FileNet ISRA *High Availability and Disaster Recovery*

WHITE PAPER
September 2006

FileNet®

Table of Contents

Overview

Application Server clusters are widely used for mission critical business applications. ISRA provides support in farmed and clustered environments with software-based load balancing for improved availability, scalability, and manageability. The purpose of this document is to provide support information to customers and partners interested in deploying and configuring ISRA in farmed and clustered environments. In addition, the document also explains the terms used in this context.

The intended audience for this document includes administrators of various application servers and/or operating systems involved in deployment and configuration of ISRA in farmed and clustered environments for high availability.

Note: The information contained in this document is subject to change.

Scope

This document describes ISRA support when deployed in a farmed and clustered environment with software-based load balancers. The steps to deploy and configure ISRA in a cluster setup are beyond the scope of this document. Please refer to the ***ISRA Deployment Guides*** for such information.

Definitions and Acronyms

Glossary

The following terms are frequently used in context of Farming and Clustering:

Client

An end user application requesting services or data from a server.

Nodes or Managed Servers

Nodes are the entities that constitute a cluster. In a Horizontal Cluster environment, nodes or managed servers are the physical machines that are used for servicing requests. In a Vertical Cluster environment, multiple instances of an application server running on the same physical machine represent nodes or managed servers.

Note: For simplicity, the term Node will be used in this document.

Session Replication

This is the phenomenon of replicating current service state of an application instance across other instances in a cluster.

There are two type of session configuration in clustered environment: Sticky Session and Non Sticky Session.

Sticky Sessions

In this scenario, one server instance communicates with the client application for a session's duration. In case of a server failure, the load balancer diverts all subsequent requests to another node in the cluster. Sticky Sessions provide improved performance in web applications. Sticky Sessions are simple and easy to handle as the load-balancer keeps the client application communicating with one server instance for the duration of the user session. This results in less overhead and better performance.

Non Sticky Sessions

In this scenario, the client application communication is not restricted to one server instance. During a user session, multiple requests can be directed to any of the available servers at any time. In Non Sticky Sessions, the load balancer does not limit the client to one server; therefore, multiple server sessions may be created for a single user session. This results in memory overhead, but ensures improved availability.

Note: ISRA Sample application is not supported with Non Sticky Session configuration. However, custom Web Application can be developed that use Non Sticky Sessions.

Farming and Clustering

What is a Cluster?

A cluster is a collection of independent servers (or nodes) that provide a single, highly available platform for hosting applications. A cluster appears to its clients as a single server. There are two types of configurations in clustering:

Vertical Clustering

Vertical Clustering refers to a configuration where multiple application server instances are running on a single machine. The two main advantages of a vertical cluster setup are:

- Failure of one application server instance running on a vertically clustered machine does not affect the full capabilities of that machine.
- Presence of multiple java processes effectively utilizes all of the resources on a powerful machine as compared to single java process.
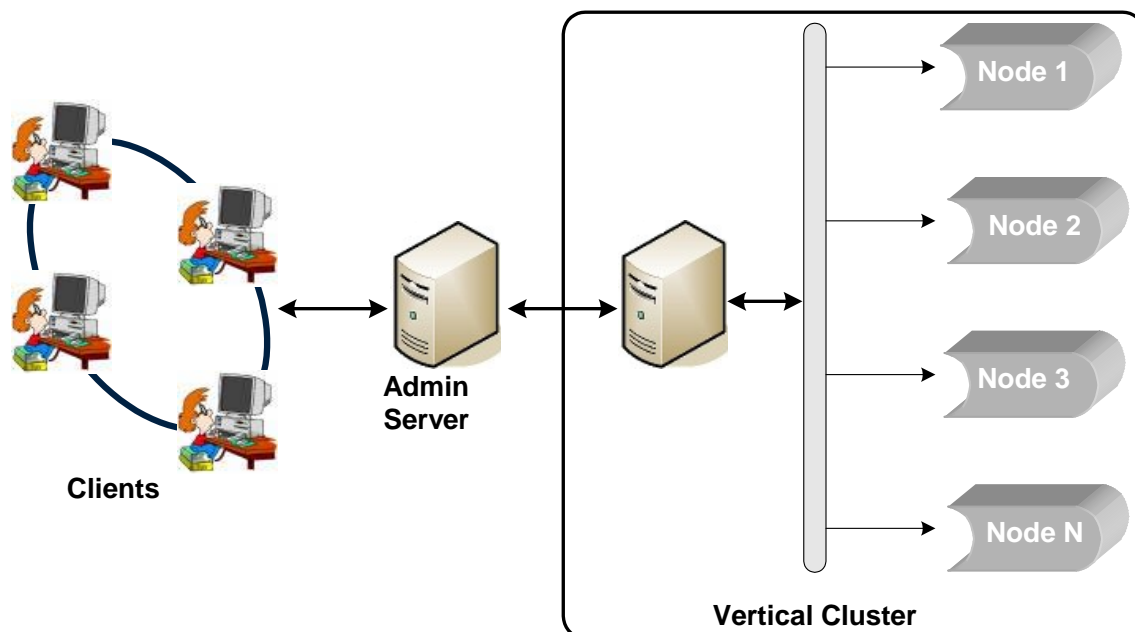
The following figure explains Vertical Clustering:

Figure 1. Vertical Clustering

Horizontal Clustering

Horizontal Clustering refers to a configuration where application server instances (in a cluster) are running across multiple physical machines. The advantage of horizontal cluster is increased performance and better availability as with multiple physical machines, the incoming requests from multiple users can be serviced faster.

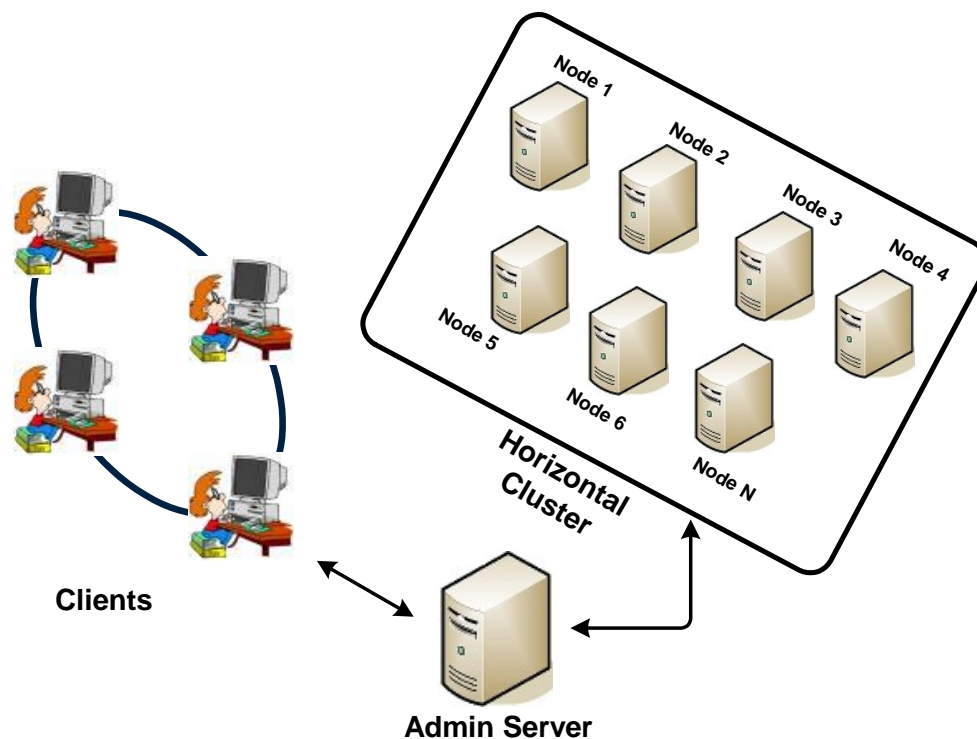The following figure explains Horizontal Clustering:

Figure 2. Horizontal Clustering

Key Capabilities of a Cluster

This section defines the key clustering capabilities that allow application and resources to be highly scalable and available:

Failover

Failover is the capability to automatically switch over from an application component, which becomes unavailable for any reason (failure or abnormal termination) to another application component available in the cluster. Failover is an integral part of mission-critical systems that must be constantly fault tolerant. The goal of a successful failover is to have the request accomplished without any interactions by the clients or administrators.

Load Balancing

Load Balancing is the distribution of load throughout the network so that no network resource is overwhelmed by too many requests. For load balancing to occur:

- There must be multiple copies of an object that can do a particular job.
- Information about the location and operational status of all objects must be available.

A Load Balancer is a logical server (e.g., Apache Web Server on JBoss and BEA WebLogic application server with Plug-in) that handles client requests by deciding which Server Instance or Node will be serving the request. If the load on a particular node increases, the incoming requests are routed to the other available nodes in the cluster. Application Servers provide various algorithms that can be configured for routing requests. In the context of this document, the Load Balancer is also referred to as the Admin Server.

