary site disaster. The script appears in the 'Select Script to Run' panel of the current Master C-System after completion of the GDD2UP13A `Recover after loss of DC1 (LDR)` script:
  - Confirms SRDF/S links are down
  - Stops ConGroup on all systems
  - Splits BCVs at the secondary site, if applicable
  - Issues ConGroup cleanup and restarts commands
  - Reestablishes BCVs at the secondary site

- GDDRP0A0A: Resumes replication after loss of DC2 — In environments which support Cascaded SRDF mode, resumes the SRDF/S link to the primary site after a secondary site disaster and restarts SRDF/A in SRDF/Star mode. The script appears in the 'Select Script to Run' panel of the current Master C-System after completion of the GDDRUP41 `Recover after loss of DC2` script:
  - Confirms SRDF/S links are down
  - Stops ConGroup on all systems
  - Splits BCVs at the primary site, if applicable
  - Issues ConGroup cleanup and restarts commands
  - Reestablishes BCVs at the primary site

Resumption operations

Resumption operations include:

- GDDRP23: Resume SRDF/S replication after ConGroup trip — Resumes SRDF/S replication and reestablishes the BCVs at the secondary DASD site if applicable, following a ConGroup trip.

- GDD2P02A: Resumes after test IPL from BCVs at DC3 — Reset clears contingency systems after a test at DC3
• Stops test business workload, if applicable
• Reset clears test system LPARs
• Reestablishes the BCVs

◆ GDD2P16A: Resumes after test IPL from R2s at DC2 — Reset clears contingency systems after a test at DC2:
  • Stops test business workload, if applicable
  • Resets clear test LPARs
  • Restarts SRDF/S to DC2

◆ GDDRPA28: Resumes SRDF/A after test IPL at DC3 — Restores the SRDF/A links to DC3 (either from DC1 or DC2 depending upon where the production workload is currently running) after a test on DC3:
  • Reset clears all systems IPL’d during the test of DC3 (at DC3)
  • Deactivates all LPARs previously activated for the test of DC3 (at DC3)
  • Restarts SRDF/Star to DC3

◆ GDDRPM29: Resumes SRDF/A in MSC mode to DC3 — Restores the SRDF/A links to DC3 in MSC mode (from either DC1 or DC2 depending upon where the production workload is currently running) after a planned or unplanned swap.

◆ GDDRPF29: Resumes SRDF/A (SRDF/Star) to DC3 — Restores the SRDF/A links to DC3 in SRDF/Star mode (from either DC1 or DC2) after a planned or unplanned stop of SRDF/A.

DC3-specific operations

DC3-specific operations include:

◆ GDDRPAAB: Abandons Sites DC1 and DC2 — In the event of a perceived threat to both DC1 and DC2, performs an orderly shutdown of DC1 and DC2 and force the C-System at DC3 as the new GDDR Master C-System:
  • Stops the business applications
  • Waits for the stop of all business applications
  • Reset clears all production systems managed by GDDR
  • Creates a consistency point on the DC3 BCVs
  • Creates a consistency point on the secondary DASD BCVs
  • Stops ConGroup on all remaining systems
Understanding GDDR

- Forces the C-System at DC3 as the new GDDR Master C-System
- GDDRPA05: Recovers at DC3 after RDR at DC1 and DC2 — In the event of a major failure that prevents the production workload from being run from either DC1 or DC2, restarts the production LPARs at site DC3 and reestablishes SRDF/A to site DC1 or DC2. There must be an SRDF/A connection to either DC1 or DC2:
  - Attempts reset clear of all systems at DC1/DC2
  - Activates all needed LPARs
  - Activates CBU (if required)
  - Creates a consistency point at DC3
  - Prepares SRDF environment
  - IPLs all needed production systems
- GDDRPA06: Restart production at DC3 SRDF/A to DC2 — In the event of a major failure that prevents the production workload from being run from either DC1 or DC2, restarts the production LPARs at site DC3 and reestablishes SRDF/A to site DC1:
  - Attempts reset clear of all systems at DC1/DC2 (except C-System)
  - Activates all needed LPARs
  - Activates CBU (if required)
  - Creates a consistency point at DC3
  - Prepares SRDF environment
  - IPLs all needed production systems
  - Performs a differential CREATEPAIR between DC3 and DC1 or DC2
  - Starts SRDF/A in MSC mode
- GDDRPA07: Recovers at DC3 after LDR at DC1 SRDFA to DC2 — In the event of a local disaster (LDR) that prevents the production workload from being run at DC1, and only if DC1 was the primary DASD site, restarts the production LPARs at site DC3 and reestablishes SRDF/A to DC2. The distinction between this script and GDDRPA06: 'Restart production at DC3 SRDF/A to DC2' script is, that the 'Star C-Ahead' indicator is interrogated to enable the workload to restart with the DC3 data, or with the DC2 data if it is more current:
Understanding GDDR

- Activates all needed LPARs including CFs at DC3
- Activates CBU (if required)
- Creates a consistency point at DC3
- Prepares SRDF environment
- IPLs all needed production systems
- Performs a differential CREATEPAIR between DC3 and DC2
- Starts SRDF/A in MSC mode

GDDR for SRDF/Star with AutoSwap (Three sites)

This section contains the following information:
- “Reconfiguration operations” on page 89
- “Planned script operations” on page 90
- “Unplanned script operations” on page 92
- “Resumption operations” on page 94
- “DC3-specific operations” on page 95

Reconfiguration operations

Reconfiguration operations include:
- GDDRPA51: Reconfigures to concurrent SRDF — Reconfigures a Cascaded SRDF/Star with AutoSwap environment to a Concurrent SRDF/Star with AutoSwap environment. The result is that site DC1 is protected at site DC2 using SRDF/S and AutoSwap, and DC1 is the source of the SRDF/A replication to DC3. The workload continues at DC1:
  - Stops SRDF/A and deletes the SRDF/A relationship between DC2 and DC3
  - Performs a differential CREATEPAIR from DC1 to DC3
  - Reestablishes the SRDF/Star with AutoSwap environment (DC1 to DC3)
- GDDRPA53: Reconfigures to concurrent SRDF with site move — Reconfigures a Cascaded SRDF/Star with AutoSwap environment to Concurrent SRDF/Star with AutoSwap with a primary processing site move. The initial state is that the workload at DC1 is protected at DC2 with SRDF/S. The result is
that DC2 is protected at DC1 using SRDF/S and AutoSwap, and
DC2 is the source of the SRDF/A replication to DC3. The
workload will be running at DC2:

- Stops the business workload at site DC1
- Swaps the DASD to DC2 (AutoSwap followed by SRDF/S
  personality swap)
- Reset clears all production systems managed by GDDR
- Resumes SRDF/S (DC2 to DC1)
- Restarts the business workload
- Reestablishes the SRDF/Star with AutoSwap environment
  (DC2 to DC3)

- GDDRP52: Reconfigures to cascaded SRDF — Reconfigures a
  Concurrent SRDF/Star with AutoSwap environment to a
  Cascaded SRDF/Star with AutoSwap environment. The result is
  that site DC1 is protected at site DC2 using SRDF/S and
  AutoSwap, and DC2 is the source of the SRDF/A replication to
  DC3. The workload continues at DC1:
  - Stops SRDF/A and delete the SRDF/A relationship between
    DC1 and DC3
  - Performs a differential CREATEPAIR from DC2 to DC3
  - Reestablishes the SRDF/Star with AutoSwap environment
    (DC2 to DC3)

Planned script operations

Planned script operations include:

- GDDRP42: Swaps production from DC1 to DC2 Protected—
  Swaps the DASD and the single-site workload from the primary
  DASD site to the secondary DASD site while maintaining
  consistent remote protection on one leg for the entire duration
  of the script. In environments with geographically dispersed
  Sysplexes supported by cross-site host-DASD channels, Couple
  Dataset and Coupling Facility Structure management is
  performed as a result of a site move, either planned or
  unplanned. These actions are controlled by the Realign Couple
  Datasets and Rebuild Couple Facility Structure parameters:
  - Stops the business workload at the primary DASD site
  - Converts from Star mode to MSC mode
• Maintains SRDF/A active in MSC mode while the synchronous leg is swapped
• Swaps the DASD to the secondary DASD site (AutoSwap followed by SRDF/S personality swap)
• Resumes SRDF/S
• Restarts the business workload
• Reestablishes the SRDF/Star with AutoSwap environment

◆ GDDRPA45: Swaps production from DC1 to DC2 Fast - Available only with Cascaded Star configurations. Swaps the DASD and the single-site workload from the primary DASD site to the secondary DASD site, completing in much less time than GDDRPA42, but having a short window where there is no remote protection at all.

◆ GDDRPA21: Swaps DASD from DC1 to DC2 Protected — Swaps only the DASD from the primary DASD site to the secondary DASD site while maintaining consistent remote protection on one leg for the entire duration of the script.

◆ GDDRPA25: Swaps DASD from DC1 to DC2 Fast - Available only with Cascaded Star configurations. Swaps only the DASD from the primary DASD site to the secondary DASD site, completing in much less time than GDDRPA21, but having a short window where there is no remote protection at all.

◆ GDD2P01A: Performs test IPL from BCVs at DC3 — IPLs contingency systems at site DC3 using BCV devices:
  • Splits BCVs; makes them R/W
  • Activates test LPARs and IPLs test z/OS systems using BCV volumes
  • Starts test business workload, if applicable

◆ GDD2P03A: Performs test IPL from R2s at DC2 — IPLs contingency systems at site DC2 using R2 devices:
  • Confirms that SRDF/S has been stopped normally via a ConGroup trip
  • Activates LPARs using R2 volumes
  • Starts test business workload, if applicable

◆ GDDRPA27: Performs test IPL from R2s at DC3 — IPLs contingency systems at site DC3 using R2 devices:
Understanding GDDR

- Confirms that SRDF/A has been stopped normally via an SRDF/A PENDDROP
- Activates LPARs using R2 volumes
- Starts test business workload, if applicable

Unplanned script operations

Operations personnel can manage unplanned events in one of two ways:

- The EMC GDDR Event Monitor prompts the operator for management confirmation of trigger events which indicate a site or DASD outage. The operator replies to the prompt in the affirmative and the GDDR recovery script is started.
- The operator may start the appropriate unplanned script and respond to prompts. The script initiates and validates that the state of the current host and storage environments matches the script prerequisites before proceeding.

Unplanned script operations include:

- GDDRUP31: Recovers after unplanned swap — Restarts processing at the secondary site after an unplanned swap has completed successfully.
  - Stops business workload on the old primary site (if applicable)
  - Triggers restart of business workload at the new primary site (if applicable)
  - Removes SRDF pairing between DC3 and the primary DASD site
  - Stops ConGroup on all systems
  - Performs a differential CREATEPAIR between the new primary DASD site and DC3
  - Starts SRDF/A in MSC mode

- GDD2U13A: Recovers after loss of DC1 (LDR) — Restarts the single site workload after the GDDR Event Monitor has detected a local disaster (LDR) at the primary site. The following events occurring concurrently are considered a local disaster: Loss of SRDF/A, Loss of SRDF/S, or a ConGroup trip, or the primary site GDDR Heartbeat Monitor has timed out:
  - Confirms that a local disaster (LDR) event occurred
  - Deactivates systems at old primary site
• Prepares the SRDF environment at secondary site
• Creates a consistent point on BCVs at DC3
• Performs a differential CREATEPAIR between the new primary site and DC3
• Starts SRDF/A in MSC mode
• Reestablishes the SRDF/Star environment
• Reestablishes BCVs at DC3
• Manages Couple Facility structures and datasets
• Activates production systems

◆ GDDRUP41: Recovers after loss of DC2 — In environments which support Cascaded SRDF mode, restores data protection for the primary site should the secondary site suffer an outage resulting in loss of SRDF/S and SRDF/A. This script is submitted by the GDDR Event Monitor when a ConGroup trip is detected and the GDDR Heartbeat Monitor at the secondary site has timed out. These events occurring concurrently are viewed by GDDR as a secondary site disaster. Following the restoration of the secondary site, the script GDDRPA0A ‘Resume replication after loss DC2’ is used to resume SRDF/S:

• Provides the option to stop primary site and secondary site workloads since all remote protection is lost
• Splits BCVs at DC3
• Removes the SRDF pairing between DC3 and the secondary DASD site
• Performs a differential re-synchronization between the primary site and DC3
• Starts SRDF/A in concurrent MSC mode
• Provides the option to start primary site workloads
• Reestablishes BCVs at DC3

◆ GDDRPA0A: Resumes replication after unplanned swap — Resumes the SRDF/S link to the secondary DASD site after an unplanned swap (due to the loss of the primary DASD). The script appears in the ‘Select Script to Run’ panel of the current Master C-System after completion of the GDDRUP31 ‘Recover after unplanned swap’ script:

• Removes the SRDF pairing between the secondary DASD site and the primary DASD site
Resumption operations

Resumption operations include:

GDDRPA23: Resume SRDF/S replication after ConGroup trip — This script resumes SRDF/S replication and reestablishes the BCVs at the secondary DASD site if applicable.

