

# Whitepaper



**iQ2880**

**Reliability, Availability and Serviceability**

### **Important Information**

© 2007 iQstor Networks, Inc. The information contained in this white paper does not constitute a contractual agreement with iQstor Networks and may be modified at anytime without advance notice. This white paper is supplied on an "as is" basis with no warranty and no support. iQstor Networks makes no express warranty, whether written or oral with respect to this white paper.

No trademark, copyright, or patent licenses are expressly or implicitly granted (herein) with this white paper.

## Introduction

In today's mission critical computing environments, service level agreements (SLAs) commit IT organizations to measurable service metrics such as application performance, end-user response time and system availability. Storage plays a vital role in an IT organization's ability to deliver IT services because it directly impacts the performance, availability and functionality of most applications. In fact the ability to consolidate storage with the latest technology has increased the costs of any storage system failure dramatically.

A storage solution must therefore include robust architectural design that withstands potential failures without impacting data availability. It is critical that even in the face of component failures, such as disk drive or power supply, storage systems be able to deliver high levels of performance and availability without compromise. A robust storage solution must have reliability, availability and serviceability (RAS) features inherent to its architectural design. Components of these RAS features should include:

- No single point of failure. All critical components (blowers, power supplies, storage controllers, drives) are redundant
- Redundant I/O paths to hard drives
- High performance RAID for protection against disk failures
- Proactive error detection and response, e.g. memory scrub and cache vaulting
- Non-disruptive operating environment upgrades
- Remote "call home" capability to enable prompt field service response
- Enclosure environmental monitoring
- Modular design enabling non-disruptive component replacement

This white paper provides an overview of the reliability, availability, and serviceability (RAS) features built into the architecture of the iQ2880 storage system.

---

## Host Data Path Redundancy

The iQ2880 storage system supports host data path redundancy to ensure uninterrupted data access in the event of component failures. Components in the host data path are the host bus adapter (HBA), Fibre Channel switch, Fibre Channel cable and storage controller. A minimum of two HBA connections per host is required to provide full redundancy. Components of the storage system that enable host data path redundancy are:

**Multi-path IO Support** - The iQ2880 storage controller supports multi-path IO and load balancing features of HBA or third-party software vendors. The supported modes include load balancing or active-active mode, which allows I/Os to be issued to both storage controllers simultaneously over multiple paths.

Another mode supported is active-passive or preferred-alternative path mode. The active or preferred path is specified and failover occurs when a link failure is detected. Failover resulting from Fibre Channel cable, HBA or controller failures are supported.

**Redundant Active-Active Controllers** - The storage controllers of the iQ2880 support active-active operation. Both storage controllers present the same storage resources (drives, raid array, pool and virtual disks) to hosts, allowing storage resources on the iQ2880 to be accessed simultaneously through either storage controller.

The storage controllers constantly monitor each other's health through a dedicated high speed communication channel and over the Ethernet connection. When a peer storage controller is "silent" for over 30 seconds on both connections, the surviving storage controller initiates a failover process to takeover all system resources by assuming ownership of all the back-end Fibre Channel drives and flushing mirrored write cache data to the designated backend storage.

Support for fail-back enables the non-disruptive replacement of a failed storage controller in the iQ2880. A fail-back command is issued to the surviving storage controller to start the fail-back process. Once initialized the fail-back process non-disruptively restores the pre-failure configuration and I/O states.

**Cache Mirroring** - Each storage controller provides cache memory to speed up read/write requests by the host. To protect write data in cache from failure of a controller, the data is mirrored over a dedicated high speed link to its peer storage controller before a write completion status is returned to the host. In the event of a controller failure, the surviving controller will flush the mirrored write cache data to back-end drives, ensuring no data loss.

---

### Power and Cooling Redundancy

The iQ2880 uses dual redundant power supplies and cooling modules to provide complete redundancy in case of failures in the regulated power supplies, cooling modules or power distribution cables. The power supplies share the load so only one operating power supply is needed to run the storage system. Fail-over of the power supply module is immediate and automatic. Similarly, only one cooling module is needed to provide sufficient cooling for the storage system. Concurrent repair is supported on both the power supplies and cooling modules.

