

# NetMotion Mobility XE Scalability

**WHITE PAPER**

## Executive Summary

This white paper serves as a general guide to the ways in which Mobility XE can scale to meet the needs of organizations for growth over time. For more specific information on capacity planning and system performance under specific load conditions, please contact a Mobility XE representative at +1-206-691-5555 or [sales@nmwco.com](mailto:sales@nmwco.com).

NetMotion Mobility™ XE was designed to support both small and large mobile device deployments. The size of the deployment and the load on the network determine whether a single server or a pool of servers should be used. Deployments that involve fewer than 1,500 simultaneously connected clients can use one Mobility XE server under reasonable load conditions; larger deployments can use a pool of up to twelve servers to increase capacity and scalability, and provide redundancy. In addition to the number of users and devices, other factors that directly impact the load, capacity and performance of a Mobility server or server pool include (a) the nature and volume of network traffic between the clients and servers, (b) the number of discrete connections between the client and server (network connections or “sockets”), and (c) the mobile device operating system. Whatever the size of the deployment, the server(s) and clients are centrally configured and managed from a unified, Web-based console.

This whitepaper discusses the options and reasons for increasing server capacity and adding new servers to scale the system.

## Scaling Up: Increasing Server Capacity

Scaling up involves increasing system resources, such as memory, processor speed, or multiple processor cores. In addition to increasing the processing power of the server, increasing a server’s capacity often requires increasing the speed of the underlying network (for example, upgrading the wired network infrastructure from 100 Mbps to 1 Gbps Ethernet). An enterprise may be able to connect up to 1,500 clients simultaneously to a dedicated Mobility server, depending on the nature of the network traffic being secured between the mobile devices and the Mobility server. To protect against server hardware failures, we recommend adding at least one additional server to provide for system redundancy and failover.

The Mobility XE client-to-server ratio depends on how many devices are connected to the Mobility XE server, how many simultaneous network connections (sockets) the applications on those devices open, how much data traffic flows during peak periods, and whether compression, encryption, and/or selective image compression (Web acceleration) are used.

In this paper we assume the following level of traffic between a client and server:

- Domain member machine without mapped drives
- Single line-of-business application — used actively two hours per day
- Light email usage (approximately 10 messages/hour)
- Light Web browsing (approximately two pages/hour)
- HTTP or FTP file transfers — two per day of 200 MB files

It is important to note that the differences between this traffic profile and the traffic profile for any given network will vary, and that the performance and scalability of Mobility XE will vary

accordingly. The total number of clients that can be supported will depend on the number of network connections (sockets) each client opens on average and the aggregate amount and type of data sent and received between the client and the server. As these numbers increase, the maximum number of clients which can be supported goes down.

## Compression

By default, Mobility is configured to turn compression on automatically when the connection speed is 2.5 Mbps or less. This optimizes throughput for clients that roam between WWAN networks and faster WLAN networks.

If Mobility client devices will be running on wireless WANs, using the compression feature improves performance and reduces the amount of data transmitted. If your mobile devices will only connect via wireless LAN, compression can be safely disabled. Each Mobility client device using compression consumes additional memory on the Mobility server for the compression dictionary — by default 75 KB of RAM. The compression dictionary can be configured to use from 9 KB to 300 KB per device (more memory yields better compression and better performance at the cost of overall connection capacity). This additional memory use does reduce the overall capacity of the system.

## Increasing a Server's Capacity and Performance

**Memory/RAM.** The number of Mobility client devices that can be connected simultaneously to a Mobility server depends directly on the amount of RAM on the server. Increasing a server's RAM to 4 GB will allow the greatest number of clients to connect. Currently, adding additional server RAM beyond 4 GB will not improve the server's capacity.

**CPU Speed.** Increasing processor speed will improve the server's performance, especially when encryption, compression, and/or Web acceleration are enabled. Increasing the performance of encryption and compression effectively increases network throughput by allowing more data to be transmitted through the server to and from the Mobility clients. In this paper we assume enterprises will deploy a dual-core 2.1 GHz Xeon processor. Faster processors will yield increases in performance and are a good investment.

**CPU Cores.** As with processor speed, adding additional processors or processor cores will increase the server's network throughput, especially in high-traffic environments with encryption or Web acceleration enabled, because the Mobility server's services support multi-threading. Additional processor cores beyond four are of less benefit because much of the core work is processed on two threads.

