

# Enabling Business Continuity for Enterprise Content Infrastructure

*July 2003*

## Contents

Introduction.....	3
1.1 Items to Consider .....	3
1.2 Level of Tolerance .....	3
1.3 Time to Recovery or Failover .....	4
1.4 Scheduled Maintenance .....	4
1.5 Initial and Ongoing Costs .....	4
1.6 Summary of Various Approaches to Business Continuity .....	5
Backup and Restore .....	5
2.1 An Overview of Backup and Restore Solutions .....	5
2.2 Advantages of a Backup and Restore Solution .....	6
2.3 Disadvantages of a Backup and Restore Solution .....	6
2.4 An Overview of the Documentum Solution for Backup and Restore .....	6
Redundancy and Failover .....	8
3.1 An Overview of Redundancy and Failover Solutions .....	8
3.2 Advantages of a Redundancy and Failover Solution .....	9
3.3 Disadvantages of a Redundancy and Failover Solution .....	9
3.4 An Overview of the Documentum Solution for Redundancy and Failover .....	9
High Availability (Clustering) .....	10
4.1 An Overview of High Availability Solutions .....	10
4.2 Advantages of a High Availability Solution .....	10
4.3 Disadvantages of a High Availability Solution .....	11
4.4 An Overview of the Documentum Solution for High Availability .....	11
Summary .....	12

## Introduction

Business continuity and disaster recovery are fundamental to the well being of any organization. The fundamental goal is to ensure the continuity of operation in the case of a system failure, accident or disaster. More than ever, company operations depend on information systems and any disruption may represent millions of dollars in lost productivity and revenue. Today, organizations in every vertical market require their outward and inward facing solutions to be available 24 hours a day, 7 days a week. This requirement combined with the complexity of high availability systems represents a major challenge for many IT organizations.

Web sites, portals, customer relationship management, supply chain management, enterprise resource planning — these are just a few examples of mission critical applications requiring a strategy for business continuity. What all these applications have in common is that they all directly use or leverage unstructured content. Protecting that content to ensure business continuity is a critical requirement, and for that reason business continuity has become one of the key decision criteria when selecting an enterprise content management solution. In this paper, we discuss the business continuity capabilities of the Documentum enterprise content management (ECM) platform, the leading solution for managing enterprise content.

### 1.1 Items to Consider

There are many aspects to consider when determining the level of system availability required for your particular application. Especially important are the needs of your employees, customers, and business partners. While very high availability — commonly referred to as “the five 9s” for 99.999 percent uptime — is often desirable, it comes at a high price. Typically, a trade-off must be made between the business continuity requirements and the costs required to design, build, and manage such system. While some users may conclude that high system availability is mandatory at any cost, business sponsors might often accept some level of risk in an attempt to minimize ongoing costs.

With that in mind, let’s look at a few criteria that can help lead you to a more informed decision relative to these issues:

### 1.2 Level of Tolerance

One obvious question that should be asked is how much uptime does your application really need? How much downtime can you afford and how much are you willing to invest in preventing any downtime beyond that? Anyone who has purchased insurance coverage can relate to this decision process. A content application used to create internal memorandums or system documentation could afford to be offline for a period of time, while a content management system serving dynamic content to an e-commerce server would require and possibly mandate zero downtime. Therefore, consider the impact of downtime in your application in terms of:

- **Productivity loss:** What impact would unplanned downtime have on your productivity and bottom line?
- **Revenue loss:** How much would unplanned downtime impact your revenue?
- **Reputation damage:** Who will notice if your system goes down — your own employees or the *Wall Street Journal*?
- **Inconvenience:** Whom would unplanned downtime upset? How much do you care about that?
- **Safety and environmental cost:** Would unplanned downtime of your system cause any hazards to your employees, customers, or the environment?
- **Legal liabilities:** Could you be sued if your system goes down?

