

Understanding High Availability

for the BlackBerry Enterprise Server





Understanding High Availability

for the BlackBerry Enterprise Server

Table of Contents

Introduction	2
Users' High Availability Expectations	2
BlackBerry Enterprise Server Overview	3
High Availability Architecture	4
Load Balancing	5
Setting Thresholds	6
Monitoring High Availability	6
Maximizing Configuration Database Availability	7
Planned Service Interruptions	7
Conclusions and Recommendations	8

Introduction

Over the last fifteen years, mobility solutions have evolved from cell phones and pagers to platforms for wireless email and mobilizing business applications. Today's road warriors depend on their mobile devices for fast, reliable, easy access to applications and corporate data. Mobile solution downtime can adversely affect customer service, productivity, sales, and revenues.

"Enterprises of all sizes demand 24/7 application delivery," notes a recent article in InformationWeek. "Server failures, maintenance downtime, and bad weather aren't excuses"¹.

In research commissioned by Neverfail, the application availability company – Osterman Research found that two of three users surveyed said they would be "upset" if e-mail was down for just one hour during the work day. More than four out of ten IT professionals surveyed by Osterman said that downtime decreases productivity by more than 10 percent².

According to Gartner, "most enterprises focus on high availability (or the reduction of unplanned downtime) before that of continuous operations (which focuses on reducing planned downtime), because unplanned outages are more catastrophic, whereas planned outages can be scheduled around periods of low business activity"³.

This white paper introduces high availability in a mobile context, describing user expectations and provides an overview of the BlackBerry® Enterprise Server and a comprehensive explanation of the architecture. Topics such as load balancing, setting failover thresholds, monitoring high availability, maximizing configuration databases and limited down-time upgrades are discussed.

Users' High Availability Expectations

Availability means that people can use their applications and receive results within an acceptable time period. If users can't get their work done on time, the system is down⁴.

Regardless of the size of an organization, high availability requirements can vary depending on the organization's business and mission. A brokerage firm might need high availability during trading hours only, while a seller of gift merchandise requires it 24/7 during the holiday season. A large global enterprise requires automatic uptime in case of failure for mission critical availability to information, systems and corporate data.

By selecting a mobile solution that has high availability built right into the system components, an organization can dramatically reduce the risk of downtime, which can help to maximize employee productivity, enhance customer service, and give both mobile users and the IT administrator greater peace of mind.

1. Marks, Howard, "When Failure Isn't an Option," Information Week, 3/2/09, p. 43

2. Based on Osterman Research, "Planning for Improved Email Availability", July 2007

3. Gartner: How to Access Your IT Service Availability Levels, Donna Scott, April 2009

4. Withers, Stephen, "High availability: Keeping it up", Technology and Business, 7/31/2002

BlackBerry Enterprise Server Overview

The BlackBerry Enterprise Server is highly secure mobile management software that acts as the centralized link between BlackBerry smartphones, enterprise applications, PBX environments through the use of BlackBerry® Mobile Voice System (BlackBerry MVS), and wireless networks.

BlackBerry Enterprise Server integrates with the enterprise messaging and collaboration systems to provide mobile users with highly secure access to email, calendar, voice, instant messaging, browser, enterprise applications, and personal information management tools. It also provides advanced security features and access to administrative tools that simplify management and centralize control of the BlackBerry® Enterprise Solution.

With BlackBerry Enterprise Server v5.0, the key components and services, including the agent, policy service, controller, dispatcher, and router – can be configured to be highly available. (see Figure 1) This results in no single point of failure helping to ensure that end users can always be up and running.