**NICs.** You can increase a server's bandwidth capacity by having the fastest possible network adapter installed in the Mobility server and by removing any bottlenecks in the network backbone. In addition, a secondary NIC on a private, non-routable network can be used to handle load balancing traffic communicated between servers in the Mobility server pool. For best performance, we recommend that the Mobility XE server be placed on a gigabit network backbone. A network backbone slower than 100 megabit should not be used as it artificially limits a server's throughput.

## Scaling Out: Using a Server Pool

Scaling out involves adding Mobility servers to a pool. A maximum of twelve servers are supported in a server pool, with two of the servers designated as backup in case failover is required. A pool of this size may support up to 15,000 concurrent device connections (1,500 clients per server for ten servers, plus two servers for failover) with a reasonable client traffic profile. A Mobility warehouse that has been updated with all the appropriate service packs and configuration settings can support up to 65,000 registered devices and users. Primary factors affecting the overall capacity of a Mobility XE server pool include:

- Mobility server and warehouse configuration
- Network configuration
- The number of devices simultaneously accessing the server pool
- The type and amount of data being transferred
- The number of network connections or sockets opened between the client and server

As the volume of devices and/or data increases beyond the capabilities of your current Mobility server, we recommend that you add Mobility servers to the server pool. A pool consists of multiple servers that (by default) automatically balance the client load, fail over as needed, and share a Mobility warehouse that stores configuration settings and client policies (see the Management section below for more on the Mobility warehouse). Adding servers also removes single points of failure, which are unacceptable in large or mission-critical deployments because of the high cost of service disruptions.

### Mobility Warehouse

The Mobility warehouse is a distributed directory server based on the Lightweight Directory Access Protocol (LDAP). It stores system configuration settings and client policies so that all servers in a pool can be centrally configured and managed. The warehouse can be collocated with the Mobility server in small deployments, or it can reside on a dedicated machine for large deployments. The warehouse should be deployed on a dedicated server in deployments with more than 400 mobile devices. A replicated, standby warehouse can also be deployed to avoid a single point of failure.

### Adding Mobility Servers

Expanding a Mobility server pool to increase capacity is as straightforward as installing another server and connecting it to the Mobility warehouse. No new configuration is required of clients or servers: the other servers in the pool immediately learn about the new server member and start sharing load with that server. The next time clients connect to any server in the pool, they also automatically learn about any new servers.

### Load Balancing

A Mobility server pool balances the client connection load among servers in the pool at the time they connect. By default, all servers in a pool participate in load balancing. Each server periodically communicates its load and status to every other load balancing server in the pool. Server load is based on available memory, CPU utilization, and available network bandwidth.

Servers in a Mobility pool use “throttling”, a feature that keeps track of available memory resources for each server in a pool and which prevents a Mobility server from accepting more client connections than it can handle. Throttling is on regardless of the number of clients in the deployment, but generally has an impact only in larger deployments. For example, if there is a flood of users connecting to a server that runs out of memory resources or one that has reached the configured maximum number of connections, that server will stop accepting new connections and automatically pass the load on to other servers in the pool.

## Failover

By default, all servers in a pool function as failover servers. If a Mobility server fails, Mobility clients automatically try to connect to an alternate server in the server pool, allowing clients to easily re-establish network connectivity.

Events that cause a Mobility client to attempt to locate a failover server include:

- Failure to establish a connection to the configured Mobility server at startup
- Failure of the current server to respond (for a client with an active connection to a Mobility server)
- Failover forced by a system administrator so that, for example, system software or hardware can be updated

## Administration, Configuration, Logging and Analytics

All Mobility XE system administration and configuration is done centrally through the Mobility server console. Security, compression and networking options can all be configured and managed from the server. NetMotion Mobility XE uses the following components to help manage your Mobility deployment.

### Mobility Console

The Mobility console is a browser-based interface used for configuring and managing the Mobility XE system. It provides tools for monitoring server status and active connections, managing mobile users and mobile devices, tracking long-term activity, and troubleshooting server and connection problems. In addition to server information, it provides information about mobile device hardware and software, battery status, IP addresses, and other connection-related statistics. Since the Mobility console is Web-based, users with appropriate permissions can configure and manage the Mobility system from a browser or on a remote computer.

### Activity Log

The Mobility activity logs displays historical information, such as connect and disconnect times, when a client roamed and to which network, and reachable status. The logs can be imported to a spreadsheet or other application.