These considerations lead to the first law of business continuity:

First law of Business Continuity

**Deploy only as much high availability as you really need but not less than you need.**

### 1.3 Time to Recovery or Failover

A completely different issue is —What happens if a system fails for any reason? Now consider the time it takes to restore operation in the event of data corruption, system failure, accident, or catastrophe. There are many technical options that guarantee recovery times from a few seconds to hours or days. As with downtime prevention, a system capable of recovering quickly from unplanned downtime will cost significantly more than one that requires a longer recovery time.

Another consideration for business continuity is your strategy for preventing downtime. Having the capability to recover fast may save your business in the case of a serious failure or disaster, but it is preferable to put a system in place that will prevent downtime in any situation. This consideration leads to the second and most important law of business continuity:

Second law of Business Continuity

**Prevention is better than any cure.**

The next issue to consider is whether your recovery system will actually be successful. Studies show, that while a large majority of businesses are performing regular backups, up to 80 percent of them would not be able to recover their data in case of serious trouble. The need to test the recovery procedures is captured in the third law of business continuity:

Third law of Business Continuity

**Only a tested system provides business continuity.**

### 1.4 Scheduled Maintenance

Can your application users tolerate planned system shutdowns? Planned shutdowns or out-of-service times may be easy to handle in a simple environment, but can become exponentially complex with the addition of distributed, or replicated systems. In any case, the maintenance downtimes need to be built into the design, and they need to be well planned and tested.

### 1.5 Initial and Ongoing Costs

Last, but potentially the most important factor to consider is cost. This includes initial cost for design, implementation, hardware and software, as well as ongoing maintenance and support costs. In an environment where enterprise content management is the core technology supporting incoming revenue, cost may not be a large factor when balanced by the lost revenue of a downed application. For these applications, business continuity is a key selection criteria, an absolute requirement for the underlying enterprise content management platform.

### 1.6 Summary of Various Approaches to Business Continuity

The following chart compares various approaches for achieving business continuity — “backup and restore,” “redundancy and failover,” and “high availability” — and their relative advantages and disadvantages. These approaches are discussed in greater detail in the following sections.

	Backup & Restore	Redundancy & Failover	High Availability Solution
Maximum Data Loss	24 hrs.	Minutes	None
Recovery Time	🕒🕒🕒	🕒	Immediate
Hardware Cost	\$	\$\$ - \$\$\$	\$\$\$ - \$\$\$\$
Software Cost	0 - \$\$	\$-\$\$	\$\$ - \$\$\$
Setup & Maintenance Time	☹️	☹️	☹️☹️☹️
Solution Complexity	☹️	☹️	☹️☹️☹️☹️
System Availability	🕒	🕒🕒🕒	🕒🕒🕒🕒
Configuration Flexibility	😊	😊😊😊😊	😊

Figure 1 – A comparison of three solutions for business continuity.

## Backup and Restore

### 2.1 An Overview of Backup and Restore Solutions

The very first, mandatory step towards ensuring business continuity is to implement a hot backup and restore solution. Even if your business tolerates possible disruptions in availability, regular backup of your content has to be a part of your daily procedures. The purpose of the backup is to make a copy of system content, metadata, configuration and application information on media distinct (and preferably separate) from the system. The purpose of a *hot* backup is to make a copy while the application is “live” without affecting functionality or end user access. “Restore” is the ability to *confidently* recover backed up information. It should be noted that any system addressing business continuity must at a minimum provide periodic backups of content and a tested restore solution.

Here are a few characteristics of a backup and restore solution for business continuity:

- Potential information loss between backups
  - Backup intervals can vary from minutes to hours to days.
- Varied recovery time
  - Depending on the solution, restoration can vary from minutes to hours.
- Performance degradation during backup
  - During the backup time, users may experience system performance degradation in the case of improperly designed solutions

Traditional approaches to performing hot backup of a content repository utilize separate phases that backup the metadata in the database and the content assets in the file store separately. The traditional hot backup cannot verify the integrity of objects, as there is no inherent knowledge of what an object is. Restoration using the traditional hot

backup files is an “all or nothing” approach, making it prohibitive for restoring one or several important objects (content).