GDD2P02A: Resumes after test IPL from BCVs at DC3 — Reset clears contingency systems after a test at DC3:

- Stops test business workload, if applicable
- Reset clears test LPARs
• Reestablishes the BCVs
  ◆ GDDR2P16A: Resumes after test IPL from R2s at DC2 — Reset clears contingency systems after a test at DC2:
  • Stops test business workload, if applicable
  • Resets clear test LPARs
  • Restarts SRDF/S to DC2
  ◆ GDDRPA28: Resumes SRDF/A after test IPL at DC3 — Restores the SRDF/A links to DC3 (either from DC1 or DC2 depending upon where the production workload is currently running) after a test on DC3:
  • Resets clear all systems IPLed during the test of DC3 (at DC3)
  • Deactivates all LPARs previously activated for the test of DC3 (at DC3)
  • Restarts SRDF/Star with AutoSwap to DC3
  ◆ GDDRPM29: Resumes SRDF/A in MSC mode to DC3 — Restores the SRDF/A links to DC3 in MSC mode (from either DC1 or DC2 depending upon where the production workload is currently running) after a planned or unplanned swap.
  ◆ GDDRPF29: Resumes SRDF/A (SRDF/Star AutoSwap) to DC3 — Restores the SRDF/A links to DC3 in SRDF/Star with AutoSwap mode (from either DC1 or DC2) after a planned or unplanned stop of SRDF/A.

DC3-specific operations

DC3-specific operations include:
  ◆ GDDRPAAAB: Abandons Sites DC1 and DC2 — In the event of a perceived threat to both DC1 and DC2, performs an orderly shutdown of DC1 and DC2 and force the C-System at DC3 as the new GDDR Master C-System:
    • Stops the business applications
    • Waits for the stop of all business applications
    • Reset clears all production systems managed by GDDR
    • Creates a consistency point on the DC3 BCVs
    • Creates a consistency point on the secondary DASD BCVs
    • Stops SRDF/Star with AutoSwap
    • Stops ConGroup on all remaining systems

GDDR supported scripts
Understanding GDDR

- Forces the C-System at DC3 as the new GDDR Master C-System
- GDDRPA05: Recovers at DC3 after RDR at DC1 and DC2 — In the event of a failure that prevents the production workload from being run from either DC1 or DC2, restarts the production LPARs at site DC3 and reestablishes SRDF/A to site DC1 or DC2. There must be an SRDF/A connection to either DC1 or DC2:
  - Attempts reset clear of all systems at DC1/DC2
  - Activates all needed LPARs
  - Activates CBU (if required)
  - Creates a consistency point at DC3
  - Prepares SRDF environment
  - IPLs all needed production systems
- GDDRPA06: Restarts production LPARs at DC3 SRDFA to DC1 or DC2 — In the event of a failure that prevents the production workload from being run from either DC1 or DC2, restarts the production LPARs at site DC3 and reestablishes SRDF/A to site DC1 or DC2:
  - Attempts reset clear of all systems at DC1/DC2 (except C-System)
  - Activates all needed LPARs
  - Activates CBU (if required)
  - Creates a consistency point at DC3
  - Prepares SRDF environment
  - IPLs all needed production systems
  - Performs a differential CREATEPAIR between DC3 and DC1 or DC2
  - Starts SRDF/A in MSC mode
- GDDRPA07: Recovers at DC3 after LDR at DC1 SRDFA to DC2 — In the event of a local disaster (LDR) that prevents the production workload from being run at DC1, and only if DC1 was the primary DASD site, restarts the production LPARs at site DC3 and reestablishes SRDF/A to DC2.
  The considerations for restarting at DC3 are:
  - There was a CG Trip before DC1 was lost. In that case DC3 will have more recent data than DC2;
• There are signs that DC2 may be in danger as well as DC1;
• DC2 is a bunker site (storage only).

The distinction between this script and the GDDRPA06 'Restart production LPARs at DC3 SRDFA to DC2' script is that the 'Star C-Ahead' indicator is interrogated to enable the workload to restart with the DC3 data, or with the DC2 data if it is more current:

• Activates all needed LPARs including CFs at DC3
• Activates CBU (if required)
• Creates a consistency point at DC3
• Prepares SRDF environment
• IPLs all needed production systems
• Performs a differential CREATEPAIR between DC3 and DC2
• Starts SRDF/A in MSC mode

The remainder of the TechBook will describe the lab environment, document the installation and configuration of GDDR, and document running scripts with more detail on the steps executed in the scripts.
This chapter deals with important considerations to ensure that your major infrastructure project is successful. Some of the issues highlighted here are best practices specific to managing technology risks. Other recommendations relate to the common but important risk factors associated with project administration and change management:

- Understanding project constraints ................................................  100
- Importance of test systems ............................................................. 101
- Minimizing operational risks during implementations ............  102
- Importance of design ....................................................................... 104
- Constructing timeline and milestones ........................................  106
- Review change management procedures................................. 107
Understanding project constraints

EMC GDDR implementations operate under the *triple constraint* model of project management. Essentially, there are three major areas of constraint on any infrastructure project:

- Time
- Quality
- Cost

In many instances the date of implementation is already fixed. Ideally, the implementation date will have been calculated from the sum total of all of the complex technical infrastructure tasks that must be accomplished. These tasks may include:

- Workload measurement
- Bandwidth assessment
- Compute resiliency (to support HA or DR etc)
- Storage layout and I/O access density (cache, device types, RAID type, etc.)
- Data center planning (floor space, power, cooling, etc.)
- Data migration planning
- Construction of test facility
- Staff training

It is a major risk factor if the target date for the completion of the project has been set without due regard to the number and complexity of tasks to achieve the goal.

In theory, the three-constraint project model is a method for the project manager to retain control of the project delivery. If the customer owns all three constraints, it is impossible for the project manager to obtain or retain control of the project. Control can be maintained by allowing the project manager to control at least one of the constraints, typically the Quality objective. In environments where time and budget are fixed, the only remaining variable is quality. Inevitably, if the time constraint is unreasonable, project quality will be affected.
Importance of test systems

As the complexity of solutions increases, so too does the requirement to allow your operations staff to become competent with new technology. This is best achieved with a test environment.

Even if the customer cannot afford a perpetual test environment, serious consideration should be given to constructing an interim test environment for the duration of the GDDR project. It will allow the operations staff to safely acquire knowledge. It will also provide a framework for formalized hands-on training offered by the EMC GDDR implementation team both during the project, and more formally, as part of the acceptance criteria towards the end of the project.

Some customers can benefit from hands-on training more than others:

- If the customer’s operations staff does not have experience with replication and automation technologies in general, and EMC solutions in particular, then the construction of a test environment should be mandatory.

- If the customer has the necessary skills in-house, then the benefit from a test environment is more limited to ensuring that software and fixes provided by EMC and other vendors operate correctly in the customer environment.
Minimizing operational risks during implementations

There are several methods for implementing GDDR into production environments. The two approaches outlined in this section are used where a customer cannot tolerate downtime in their existing production workflow:

- “Parallel infrastructure” on page 102
- “Partial production” on page 103

The techniques used now allow for both continuous operation of production systems and also maintenance of the existing DR posture.

Parallel infrastructure

The parallel infrastructure approach is a technique where new network, compute, and storage infrastructure is commissioned adjacent to the existing production systems. Data is migrated from the existing production infrastructure into the new GDDR-managed environment to allow for infrastructure and applications orientated testing to be conducted.

The testing process is destructive to the data image, so the data migration process is repeated for each test cycle. Once the infrastructure has been proven to operate successfully in the new environment, one final data migration is conducted.

The GDDR-managed infrastructure becomes the new production environment and the older infrastructure is decommissioned.

Another significant advantage of this technique over other techniques is that not only can the resiliency of the compute infrastructure be preserved, but the existing HA/DR posture can be retained until the new infrastructure can be deployed.

The parallel infrastructure approach is an EMC best practice for deploying GDDR management infrastructures into customer environments with both minimal risk and minimal disruption to existing application workloads.
Partial production

The *partial production* approach is where a small subset of DASD volumes on the production system is used to prove the entire infrastructure. This allows for the validation of GDDR and prerequisite technologies into the production environment without actually using DASD that contains production data.

When the time comes to increase the scope of the parameters to address the full environment, the first time it is executed is live in the production environment.

This partial production approach is inferior to the parallel infrastructure approach in that the full function and scope of functionality cannot be fully tested without risking impact to production environments. The emphasis here is to eliminate (in particular) typing errors in parameter libraries, security issues, issues with gate keeper devices, firewalls, etc., all which can be debugged without impact using the parallel infrastructure approach, but which cannot be fully detected in the partial production approach until actually encountered live in the production environment.

The importance of a test environment to allow an operations staff to obtain competency cannot be overstressed when the partial production approach is used. The partial production approach is a good compromise method of implementing where the customer is able to provide sufficient test windows to safely test the new infrastructure *in place*. It is also commonly used where financial or other constraints on a project make the *parallel infrastructure* implementation approach impractical.
Planning for your GDDR Implementation

Importance of design

This section provides information on the important of design, including:

◆ “Sizing” on page 104
◆ “Storage layout” on page 104
◆ “GDDR Solutions Support Team” on page 104

Sizing

A significant area of addressable risk with GDDR projects is in the correct sizing of the entire solution. This book provides guidelines for properly sizing the proposed solution. Particular attention must be paid to the sizing of network infrastructure. See Chapter 7, “Network Design Best Practices,” for more details.

Storage layout

A second major area of design risk is in the storage layout. Storage must be configured not only for raw storage capacity, RAID groups, etc., but also for throughput capacity. The solution must be designed to sustain not only I/O loads from the production environment, but also to have sufficient reserve capacity in terms of cache and throughput to operate the replication infrastructure and (if required) restart operations. See Chapter 6, “Storage Infrastructure Design Considerations,” for more details on this important topic.

GDDR Solutions Support Team

EMC recommends early contact (pre-sales) with GDDR solution design experts to ensure that any solution proposal meets design standards to ensure optimal operational performance. The GDDR Solutions Support Team operates globally and is centrally managed from EMC Corporate Headquarters in Hopkinton, Massachusetts.

The team can be contacted via email at the following addresses:
gddrsolutionssupport@emc.com
Conventional mail:
GDDR Solutions Support Team
Mailstop 1/P-36
176 South Street
Hopkinton, MA 01748 USA
Constructing timeline and milestones

Setting a too aggressive timeline can cause a project to fail. As discussed in “Understanding project constraints” on page 100, if the timeline is fixed and the budget is fixed that leaves only one variable: quality. In other words, to meet strict deadlines, shortcuts may be taken in quality assurance measures. The result may be a sequence of failed changes or service interruptions to the production environment.

One of the ways to avoid placing undue or unreasonable time pressure on a project is to make sure to involve the technical architecture and implementation people in the construction of the implementation plan time outline.
Review change management procedures

The GDDR Project is typically much larger than just installing the GDDR software and customizing parameter libraries. As discussed earlier, there may be changes to physical infrastructure, utilities, air-conditioning, as well as IT infrastructure such as network provisioning, new compute and storage arrays, switches, routers, etc. In the space of a few months a large number of potentially high-risk changes need to be executed.

Change scheduling is important to avoid conflicts between infrastructure provisioning groups and to align with the business calendar.

However, much more important than change scheduling is change management. This is a much broader topic and addresses issues such as:

- Change scheduling
- Backout/fallback planning
- Risk mitigation strategies
- Documentation
- Root cause analysis (for failed changes)

When trying to determine the health of a change management system, look beyond the number of change tickets to see how many changes are rejected because of scheduling conflicts, inadequate documentation, risk mitigation, etc. A robust change management system will exhibit much more than just change scheduling.

Assess the health of your change management systems and, if necessary, adjust to deal with the large volume of changes typically encountered during these infrastructure projects.
EMC provides many hardware and software products that support applications in the z/OS environment. The following products, some required and some configuration-specific, are used in EMC GDDR solutions:

- EMC Symmetrix VMAX with EMC Enginuity ................. 110
- EMC ResourcePak Base for z/OS ........................................ 115
- EMC TimeFinder ................................................................ 118
- EMC Symmetrix Remote Data Facility (SRDF) .............. 122
- EMC SRDF consistency groups ............................................ 132
- EMC AutoSwap .................................................................... 136
- EMC SRDF/Star .................................................................. 138
EMC Symmetrix VMAX with EMC Enginuity

EMC Symmetrix hardware architecture and the EMC Enginuity™ operating environment are the foundation for the Symmetrix storage platform. This environment consists of the following components:

- Symmetrix hardware
- Enginuity-based operating functions
- Mainframe Enablers
- Symmetrix application program interface (API) for mainframe
- Symmetrix-based applications
- Host-based Symmetrix applications
- Independent software vendor (ISV) applications

This section provides information on the following:

- “Symmetrix VMAX hardware platform” on page 110
- “EMC Enginuity operating environment” on page 112
- “Symmetrix features for mainframe” on page 112
- “EMC Mainframe Enablers” on page 113

Symmetrix VMAX hardware platform

The Symmetrix VMAX™ design is based on a highly-available VMAX Engine with redundant CPU, memory, and connectivity on two directors for fault tolerance. Symmetrix VMAX Engines connect to and scale out linearly through the Virtual Matrix Architecture®, which allows resources to be shared across Symmetrix VMAX engines. To meet growth requirements, additional engines can be added non-disruptively for efficient and dynamic scaling of capacity and performance.
Figure 24 shows the Symmetrix VMAX hardware architecture.
EMC Enginuity operating environment

The Enginuity operating system provides controls for all components in a Symmetrix array. Enginuity coordinates real-time events related to the processing of production data providing the following services:

- Manages system resources to intelligently optimize performance across a wide range of I/O workload requirements.
- Ensures system availability through advanced fault monitoring, detection and correction capabilities and provides concurrent maintenance and serviceability features.
- Provides the foundation for specific software features available through EMC disaster recovery, business continuance, and storage management software.
- Supports functional services for both Symmetrix-based functionality and for a large suite of EMC storage application software.
- Defines priority of each task, including basic system maintenance, I/O processing, and application processing.
- Provides uniform access through APIs for internal calls and provides an external interface to allow integration with other software providers.