For full power loss situations, the iQ2880 supports optional uninterruptible power supply (UPS) in addition to, or in place of the battery backup module (BBM).

## Battery Backup

A Battery Backup Module (BBM) protects the data in the write-back cache during power outages. If system power is lost before the storage controller completes a cache write/flush operation, the BBM supplies auxiliary power to the storage controller cache memory to safely maintain the data in cache for up to 72 hours. On restoration of power, the contents of the write-back cache are written to the back-end drives before new read/write commands are accepted.

The battery backup module is periodically tested to ensure there is sufficient power to sustain data in the cache memory in the event of power failure. The battery is re-charged if the system detects a load below the re-charge threshold. If the battery load cannot be maintained above the re-charge threshold, a warning is issued to replace the BBM module. The caching policy will also be changed to write-through until the BBM module is replaced and the battery is recharged.

---

## External UPS

For full power loss situations, an optional UPS provides power to keep the storage system operating. UPS monitoring is performed by the storage controllers through the Ethernet network port. When the storage controller detects that input power has been switched to UPS power, the write cache memory contents is flushed to the back-end hard drives and the caching policy is set to write through.

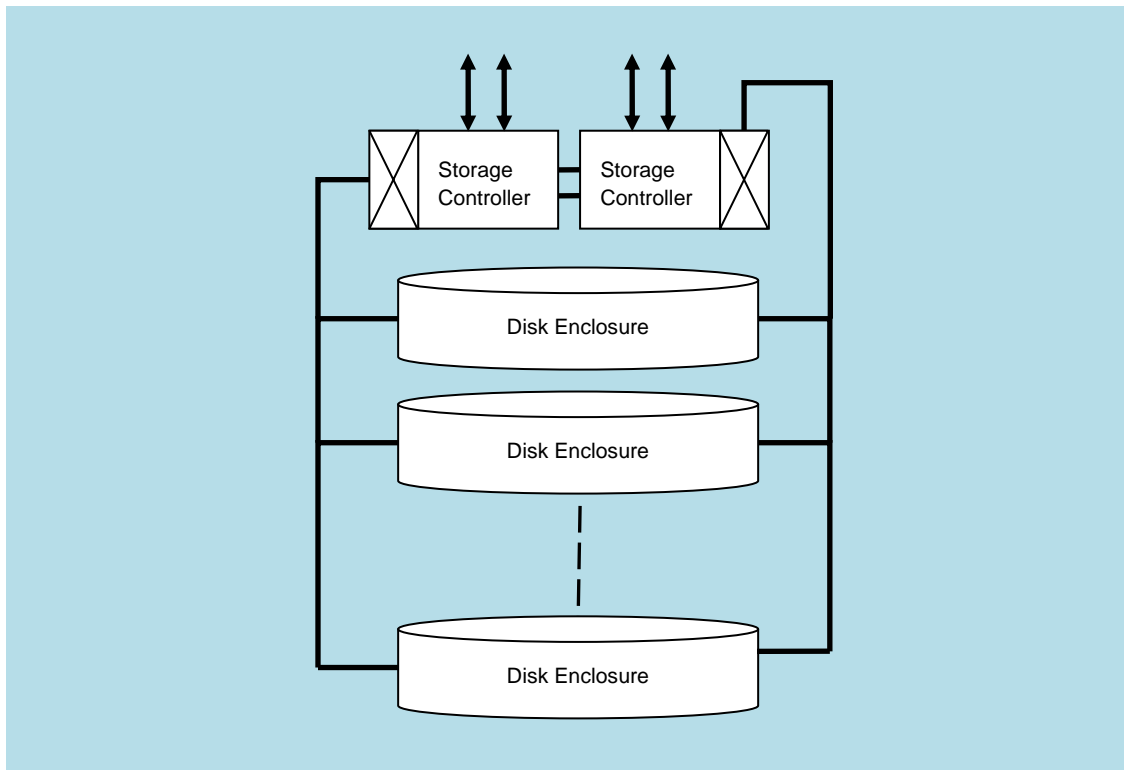
---

## Fibre Channel and Back-End Redundancy

The iQ2880 architecture incorporates 4 Gb/s Fibre Channel interfaces to ensure high performance and redundancy. Redundant paths to the disk drives provide non-stop access to information, even in the event of a component failure or replacement. The redundant components on the Fibre Channel back-end are:

- **Redundant Storage Controllers** – Dual storage controllers provide access to each disk drive. Each controller operates a separate physical path to the drive, while at the same time synchronizing write caches over dedicated redundant links to its peer controller. The dual paths to data eliminate single point of failure.
- **Redundant Loop Switches** – Switched Loops on each storage controller deploy Fibre Channel switch technology to minimize latency. FC frames destined for a drive within an enclosure are immediately delivered through the “cut-through switch” to the frame recipient. Frames not destined for the enclosure are switched to the next enclosure on the loop, eliminating the latency associated with traditional FC drive loops. In the event of disk failures, the cut-through switching also provides unique disk isolation features that increase total system availability, reliability and serviceability.

- **Redundant Cable Paths** – A separate, independent cable assembly attaches each storage controller to its associated set of Fibre Channel disk drives. If the cable path of one controller fails and there are pending I/Os to disconnected drives, a controller failover is triggered so that the peer controller completes the pending I/O. Working with multi-path failover driver on a host, the controller failover mechanism and redundant cable paths ensure all host I/Os are successfully completed. When a failed cable path is restored, the storage controller will automatically determine if all drives are available. If accessible, a controller fallback is facilitated.
- **Redundant Disk Ports** – Each Fibre Channel disk drive has two fully independent ports that are connected through separate cable assemblies to two different storage controllers. The dual pathways from the storage controllers to the disk ports operate completely independent of one another.



## Data Protection

The iQ2880 RAID data protection feature ensures uninterrupted access to data in the event of a disk failure. Support for different RAID levels below allows performance availability and capacity choices to be matched to application requirements.

<b>Data Protection Feature</b>	<b>Description</b>
Mirroring (Raid 1)	Duplicate copy of each volume on two disk drives for high performance and availability.
Parity RAID (Raid 3, 5)	Provides parity protection with a fixed parity drive for RAID 3 and distributed parity for RAID 5 to optimize available capacity while providing data protection.
Double Parity RAID (Raid 6)	Double parity protection enhances ability to provide continuous data availability across two drive failures.
Striped Mirrors (Raid 1+0)	Striping across multiple mirrored arrays provides the highest performance and availability.
Global Sparing	Increases data availability by allowing a failed disk drive to be automatically replaced by a spare disk drive.

Beyond hardware failures, disruptions to data access can also come from human errors, data corruption or natural disasters. As a container of corporate data assets, storage systems should be viewed as more than just a receptacle. Storage is also a platform for data services to manage and mitigate risk to data access. The data services options listed below are available to provide a higher level of data protection, recoverability and availability.

<b>Data Protection Feature</b>	<b>Description</b>
Volume Manager Virtualization	Enables virtual volumes to be created, expanded, deleted or moved on-demand. Eliminates downtime and disruption when scaling storage or expanding volumes.
Volume Copy Services	Local mirroring avoids costly interruptions to production volumes and mitigates the loss of any kind of outage by enabling data to be replicated.
Managed Snapshot Services	Provides the ability to recover from logical errors such as corrupted data or deleted files by preserving data at multiple points in time.
Remote Replication Synchronous	Provides synchronized copies of data to remote sites to address business continuity requirements.