Traditional approaches for hot backup can be implemented by custom scripts including various modules from popular backup vendors that can backup a database while it’s online. These scripts or programs must be built by the internal IT staff (database and system administrators) or by consultants. And they must then be put through quality testing and release including configuration management. The final solution must be maintained internally as well. All of this work adds to the cost of ownership for the application. A preferable solution would be for an off-the-shelf product that handles the backup of the metadata and content together in a single transaction.

## 2.2 Advantages of a Backup and Restore Solution

Here are some of the advantages to performing hot backups of a content repository.

- Application remains online during backup to service end users.
- Repository data can be restored to a specific state in the event of data loss.
- Can be applied to existing infrastructure, thus leveraging existing investment
- Backup media can be kept offsite

## 2.3 Disadvantages of a Backup and Restore Solution

This section highlights disadvantages to performing hot backups of a Documentum repository.

- Some added cost, both recurring and non-recurring from a conventional cold backup approach
- More complex recovery using mutually exclusive phases to backup metadata and content.
- Some data loss is possible between backup runs.

## 2.4 An Overview of the Documentum Solution for Backup and Restore

To provide a sophisticated backup and restore solution that maintains the integrity between content assets and metadata, Documentum resells a product called HOTBackup from CYA Technologies, Inc. CYA HOTBackup expands Documentum business continuity capabilities, complementing the native capabilities of the Documentum ECM platform, such as redundancy, failover, and clustering. With CYA HOTBackup, Documentum can provide its customers with an enterprise solution for business continuity, disaster recovery, and data integrity management. Designed specifically for the robust and complex business processes and content relationships within the Documentum platform, CYA HOTBackup is a unique solution that allows users to reliably backup their content, metadata and system information with point-in-time and object-level restoration. CYA HOTBackup provides a cost-effective means for Documentum backups while allowing continuous access to business-critical information. And because CYA HOTBackup is tightly integrated with the Documentum platform, the deployment of a hot backup solution for Documentum is easy and effortless.

Key benefits of CYA HOTBackup:

- Enables business continuity by providing an effective data recovery solution
- Reduces planned and unplanned downtime
- Verifies data integrity
- Detects corruption at the application level and provides alerts

- Provides incremental synchronized hot backup
- Maintains object cross-references
- Allows granular restoration at the object level
- Provides point-in-time restoration

CYA HOTBackup provides a safe and reliable solution to full and incremental backups of Documentum repositories. And because it executes while allowing read/write access to the repository, CYA HOTBackup allows users around the world to access a Documentum repository, view content, and even author new content while the backup is being performed (a hot backup). This key functionality provides true 24x7 access to a global repository. With CYA HOTBackup, the Documentum administrator has a single-point solution that can be managed from a workstation on the network. No special hardware is required for CYA HOTBackup. CYA HOTBackup handles the full backup of metadata stored for objects in the relational database without affecting other applications using the database. CYA HOTBackup also handles the backup of content files associated with metadata: the backup of metadata and content are synchronized to minimize any object-to-content inconsistencies. CYA HOTBackup provides the unique capability of a continuous, incremental backup of Documentum repository changes, capturing the modifications of related objects. This functionality provides a Documentum administrator with the following capabilities:

- Secure, reliable, nonproprietary backup format
- Object recovery (for deleted objects, their relationships, and their states)
- Point-in-time recovery (to recover from hardware failure or disaster)

A backup solution is only as good as its ability to recover data reliably and quickly. CYA HOTBackup is the only product that will backup and restore an object, its content (all renditions, media type transformations, or renditions in different languages), and all of its relationships (standard Documentum relationships as well as custom relationships). Using CYA HOTBackup, the Documentum administrator can — from a single GUI-based interface — restore deleted objects and their state or roll the repository (metadata and content) forward to a point in time. Figure 2 depicts a high level architectural diagram of CYA HOTBackup, including the restore component.