Symmetrix features for mainframe

Symmetrix storage systems appear to mainframe operating systems as any of the following control units: IBM 3990, IBM 2105, and IBM 2107. The physical storage devices can appear to the mainframe operating system as any mixture of different sized 3380 and 3390 devices. Mainframe host connectivity is supported through serial ESCON and FICON channels.

Fibre Channel and GigE are supported options in SRDF® environments. Symmetrix GigE directors in an SRDF environment provide direct TCP/IP connectivity end-to-end for remote replication solutions over extended distances. This negates costly FC to IP converters and helps utilize the existing IP infrastructure without major disruptions.
EMC supported IBM compatibility features include:

- Channel Command Emulation for IBM ESS 2105/2107
- Multiple Allegiance (MA)
- Parallel Access Volume (PAV) Static and Dynamic
- HyperPAV
- Concurrent Copy
- Dynamic Channel Path Management (DCM)
- Dynamic Path Reconnection (DPR) support
- Host Data Compression
- Logical Path and Control Unit Address Support (CUADD)
- Mainframe systems hypervolumes
- Partitioned Dataset (PDS) Search Assist
- FlashCopy
- GDPS
- GDPS HyperSwap
- Peer to Peer Remote Copy (PPRC)
- Extended Remote Copy (XRC)
- Extended Address Volume
- Extended Distance FICON
- High Performance FICON (zHPF)
- DFSMS support for Flash Drives
- 8 Gb/s FICON

**EMC Mainframe Enablers**

EMC Mainframe Enablers (MFE) is a software suite that includes the Symmetrix API runtime libraries for all EMC mainframe software. These software packages can be used to monitor device configuration and status and to perform control operations on devices and data objects within a storage complex.
The following components are included in Mainframe Enablers (MFE) 7.0 and above:

- ResourcePak® Base for z/OS
- SRDF Host Component for z/OS
- TimeFinder®/Clone Mainframe SNAP Facility
- TimeFinder/Mirror for z/OS
- Consistency Groups for z/OS
- TimeFinder Utility for z/OS
- AutoSwap
- SRDF/Star

These are further discussed in the following sections:

- “EMC ResourcePak Base for z/OS” on page 115
- “EMC TimeFinder” on page 118
- “EMC Symmetrix Remote Data Facility (SRDF)” on page 122
- “EMC SRDF consistency groups” on page 132
- “EMC AutoSwap” on page 136
- “EMC SRDF/Star” on page 138
EMC ResourcePak Base for z/OS

EMC ResourcePak Base delivers EMC Symmetrix Control Facility (EMCSCF) for IBM and IBM-compatible mainframes. EMCSCF provides a uniform interface for EMC and ISV software products. EMCSCF delivers a persistent address space on the host that facilitates communication between the host and the Symmetrix as well as other EMC-delivered and partner-delivered applications. In GDDR managed configurations this host is normally the GDDR Master C-System.

This section includes the following information:
- “Tasks” on page 115
- “Functionality” on page 116

Tasks

ResourcePak Base performs tasks such as the following:
- Maintains an active repository of information about EMC Symmetrix devices attached to z/OS environments and making that information available to other EMC products.
- Performs automation functions.
- Handles inter-LPAR (logical partition) communication through the Symmetrix storage system.

Figure 25  z/OS SymmAPI architecture
ResourcePak Base provides faster delivery of new Symmetrix functions by EMC and ISV partners, along with easier upgrades. It also provides the ability to gather data when using tools such as TimeFinder/Mirror query because device status information is now cached along with other important information.

ResourcePak Base for z/OS is a prerequisite for EMC mainframe applications like the TimeFinder/Clone Mainframe SNAP Facility or SRDF Host Component for z/OS. As of release 7.0, these products are packaged and installed together as Mainframe Enablers, along with TimeFinder/Mirror for z/OS, Consistency Groups for z/OS, TimeFinder Utility for z/OS, and AutoSwap.

ResourcePak Base provides the following functionality through EMCSCF:

**Cross-system communication**
Inter-LPAR communication is handled by the EMCSCF cross-system communication (CSC) component. CSC uses a Symmetrix storage system to facilitate communication between LPARs. Several EMC Symmetrix mainframe applications use CSC to handle inter-LPAR communications.

**Non-disruptive SymmAPI-MF refreshes**
As of version 5.3, EMCSCF allows the SymmAPI-MF to be refreshed non-disruptively. Refreshing SymmAPI-MF does not impact currently executing applications that use SymmAPI-MF; for example, SRDF Host Component for z/OS or TimeFinder/Clone Mainframe SNAP Facility.

**SRDF/A Monitor**
The SRDF/A Monitor in ResourcePak Base is designed to:
- Find EMC Symmetrix controllers that are running SRDF/A.
- Collect and write SMF data about those controllers.

After ResourcePak Base is installed, the SRDF/A Monitor is started as a subtask of EMCSCF.

**Group Name Service support**
ResourcePak Base includes support for Symmetrix Group Name Service (GNS). Using GNS, you can define a device group once and then use that single definition across multiple EMC products on multiple platforms. This means that you can use a device group defined through GNS with both mainframe and open systems-based EMC applications. GNS also allows you to define group names for volumes that can then be operated upon by various other commands.
Pool management

With ResourcePak Base V5.7 or higher, generalized device pool management is a provided service. Pool devices are a predefined set of devices that provide a pool of physical space. Pool devices are not host-accessible. The CONFIGPOOL commands allow management of SNAPPOLLS or DSEPOOLS with CONFIGPOOL batch statements.

SRDF/A Multi-Session Consistency

SRDF/A Multi-Session Consistency (MSC) is a task in EMCSCF that ensures remote R2 consistency across multiple Symmetrix storage systems running SRDF/A. MSC provides the following:

- Coordination of SRDF/A cycle switches across systems.
- Up to 24 SRDF groups in a multi-session group.
- One SRDF/A session and one SRDF group per Symmetrix storage system when using Enginuity release level 5X70.
- With Enginuity release level 5X71 and later, SRDF groups are dynamic and are not limited to one per Symmetrix storage system. Group commands of ENABLE, DISPLAY, DISABLE, REFRESH, and RESTART are now available.

SWAP services

ResourcePak Base deploys a SWAP service in EMCSCF. It is used by EMC AutoSwap for planned outages with the ConGroup Continuous Availability Extensions (CAX).

Recovery services

Recovery service commands allow you to perform recovery on local or remote devices (if the links are available for the remote devices).

Licensed Feature Code management

EMCSCF manages Licensed Feature Codes (LFCs) to enable separately chargeable features in EMC software. These features require an LFC to be provided during the installation and customization of EMCSCF. LFCs are available for:

- Symmetrix Priority Control
- Dynamic Cache Partitioning
- AutoSwap (ConGroup with AutoSwap Extensions); separate LFCs are required for planned and unplanned swaps
- Geographically Dispersed Disaster Restart
- EMC z/OS Storage Manager
- SRDF/Asynchronous (MSC)
- SRDF/Automated Replication
- SRDF/Star
- TimeFinder/Clone (TARGET)
- TimeFinder/Consistency Group (CONSISTENT)
- TimeFinder/Snap (VDEV)
EMC TimeFinder is a family of products that enables both volume-based replication and data set level replication within a single Symmetrix system. Data is copied from Symmetrix devices using array-based resources without using host CPU or I/O. The source Symmetrix devices remain online for regular I/O operations while the copies are created. GDDR automation extends to managing TimeFinder within an array or across multiple arrays; for example, to provide gold copy enterprise data images for DR testing.

The following information is included in this section:

- “TimeFinder and GDDR” on page 118
- “TimeFinder/Clone” on page 119
- “TimeFinder/Consistency Group” on page 119
- “TimeFinder/Mirror for z/OS” on page 119

TimeFinder and GDDR

TimeFinder enables customers to create multiple point-in-time copies of data, allowing simultaneous execution of business tasks that were previously sequential. GDDR exploits the following TimeFinder modes of operation:

- TimeFinder/Mirror enables users to configure special devices called business continuance volumes (BCVs) to create a mirror image of Symmetrix standard devices. Using BCVs, TimeFinder creates a point-in-time copy of data that can be repurposed. The TimeFinder/Mirror component extends the basic API command set of Mainframe Enablers to include commands that specifically manage Symmetrix BCVs and standard devices.

- TimeFinder/Clone enables users to make copies of data from source volumes to target volumes without consuming mirror positions within the Symmetrix. The data is available to a target's host immediately upon activation, even if the copy process has not completed. Data may be copied from a single source device to multiple target devices. A source device can be either a Symmetrix standard device or a BCV device.
**TimeFinder/Clone**

TimeFinder/Clone for z/OS produces point-in-time copies of full volumes or individual datasets. TimeFinder/Clone operations involve full volumes or datasets where the amount of data at the source is the same as the amount of data at the target.

TimeFinder/Clone provides significant configuration flexibility because clone copies do not require Symmetrix mirror positions. TimeFinder/Clone source and target devices can have any form of RAID protection. The clone copies can also be configured as a standard device or as a Business Continuance Volume (BCV).

A Full-Volume Snap captures a complete replica of the source volume on the target volume in the local Symmetrix system. TimeFinder/Clone requests that span control units can invoke an external datamover to accomplish the request.

**TimeFinder/Consistency Group**

TimeFinder/Consistency Group, using the Enginuity Consistency Assist (ECA) feature, provides consistent snap operations on multiple volumes so that the targets are dependent-write consistent. TimeFinder/Consistency Group is available for full device, virtual device, and remote full device snaps. The source and target device pairs must reside in the same Symmetrix system. Consistency can be preserved over multiple volumes.

**TimeFinder/Mirror for z/OS**

EMC TimeFinder/Mirror is a business continuance solution. GDDR uses TimeFinder/Mirror to make full-volume copies of production data from a standard Symmetrix device (which is online for regular I/O operations from the host) to a Business Continuance Volume (BCV) with which the standard device is paired. The BCV was a specially tagged volume established when the Symmetrix unit was configured. The BCV functioned as a mirror controlled with the TimeFinder/Mirror ESTABLISH, SPLIT, RE-ESTABLISH, and RESTORE commands. The BCV device can then be separated (split) from the standard device and used for backup, restore, decision support, or applications testing.
In GDDR controlled environments GDDR will re-establish split BCVs with their associated devices at the conclusion of test windows or otherwise as required during disaster recovery operations.

Under Enginuity 5773 and earlier, TimeFinder/Mirror supports two local-replication technologies as shown in Figure 26.
Starting with Enginuity 5874, TimeFinder/Mirror uses clone emulation for all operations, as shown in Figure 27.

There are no specific steps that need to be taken to exploit Clone Emulation. Whenever TimeFinder/Mirror internally detects a Symmetrix controller running at Enginuity level 5874 and later, TimeFinder/Mirror automatically sets the mode to clone emulation. All TimeFinder/Mirror commands will be converted to clone emulation. GDDR uses TimeFinder/Mirror command syntax. For Enginuity 5874 and above code releases Clone Emulation will be used-this is transparent to GDDR. So GDDR generated scripts continue to operate as they did on TimeFinder/Mirror environments.

**Figure 27** TimeFinder/Mirror under Enginuity 5874 and later
EMC Symmetrix Remote Data Facility (SRDF)

The SRDF family of products provides synchronous and asynchronous remote replication capabilities for Symmetrix storage systems. At the conceptual level, SRDF is mirroring (RAID level 1) one logical disk device (the primary source/R1 within the primary Symmetrix storage system) to a second logical device (the secondary target/R2 within a physically separate secondary Symmetrix storage system) over Fibre Channel or GigE high-speed communication links. SRDF’s basic premise is that a remote mirror of data in a different Symmetrix storage system can serve as a valuable resource for:

- Protecting data using geographical separation.
- Giving applications a second location from which to retrieve data should the primary location become unavailable for any reason.
- Providing a means to establish a set of volumes on which to conduct parallel operations, such as testing or modeling.

GDDR supports SRDF Synchronous, SRDF/ Asynchronous and SRDF/Star solutions in various forms of two and three-site solutions.

