Datasheet

IRIS FailSafe™ 2.0

Features

- High availability at a low cost
- Supports up to eight node clusters
- Easy administration with Java language-based Cluster Manager
- Dynamic cluster reconfiguration to reduce planned downtime
- Fine-grained failover for minimal disruption to high-availability services

The High-Availability Solution from SGI

IRIS FailSafe 2.0 is the next-generation high-availability solution for your business-critical applications. Using scalable NUMA-based SGI™ Origin™ family servers, IRIS FailSafe 2.0 provides a highly available application platform at a fraction of the cost of specialized fault-tolerant systems. IRIS FailSafe runs in a cluster environment. In the event of a failure, IRIS FailSafe automatically fails over applications from one system in the cluster to the other. In combination with a RAID or mirrored disk configuration, an IRIS FailSafe cluster provides resilience from any single point of failure and acts as insurance against unplanned outages. The applications do not need to be modified in order to realize the high level of availability provided by IRIS FailSafe.

High Availability at a Low Cost

IRIS FailSafe 2.0 uses advanced distributed software technology and standard off-the-shelf hardware to provide high availability at a low cost. In normal operation all systems in a cluster can be active, working as if they were independent servers. In the event of a failure, one of the other systems will take over the services of the failed system, transparently fulfilling requests from clients on the network. IRIS FailSafe supports a mix of SGI servers within one cluster, preserving your investment in your existing computing infrastructure.

Easy Setup and Administration

An IRIS FailSafe cluster is configured and managed using an intuitive Java™ language-based GUI that can be run from any Java language-compliant browser, giving you the utmost flexibility in choosing your cluster management platform. A text-based command line interface is also provided for administration over slow connections and script-based automation.

IRISconsole™ provides a single view for console services on all systems in the cluster, simplifying system administration. You can run the IRIS FailSafe Cluster Manager GUI and IRISconsole from the same workstation.

The Performance Co-Pilot™ monitoring tool helps locate and visualize trouble spots, preventing failures before they affect the system. A special-purpose Performance Co-Pilot agent for IRIS FailSafe provides integrated cluster-wide performance management.

IRIS FailSafe Architecture

In an IRIS FailSafe 2.0 cluster, up to eight servers are connected to both a public and a private network. Clients use the public network to access services from the cluster. The IRIS FailSafe software uses the private

network to monitor the cluster members and exchange control messages. In the event of a server or application failure, one of the other cluster members will assume the public network address of the failed system and respond to client requests.

The distributed and modular IRIS FailSafe 2.0 architecture enables efficient run-time addition and deletion of systems and applications in a cluster, minimizing the need for planned system downtime for cluster reconfiguration or upgrades.

The clustered nodes share storage either on RAIDs or mirrored disks. In the event of a failure, a shared storage subsystem allows multiple servers to assume control of the data; the filesystems are automatically made available on the system(s) where their corresponding applications are resumed.

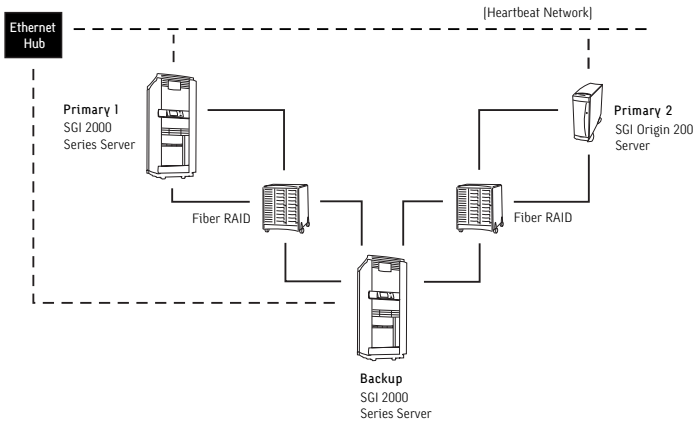
An IRIS FailSafe 2.0 cluster can be configured to best suit the needs of your environment either in an N x 1 configuration where one system is dedicated as backup for N active systems or in an N x N mode where all servers are running business applications while also acting as backup servers. This enables efficient capacity planning and flexible workload distribution to minimize any performance degradation due to failures.

Highly Available Application Agents

IRIS FailSafe 2.0 is flexible enough to meet the needs of almost any application that must be made highly available. SGI provides prepackaged agents for some popular applications and services such as NFS™, Web and e-mail serving, and databases (Oracle® and Informix®) for rapid deployment. These agents can be modified to meet the needs of individual customer environments.

For integrating other custom or third-party applications into the IRIS FailSafe high-availability framework, the highly qualified SGI Professional Services team is available to develop custom agents. Additionally, customers can develop simple scripts to integrate their own crash-tolerant applications with IRIS FailSafe, extending the high-availability functionality to a wide range of additional applications.

Configuration Diagram



IRIS FailSafe 2.0
Technical Specifications

Hardware/Software

- IRIS FailSafe 2.0 is available as an optional product for SGI™ 2000 series, Origin™ 200, Challenge®, Power Challenge™, Onyx®, Onyx2®, and Power Onyx™ systems

Storage I/O Devices

- The shared storage subsystem must be mirrored disk or SGI RAID; third-party RAID is not supported

Highly Available Service Agents

- NFS, Web, Oracle, Informix, DMF, Samba, and other third-party agents are supported

Public Network I/O Devices

- Ethernet, FDDI, HIPPI, and ATM LAN emulation are supported
- IRIS FailSafe 2.0 requires IRIX 6.5.2 or above; mixed clusters of different servers are supported unless noted*

*Challenge S servers are not supported in IRIS FailSafe 2.0 clusters.

Corporate Office
1600 Amphitheatre Pkwy.
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America 1(800) 800-7441
Latin America 1(650) 933-4637
Europe (44) 118.925.75.00
Japan (81) 3.5488.1811
Asia Pacific (65) 771.0290

2778 [7/00]

JI1531