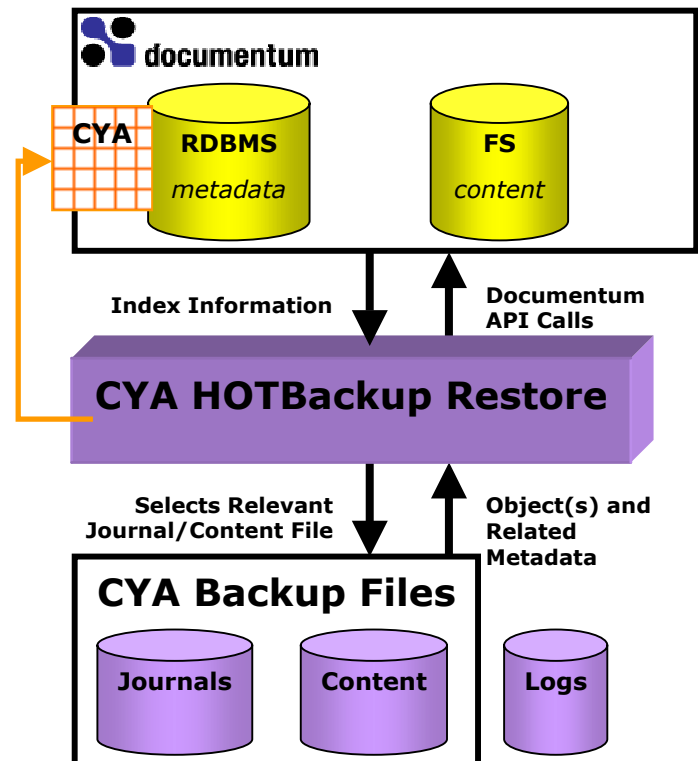


Figure 2 – CYA HOTBackup Architecture

## Redundancy and Failover

### 3.1 An Overview of Redundancy and Failover Solutions

To mitigate the risk of information loss, a company must consider many types of failures, accidents, and disasters — and design solution to help the company protect itself from these risks. Figure 3 shows the different levels of incidents that can disrupt the business operation. The causes can vary from system failures, human errors, and power outages, to natural catastrophes such as floods or tornadoes. To protect against these types of failure or disaster, many companies implement a solution for redundancy and failover.