The following sections describe the SRDF features exploited by GDDR:

- “SRDF/Synchronous” on page 122
- “SRDF/Asynchronous” on page 123
- “SRDF/A Multi-Session Consistency (MSC)” on page 126
- “Concurrent SRDF” on page 128
- “Cascaded SRDF” on page 129
- “SRDF/Extended Distance Protection” on page 130

SRDF/Synchronous

SRDF synchronous (SRDF/S) mode facilitates disaster recovery within the customer’s campus or metropolitan area network through real-time synchronous remote mirroring from one Symmetrix system to one or more Symmetrix systems.
SRDF/Asynchronous

SRDF asynchronous (SRDF/A) mode provides a long distance disaster restart data image with minimal impact on performance. Symmetrix systems implement asynchronous mode host writes from the primary Symmetrix to the secondary Symmetrix system using dependent-write consistent delta sets transferred in cycles:

- Each delta set contains groups of write I/Os for processing, which are managed for dependent-write consistency by the Enginuity operating environment.
- SRDF/A transfers these sets of data using cycles of operation, one cycle at a time, between the primary Symmetrix system and the secondary Symmetrix system.

Write Folding improves the efficiency of the SRDF network links. If there are multiple data updates in the same cycle, the systems send the most current data across the SRDF links. This is called Write Folding and is major advantage over competitive asynchronous replication solutions as it decreases network bandwidth consumption and the number of I/Os the SRDF director processes, thereby reducing the system overhead per host I/O.

Figure 28 illustrates SRDF/A delta sets and its relationships to cycle processing.

![Figure 28 SRDF/A delta sets](image)
In single session mode, the Symmetrix SRDF director ensures dependent-write consistency within SRDF/A by obtaining the active cycle number from a single location in global memory and assigning it to each I/O. The director retains that cycle number even if a cycle switch occurs during the life of that I/O. This results in an atomic cycle switch process for dependent-write sequences, even though it is not physically an atomic event across a range of volumes. As a result, two I/Os with a dependent relationship between them can be in the same cycle, or the dependent I/O can be in a subsequent cycle.

Delta set switching is at the core of the SRDF/A active session operation. The following is required before a primary Symmetrix system cycle switch can occur:

- The transmit delta set must have completed data transfer to the secondary Symmetrix system.
- The minimum cycle time (single session mode) or minimum cycle target time (MSC mode) must be reached.
- The previous Apply delta set must have been completed.

During the delta-set switching process:

1. Write I/Os are collected in the capture delta set on the primary Symmetrix system. The previous cycle’s transmit delta set is completing the SRDF transfer to the secondary Symmetrix system receive delta set, which is the N-1 copy. The secondary Symmetrix system apply delta set (N-2) is written to global memory, so that data is marked write pending to the secondary devices.

2. The primary Symmetrix system waits for the minimum cycle time to elapse and the transmit delta set to empty, meaning that all data has been transferred to the secondary Symmetrix system.

3. Once these conditions are satisfied, the primary Symmetrix system sends a commit message to the secondary Symmetrix system to begin the secondary Symmetrix system cycle switch.

4. On the primary system, the new capture delta set is available to collect new host I/Os. Before the secondary Symmetrix system cycle switch can occur, the following must be true:
   - The secondary Symmetrix system receives the commit message from the primary Symmetrix system.
   - Apply delta set (N-2 copy) completes its write process marking the data write pending to the secondary devices.
5. Once the secondary Symmetrix system receives the commit message from the primary Symmetrix system, the secondary Symmetrix system verifies the apply delta set has been written. This occurs while the primary Symmetrix system is performing the cycle switch between the capture and transmit delta sets.

6. The next step is a delta set cycle switch on the secondary Symmetrix system between the receive (inactive) and apply (active) delta sets. This preserves the dependent-write consistent copy at the secondary Symmetrix system prior to receiving the next dependent-write consistent delta set.

7. The secondary Symmetrix system sends an acknowledgement to the primary Symmetrix system. The data in the apply delta set is written to disk.

Figure 29 shows the delta set switching process.

---

**Figure 29** Delta set switching
SRDF/A Multi-Session Consistency (MSC)

SRDF/A Multi-Session Consistency (MSC) supports SRDF/A operations in configurations where there are multiple primary Symmetrix systems or multiple primary Symmetrix system SRDF groups connected to multiple secondary Symmetrix systems or multiple secondary Symmetrix systems SRDF groups. SRDF/A MSC configurations can also support mixed open systems and mainframe data controlled within the same SRDF/A MSC session.

GDDR managed solutions using SRDF asynchronous replication require MSC because the interface to control asynchronous replication is provided via MSC.

Achieving data consistency across multiple SRDF/A groups requires the cycle switch process be coordinated among the participating Symmetrix system SRDF groups or systems, and that the switch occur during a very brief time period when no host writes are being serviced by any participating Symmetrix system. SRDF control software running on the host provides a single coordination point to drive the cycle switch process in all participating Symmetrix systems.

I/Os are processed exactly the same way in SRDF/A MSC mode as they are in single session mode:

1. The active cycle on the primary Symmetrix system contains the current host writes or N data version in the capture delta set.

2. The inactive cycle contains the N-1 data version that is transferred using SRDF/A from the primary Symmetrix system to the secondary Symmetrix system. The primary inactive delta set is the transmit delta set and the secondary Symmetrix system’s inactive delta set is the receive delta set.

3. The active cycle on the secondary Symmetrix system contains the N-2 data version of the apply delta set. This is the guaranteed dependent-write consistent image in the event of a disaster or failure.
Figure 30 illustrates the delta sets and their relationships to SRDF/A cycles.

For the host to control the cycle switch process, the Symmetrix systems must be aware that they are running in multi-session consistency mode. This is done using the SRDF control software running on the host. The host software:

1. Coordinates the cycle switching for all SRDF/A sessions comprising the SRDF/A MSC configuration.
2. Monitors for any failure to propagate data to the secondary Symmetrix system devices and drops all SRDF/A sessions together to maintain dependent-write consistency.
3. Performs MSC cleanup if able.

As part of the process to enter MSC mode, and with each cycle switch issued thereafter, Enginuity assigns a cycle tag to each new capture cycle. That cycle tag is retained throughout that cycle's life. This cycle tag is a value that is common across all participating SRDF/A
sessions and eliminates the need to synchronize the cycle numbers across them. The cycle tag is the mechanism by which dependent-write consistency is assured across multiple MSC controlled participating Symmetrix system SRDF groups or systems.

SRDF/A MSC mode performs a coordinated cycle switch during a very short window of time referred to as an SRDF/A window, which is actually a flag (open/closed) indicating when there are no host writes being completed. The SRDF/A window flag is an attribute of the SRDF/A group and is checked at the start of each I/O, imposing no additional overhead because the front-end director is already obtaining the cycle number from global memory as part of the existing SRDF/A cycle switch operations.

When the host software discovers that all the SRDF groups and Symmetrix systems are ready for a cycle switch, MSC opens the SRDF/A window and issues a single command to each SRDF group to perform a cycle switch. In multi-session mode, the front-end director obtains the cycle number at the start of each write and also checks the SRDF/A window flag, if is the flag is on (an open window) the front-end director disconnects upon receiving host write I/O and begins polling to determine when the SRDF MSC host control software has closed the window. While the window is open, any write I/Os that start are disconnected and as a result no dependent-write I/Os are issued by any host to any devices in the SRDF/A MSC session.

The SRDF/A window remains open on each SRDF group and Symmetrix system until the last SRDF group and Symmetrix system in the multi-session group acknowledges to the SRDF MSC host control software that the open and switch command has been processed indicating a successful cycle switch. At this point the SRDF MSC host control software issues a close command for each SRDF/A group under MSC control. As a result, dependent-write consistency across the SRDF/A MSC session is ensured.

**Concurrent SRDF**

Enginuity version 5567 and later support the ability for a single primary device to be remotely mirrored to two secondary devices concurrently. This feature is called concurrent SRDF and is supported on Fibre Channel, Gigabit Ethernet (GigE) topologies. In concurrent SRDF, the primary (R11) devices must be assigned to two different
SRDF groups. This is because each remote mirror can be assigned to only one SRDF group, even if the two secondary devices reside in the same Symmetrix system.

The secondary devices in a concurrent SRDF configuration can be operating in synchronous, adaptive copy, or asynchronous mode. The only unsupported combination is for both secondary devices to be operating in asynchronous mode. Concurrent SRDF/S with SRDF/A provides the ability to remotely mirror a group of devices in synchronous mode to one secondary site and in asynchronous mode to a different secondary site which may be an extended distance site.

Normal I/O operational rules for SRDF also apply to Concurrent SRDF configurations. When operating in synchronous mode, ending status for an I/O is not presented to the host until the remote Symmetrix system acknowledges receipt of the I/O to the primary Symmetrix system. If both secondary devices are operating in synchronous mode, ending status is not presented to the host until both devices acknowledge receipt of the I/O. If one remote mirror is in synchronous mode and one remote mirror is in adaptive copy or asynchronous mode, ending status is presented to the host when the synchronous device acknowledges receipt of the I/O. GDDR automation provides scripts that allow customers to convert from concurrent to cascaded mode provided that the initial configuration is installed as cascaded mode.

Cascaded SRDF

Cascaded SRDF is a three-site disaster recovery solution where data from a primary site is synchronously replicated to a secondary site, and then asynchronously replicated from the secondary site to a tertiary site, as shown in Figure 31 on page 130. The core benefit behind a cascaded configuration is its inherent capability to continue replicating, with minimal user intervention, from the secondary site to the tertiary site in the event that the primary site fails. This enables a faster recovery at the tertiary site, provided that the tertiary site is where the customer wishes to restart production operations.
Cascaded SRDF uses dual-role SRDF devices (R21 devices) on the secondary site which acts as both an R2 to the primary site and an R1 to the tertiary site as shown in Figure 31.

The following SRDF modes are allowed in Cascaded SRDF:
- R21 -> R2: SRDF/A, Adaptive copy disk mode

The most common implementation is for the first hop to be in SRDF/S mode, and the second hop to be in SRDF/A mode.

The following limitations apply to Cascaded SRDF:
- Only one hop (R1 -> R21, or (R21 -> R2) can be asynchronous at a time
- Like concurrent SRDF, the two SRDF mirrors of an R21 cannot be configured using the same SRDF group. GDDR automation provides scripts that allow customers to convert from cascaded to concurrent mode provided that the initial configuration is installed as cascaded mode.

SRDF/Extended Distance Protection

SRDF/Extended Distance Protection (EDP) is a three-site configuration that requires Enginuity 5874 and later running on the secondary Symmetrix system and Enginuity 5773 or 5874 and later running on the primary and tertiary systems. Figure 32 on page 131 shows an example of an SRDF/EDP basic configuration.
SRDF/EDP is achieved through a cascaded SRDF setup, where a Symmetrix VMAX system at a secondary site uses diskless R21 devices to capture only the differential data that would be owed to the tertiary site in the event of a primary site failure.

Figure 32  **SRDF/EDP basic configuration**

SRDF/EDP provides a long distance replication solution with the ability to achieve zero RPO at the tertiary site. This is also a lower cost solution when compared to three-site solutions such as concurrent and cascaded SRDF.
GDDR Managed Technologies

EMC SRDF consistency groups

GDDR manages a single consistency group. Devices in the group can be FBA (Open Systems) or CKD (mainframe) format. The design of ConGroup and GDDR allows for customers to build HA and DR solutions around the concept of the enterprise. Since GDDR is installed outside of the scope of a Sysplex-GDDR, controlled solutions can also span multiple Sysplex environments.

An SRDF consistency group is a collection of related Symmetrix devices that are configured to act in unison to maintain data integrity. The devices in consistency groups can be spread across multiple Symmetrix systems. Consistency Groups for z/OS (ConGroup) is an SRDF product offering designed to ensure the dependent-write consistency of the data remotely mirrored by SRDF/S operations in the event of a rolling disaster. Most applications, and in particular database management systems (DBMSs), have dependent-write logic embedded in them to ensure data integrity if a failure occurs, either hardware or software.

ConGroup is based on the concept of dependent-write operations. A dependent-write is a write not issued by an application until a prior, related write I/O operation is completed. An example of dependent-write activity is a database update:

1. The DBMS writes to the disk containing the transaction log.
2. The DBMS writes the data to the actual database.
3. The DBMS writes again to the log volume to indicate that the database update was made.

In a remotely mirrored environment, data consistency cannot be ensured if one of the writes is remotely mirrored, but its predecessor write was not. This could occur, for example, in a rolling disaster where a communication loss occurs and affects only a subset of the devices involved in the remote copy function.

ConGroup prevents a rolling disaster from affecting data integrity at the secondary site. When ConGroup detects any write I/O to a volume that cannot communicate with its R2 (secondary) mirror, it suspends the remote mirroring for all volumes defined to the consistency group before completing the intercepted I/O and returning control to the application. This is referred to as a ConGroup trip. In this way, ConGroup ensures a dependent-write consistent
image is available at the secondary site by preventing a dependent-write I/O from reaching the secondary site if the previous I/O only gets as far as the primary mirror.

I/O to the primary devices in the consistency group can still occur even when the devices are Not Ready on the SRDF links. Such updates are not immediately sent to the secondary site. However, they are propagated after the affected links are again operational, and data transfer from the primary devices to the secondary devices resumes.