The following figure explains software-based Load Balancing in a cluster setup. The cluster nodes can be on the same as well as different machines.
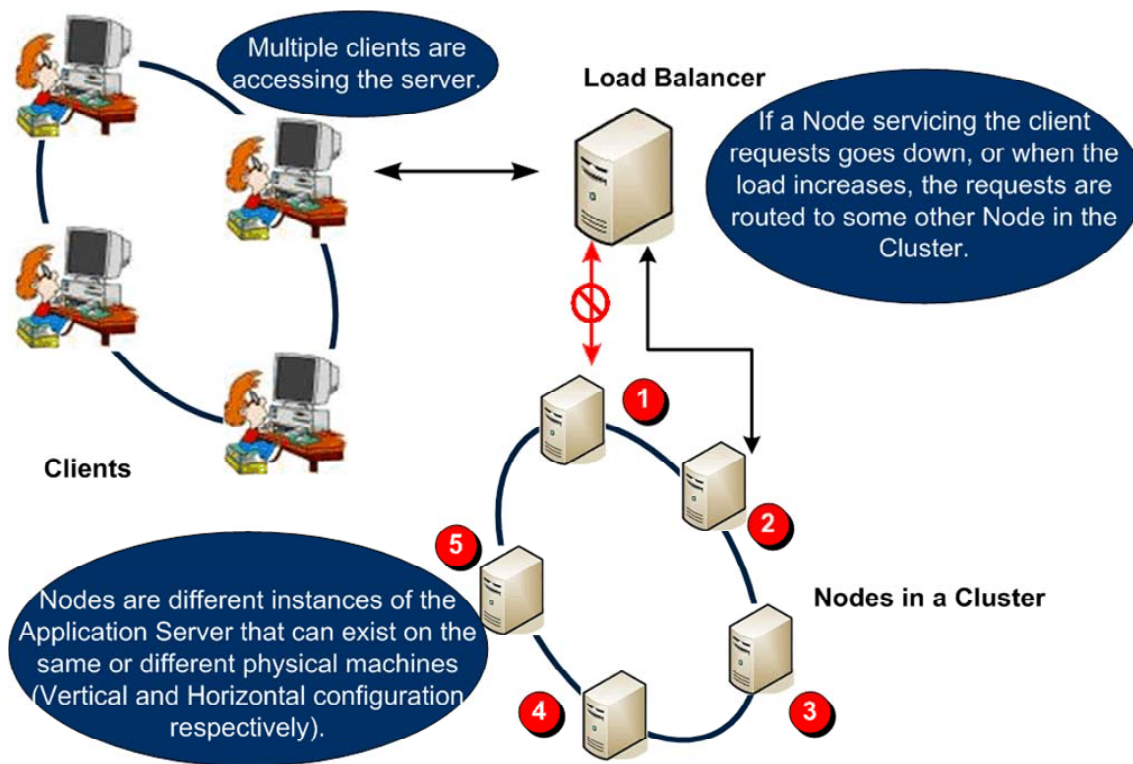
Figure 3. Failover and Load Balancing

Server Clustering provides the following benefits:

Availability	Availability refers to an application's ability to run continuously in the event of planned downtime due to maintenance or unplanned downtime due to failures. A server cluster ensures availability of resources during a failover scenario.
Scalability	Scalability refers to an application's ability to support increasing numbers of users and the load on available resources to handle the requests. Additional servers can be added in a cluster, thus ensuring scalability.

ISRA Support for Farming and Clustering

In a Farm and Cluster environment, ISRA 3.2.1 has already been qualified for WebSphere Application Server on AIX and Solaris platform. With ISRA 3.3.0 release Farming and Clustering (both Horizontal and Vertical with fail-over) will be supported on the following application servers:

- IBM WAS 5.1.1.11, 6.0.2, and 6.1
- BEA WebLogic 8.1 SP6, 9.0, and 9.1
- JBoss 4.0.x

The *ISRA Deployment Guides* provide help and guidance for deployment and configuration of ISRA in a Farm and Cluster environment on supported Application Servers. For detailed information about supported Application Servers and Operating Systems, please visit the FileNet CSS website

References

- J2EE Clustering, Part 1
- J2EE Clustering, Part 2
- Monitoring Session Replication in J2EE Clusters
- High Availability for J2EE Platform-Based Applications
- HttpSession replication in a clustered environment

About FileNet Corporation

FileNet Corporation (NASDAQ: FILE) helps organizations make better decisions by managing the content and processes that drive their business. FileNet's Enterprise Content Management (ECM) solutions allow customers to build and sustain competitive advantage by managing content throughout their organizations, automating and streamlining their business processes, and providing the full spectrum of connectivity needed to simplify their critical and everyday decision making.

FileNet ECM solutions deliver a comprehensive set of capabilities that integrate with existing information systems to provide cost-effective solutions that solve real-world business problems.

Since the Company's founding in 1982, more than 4,000 organizations, including three quarters of the FORTUNE 100, have taken advantage of FileNet solutions for help in managing their mission-critical content and processes.

Headquartered in Costa Mesa, Calif., the Company markets its innovative ECM solutions in more than 90 countries through its own global sales, professional services and support organizations, as well as via its ValueNet® Partner network of resellers, system integrators and application developers.