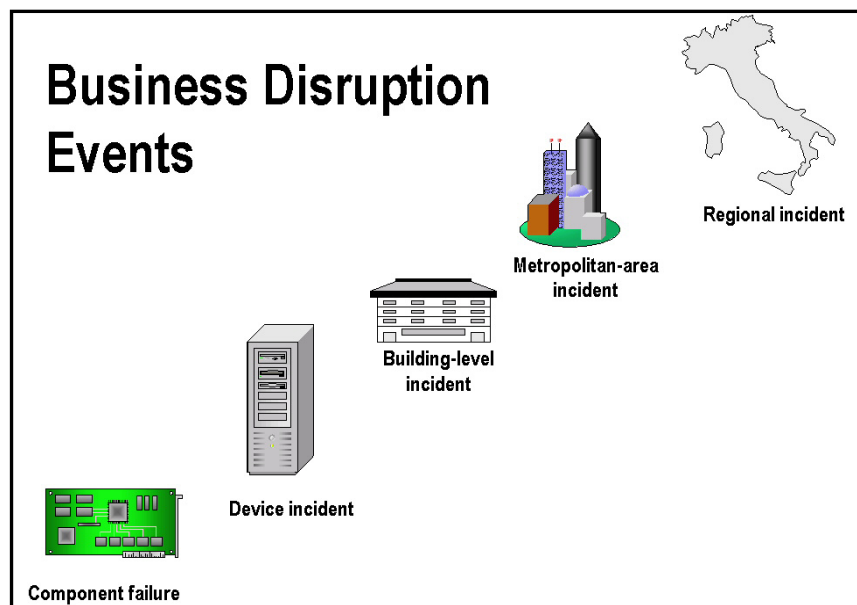


Figure 3 – Business Disruption Events

One of the advantages of data redundancy is the capability to handle any level of incident, from a component failure to a regional disaster. A redundancy solution creates a secondary standby system at a different location — in a different room, a different building, a different city, or even on a separate continent. The standby system contains an identical copy of data from the original system, ready to take over whenever the original system fails to respond, and is designed to handle either the same level of load as the original system, or a level of performance considered acceptable given the level of risk and budgetary constraints. The failover — or data recovery — can be performed manually or automatically.

*Note: Deployment of a redundancy and failover solution does not make backup and restore obsolete or unnecessary. Even a system with built-in redundancy needs to be backed up regularly.*

### 3.2 Advantages of a Redundancy and Failover Solution

These are the advantages of providing a redundant and failover solution for a Documentum repository:

- Recovery time is quick
- Setup and maintenance are easy
- Standby server can be offsite, allowing protection against events such as power failures or local disasters
- Maximum data loss is relatively small
- Backup can be run on a standby server, eliminating any concerns about performance degradation during backup procedure

### 3.3 Disadvantages of a Redundancy and Failover Solution

This section highlights disadvantages of providing a redundant and failover solution for a Documentum repository.

- Requires redundant platform, which means additional hardware and software cost
- Requires changes to existing architecture
- Although relatively small, the maximum potential data loss may still be significant
- Requires procedures for failover (in other words, determining who, how, and under what circumstances someone can “flip the switch” over to a redundant system)

### 3.4 An Overview of the Documentum Solution for Redundancy and Failover

The Documentum solution for redundancy and failover is based on leveraging partner solutions. For the Documentum content repository, integration with the storage infrastructure is fully transparent, enabling Documentum to take advantage of the storage platform’s business continuity capabilities. Documentum partners — including EMC and Network Appliance — provide solutions with very sophisticated storage capabilities for business continuity, including mirroring, snapshotting (point-in-time copies), replication, and clustering. Leveraging the storage system’s capabilities allows corporations to create a completely replicated solution for content and metadata, either within the same data center or at a remote site. The storage system’s replication combined with a standby Documentum Content Server creates a complete standby content management architecture.

Here are examples of some of the partner solutions that Documentum leverages to deploy a redundancy and failover architecture:

- **Documentum partner EMC Corporation** is the world leader in information storage systems, software, networks, and services, providing the information infrastructure for a connected world. Storing valuable information on EMC Symmetrix, CLARiiON, and Centera storage systems allows Documentum customers to consolidate any and all disparate data and platforms quickly and securely. EMC Symmetrix Remote Data Facility (SRDF) and EMC MirrorView enable failover to a remote site for complete business continuity in the case of an outage or disaster. EMC TimeFinder and SnapView provide the capability to make multiple online copies of data for various purposes. EMC data recovery product EMC Data Manager (EDM) allows Documentum customers to perform frequent backups and restores of the data without affecting production system availability. Further, EMC offers a complete set of connectivity and storage management products providing a complete business continuity solution.
- **Documentum partner Network Appliance** is a leader in network data storage and content delivery and provides a complete solution for business continuity. The Snapshot feature of Network Appliance’s innovative Data ONTAP operating system stores up to 31 read-only versions of each volume or file system containing

the database metadata or content assets of a Documentum repository. Snapshot backs up terabytes in seconds and each Snapshot copy takes up minimal storage overhead. The SnapMirror software replicates data volumes at high speeds over the LAN or WAN and the SnapRestore software restores even extremely large volumes in just minutes. NetApp filers typically enjoy greater than 99.99% availability and NetApp Clustered Failover ensures even higher data availability by transferring the data service of an unavailable filer to another filer in the cluster. Network Appliance also provides a very comprehensive set of solutions for business continuity, including the NearStore and SnapLock product lines, offering fast near-line storage for business continuance and backup consolidation. These NetApp solutions complement the capabilities provided by a Documentum standby Content Server and the object-level restoration capabilities of CYA HOTBackup.