Assume an SRDF configuration consists of three Symmetrix systems with primary devices, and two additional Symmetrix systems with secondary devices. The systems with primary devices send data to the systems with secondary devices as shown in Figure 33 on page 133.
Next, assume that the links between primary system 2 and secondary system 1 fail. Without a consistency group, primary systems 1 and 3 continue to write data to the secondary site systems 1 and 2 while primary system 2 does not, as shown in Figure 34 on page 134. The copy of the data spread across secondary systems 1 and 2 becomes inconsistent.

However, if primary systems 1, 2, and 3 belong to a consistency group, as shown in Figure 35 on page 135, and the links between primary system 2 and secondary system 1 fail, the consistency group automatically stops primary systems 1 and 3 from sending data to secondary systems 1 and 2. Thus, the dependent-write consistency of the data spanning secondary systems 1 and 2 remains intact.
Figure 35  Primary systems 1, 2, and 3 in a consistency group
EMC AutoSwap

EMC AutoSwap provides the ability to move (swap) workload I/Os transparently from volumes in one set of Symmetrix storage systems to volumes in other Symmetrix storage systems without operational interruption. An example is shown in Figure 36. AutoSwap, with SRDF and EMC Consistency Groups, dramatically increases data availability.

![AutoSwap diagram]

**Figure 36  AutoSwap example**

AutoSwap differs from IBM HyperSwap in that AutoSwap is not confined to operating completely within a Sysplex. AutoSwap can coordinate swaps for multiple monoplex or Sysplex systems in a single swap event.

Swaps may be initiated either manually as planned events or automatically as unplanned events (upon failure detection).

- Planned swaps facilitate operations such as non-disruptive building maintenance, power reconfiguration, DASD relocation, and channel path connectivity reorganization.
Unplanned swaps protect systems against outages in a number of scenarios. Examples include: power supply failures, building infrastructure faults, air conditioning problems, loss of channel connectivity, entire DASD system failures, operator error, or the consequences of intended or unintended fire suppression system discharge.

In GDDR managed environments planned swap events are initiated and controlled by GDDR. Unplanned swap events are initiated and controlled by AutoSwap. GDDR monitors for unplanned swap events and will re-configure surviving infrastructure to deal with the new source of write I/O activity after the successful unplanned swap event.

Swaps are concurrently performed while application workloads continue in conjunction with EMC Consistency Groups. This option protects data against unforeseen events, and ensures that swaps are unique, atomic operations that maintain dependent-write consistency.
EMC SRDF/Star

SRDF/Star is a three-site disaster recovery solution consisting of a primary Symmetrix system (Site A) and two remote Symmetrix systems (Sites B and C). SRDF/Star provides advanced multisite business continuity replication that augments concurrent SRDF/S (synchronous) and SRDF/A (asynchronous) operations for the same primary volumes with the ability to incrementally establish or reestablish SRDF/A replication between the primary (R1) and remote (R2) sites in the event of a primary site outage.

This section provides information on the following:
- “SRDF/Star and GDDR” on page 138
- “Concurrent SRDF/Star” on page 139
- “Cascaded SRDF/Star” on page 140
- “SRDF/EDP” on page 142

SRDF/Star and GDDR

IMPORTANT

In mainframe environments, EMC GDDR is a mandatory requirement for all SRDF/Star deployments to automate site swap activity and/or disaster restart.

In the event of a primary site outage, the EMC GDDR SRDF/Star solution allows customers to quickly move operations and re-establish protection between the remaining sites. Once conditions permit, customers can rejoin the primary site to the configuration, resuming SRDF/Star protection.

SRDF/Star can operate in concurrent or cascaded environments, providing the appropriate differential synchronization. These environments address different recovery and availability objectives:
- Concurrent SRDF/Star positions the secondary site or the remote site as potential recovery sites, and provides differential resynchronization between the secondary and remote sites. To achieve this positioning, some level of reconfiguration intervention is required to access point-of-disaster data.
◆ Cascaded SRDF/Star positions only the remote site as the recovery site with minimal intervention to access point-of-disaster data. This solution differentially synchronizes data between the primary site and the remote site.

◆ SRDF/EDP can also be used in an SRDF/Star configuration. By design, SRDF/EDP only allows recovery at the tertiary site.

SRDF/Star provides rapid reestablishment of cross-site protection in the event of primary site failure. Rather than a full resynchronization between the remote sites, SRDF/Star provides differential synchronization between the remote sites, and dramatically reduces the time it takes to remotely protect the new production site. SRDF/Star also provides a mechanism to determine which remote site has the most current data in the event of a rolling disaster that affects the primary site. In all cases, users maintain the ability to choose which site to operate from and which site's data to use when recovering from a primary site failure.

The host-based Multi-Session Consistency (MSC) task at the primary (R1) site controls normal SRDF/Star operation. MSC performs session management at SRDF/S site B and when necessary at SRDF/A site C. The MSC session management task maintains the information needed to perform differential synchronization between site B and site C. Automation for SRDF/Star mainframe configuration management is provided by EMC GDDR.

Concurrent SRDF/Star

In Concurrent SRDF/Star operations, site B serves as the secondary site and the target of the SRDF/S links from site A. Site C serves as the tertiary site and the secondary site of the SRDF/A links from site A. The recovery links are the SRDF/A links between site C and site B. Figure 37 on page 140 shows a Concurrent SRDF/Star configuration.
SRDF/Star in a cascaded SRDF configuration has the ability to incrementally establish an SRDF/A session between the primary and the asynchronous site in the event the synchronous site fails. With cascaded SRDF/Star, the synchronous secondary site is always more current than the asynchronous secondary site.
Figure 38 shows a basic cascaded SRDF/Star configuration. Site B serves as the secondary site and the target of the SRDF/S links from site A. Site C serves as the out of region site and the target of the SRDF/A links from site B. The recovery SRDF/A links are between site C and site A.
Figure 39 illustrates a cascaded SRDF / Star EDP diskless environment with the remote devices at the asynchronous secondary site. The diskless R21 device streamlines the link connections to the asynchronous site in cascaded mode. No data copies are available at the synchronous target site B because that site's Symmetrix use diskless cache only R21 devices.
This chapter contains the following information:

- Overview ........................................................................................................... 144
- Ignore, Reset Clear, or IPL in place.............................................................. 146
- Contingency systems and workload movement .......................................... 147
- HMC-Only systems .................................................................................... 148
- LPAR and CEC recovery ............................................................................ 152
- Automated Configuration Check – DASD ................................................. 155
Overview

One of the many decisions that must be made as part of a GDDR implementation is whether or not customers want GDDR to manage production systems, production LPARs, and/or production CPCs from GDDR. It is important to understand that systems management in GDDR is optional. Some customers choose to have GDDR automation extend into managing zSeries systems, LPARs, and workloads. Other customers want some systems managed by GDDR and others excluded. There are also customers who do not want GDDR to do anything to their environment other than to manage EMC infrastructure software such as SRDF, TimeFinder, and AutoSwap. The GDDR systems management functions are flexible and can accommodate this wide ranging set of customer operational requirements.

If a decision is made for GDDR to manage system recovery beyond just the storage layer, a further decision needs to be made whether to include systems management into the scope of the implementation project or defer systems management to a later date. Deferring systems management to a second phase of implementation does make the implementation project simpler and shorter. When a decision is made to implement GDDR systems management, the following features/functions must be considered for each system to be managed by GDDR:

- Ignore, Reset Clear, or IPL in place
- Contingency systems and workload movement
- HMC Only
- LPAR and CPC recovery

Each of these functions are further discussed in this chapter.

Production systems under GDDR control are known as P-Systems or sometimes managed systems. These are normally the customer systems that run application workload, but may also include test systems and systems dedicated to Coupling Facility LPARs.

Commencing with GDDR 3.1, EMC began introducing more functionality into CEC, LPAR, and system recovery options. Some customers use GDDR to manage their P-Systems and others chose to manage these systems themselves. GDDR has been architected to provide significant flexibility when it comes to P-System management.
For customers that have more complex environments, it helps to build a GDDR managed system recovery map similar to Table 6. This documents the actions that GDDR is to take for event management on any of the managed systems.

**Table 6  GDDR managed system recover map**

<table>
<thead>
<tr>
<th>LPAR NAME</th>
<th>E01</th>
<th>E02</th>
<th>E03</th>
<th>E04</th>
<th>F01</th>
<th>F02</th>
<th>F03</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>DC1</td>
<td>DC1</td>
<td>DC1</td>
<td>DC1</td>
<td>DC2</td>
<td>DC2</td>
<td>DC2</td>
</tr>
<tr>
<td>CPC</td>
<td>SYSTEM E IBM 2096-S07-R04 #012345</td>
<td>SYSTEM E IBM 2096-S07-R04 #012345</td>
<td>SYSTEM E IBM 2096-S07-R04 #012345</td>
<td>SYSTEM E IBM 2096-S07-R04 #012345</td>
<td>SYSTEM E IBM 2096-S07-R04 #6789A</td>
<td>SYSTEM E IBM 2096-S07-R04 #6789A</td>
<td>SYSTEM E IBM 2096-S07-R04 #6789A</td>
</tr>
<tr>
<td>Workload type</td>
<td>Production</td>
<td>Production</td>
<td>Development</td>
<td>Test</td>
<td>Production</td>
<td>Production</td>
<td>Development</td>
</tr>
<tr>
<td>Sysplex Name</td>
<td>Prod</td>
<td>Prod</td>
<td>Prod</td>
<td>Prod</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Name</td>
<td>MVSA</td>
<td>MVSB</td>
<td>MVSC</td>
<td>MVSD</td>
<td>MVSE</td>
<td>MVSF</td>
<td>MVSG</td>
</tr>
<tr>
<td>Desired State</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
</tr>
<tr>
<td>Recover</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Recovery type</td>
<td>Workload Move</td>
<td>Workload Move</td>
<td>Recovery LPAR</td>
<td>Reset Clear</td>
<td>IPL in place</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recovery location</td>
<td>F01</td>
<td>F02</td>
<td>F03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activate CBU</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ignore, Reset Clear, or IPL in place

When a system is defined to GDDR systems management a decision must be made regarding GDDR treatment of a failure of the defined system. The options are:

- Ignore — The system is defined to GDDR but this instructs GDDR to take no action if GDDR observes that this system has failed. Failure is typically “noticed” by GDDR when the EMC SCF product “de-registers” the system typically because of polling time-out.

- Reset Clear — If GDDR detects that the system has crashed (such as SCF de-registration), GDDR will access the HMC and RESET CLEAR the LPAR but take no other action.

- IPL in place— If GDDR detects that the system has crashed, GDDR will access the HMC, RESET CLEAR the system and then IPL in place.
Contingency systems and workload movement

Contingency systems are *hot standby* systems that are IPLed and active and (typically) are located in a different data center.

Often, Contingency Systems are the *remote half* of a geographically separated, single site workload, parallel Sysplex.

In the event of a primary site failure (for example, system MVSA fails in LPAR E01 in Table 6 on page 145) GDDR will trigger workload restart in system MVSE at site DC2 in LPAR F01. GDDR does this by sending message GDDX191I to appear on the MVSE console.

A message rule must be written and executed for MVSE to intercept the GDDX191I message and fire a REXX rule for that system to carry out the actions necessary to start the production workload. This message interception REXX rule would be written for whatever automation package runs on MVSE.
HMC-Only systems

For systems defined as “HMC-Only” GDDR can be used in place of native access to the HMC to perform system options. Interaction with HMC-Only systems is obtained through the “HMC LPAR Actions” panel, shown in Figure 40, by choosing option L.

--- GDDR - Actions Menu ---

Option =>

H Perform HMC Discovery
L Perform HMC LPAR actions
CSU Perform HMC CSU actions
S Manage Couple Datasets

This System: SYS1
This Site: DC1
Master-C: SYS1
Primary Site: DC1
Primary DASD: DC1

Automation: ON
Planned script: None
Unplanned script: None

EMC Geographically Dispersed Disaster Restart 04.00.00
Copyright © 2007-2011 EMC Corporation.

Select an option and press <Enter>
Press <F3> to return to the GDDR Primary Options Menu

Figure 40 GDDR Actions Menu

After choosing option L, a screen similar to Figure 41 on page 149, is displayed. Systems defined as HMC Only will appear. The “T” column in Figure 41 shows the GDDR system type. OS in the display indicates that this is a “HMC Only” system with Standard IPL Parms. For a full description of fields in this display, refer to the GDDR product guide.
GDDR Managed Systems Design Considerations

Figure 41  GDDR Perform HMC LPAR Actions

All of the functions displayed on this screen are protected by SAF profiles. This offers more controls than are available from direct HMC access. Customers choose to add “HMC Only” systems so that they have more control over access controls to these powerful system commands.

The following are actions you can request:

- **A — Activate LPAR**

To activate the LPAR of the selected system, type A next to the selected system and press Enter. A confirmation pop-up will be displayed. Confirm the request by entering Y and pressing Enter, or cancel the request by pressing F3.
D — Deactivate LPAR
To deactivate the LPAR the selected system runs in, type D next to the selected system and press Enter. A confirmation pop-up will be displayed. Confirm the request by typing Y and pressing Enter, or cancel the request by pressing F3.

