

Successful Application of NEC's High Availability Cluster Product to the Archives of a Shanghai District NEC Archive Retrieval System High Availability Solution

Project Background

Shanghai's District Archives Bureaus are cultural organizations directly affiliated with district governments, and are in charge of centralized and consolidated management of historical archive materials with preservation value throughout the district, dating from the founding of the People's Republic of China onwards. They are charged with providing usage services to all parts of society.

To help create an information society, this district Archives Bureau launched a project to develop a District-level Archives Directory Center and a Current File Sharing Retrieval System, and establish an Archives Directory Center for the city's archive system. Beginning in July, 2002, 2.1 million data items and 950 thousand full-text files were digitized and put online, making them computer retrievable, with important archives, lifestyle archives, and photographic archives fully digitized and available for online browsing.

The Bureau's archive retrieval system is composed of three main parts - an archive retrieval system, a database service, and a file service - which allow people to conveniently search and retrieve archived contents using a web browser.

The Archives Bureau provides archive retrieval services to the general public, so the retrieval system's stability is a critical concern. On work days, normal services must be available, and the following problems must be avoided:

1. Server hardware failures: single points of failure, especially hard disk failures
2. Difficult, time consuming system maintenance work, such as OS patches
3. Archive retrieval system failures: database abnormalities and slow exception checking that obstruct system usability
4. Long incident recovery time of at least one day that seriously affects normal use of the system

As a consequence of the above, the Archives Bureau decided to use high availability cluster software to protect archive retrieval systems, MS SQL databases, and FTP file services, and solve the problems associated with data security and high availability. Before, they had used file system level cluster software with mediocre high availability performance, long switching time, time difference during data synchronization, weak monitoring functions, and poor failure detection capability. Only later did they introduce the NEC ExpressCluster products and establish integrated shared and mirrored Windows based cluster systems to provide high availability archive retrieval services.

NEC High Availability Cluster Software Product – ExpressCluster X1.0

NEC's new-generation high availability cluster software - ExpressCluster X1.0 is capable of correcting single points of system failure and guaranteeing the continuous and reliable operation of the archive retrieval system. It can detect various software and hardware failures on all of the cluster's servers, and, once the running service or server has failed, can switch the service to another normal server and guarantee uninterrupted operation 24 hours a day, 365 days a year, as well as synchronizing data in real time to guarantee its consistency. (For product details, please visit our website: <http://www.nec-as.com.cn/>)

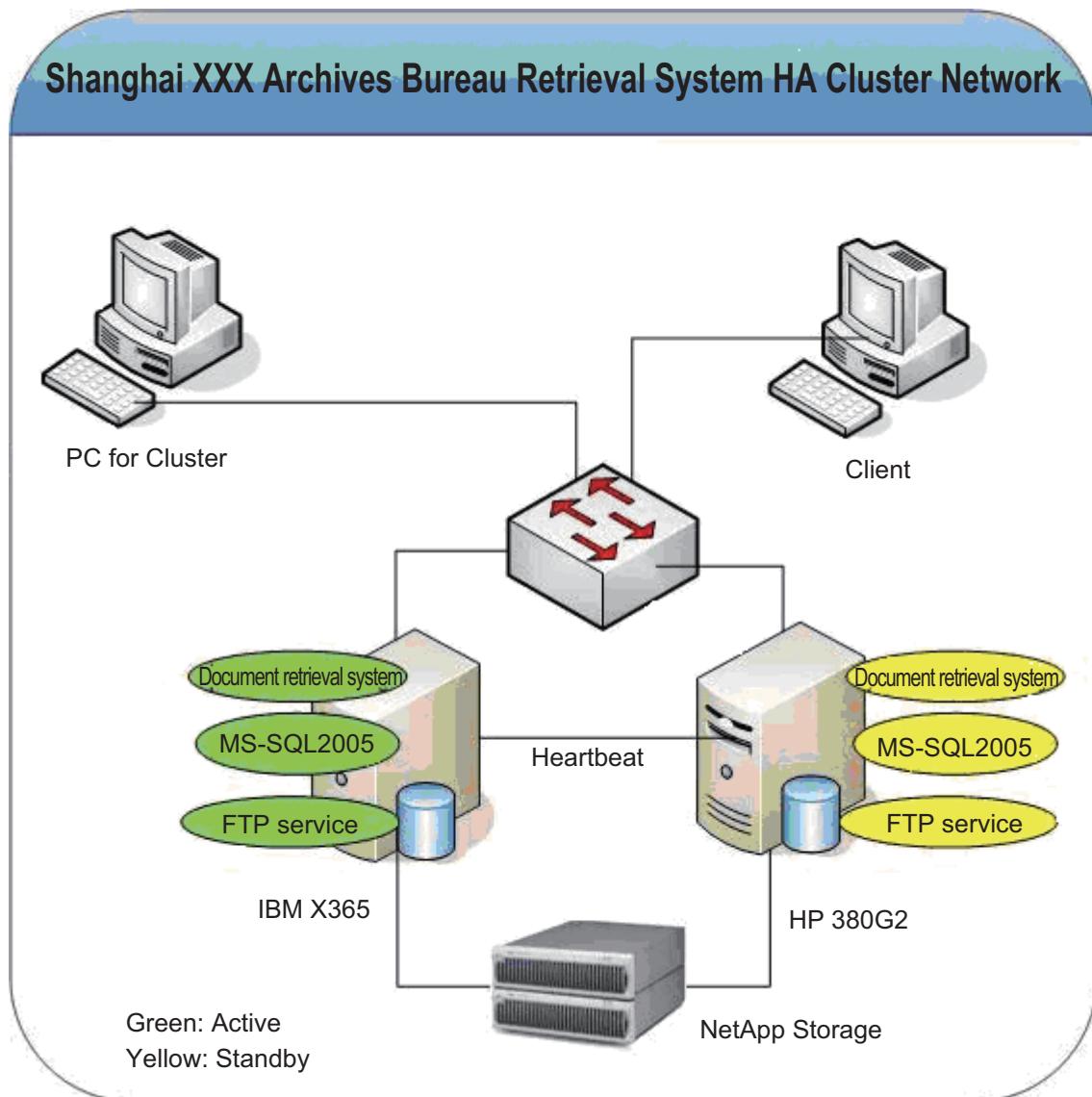
NEC Archive Retrieval System High Availability Solution