In addition to these solutions and complementary to them, CYA HOTBackup works as a filter on top of any existing hardware installed across the enterprise. Utilizing a series of products from CYA including CYA HOTBackup Enterprise Edition, CYA Virtual StandBy, and CYA iCommand, users can achieve true business continuity, allowing them to roll back to a specific point in time, ensure data integrity, maintain a secondary production-ready site at a remote location and administer CYA products on multiple Documentum repositories from a single location. In the event of a disaster, the system would simply redirect users by changing an IP address and kicking off backups and restorations via CYA iCommand on either the failed or existing repositories, thus enabling continued access to critical information.

## High Availability (Clustering)

### 4.1 An Overview of High Availability Solutions

High availability refers to a system or component that is continuously operational for a desirable length of time. Accomplished through “clustering” technology, high availability is a valuable strategy for achieving the goal of business continuity. The principle of high availability solutions is to multiply all components of the architecture through active clusters, in which multiple server and storage components appear to users as a single system. In a server cluster, redundant components are ready to take complete control over transaction requests in the event of failure of one or more system components. Clusters are also used to load balance an application, improving overall system performance. Unlike redundant architecture, all components are actively engaged at all times in a clustered solution and the workload is balanced between all components, which explains why many companies use clustering for performance as much as for business continuity. In fact, performance is a big advantage of high availability solutions. The disadvantage, on the other hand, is their relatively high cost and high complexity.

In a clustering system, two or more servers run the same software components in unison in addition to the clustering software that watches for the availability of each server. An internal or external load balancer distributes the load between the different servers, thus increasing the performance of the system. When one server fails, the clustering software detects the failure and allows the remaining servers take over the load virtually instantly. Although the failure results in a slight loss of performance for the system, users usually don't notice either the failure or the temporary difference in performance.

*Note: Using a high availability solution does not make backup and recovery or redundancy and failover solutions obsolete or unnecessary. Even a high availability solution requires regular backups to protect against data inconsistency, object-level restoration requests, and potential major disaster. Additionally, high availability solutions typically do not permit to keep any significant distance between the clustered servers. Therefore, many corporations chose to combine high availability with a redundant architecture to build an extremely robust architecture.*

### 4.2 Advantages of a High Availability Solution

These are some of the advantages of providing a high availability solution for the Documentum platform:

- Nearly 100% uptime
- Contributes to performance and scalability
- No data loss in a system failure
- Backups can be performed on a live system

#### 4.3 Disadvantages of a High Availability Solution

Here are the disadvantages of providing a high availability solution for the Documentum platform:

- High complexity and high cost
- Effective only when supported by the entire architecture
- Special training requirements
- Requires redundant hardware components

#### 4.4 An Overview of the Documentum Solution for High Availability

While many companies implement a high availability solution for performance reasons, the business continuity benefits of high availability are perhaps more significant. When designing a high availability solution for content applications, companies need to consider all the relevant levels of the infrastructure, including hardware, data, operating systems, network, applications, and more. As with the solutions described above, Documentum leverages partners for the high availability of many solution components. Specifically, the enterprise content management solution from Documentum, requires clustering consideration for the following systems:

- Documentum content repository
- Documentum Content Server
- Presentation and delivery components

Now, let's discuss the details of these components:

**Documentum Repository:** The Documentum content repository consists of two components: a file system storing the physical content assets and relational database for storing content properties and metadata. In both instances, Documentum fully leverages the clustering capabilities of the relational database and underlying operating system. Since the content repository is typically deployed on its own set of servers, Documentum customers can simply leverage the existing capabilities of their database of choice — Oracle, Sybase, IBM DB2, or Microsoft SQL Server. Similarly, Documentum partners such as Network Appliance, EMC, IBM, HP, Sun or Microsoft provide robust clustering capabilities for the Unix and Windows file systems.