## Modular Serviceability

The iQ2880 provides full component redundancy to protect against component failure and ensures uninterrupted access to information. The modular design improves serviceability by allowing non-disruptive replacement of components should a failure occur. Concurrent maintenance of all major components are supported, including:

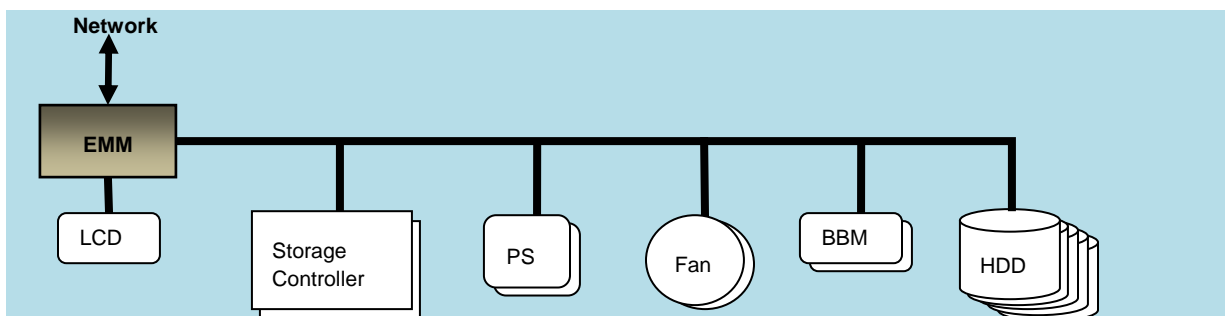
- Disk drives
- Storage controllers
- Enclosure management module
- Cooling modules
- Power supplies
- Battery backup module

## Enclosure Management

Each iQ2880 has an integrated Enclosure Management Module (EMM) that provides communications and environmental monitoring functions to the storage system. An Ethernet interface on the EMM enables remote support which is an integral part of the iQstor customer service and support. On detection of component failure or environmental events which are out of range, Customer Support can be automatically notified to respond to the source of a problem before it potentially becomes critical.

The EMM monitors all critical components and forwards environmental events for logging on the storage controllers. The critical components monitored include power supplies, fan modules, disk drives and battery backup module. AC failures, DC failures, enclosure operating temperature, battery charge status are also continuously monitored. Additionally, the EMM drives a LCD display to provide status of all devices in the enclosure locally. Failure events when they occur are also displayed. The LCD also supports FRU management by displaying system serial numbers and firmware revisions.

In the event of failure, the EMM can be non-disruptively replaced so data access requests from hosts are not affected.



## Non-disruptive Operations

To eliminate the business costs associated with scheduled downtime when performing a firmware upgrade, the storage controller operating environment allows firmware to be upgraded non-disruptively. This feature takes advantage of the multi-path software and failover capabilities of many applications and operation systems. While firmware is being upgraded on a controller, all I/O is directed to its peer controller. When the firmware upgrade is completed, all I/O traffic is directed to the upgraded controller while its peer controller is upgraded and rebooted.

Support for on-line capacity expansion allows new storage space to be made available to applications without disruptive operations on the host. The host operating system is able to detect the new capacity, format and partition it for immediate use. The on-line capacity expansion is supported at all four layers of the storage hierarchy; disk drive, disk array, storage pool and virtual disks (vdisk) layers. New physical drives can be added at anytime. RAID arrays can be created dynamically as new drives are made available. The newly created RAID arrays can be added to the capacity of an existing pool of storage blocks. The capacity of virtual disks can be expanded on demand as long as there are spare blocks in the storage pool.

Other non-disruptive operations supported by the operating environment are:

- Hardware upgrades, such as adding additional disk drives or disk enclosures.
- Configuration changes to redeploy storage resources, e.g. freeing up disk space under use and moving the storage space to new hosts or applications.
- Configuring data services options, such as local mirrors or snapshots dynamically to support zero window backups, application testing, etc. without disrupting access to the primary application volume.

---

## Summary

Data centers, large and small, require storage solutions that provide uncompromised levels of service, backed by quality support. Such a storage solution must include robust architectural design that withstands potential failures without impacting data availability. The reliability, availability and serviceability features of the iQ2880, together with embedded data protection services, make the iQ2880 an ideal choice for critical applications and 24x7 environments that demand continuous access to information.

Compliments of: 



3835R East Thousand Oaks BLVD. #315  
Westlake Village, CA 91365

Tel 877.230.2837 / Fax 805.435.2500 / [www.ess-direct.com](http://www.ess-direct.com)

iQ2880: Reliability, Availability & Serviceability page 9