L — Load Clear
To clear and load a selected system using the displayed load address and load parameters, type L next to the selected system and press Enter after ensuring the value of the Desired State field (DS) is consistent with the Load action. A confirmation pop-up will be displayed. Confirm the request by typing Y and pressing Enter, or cancel the request by pressing F3.

R — Reset Clear
To reset and clear a selected system, type R next to the selected system and press Enter after ensuring the value of the Desired State field (DS) is consistent with the Reset action. A confirmation pop-up will be displayed. Confirm the request by typing Y and pressing Enter, or cancel the request by pressing F3.

X — Load Recovery
To IPL at the system’s alternate location, type X next to the selected system and press Enter. If this option is available, the Site value will display a fourth character after DCx; either H or A, meaning the system is currently running at DCx in its home (H) location or its alternate (A) location.

W — CPC Swap
To perform a planned CPC swap, type W next to the selected system and press Enter.

S — Start LPAR
To start a selected system, type S next to the selected system and press Enter. A confirmation pop-up will be displayed. Confirm the request by typing Y and pressing Enter, or cancel the request by pressing F3.

P — Stop LPAR
To stop a selected system, type P next to the selected system and press Enter. A confirmation pop-up will be displayed. Confirm the request by typing Y and pressing Enter, or cancel the request by pressing F3.
◆ Q — Query LPAR
To query the LPAR of the selected system, type Q next to the selected system and press Enter. A confirmation pop-up will be displayed. Confirm the request by typing Y and pressing Enter, or cancel the request by pressing F3.

◆ E — Show Events
Display any relevant active events for the system in the Message field when you type E next to the selected system. System related events are IPL, (planned IPL), MXS, (unplanned MXS), and MHB, (unplanned MHB).

◆ T — Toggle desired state (U to D or D to U)
To toggle the state of a system, type U (up) or D (down). The default value is U.
GDDR Managed Systems Design Considerations

LPAR and CEC recovery

The following information is contained in this section:

- “C-System” on page 152
- “Production system options” on page 152
- “Managed system options” on page 153

The descriptions and actions in this section refer to Figure 41, “GDDR Perform HMC LPAR Actions” on page 149.

C-System

GDDR can tolerate the loss of a C-System. In the event of a C-System failure, GDDR will report *degraded mode* errors periodically on the Perform Health Check panel, as described in *Perform Pre-Script Checkup* on page 59 until the C-System rejoins the GDDR complex. In the case of the loss of a GDDR Master C-System, another GDDR C-System will take over the master function. Again, *degraded mode* will be reported until the missing C-System returns. LPAR Recovery can be specified for GDDR C-Systems, which may help minimize the amount of time GDDR operates in degraded mode.

Production system options

For production systems, the following options are available:

- Ignore
- IPL in place
- Reset Clear
- Start workload in contingency systems

The *desired state* (DS) indicator can be useful if a system needs to be down for maintenance ensuring that GDDR will not kick into action. If SCF is topped normally on that system, GDDR will not kick into action as the SCF controls the de-registration process and GDDR will know that the system has been shut down in a controlled manner.
Managed system options

For managed systems, the site field’s location indicator is either set to “H” indicating “H”ome location for this system or “A” indicating “A”way, meaning that this system is operating in its recovery location. It is possible that the Home and Away locations are in the same data center, but commonly the Away location is a different data center.

For managed systems, the following should be considered:

- Multi-site workloads do not have CONTingency systems.
- A system can have LPAR-RECOVERY specified indicating that the workload can be restarted in a different LPAR to where the system is currently running. The recovery LPAR can be in the same or different data center.
- Recovery LPARs defined to GDDR as REGULAR LPARs, but are defined a second time, on the system Recovery Attributes panel, as recovery for a protected systems.
- A protected system can have a CONTingency system as a Recovery LPAR.
- Sysplex systems with more than one LPAR can have Recovery LPARs defined, which could be in the primary or secondary data center.
- So for systems protected with LPAR recovery, they are either Home or Away but their SITE value does not change when they are “Away”.
- Systems NOT protected by LPAR Recovery can (optionally) have a second regular LPAR defined for them, on a different site. These second LPARs could even be the home location of a system protected with LPAR_Recovery.
- The difference between a Recover LPAR definition and a second LPAR definition is subtle:
  - The second LPAR definition is NOT considered as a recovery option, and will never lead to a system running in that location to be displaced.
  - The SITE value for a system NOT protected with LPAR_RECOVERY changes depending on whether it is running in its first or second LPAR.
GDDR Managed Systems Design Considerations

- LPAR Recovery for C-systems: avoid degrade mode when a C-system is lost.
- GDDR LPAR Recovery supports repeated failures: LPAR in Away location is protected with LPAR Recovery in Home location.
Automated Configuration Check – DASD

Traditionally, GDDR parameter validation has been very much limited to checking syntax and internal consistency, with relatively few verifications against real world data.

GDDR 3.2 introduced the Automated Configuration Check—DASD function with the GDDR ACCD utility.

One of the more challenging parts of defining an environment to GDDR has always been to ensure correct definition of large and complex SRDF configurations.

GDDRACCD removes the uncertainty about the correctness of an existing GDDR SRDF device definition by checking the real configuration found in the GDDR managed EMC Symmetrix storage.

GDDR variables influencing the expected configuration

GDDRACCD works is by formulating a detailed expectation about the configuration based on a whole series of GDDR global variables. This expectation is then compared against what we find in the EMC Symmetrix DMX.

5 phases:

Phase 1
Validation for internal consistency and exit if errors: no point checking the DMX if we already know the parms are bad.

Phases 2 and 3
Query the box and report deviations.

Phase 4
Provides information allowing the user to verify what GDDRACCD based its decisions on.

Phase 5
RDF.DEVICES parms allow comparison to existing parms, but differences are to be expected:

- Real parms could be a subset of generated parms
- Ranges in real parms could for a split of generated ranges
- Gatekeepers will certainly be different
This chapter contains the following information:

- Storage configuration and layout .................................................. 158
- Infrastructure recommendations ................................................... 159
- EMC software levels ................................................................. 164
- LPAR configuration specifics ....................................................... 165
- Sysplex and coupling facility considerations............................. 167
Storage configuration and layout

Symmetrix storage systems incorporate many standard features that ensure a higher level of recoverability and availability than other storage systems. They are configurable at the logical volume level so different protection schemes can be applied to different classes of data within the same Symmetrix storage system on the same physical device. Customers choose data protection options, such as the following, to match their data requirements:

- Mirroring (RAID 1) or RAID 10
- RAID 6 (6+2) and RAID 6 (14+2)
- RAID 5 (3+1) and RAID 5 (7+1)
- Symmetrix Remote Data Facility (SRDF)
- TimeFinder
- Dynamic Sparing
- Global Sparing
Infrastructure recommendations

There is benefit to standardizing configurations best practices across the install base of GDDR worldwide. This makes the support function more efficient in that support people will be able provide guidance aided by common naming conventions and infrastructure layout.

The recommendations in this section are designed to minimize the overheads associated with large infrastructure deployments. A good example is the number of device ranges. Having very small device ranges increases the overheads associated with infrastructure manipulations, consuming more C-System memory, more worker tasks, more commands issued to the arrays, and so on. Engaging the GDDR Solution Support team very early in your design phase will ensure an appropriate trade-off between configuration flexibility and associated overheads.

The following recommendations are discussed in this section:

- “Gatekeeper device list” on page 159
- “Started tasks names” on page 161
- “Important dataset names” on page 163

Gatekeeper device list

A Gatekeeper device is a reserved device used to direct commands to a particular array(s). Due to the volume of the commands during certain activities, such as data center swaps, the volume of commands to these devices can be quite high. GDDR does not require gatekeeper devices for itself but some underlying technologies do require these devices. For performance and to avoid potential conflicts, gatekeeper devices for SRDF/S, SRDF/A, MSC/Star etc must be defined and must be separate devices.

It is important to understand the layout of gatekeeper devices and which component uses gatekeepers. It is recommended to create and maintain charts similar to these for each array in the solution.

The example in this section shows a list of the SRDF Gatekeeper devices and RDF groups for a SRDF/Star with AutoSwap configuration with one VMAX frame located at each site. The following list of SRDF Gatekeeper devices and RDF groups used at each location describes the intended purpose for use in commands.
The required Host Component (H.C.) command prefix is also provided to help facilitate command generation for each specific RDF group. This additional information can be helpful in a configuration where multiple SRDF Host Component tasks are present.

**DC1 Complex:**

- **GDDRPARM:** Gatekeeper = 803D-8075 (00E1-0119)
- **H.C. Prefix:** ##

**Customer-Specific-Name VMAX# 2455:**

- **SRDF/ S - STAR:** Group = 10
- **R/W - AD:** Group = 21
- **SRDF/ A - Concurrent:** STAR::Group = 20
- **MSC/STA:** Gatekeeper = 8037 (00DB)
- **SRDFA/MSC:** Gatekeeper = 8076 (011A)
- **SCF:** Gatekeeper = 803B-803A (00DC-00DE)
- **CSC:** Gatekeeper = 903B-903C (00DF-00E0)

**DC2 Complex:**

- **GDDRPAR:** Gatekeeper = 903D-9075 (00E1-0119)
- **H.C. Prefix:** ##

**Customer-Specific-Name VMAX# 2457:**

- **SRDF/S-STAR:** Group=10
- **STAR-Recovery:** Group=30
- **MSC/STAR:** Gatekeeper=9037 (00DB)
- **SRDFA/MSC:** Gatekeeper=9076 (011A)
- **SCF:** Gatekeeper=903B-903A (00DC-00DE)
- **CSC:** Gatekeeper=903B-903C (00DF-00E0)
DC3 Complex:

<table>
<thead>
<tr>
<th>STC</th>
<th>Parmlib</th>
<th>Gatekeeper</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDDRPARM</td>
<td></td>
<td>803D-8075 (00E1-0119)</td>
</tr>
<tr>
<td>H.C.Prefix</td>
<td></td>
<td>##</td>
</tr>
</tbody>
</table>

**Customer-Specific-Name VMAX# 2465:**

<table>
<thead>
<tr>
<th>STC</th>
<th>Parmlib</th>
<th>Gatekeeper</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRDF/A - Concurrent STAR:</td>
<td></td>
<td>Group = 20</td>
</tr>
<tr>
<td>R/W - AD:</td>
<td></td>
<td>Group = 21</td>
</tr>
<tr>
<td>STAR - Recovery:</td>
<td></td>
<td>Group = 30</td>
</tr>
<tr>
<td>SRDFA/MSC</td>
<td></td>
<td>Gatekeeper = 8076 (011A)</td>
</tr>
<tr>
<td>SCF</td>
<td></td>
<td>Gatekeeper = 8038-803A (00DC-00DE)</td>
</tr>
<tr>
<td>CSC</td>
<td></td>
<td>Gatekeeper = 803B-803C (00DF-00E0)</td>
</tr>
</tbody>
</table>

**Started task names**

EMC recommends the following conventions are adopted for started task names.

**STCs in DC1**

**GDDR managed Production LPARs:**

- EMCBASE
  - SYISVP.GDDR.PARMLIB(SCFINI00)
- EMCCGRP
  - SYISVP.GDDR.PARMLIB(CONGRP34)
- EMCSRDF
  - SYISVP.GDDR.PARMLIB(EMCCMD00)
- GDDRMAMIN
  - SYISVP.GDDR.PARMLIB(GDDMPARM)

**GDDR Control LPAR Customer-Specific-Name:**

- GDDRSCF
  - SYISVP.GDDR.PARMLIB(SCFINI00)
- GDDRSRDF
Storage Infrastructure Design Considerations

- SYSISVP.GDDR.PARMLIB(EMCCMD00)
- GDDRCGRP
  - SYSISVP.GDDR.PARMLIB(CONGRP34)
- GDDRMAIN
  - SYSISVP.GDDR.PARMLIB(GDDMPARM)
- GDDREVMP
  - SYSISVP.GDDR.PARMLIB(GDDREVMP)
- GDDRHBMP
  - SYS1.GDDR.PARMLIB(GDDRHBMP)
- GDDRWORK
- GDDRPARMS
  - SYSISVP.GDDR.PARMLIB

STCs in DC2  
**GDDR Control LPAR Customer-Specific-Name:**
- GDDRSWF
  - SYSISVP.GDDR.PARMLIB(SCFINI00)
- GDDRSRDF
  - SYSISVP.GDDR.PARMLIB(EMCCMD00)
- GDDRCGRP
  - SYSISVP.GDDR.PARMLIB(CONGRP34)
- GDDRMAIN
  - SYSISVP.GDDR.PARMLIB(GDDMPARM)
- GDDREVMP
  - SYSISVP.GDDR.PARMLIB(GDDREVMP)
- GDDRHBMP
  - SYSISVP.GDDR.PARMLIB(GDDRHBMP)
- GDDRWORK
- GDDRPARMS
  - SYSISVP.GDDR.PARMLIB
STCs in DC3  

GDDR Control LPAR Customer-Specific-Name:

- GDDRSCF
  - SYSISVP.GDDR.PARMLIB(SCFINI00)
- GDDRSRDF
  - SYSISVP.GDDR.PARMLIB(EMCCMD00)
- GDDRMAIN
  - SYSISVP.GDDR.PARMLIB(GDDMPARM)
- GDDREVM
  - SYSISVP.GDDR.PARMLIB(GDDREVMP)
- GDDRWORK
- GDDRHBMP
  - SYSISVP.GDDR.PARMLIB(GDDRHBMP)

Important dataset names

The following is a list of the dataset names which are used by the various LPARs for GDDR control. These are the names as supplied by EMC, regardless of the naming convention used it is important to ensure that any support staff are familiar with the naming convention:

- SYSISVP.GDDR.*
- SYS1.GDDR.*
- SYS1.EMC.*
EMC software levels

Maintain a list of software levels, as shown in Table 7.