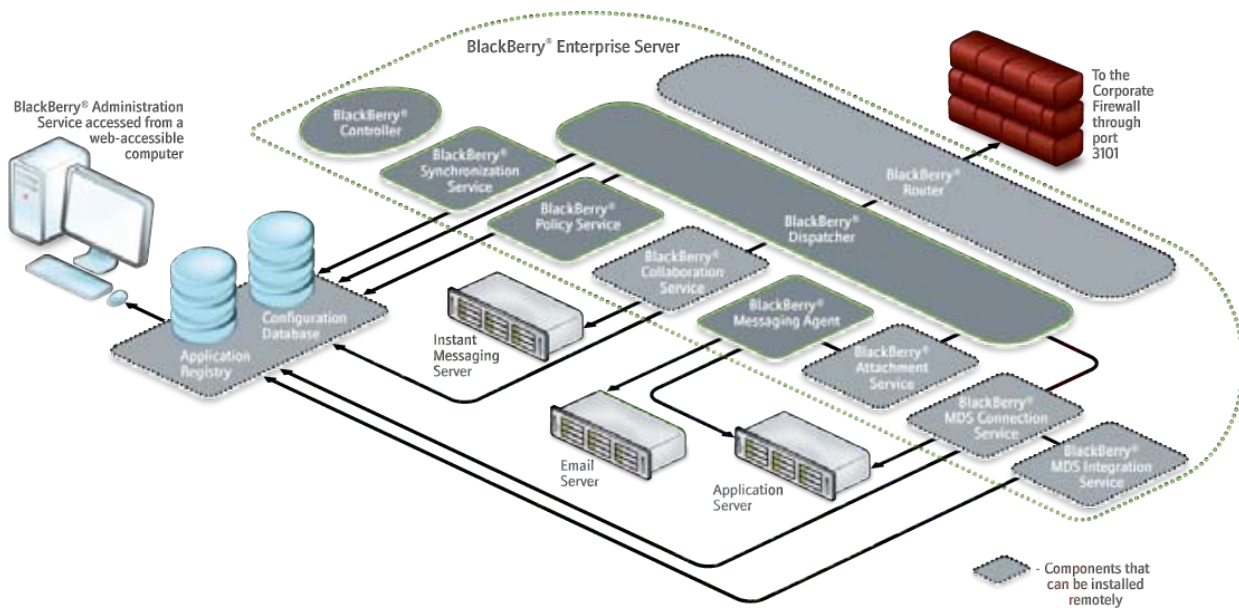


Figure 1. The BlackBerry Enterprise Server contains components and services for managing and monitoring mobile solutions.

High Availability Architecture

The BlackBerry Enterprise Server high availability solution is based on a component level architecture which includes a primary server and a standby server. Either server is capable of running the mobile solution without the other. The solution is designed to support deployment options on both physical and virtual hosts (specifically VMware®), as well as a mix of both.

When the BlackBerry Enterprise Server is configured for high availability, the IT administrator installs a primary BlackBerry Enterprise Server and a standby BlackBerry Enterprise Server on different computers. These BlackBerry Enterprise Server instances create a BlackBerry Enterprise Server pair. Both BlackBerry Enterprise Server instances use the same SRP credentials and BlackBerry Configuration Database. An additional BlackBerry Enterprise Server license is not required for the standby instance. Since both the primary and the standby server share the same SRP, lockouts are avoided.

Both BlackBerry Enterprise Server instances in the BlackBerry Enterprise Server pair include, by default, the BlackBerry Attachment Service, BlackBerry Dispatcher, BlackBerry® Mobile Data System (BlackBerry MDS) Connection Service, BlackBerry Messaging Agent, BlackBerry Policy Service, BlackBerry Router, and BlackBerry Synchronization Service components. By default, if the BlackBerry administrator chooses to install the BlackBerry Collaboration Service on both instances, the BlackBerry Collaboration Service is included in the BlackBerry Enterprise Server pair.

To administer the BlackBerry Enterprise Server pair, you can install the BlackBerry Administration Service with both BlackBerry Enterprise Server instances and configure high availability separately for the BlackBerry Administration Service.

In a large-scale environment, any number of BlackBerry Enterprise Server pairs that use the same BlackBerry Configuration Database can be added.