**Content Server:** Documentum provides the ability to cluster the Content Server to create a high availability solution. With this approach, the system uses Documentum Docbroker to distribute user requests between a cluster of Content Servers. This task distribution balances the load evenly between servers using a round robin algorithm and hence contributes significantly to the high scalability of the system. Should one or more of the Content Servers ever fail, the remaining servers take over the load and the entire system continues operation without interruption. Documentum uses a persistent transaction technique to validate that any content asset that has been checked out on a failed server is still checked out when another server takes over. Additionally, the self-healing capabilities of Content Server help to repair any damages to content once the failed server is restored.

To protect the operation of Docbroker against a possible incident, Documentum provides the ability to cluster the Docbrokers as well. That way, two or more servers can run a Docbroker cluster and distribute the load to a cluster of Content Servers.

**Presentation and Delivery System:** The components delivering content to users very often require high availability to provide an adequate and seamless experience. Particularly with Web content management, systems are often required to serve millions of user requests on a 24x7 basis. For that reason, Documentum provides a powerful solution for clustering the Web-based content presentation and delivery components. Since the majority of Web applications today are based on an application server infrastructure, Documentum Web delivery components support the major application servers. These servers, including BEA WebLogic, IBM WebSphere, ATG Dynamo, Sun ONE, or Oracle 9i Application Server provide a solid application infrastructure based on the J2EE standard. Similarly, Documentum supports the Microsoft infrastructure, leveraging its COM and .NET environments. What all these application servers provide is a sound clustering architecture for Documentum delivery systems, including Site Caching Services and Site Deployment Services, as well as presentation logic based on Documentum Web Development Kit, including Webtop, Digital Assets Manager, and Web Publisher.

## Summary

As content management becomes increasingly more mission-critical for corporations in all industries, business continuity receives increased attention from the executive management. The task of each corporation's management is to gauge the risks involved with loss of productivity, loss of revenue, and opportunity cost related to system downtime. These risks need to be assessed against the cost of implementing solutions to prevent business disruption from dangers to business continuity such as system failures, human errors, accidents, and Web delivery component disasters. The solutions provided by Documentum and its partners — including backup and recovery, redundancy and failover, and high availability — offer the most comprehensive set of options, meeting the requirements of even the most risk-sensitive companies.

## About Documentum

Documentum provides enterprise content management (ECM) solutions that enable organizations to unite teams, content, and associated business processes. Documentum's integrated set of content, compliance, and collaboration solutions support the way people work, from initial discussion and planning through design, production, marketing, sales, service, and corporate administration. With a single platform, Documentum enables people to collaboratively create, manage, deliver, and archive the content that drives business operations, from documents and discussions to e-mail, Web pages, records, and rich media. The Documentum platform makes it possible for companies to distribute all of this content in multiple languages, across internal and external systems, applications, and user communities. As a result, Documentum customers, which include thousands of the world's most successful organizations, harness corporate knowledge, accelerate time to market, increase customer satisfaction, enhance supply chain efficiencies, and reduce operating costs, improving their overall competitive advantage.

For more information about Documentum, visit [www.documentum.com](http://www.documentum.com) or call **800.607.9546** (outside the U.S.: +1.925.600.6754).

**Documentum, Inc.**  
6801 Koll Center Parkway  
Pleasanton, CA 94566-7047  
phone (925) 600-6800  
fax (925) 600-6850

[www.documentum.com](http://www.documentum.com)

© 2003 Documentum, Inc. All rights reserved. Documentum, and the corporate logo are trademarks or registered trademarks of Documentum, Inc. in the United States and throughout the world. All other company and product names are used for identification purposes only and may be trademarks of their respective owners. Documentum cannot guarantee completion of any future products or product features mentioned in this document, and no reliance should be placed on their availability. Printed in the U.S.A. 60200703V3