<table>
<thead>
<tr>
<th>Table 7</th>
<th>Software levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Customer-Specific-Name SOFTWARE LEVELS</td>
</tr>
<tr>
<td>Updated</td>
<td>Version</td>
</tr>
<tr>
<td>Mainframe Enablers</td>
<td>7.x</td>
</tr>
<tr>
<td>Resource Pack Base</td>
<td>7.x</td>
</tr>
<tr>
<td>Consistency Group</td>
<td>7.x</td>
</tr>
<tr>
<td>SRDF Host Component</td>
<td>7.x</td>
</tr>
<tr>
<td>TimeFinder Mirror</td>
<td>7.x</td>
</tr>
<tr>
<td>TimeFinder Snap</td>
<td>7.x</td>
</tr>
<tr>
<td>TimeFinder Utility</td>
<td>7.x</td>
</tr>
<tr>
<td>GDDR</td>
<td>3.x</td>
</tr>
</tbody>
</table>

**Microcode**

- VMAX base code: 5874.230.183
- VMAX E-Packs: RB/23 with 50853, 50854
- DMX-3/DMX-4 base code: 5773.155.107
- E-Packs: RB/53 with 49123, 49125
**LPAR configuration specifics**

These are guidelines/best practices for GDDR C-System LPAR configurations.

**Each GDDR LPAR requirements:**
- 2 GB storage
- 2 Shared CPs
- Minimum 25 Mod-9 DASD Units (or equivalent if different volume sizes are used)

**GDDR LPAR configuration restrictions and considerations:**
- GDDR Control LPAR DASD volumes are not shared with any LPAR
- No coupling facility or specialty engine requirements
- No DR Requirements for GDDR Control LPARs
- Initial GDDR LPAR to be modeled from CMCX LPAR
- Maintain recommended EMC Symmetrix microcode levels
- Initial security databases created from customer-environment LPAR database
- LPAR access restricted to Infrastructure Support Staff

**GDDR - Required automation and software elements:**
- z/OS 1.x Operating System
- CA/ACF2 Security
- EMC SRDF Product Suite

**Example customer infrastructure management components:**
- IOC Console operations support/automation (Tivoli Netview, SA for z/OS)
- Storage (SMS, DFHSM DASD space management, tape management)
- Network (FTP, Firewall, TN3270, NDM)
- Systems management (Incident, change (ECMS))
- Change management (Version and change control)
- SMF and Logrec data collection (Security, billing, performance)
Storage Infrastructure Design Considerations

- Output archival/retrieval/spool management (product)
- Infrastructure job scheduling (Data collection, infrastructure backups)

Cross system communications (CSC) requirements:

Update the SCFINIxx member in SYSISVP.GDDR.PARMLIB with the following updates on all LPARs (PROD and GDDR C-Systems):

- SCF.CSC.IDLEPOLL=5
- SCF.CSC.EXPIRECYLE=20
- SCF.CSC.MITPERIOD=60
Sysplex and coupling facility considerations

GDDR operations, system management, and enablement of data consistency are compliant with z/OS Sysplex availability, design, and operations. GDDR extends Sysplex availability by addressing a wide range of planned and unplanned operations for Sysplex systems, DASD resources, Sysplex resources, Central Processors, processor capacities, and sites.

Sysplex systems require a single time source and synchronous mirroring between the source DASD and recovery DASD. Asynchronous mirroring does not support Sysplex operations; therefore Sysplex linkage between asynchronous operational sites is not a consideration for sites using SRDF/A replication, either as a two-site GDDR SRDF/A or to the third site in STAR configurations.

Sysplex configuration is a key component of business continuity capability of a proposed solution. The design of the Sysplex, its required connectivity and the topology of shared Sysplex resources has a role in the solution’s redundancy, performance, and business continuity. Typically, solutions designed for higher availability and continuous availability have increasingly complex Sysplex considerations. However, GDDR ConGroup solutions designed for site switch and system restart require very modest Sysplex considerations between the sites.

GDDR’s C systems do not have any Sysplex operational requirement. The C systems are stand-alone systems outside any of the managed Sysplexes. This isolates the C systems from Sysplex wide issues, simplifies C system implementation, and allows the C systems to manage multiple Sysplexes simultaneously. It also allows asynchronous mirroring solutions and some synchronous mirroring solutions without any cross site required Sysplex connectivity and without degradation of Recovery Time Objectives.

When data is shared within a Sysplex data, consistency is presumed and GDDR’s architecture should align with Sysplex requirements for data consistency. In most instances, a Sysplex’s mirrored DASD resources required for the restart of systems and workload should belong to a single ConGroup. There is no limitation to the number of Sysplexes encompassed by a single ConGroup and managed by GDDR.
Significant variations in the required Sysplex connectivity depend on the Sysplex topology, the models of the z/series hardware, Sysplex timing mechanisms and the desired level of business continuity. Careful planning of the cross site linkages, supported distances, sufficient redundancy and considerations for performance impact must be part of the early design and project planning.

Typically, when architecting a multisite Sysplex, DWDMs are employed to coalesce, aggregate, and manage the Sysplex connectivity as well as cross site DASD and mirroring connectivity. Solutions including AutoSwap are easily facilitated by cross site DASD connectivity over DWDMs. Furthermore, many of the z/series linkage protocols can not employ channel extension, requiring dark fiber or DWDM connectivity.

When the solution includes AutoSwap, regardless of the Sysplex topology all DASD is required to stay up:
- Must be mirrored in SRDF/S
- Must be available to the systems at DC1 and DC2
- Must have all R1-devices on 1 site, and all R2-devices on the opposite site
- If it has a UCB on the R1-side, must also have a UCB on the R2-side

When employing Sysplex LOGR, the Sysplex LOGGER data sets should be mirrored with SRDF/S and SRDF/A in SRDF/Star configurations.

**Single Site Sysplex**
All active production systems reside at primary site, the systems comprising the Sysplex can not span DC1 and DC2. This is usually representative of GDDR ConGroup or GDDR SRDF/A solutions where a complete Sysplex restart is required to restart systems and workload at the recovery site.

No Sysplex connectivity is required between sites. Sysplex time is not an issue, duplexing of structures between sites is not indicated. Structures would be rebuilt when the systems are IPLed at the recovery site. Restart of systems can be facilitated by the occasional mirroring of system resources including the Sysplex couple data sets. The Sysplex couple data should not be mirrored using SRDF/S and not in the ConGroup (consider mirroring with Adaptive Copy).
Alternatively, with minimal cross site DASD connectivity the Sysplex couple data sets can be mirrored with Sysplex couple data set duplexing.

In the event of a ConGroup trip, GDDR does not attempt to stop all systems to preserve the consistency of the target SRDF/S DASD and the contents of the target duplex Coupling Facility structures. This technique has only limited value during the restart of systems.

**Multi-Site Sysplex**

This architecture is consistent with solutions requiring high or continuous levels of availability, when some subset of operations is expected to continue in the event of systems, DASD, site failures or multiple failures.

The z/series systems hosting z/OS systems in the multisite Sysplex at both sites must share a single time consistency in a Common Time Network (CTN). The CTN network topology, redundancy, and protocols depend upon the timer technology employed. When employing STP, placement of the PTS and BTS is an important consideration.

Base Sysplex requires cross site CTC connectivity between all z/series CPCs, single time consistency, and common Sysplex couple data sets.

Parallel Sysplex employs Coupling Facilities to facilitate system to system communications. CFs can be Internal or External Coupling Facilities and all systems require connectivity to the CFs. These Sysplex links should be redundant. CFs should be deployed in a manner consistent with the business continuity objectives. If structures are duplexed these links should be redundant.

AutoSwap is usually included in the implementation to eliminate DASD as a single point of failure. If so, all DASD required for operations must be mirrored. The Sysplex couple data sets must not be mirrored with SRDF/S and must rely upon Sysplex couple data set duplexing. Cross site DASD connectivity consistent with cross site operations is typically a consideration.
This chapter contains the following information:

- Overview ................................................................. 172
- Understanding bandwidth requirements .................. 173
- Synchronous network design considerations .......... 176
- Asynchronous network design considerations ........ 179
Overview

After the synchronous and/or asynchronous bandwidth assessment has been conducted, attention turns to the design and construction of the network. There are many different types of networks but for this discussion we will differentiate between networks designed to carry synchronous workloads and those designed to carry asynchronous workloads.

Delays in the provisioning, debugging, and load testing of network infrastructure are common causes of delay for GDDR projects. Remember the constraint framework for the project. If your network provisioning is late, be prepared to adjust the implementation timelines around new and realistic dates.
Understanding bandwidth requirements

Before a network topology can be designed, the data change rate must be measured. This change rate will be used to determine both synchronous and asynchronous network bandwidth requirements.

EMC Technical Support (TS) has toolsets and methods to measure existing I/O flows to construct data replication requirements over time. Replication requirements not only vary throughout the 24 hour day (batch vs. online) but also throughout the month/year, with most enterprises having at least one peak processing period. Therefore, the data sampling for modeling must span a critical time period, or at least be factored from the observation period based on the known difference between the observation period and the peak period.

Both SRDF/S and SRDF/A use block level I/O for replication; therefore the collected data can be used for both modeling exercises. The actual bandwidth requirement will typically vary between the two types of networks (even for combined Synchronous and Asynchronous:

- For HA networks, page volumes and temporary data sets are included in the replication set.
- For asynchronous networks, these data types are typically excluded since in the event that the DR site copy is required the contents of the page volumes and temporary data set volumes are not required for system restart.

Note: Figure 42 on page 174 shows a measured workload with significant write I/O imbalance across the three Symmetrix. Part of the deployment of the project will be to move workloads between the arrays to achieve a more balanced I/O profile.
The raw (uncompressed) bandwidth requirement is determined by multiplying the number of writes observed in a specific interval by the size of the write. Earlier design tools use the Connect Time to determine transmission interval and then calculate blocksize. This method does not work for FICON attached arrays, so an average block size is determined by dividing the total bytes transferred in the time interval by the number of writes observed. The method also varies based on data source, such as SMF records versus internal EMC Symmetrix STP data.

For synchronous workloads, performance is critical. Therefore, this analysis includes a check of hot volumes, if detected. Hot volumes is a list of volumes produced where the overhead of synchronous replication is likely to be problematic. Further analysis of these volumes is required to determine the contents of the volume and identify mitigation techniques, which may include:

- Reblocking data sets
- Separating hot data sets
• Deploying SMS data striping
• Changing ACS routines
• Balancing workload across Symmetrix arrays

At the conclusion of the study period, a report is produced which will include a profile of the observed bandwidth requirement as shown in Figure 43.

**Note:** Many customers are accustomed to working with their local account teams for the sizing of infrastructure deployments. Arguably though, your account team will not have experience designing for this type of complex solution. The GDDR Solution Support Team can help to direct EMC Corporate resources to ensure that this vital piece of supporting infrastructure meets best practice standards.

![Figure 43 Bandwidth requirement](image)
Synchronous network design considerations

All array-based synchronous data replication creates an overhead on write response times. For this reason, it is absolutely essential to understand the impact of synchronous data replication before building out any infrastructure.

EMC has sophisticated modeling tools which can be used to apply the overheads of SRDF/S synchronous mirroring at zero distance, and then apply latency into the model as the distance between the primary (R1) and secondary (R2) storage is increased. At some point in these modeling studies, the maximum amount of separation between the R1 and R2 becomes known when the increase I/O response time begins to slow write throughput to the point that the observed write activity to the volume can no longer be sustained. Post implementation, SRDF/S overheads can be seen as disconnect time when looking at I/O response times.

Figure 44 on page 177 shows the results of one part of a typical synchronous design study output: DASD volumes sorted into descending write activity.

Looking more closely at volume CI9003, it can be seen to have a peak write activity of 300 writes per second. In effect, this volume could have an average response time of 3.3 milliseconds and sustain a throughput of 300 writes per second.