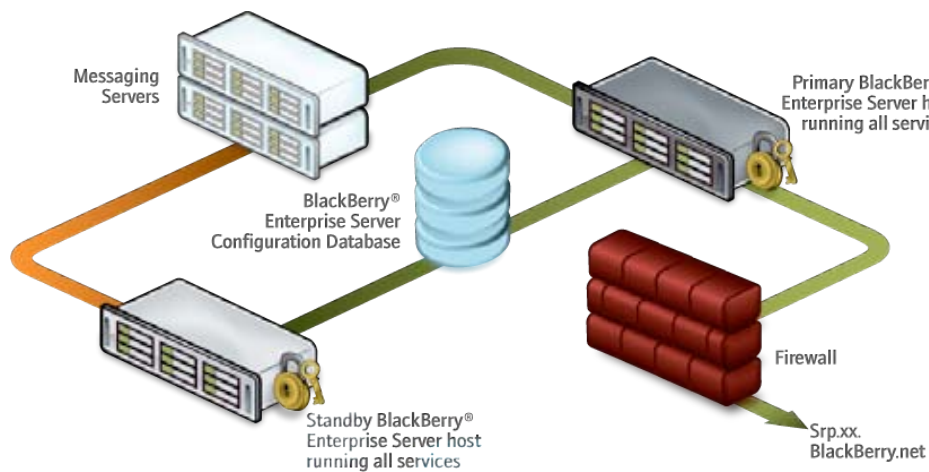


Figure 2. The system automatically fails over from a primary to a standby BlackBerry Enterprise Server.

With the component level architecture, health metrics are continually monitored by the BlackBerry Enterprise Server. BlackBerry administrators can set failover thresholds, which when exceeded, are designed to trigger the BlackBerry Enterprise Server to automatically switch over to the standby server. For example, if the primary server loses its connection to the mail server, automatic failover would occur to the standby server, helping to eliminate the delay of switching over manually. The administrator acknowledges when an automatic failover has occurred, fixes the problem on the originating server, and then manually sets the systems back, ensuring that failover loops are avoided.

Important: Before a standby server attempts to promote itself in an automatic failover, it will perform an out-of-band check to the BlackBerry® Infrastructure to see if the primary server is still connected and operational. The standby server will only promote itself if the BlackBerry Infrastructure reports that the other server is not connected.

How does the high availability architecture for the BlackBerry Enterprise Server compare with clustering? To achieve high availability with server clusters often requires specialized hardware. No special hardware (such as special network cards) or software is required to achieve high availability with the BlackBerry Enterprise Server v5.0. It is built into the server software at the application (software) level.

The BlackBerry administrator can also choose to manually failover the primary server for planned service interruptions.

Load Balancing

The flexibility of the component level architecture allows administrators to distribute components across multiple servers. Customers with a multi-server deployment who distribute BlackBerry Enterprise Server components, such as the BlackBerry Attachment Service or the BlackBerry MDS Connection Services, can configure these distributed components into highly available service pools that support load balancing.

In load balancing, each component is configured with a virtual IP address. When a request is made on this virtual IP address, a network driver intercepts it. The request is then rerouted to one of the distributed components according to rules set by the IT administrator. In this way, the BlackBerry environment can handle increases in traffic by distributing the load among multiple components⁵.

Other distributed BlackBerry Enterprise Server components, including the BlackBerry Attachment Service, BlackBerry MDS Connection Service Pull, and BlackBerry Router – can achieve high availability through the use of shared pools. For example, the administrator can configure a list of routers the BlackBerry Dispatcher can connect to. When the primary router fails, the dispatcher chooses a secondary router to utilize.