## Event Logs

Each Mobility server keeps an event log with detailed information, including warning, error and debug information (depending on how it is configured.) You can use the Mobility console to control the information that is recorded in the server event log files and the information that is displayed in the console.

## Analytics Module

Beyond activity and event logging, the optional Analytics Module delivers intelligence about the performance of a Mobility XE deployment, guiding decisions about when and how to scale. The Analytics Module adds two additional components to the Mobility XE deployment: a reporting server and a reporting database. In all but the largest deployments, these typically reside on the same physical server separate from the Mobility servers and warehouse.

A copy of Microsoft SQL Server 2005 Express Edition ships with Mobility XE. Microsoft SQL Server 2005 Express Edition will store a maximum of 4 GB of data, under reasonable load conditions – about 13 months of data generated by a single Mobility server supporting 1000 simultaneously connected clients. The Mobility reporting database should be hosted on a platform with a processor core of at least 2 GHz, at least 4 GB of system memory, and at least 15 GB of free space on the hard disk on which the Mobility reporting database data files reside.

Larger deployments (over 1,000 simultaneously connected clients) may benefit from hosting the Mobility reporting server and the Mobility database on an enterprise-grade database server. Microsoft SQL Server 2005 (standard or enterprise edition) and Microsoft SQL Server 2008 (standard or enterprise edition) are both supported. The licensing and optimal hardware requirements for hosting analytics on these platforms are the responsibility of the customer and should take into account the load imposed by all of the database applications sharing the platform.

The minimum server hardware requirements for this type of installation are a dual-core processor (2 GHz or faster), 4 GB of system memory, and 60 GB of free space on the hard drive used by the Mobility database.

For deployments exceeding 6,000 simultaneously connected clients, the reporting database and server should each reside on their own dedicated server hardware with four or more 2 GHz processor cores, at least 8 GB of system memory, and at least 120 GB of free disk space on the hard drive on which the reporting database is installed. To achieve the highest level of database scalability, the disk subsystem for the Mobility reporting database server should use hardware RAID (level 10) and write-back caching must be disabled for the drives on which the Mobility reporting database files are stored. The following elements of the system should each be on a separate physical hard drive:

- SQL Server program files and Windows operating system
- Mobility reporting database data file (*MobilityDB.MDF*)
- Mobility reporting database log file (*MobilityDB.LDF*)
- SQL Server *tempdb* database files

# Mobility Client Software Deployment

There are several methods that can be used to deploy the Mobility client software:

1. Most large mobile device deployments use mobile device-management software. These solutions generally automate updates and upgrades of software packages on the mobile devices.
2. Using Active Directory domain policies, a client can automatically download and install the Mobility client without user intervention.
3. Client deployment can be automated using an .MSI file using Microsoft SMS or other deployment tools that support MSI files.
4. The Policy Management module provides the ability to automatically upgrade clients. See Technote 2194 “Policy Management Example – Upgrading Clients Automatically” on [www.netmotionwireless.com](http://www.netmotionwireless.com) for more information.

In a corporate wireless LAN with DHCP services, no client configuration is required at all: the NetMotion Mobility server’s IP address is obtained by the client via DHCP. Even the server can be used “out of the box” — all the administrator needs to do is set up a User Group to specify who is allowed access to the Mobility server.

## Summary

NetMotion Mobility XE is designed to scale from the smallest test deployment to the largest enterprise. From server capacity and load balancing, to deployment and management, Mobility XE allows organizations to maximize resources and increase mobile-computing capacity in a straightforward, intelligent and informed manner.

© 2009 NetMotion Wireless, Inc. All rights reserved. NetMotion and NetMotion Mobility are registered trademarks of NetMotion Wireless, Inc., and Mobility XE, Roamable IPsec, InterNetwork Roaming, Best-Bandwidth Routing and Analytics Module are trademarks of NetMotion Wireless, Inc. Microsoft, Microsoft Windows, Active Directory, ActiveSync, Internet Explorer, Windows Mobile, Windows Server, Windows XP, SQL Server, Windows XP Tablet PC Edition and Windows Vista are registered trademarks of Microsoft Corporation. All other trademarks, trade names or company names referenced herein are used for identification purposes only and are the property of their respective owners. NetMotion technology is protected by one or more of the following US Patents 6,198,920; 6,418,324; 6,546,425; 6,826,405; 6,981,047; 7,136,645; 7,293,107; 7,574,208; Cdn. Pat. 2,303,987. Other US and foreign patents pending.