A separate analysis (not show) reveals the average response time to this volume as 0.5 milliseconds. Subtracting 0.5 from 3.3 reveals that the overheads of synchronous replication cannot impose more than 3.3 – 0.5 = 2.8 milliseconds to this I/O before the application would experience performance degradation (because the throughput requirement of 300 writes cannot be maintained if the average response time exceeds 3.3 milliseconds).
Applying this logic to all volumes in the study group produces a chart similar to Figure 45. The chart shows the existing throughput requirement (in blue) and the likely throughput requirement (in red) when synchronous overheads at the customer-specific distance between the R1 and R2 devices is applied to the existing write I/O response time.
Tuning for synchronous replication can be done well in advance of the implementation of GDDR. The following techniques are typically used to address performance issues:

- Use sequential data striping.
- Change the block size of the data set write. Many of the overheads associated with synchronous replication are fixed, such as the speed of light problem. This is the case whether the block being transmitted is small or large. Therefore, large block writes are more efficient in terms of write throughput over time in synchronous replication than smaller block sizes.
- Separate hot data sets — Two or more write I/O intensive data sets can be moved from the same volser and separated.
- Change ACS routines — This may be needed to ensure that separated data sets remain separated.

Most customers deploy a dark fiber network to support synchronous replication. The Network Terminating Equipment (NTE) for these networks is usually a Dense Wave Division Multiplexer (DWDM). DWDMs are able to concentrate multiple wavelengths onto single fiber(s) laid between data centers. DWDMs are very low latency so the greatest impact in terms of write I/O performance is typically the effect of the speed-of-light transmission of data through the fiber and is directly proportional to the distance of the fiber route.

Another advantage to the DWDM approach is for customers looking to build HA infrastructures. In addition to the Fibre Channel packets generated by SRDF/S replication, the same fiber can carry FICON signals necessary to support AutoSwap.

In some instances customers will build HA DASD infrastructures where the R1 and R2 devices are in the same data center (perhaps separated by a thermal firewall), in which case direct connect or connect via cascading directors is common.
Asynchronous network design considerations

Asynchronous replication is provided by SRDF/A. The interface to SRDF/A is via EMC Multi-Session Consistency (MSC) software.

MSC provides an API interface used by GDDR to control both SRDF/A running in MSC mode and also SRDF/A when it is running in Star mode. The GDDR for SRDF/A configuration (two-site) controls SRDF/A replication using the MSC API. The same MSC API interface is used by GDDR to control GDDR for SRDF/Star solutions (three-site). For more specific information on these control interfaces, refer to the following sections:

- “SRDF/Star with AutoSwap” on page 44
- “GDDR for SRDF/S with ConGroup (Two sites - DC1 and DC2)” on page 76
- “GDDR for SRDF/A (Two sites - DC1 and DC3)” on page 81

SRDF/A is designed as a cache-to-cache replication infrastructure. It typically supports cycle times down to three second durations, assuming that the network infrastructure is designed to carry the workload. Since it is theoretically possible to lose up to two SRDF/A cycles in a DR event, the RPO is calculated by multiplying the SRDF/A cycle time by two (2). Thus, a three-second cycle time is said to support an RPO of six (6) seconds. It is important to understand that the RPO is an objective. At times it may not be possible for the replication infrastructure to maintain the objective time. Examples include:

- Program loops that generate large amounts of write I/O
- Unusually high write activity, such as during data base reorgs
- Compromised network (for example, a failed switch)

EMC has instituted a number of measures to improve resiliency of SRDF/A replication:

- **Transmit Idle** — Temporarily suspend SRDF/A data transmission to allow SRDF/A replication to survive temporary link problems such as link bounces.
- **Delta Set Extension (DSE)** — Elongates the SRDF/A capture cycle during longer duration temporary link problems, instead of terminating SRDF/A replication. DSE uses a disk pool to buffer the current capture cycle data until such time as it can be destaged across the link.
◆ **Write Pacing** — Slowing the arrival rate of write activity into the capture cycle by accepting the write to cache, disconnecting for a specified period of time (milliseconds), and then reconnecting to complete the I/O.

None of these improvements should be assumed to be substitutes to provisioning the network to the observed peak bandwidth requirement. To build one- or two-year operational resiliency into the network, it should be provisioned at the peak observed rate plus the annual rate of increase of observed throughput. This can be calculated from the known growth in storage. Assuming a growth rate of 20% over the past 12 months, the allowance for capacity increase can be calculated as:

\[
\text{Bandwidth}_{\text{growth}} = \frac{\text{Total Write IOs} \times \text{I/O Size} \times 20\%}{\text{Compression Ratio}}
\]

Many of the GDDR supported configurations are very large. The designer must consider the scalability of the proposed network and also troubleshooting. For mainframe workloads, EMC has observed FCIP networks to be reliable and scalable and relatively easy to troubleshoot and tune. Consider operational complexity when looking to tune or balance network workloads. Gig/E networks are becoming more popular but do introduce more complexity for workload balancing and, if used, WAN Optimizers can complicate the network architecture and impact the throughput, latency, and scalability of the network.

EMC has observed the following type of network topology, shown in Figure 46 on page 181, to be very successful in large scale SRDF/A implementations. EMC therefore recommends the following network topology for SRDF/A replication. Figure 46 shows 15 storage arrays in three-sites (DC1, DC2, and DC3) deploying multiple switches to aggregate bandwidth up to 10 GB for long-distance transmission and then splitting back through switches on the target side to the Symmetrix at the remote location. Note how readily this architecture allows for scale of communications paths, storage devices, and switches.
Figure 46  Recommended asynchronous network topology

Asynchronous network design considerations
This glossary contains terms related to disk storage subsystems. Many of these terms are used in this manual.

A

**alternate track**  A track designated to contain data in place of a defective primary track. See also "primary track."

**actuator**  A set of access arms and their attached read/write heads, which move as an independent component within a head and disk assembly (HDA).

**adapter**  Card that provides the physical interface between the director and disk devices (SCSI adapter), director and parallel channels (Bus & Tag adapter), director and serial channels (Serial adapter).

C

**cache**  Random access electronic storage used to retain frequently used data for faster access by the channel.

**cache slot**  Unit of cache equivalent to one track.

**channel director**  The component in the Symmetrix subsystem that interfaces between the host channels and data storage. It transfers data between the channel and cache.
<table>
<thead>
<tr>
<th><strong>Term</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>controller ID</strong></td>
<td>Controller identification number of the director the disks are channeled to for EREP usage. There is only one controller ID for Symmetrix.</td>
</tr>
<tr>
<td><strong>CKD</strong></td>
<td>Count Key Data, a data recording format employing self-defining record formats in which each record is represented by a count area that identifies the record and specifies its format, an optional key area that may be used to identify the data area contents, and a data area that contains the user data for the record. CKD can also refer to a set of channel commands that are accepted by a device that employs the CKD recording format.</td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>Direct access storage device, a device that provides nonvolatile storage of computer data and random access to that data.</td>
</tr>
<tr>
<td><strong>data availability</strong></td>
<td>Access to any and all user data by the application.</td>
</tr>
<tr>
<td><strong>delayed fast write</strong></td>
<td>There is no room in cache for the data presented by the write operation.</td>
</tr>
<tr>
<td><strong>destage</strong></td>
<td>The asynchronous write of new or updated data from cache to disk device.</td>
</tr>
<tr>
<td><strong>device</strong></td>
<td>A uniquely addressable part of the Symmetrix subsystem that consists of a set of access arms, the associated disk surfaces, and the electronic circuitry required to locate, read, and write data. See also “volume.”</td>
</tr>
<tr>
<td><strong>device address</strong></td>
<td>The hexadecimal value that uniquely defines a physical I/O device on a channel path in an MVS environment. See also “unit address.”</td>
</tr>
<tr>
<td><strong>device number</strong></td>
<td>The value that logically identifies a disk device in a string.</td>
</tr>
<tr>
<td><strong>diagnostics</strong></td>
<td>System level tests or firmware designed to inspect, detect, and correct failing components. These tests are comprehensive and self-invoking.</td>
</tr>
<tr>
<td><strong>director</strong></td>
<td>The component in the Symmetrix subsystem that allows Symmetrix to transfer data between the host channels and disk devices. See also “channel director.”</td>
</tr>
<tr>
<td><strong>disk director</strong></td>
<td>The component in the Symmetrix subsystem that interfaces between cache and the disk devices.</td>
</tr>
</tbody>
</table>
dual-initiator  A Symmetrix feature that automatically creates a backup data path to the disk devices serviced directly by a disk director, if that disk director or the disk management hardware for those devices fails.

dynamic sparing  A Symmetrix feature that automatically transfers data from a failing disk device to an available spare disk device without affecting data availability. This feature supports all non-mirrored devices in the Symmetrix subsystem.

E

ESCON  Enterprise Systems Connection, a set of IBM and vendor products that connect mainframe computers with each other and with attached storage, locally attached workstations, and other devices using optical fiber technology and dynamically modifiable switches called ESCON Directors. See also "ESCON director."

ESCON director  Device that provides a dynamic switching function and extended link path lengths (with XDF capability) when attaching an ESCON channel to a Symmetrix serial channel interface.

F

fast write  In Symmetrix, a write operation at cache speed that does not require immediate transfer of data to disk. The data is written directly to cache and is available for later destaging.

FBA  Fixed Block Architecture, disk device data storage format using fixed-size data blocks.

FRU  Field Replaceable Unit, a component that is replaced or added by service personnel as a single entity.

frame  Data packet format in an ESCON environment. See also "ESCON."

G

gatekeeper  A small logical volume on a Symmetrix storage subsystem used to pass commands from a host to the Symmetrix storage subsystem. Gatekeeper devices are configured on standard Symmetrix disks.

GB  Gigabyte, $10^9$ bytes.
Glossary

H

head and disk assembly
A field replaceable unit in the Symmetrix subsystem containing the disk and actuator.

home address
The first field on a CKD track that identifies the track and defines its operational status. The home address is written after the index point on each track.

hyper-volume extension
The ability to define more than one logical volume on a single physical disk device making use of its full formatted capacity. These logical volumes are user-selectable in size. The minimum volume size is one cylinder and the maximum size depends on the disk device capacity and the emulation mode selected.

I

ID
Identifier, a sequence of bits or characters that identifies a program, device, controller, or system.

IML
Initial microcode program loading.

index marker
Indicates the physical beginning and end of a track.

index point
The reference point on a disk surface that determines the start of a track.

INLINES
An EMC-provided host-based Cache Reporter utility for viewing short and long term cache statistics at the system console.

I/O device
An addressable input/output unit, such as a disk device.

K

Kilobyte, 1024 bytes.

L

least recently used algorithm (LRU)
The algorithm used to identify and make available the cache space by removing the least recently used data.

logical volume
A user-defined storage device. In the Model 5200, the user can define a physical disk device as one or two logical volumes.
### Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>long miss</strong></td>
<td>Requested data is not in cache and is not in the process of being fetched.</td>
</tr>
<tr>
<td><strong>longitude redundancy code (LRC)</strong></td>
<td>Exclusive OR (XOR) of the accumulated bytes in the data record.</td>
</tr>
<tr>
<td><strong>MB</strong></td>
<td>Megabyte, $10^6$ bytes.</td>
</tr>
<tr>
<td><strong>mirrored pair</strong></td>
<td>A logical volume with all data recorded twice, once on each of two different physical devices.</td>
</tr>
<tr>
<td><strong>mirroring</strong></td>
<td>The Symmetrix maintains two identical copies of a designated volume on separate disks. Each volume automatically updates during a write operation. If one disk device fails, Symmetrix automatically uses the other disk device.</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>Physical identification number of the Symmetrix director for EREP usage. This value automatically increments by one for each director installed in Symmetrix. This number must be unique in the mainframe system. It should be an even number. This number is referred to as the SCU_ID.</td>
</tr>
<tr>
<td><strong>primary track</strong></td>
<td>The original track on which data is stored. See also “alternate track.”</td>
</tr>
<tr>
<td><strong>promotion</strong></td>
<td>The process of moving data from a track on the disk device to cache slot.</td>
</tr>
<tr>
<td><strong>read hit</strong></td>
<td>Data requested by the read operation is in cache.</td>
</tr>
<tr>
<td><strong>read miss</strong></td>
<td>Data requested by the read operation is not in cache.</td>
</tr>
<tr>
<td><strong>record zero</strong></td>
<td>The first record after the home address.</td>
</tr>
<tr>
<td><strong>scrubbing</strong></td>
<td>The process of reading, checking the error correction bits, and writing corrected data back to the source.</td>
</tr>
</tbody>
</table>

*GDDR Solution Design and Implementation Techniques*
### Glossary

- **SCSI adapter**: Card in the Symmetrix subsystem that provides the physical interface between the disk director and the disk devices.
- **short miss**: Requested data is not in cache, but is in the process of being fetched.
- **SSID**: For 3990 storage control emulations, this value identifies the physical components of a logical DASD subsystem. The SSID must be a unique number in the host system. It should be an even number and start on a zero boundary.
- **stage**: The process of writing data from a disk device to cache.
- **storage control unit**: The component in the Symmetrix subsystem that connects Symmetrix to the host channels. It performs channel commands and communicates with the disk directors and cache. See also “channel director.”
- **string**: A series of connected disk devices sharing the same disk director.
- **unit address**: The hexadecimal value that uniquely defines a physical I/O device on a channel path in an MVS environment. See also “device address.”
- **volume**: A general term referring to a storage device. In the Symmetrix subsystem, a volume corresponds to single disk device.
- **write hit**: There is room in cache for the data presented by the write operation.
- **write miss**: There is no room in cache for the data presented by the write operation.