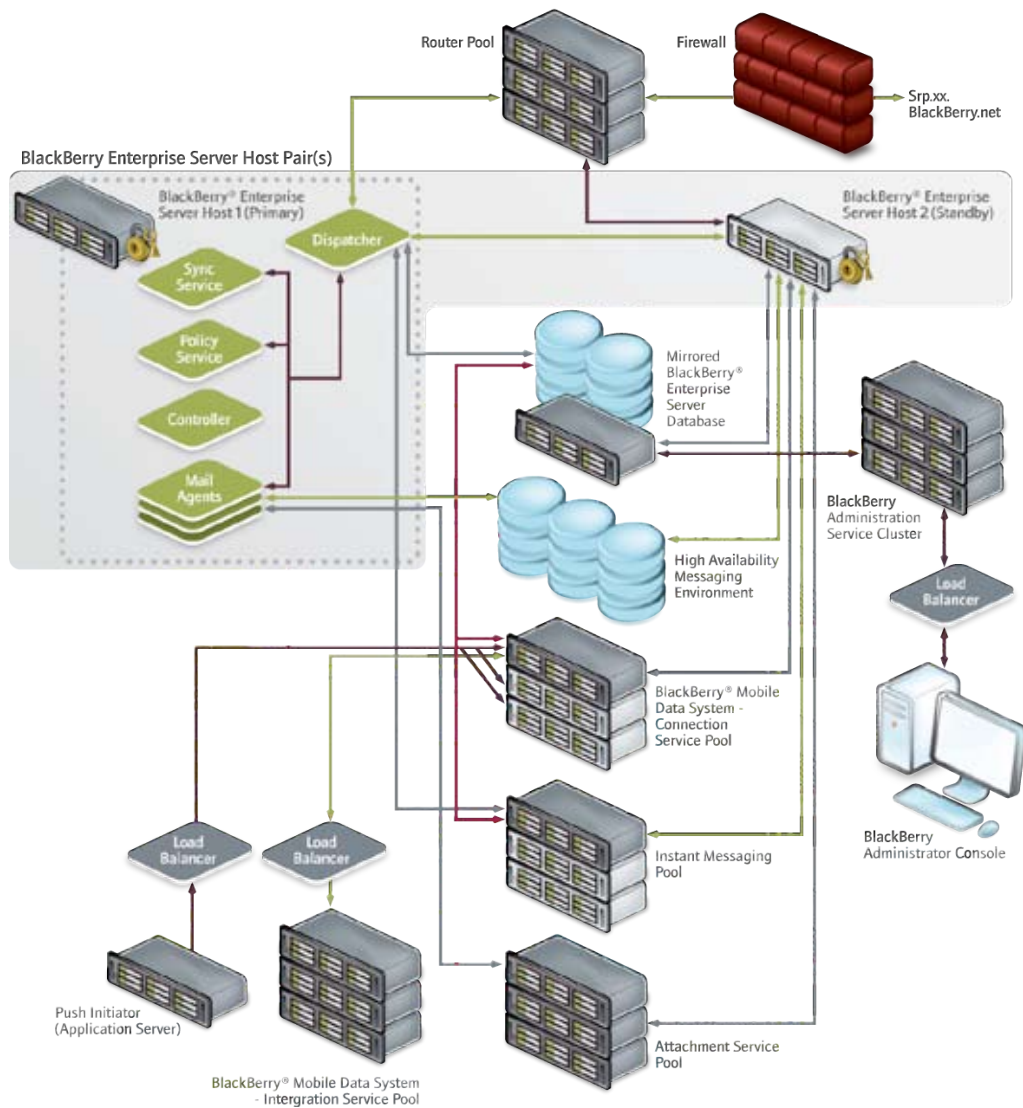


Figure 3. Load balancing maximizes availability of the BlackBerry Attachment Service and other distributed components.

5. Components that can be fronted with a load balancer include BlackBerry MDS Push, BlackBerry Administration Service and BlackBerry MDS-IS

Setting Failover Thresholds

BlackBerry Enterprise Server components calculate a health score that indicates how well the components can provide specific services. The components send their health scores to the BlackBerry Dispatcher, which combines the health scores of the components to calculate the overall health score of the BlackBerry Enterprise Server. The BlackBerry Dispatcher writes the information to the BlackBerry Configuration Database, and it provides the information to a BlackBerry Enterprise Server that requests it.

The BlackBerry Enterprise Server components calculate their health scores by examining their operating health, the stability of their connections to other components, and the health scores of the other components.

The health score for the BlackBerry Enterprise Server consists of various health parameters. Each health parameter indicates whether a particular service or feature is available. The IT Administrator can turn on the automatic failover feature and configure health parameters (i.e. failover and promotional thresholds) so that the BlackBerry Enterprise Server fails over automatically when a particular service or component is functioning below an acceptable level.

Health parameters for which the IT administrator can set thresholds include wireless network access, connection to the messaging server, access to web content, access to enterprise application content, address look-up, calendar synchronization, attachment viewing, and the ability to push application data to BlackBerry smartphones.

By default, the threshold levels are set so that if the primary BlackBerry Enterprise Server loses its connection to the BlackBerry infrastructure or messaging server connection, or cannot browse the Internet, it is designed to fail over to the standby automatically. For this automatic fail-over to take place, the standby server must be connected to the BlackBerry Infrastructure and messaging server (see Figure 1).

Monitoring High-Availability

For the BlackBerry administrator, understanding what is happening within their BlackBerry deployment can be almost as important as knowing whether the system is up. The worst way for IT to be alerted to an issue is to have an irate user call to complain and be caught unaware.

Log files record the activity of BlackBerry Enterprise Server components. These log files are saved on the computer that hosts the BlackBerry Enterprise Server and can be used to troubleshoot issues with component performance.