Description of System

The system uses use Microsoft Windows Server 2003 as a platform, and uses SCSI disk arrays and cluster software, storing the Archives Bureau's retrieval system data in the disk array and installing local system files and cluster software on the application servers only.

Cluster High Availability Operation Mode

Two servers are installed with the Archives Bureau's retrieval systems, which include Microsoft SQL Server 2005 databases, FTP services, and ExpressCluster cluster software. They are connected to a cable shared disk array via SCSI, so that Microsoft SQL Server 2005, if the services running on one server fail, they are automatically switched to another. The Archives Bureau's retrieval system data is stored on the bare drive of the hard disk provided by the disk array, and the MS SQL databases are stored on the local disks to facilitate the switching between the Archives Bureau's retrieval systems and the FTP services via host switching scripts and to guarantee uninterrupted services and data security of application services. Meanwhile, the database programs and the Archives Bureau's retrieval systems are monitored via Ethernet, and can be quickly switched if there is a failure to ensure the continuous operation of foreground services.

Network Topology Map**System Features**

- Integration of hardware that actually separates data and systems.
- Low hardware configuration requirements so servers can use different or mostly different configurations.
- A short average system switching time of 40 seconds; "best of breed".
- High system efficiency, since read/write, management, and fault tolerance of data in the system are accomplished by disk arrays and failure correction is completed by the cluster software, making the two subsystems relatively independent. The cluster fault tolerance monitor path is a 10/100M adaptive network card line that does not occupy CPU resources of the host or the bandwidth on the basic network. This makes the systems highly efficient, and has won high praise from users of practical applications.

Main Functions

Automatic Detection:

The cluster software provides uni-directional real-time failure monitoring to monitor the service status of the network, storage, servers, OS, applications, the cluster software itself, and the software/hardware resources of standby servers.

The network heartbeat of the cluster software is propagated by the core space of systems without being influenced by system load. Besides supporting the network heartbeat, the system is able to check the server status through COM ports, specific partitions in the storage devices, and the pinging of third parties.

Besides monitoring applications and services at the normal level (i.e. monitoring whether processes are running or dead), the cluster software also provides a series of optional monitoring products including the Database Agent, Internet Server Agent, Application Server Agent, and File Server Agent, which periodically perform practical access processing and detect status indices including the application response time and return results, to detect stalled applications and services and protect key service systems for higher level users. The optional monitoring products are capable of supporting the majority of mainstream applications used with database servers, internet servers, file servers, and application servers.

Users within the scheme can detect hardware errors by connecting to the array cabinet and cluster jointly, and can inter-monitor the running status of two production servers by establishing a private network heartbeat line between cluster servers.

Automatic Switching:

Once a hardware error or server failure is detected, the cluster automatically transfers resources (including virtual IP addresses, MS SQL services, FTP services, and virtual computer names) and data in the shared disk to the standby machine, and keeps the MS SQL services and FTP services running continuously. For the client, the switchover is imperceptible. Users will return to normal operation soon after confirmation of the normal status and the beginning of the failure recovery process.

Easy Maintenance:

The standard GUI user interface is based on a B/S structure using a tree hierarchy to make the cluster architecture clear. The Java based development technology supports remote management and operation to facilitate simultaneous management of clusters (both Windows and Linux) running on a number of different platforms, and allows managers to view detailed management information and management functions on each node of the cluster. When a failure occurs or switching is performed, the system can automatically send a failure notification email to the administrators, simplifying operation and administration.

Simple and Easy-To-Use Cluster Configuration Tools:

Cluster Builder is a simple and easy-to-use cluster management tool that uses tree diagrams to give managers a clear overview of the cluster status. The cluster configuration files need only be read into one of the cluster's servers during the cluster's creation, and other servers can obtain and add information in the cluster automatically. This achieves simultaneous configuration of numerous servers and facilitates smooth importing and convenient operation. Moreover, installation and configuration are separate, so configuration files can be copied, backed up, and edited offline, regardless of the connection status, and administrators can modify cluster configuration information conveniently, and import configuration files to failed servers quickly to speed up recovery.

Advantages:

- Minimal downtime, protecting data and maximizing utilization rate
- Remote management and OS support that reduce the time investment and minimize maintenance costs
- Ability to create online records of Active status
- Calculation of IP address that supports redundant Ethernet interface (for NT version only)
- Users can set up a startup sequence with flexible configuration of clusters, simple editing operations, and convenient cluster construction
- Automatic database startup procedure/engine
- Failure alerts sent by email and acousto-optic alerts available
- No data transmission bottleneck of the background storage for the disk array

Configuration List of Server

- Hardware configuration

Device	Model	Number
Server	IBM X365 and HP3 80G2	3
Network card	Intel 1000PRO	2 per server

- Software configuration

OS	Windows Server 2003: Enterprise x32 Edition
HA software	ExpressCluster X 1.0 for Windows ExpressCluster X Replicator 1.0 for 32-bit Windows
Applications	MS-SQL 2005, 32-bit FTP