With the BlackBerry® Monitoring Service, BlackBerry Enterprise Server Alert Tool, or another SNMP monitoring tool, IT administrators can monitor their BlackBerry environment for automatic fail-over events and receive notification when a switch from the primary to standby server takes place.

When the primary and standby server switch roles, both servers write the time and reason in the log files for the BlackBerry Dispatcher, Controller, and Messaging Agent, while the BlackBerry Enterprise Server Alert Tool issues an SNMP alert.

Maximizing Configuration Database Availability

In the high availability configuration (Figure 1), both the primary and standby BlackBerry Enterprise Servers are connected to a BlackBerry Configuration Database. Although this database can reside on any server, it is recommended that for high availability, the Configuration Database is stored on a remote SQL server and not on the same machine as the primary server. In the event of a hardware failure in the primary BlackBerry Enterprise Server, the standby server may be unable to access the Configuration Database and therefore unable to promote itself to the primary server.

SQL 2005 mirroring duplicates the configuration database in real-time, as long as the same version of Microsoft® SQL Server® is installed on both the primary database and the mirrored configuration database. The same permissions must also be configured on both databases.

SQL 2005 mirroring requires the use of a third server instance, known as the witness. This witness is an optional instance of SQL Server that enables the mirror server to recognize whether to initiate an automatic failover. Unlike the two partners, the witness does not serve the database. Supporting automatic failover is the only role of the witness. Should the principle configuration database fail, the BlackBerry Enterprise Server components are designed to automatically connect with the mirrored database.

Planned Service Interruptions

High availability is designed to protect mobile users against both planned and unplanned service interruptions. Unplanned service interruptions are typically triggered by technical problems within the system. Planned service interruptions, such as upgrades, are often necessary for maintaining peak performance and service.

The high availability architecture in the BlackBerry Enterprise Server provides the ability to perform limited-downtime upgrades. While users are still being serviced by the primary server, the standby server can be upgraded to a new version of software. The administrator can then

perform a manual failover to the standby server in order to upgrade the primary server to the new version of software. This flexibility results in minimal impact to the users during server software upgrades. In addition, it provides a back-out strategy where an IT administration can manually fail back to the server running the original version if required.

Conclusions and Recommendations

As mobility has become mission critical for enterprises, an increasing importance has been placed on high availability.

Built into BlackBerry Enterprise Server, the high-availability (Figure 2) is designed to minimize downtime from both planned and unplanned service interruptions, helping to ensure that mobile users have continuous access to e-mail, instant messaging, and other critical applications, systems and data they require to do their jobs effectively.

was built for the enterprise and can provide peace of mind – one that is trusted by some of the highest security conscious organizations in the world. High availability of the BlackBerry Enterprise Server helps maximize mobile worker productivity, improve customer service, and reduce calls to the IT help desk.

Over 28.5 million people use BlackBerry smartphones with approximately 475 carriers and channels in over 165 countries around the world. Wireless technology gives an organization the potential to realize a clear competitive advantage through improved communication, responsiveness and productivity. Make sure to choose a solution that

Additional Resources

For more detailed information on the high availability architecture in BlackBerry Enterprise Server, visit www.blackberry.com/go/serverdocs.

A number of BlackBerry partners now offer third-party solutions that further augment the high system availability of BlackBerry mobile solutions. These include tools that monitor mobile applications, rapidly restore mobile messaging and data services, and proactively find and fix mobile issues.

For more information, visit www.blackberrysolutionscatalog.com.





Check with your service provider for roaming arrangements, service plans and supported features and services.

When you subscribe for, acquire, or use third party products and services with RIM products or services you accept that: 1. It is your sole responsibility to (a) ensure that your airtime service provider will support all of their features; (b) identify and acquire all required intellectual property and other licenses prior to installation or use and to comply with the terms of such licenses; 2. RIM provides such products and services on an "AS IS" basis with no express or implied conditions, endorsements, guarantees, representations or warranties of any kind, and assumes no liability whatsoever with respect to them.

© 2009 Research In Motion Limited. All rights reserved. BlackBerry®, RIM®, Research In Motion®, SureType®, SurePress™ and related trademarks, names and logos are the property of Research In Motion Limited and are registered and/or used in the U.S. and countries around the world. Microsoft and SQL Server are trademarks of Microsoft Corporation. VMware is a trademark of VMware, Inc. All other trademarks are the property of their respective owners. MKT-27884-001