

GWAVA Inc.

# GWAVA Retain Worker

For GroupWise®

---

## Netware Install Guide

Version 1.7

## Contents

<b>IMPORTANT NOTES.....</b>	<b>4</b>
Overview .....	6
What Retain does.....	6
How Retain works .....	6
Architecture .....	6
Best Practices for Component Placement .....	7
How Retain Stores the Archives.....	7
Other components that Retain depends on: .....	7
System Requirements .....	8
Novell GroupWise Requirements: .....	8
Server hardware, memory, and storage .....	9
Known Incompatibilities .....	9
Design Considerations .....	10
Retain Worker .....	10
GroupWise Address Book .....	10
GroupWise .....	10
Tomcat .....	11
Installation Preparation .....	11
Preparing Your Server(s) .....	11
Java.....	11
Apache .....	11
Tomcat .....	11
Retain Installation .....	12
Appendix A.....	13
Tomcat Memory Tuning.....	13
Appendix B .....	15
Estimating Storage Requirements .....	15
Appendix C .....	17
Uninstalling Retain .....	17

## Technical Support

If you have a technical support question, please consult the GWAVA Technical Support section of our website at [www.gwava.com](http://www.gwava.com).

## Sales

To contact a Beginfinite sales team member, please e-mail [info@gwava.com](mailto:info@gwava.com) or call Tel: 866-GO-GWAVA (866-464-9282) in North America or +1 514 639 4850.

100 Alexis Nihon Blvd. suite 500  
Montreal, Quebec  
Canada. H4M 2P1

## Intended Audience

This manual is intended for IT administrators in their use of Retain or anyone wanting to learn more about Retain. It includes installation instructions and feature descriptions.

## Organization

The guide is organized into two sections, Overview and Installation. The Overview section covers all the basic information needed to plan where to implement and install the different parts of Retain in your System. The Installation section briefly goes over the necessary steps to install Retain.

General use and management of Retain is covered in the User's Guide, please read that guide for setup and configuration steps and information.

## Copyright Notices

The content of this manual is for informational use only and may change without notice. Beginfinite Inc. assumes no responsibility or liability for any errors or inaccuracies that may appear in this documentation. GroupWise is a registered trademark of Novell, and copyrighted by Novell. Windows is copyrighted by Microsoft. © 2005 GWAVA Inc. All rights reserved. ® GWAVA is a registered trademark.

Retain utilizes a SQL Database to store email information. **At no time should the database system incorporate a NFS mount. NFS locking is insufficient for database requirements and will result in corruption.** GWAVA inc. will not, and cannot support database systems incorporating NFS mounts.

## IMPORTANT NOTES

Retain 1.7 contains several important changes from Retain 1.5, and performs these updates the first time 1.7 is run. To run the update, run Retain 1.7 after the update and login as administrator. Retain will then perform the update. **DO NOT INTERRUPT THE UPDATE OR DATABASE CORRUPTION MAY RESULT.** Wait until the Retain console notifies you that the update has completed, then restart Tomcat.

### Updates

#### Deletion Manager

Retain 1.6 and above include a new Deletion Manager option. This option will not be visible unless the administrator account is granted the Deletion Manager rights in user management.

#### There are two major updates to the system:

##### ➤ **Indexer updated**

Lucene, the indexing engine that Retain uses to catalog and index the mail archive, has been updated from version 2.2 to 2.3. The update changes the index schema, providing performance enhancements. Because of the update, once 1.7 has been installed and run, it is impossible to return to a previous version, as the previous version will view the index as broken. **The upgrade is permanent.**

##### ➤ **Database update**

There is a database update between 1.7 and previous versions. With large databases, the update can take hours. When you first connect to the Retain Server console after the upgrade, you will be prompted to initiate the database update. This is the only screen available until the database update is complete. **DO NOT INTERRUPT THE UPDATE OR DATABASE CORRUPTION MAY RESULT.** If any errors occur, Retain lists them in the Server logs. The Retain Server logs are nested with the embedded Tomcat 5 logs, which differs depending on installation platform.

Linux: /opt/beginfinite/retain/tomcat/logs

Windows: C:\Program Files\Apache Software Foundation\Tomcat 5.5\logs

### **Retain Server moving to 64-bit architecture**

Retain Server, to accommodate the needs of large systems and stability, is moving to 64-bit architecture. 32-bit Server support will end with the Retain 1.x tree and **Retain 2.x will only support 64-bit server operation.** It is suggested that administrators begin planning their systems accordingly. The Retain Worker will still be supported on both 32-bit and 64-bit systems. The list of supported operating systems is as follows:

SuSE Enterprise Linux 9 SP4 32-bit

SuSE Enterprise Linux 10 SP2 32-bit

Open Enterprise Server 1, 2 (Linux) 32-bit versions

SuSE Enterprise Linux 10 SP2 64-bit

Open Enterprise Server 2 64-bit

NetWare 6.5 (Retain Worker only)

Microsoft Windows Server 2000, 2003 32-bit

Microsoft Windows Server 2003 64-bit

#### New Supported Software or features:

Retain also now supports 64-bit JVM and Centera for SLES 10 SP2, OES2, and Windows Server 2003.

Retain also supports NetApp.

Retain now performs LDAP email address lookups. For large eDirectory trees, email addresses will need to be indexed for this feature to work. GroupWise does not index these addresses by default.

Retain also supports auto-creation of users in Livelink.

Retain supports the Hitachi Content Archive Platform, (HCAP).

#### GroupWise 8 Support

Retain 1.7 provides full support for GroupWise 8

#### Stubbing Server

Retain now has an integrated server to take advantage of the stubbing abilities of GroupWise 8. The Stubbing feature in GroupWise 8 allows Retain to replace archived messages and items in the GroupWise Post office that exceed a certain size with smaller 'stubs' that link back to the full message or item in the Retain Archive. The end user experience is not changed. Stubbing is not appropriate for every system. **Once enabled, the stubbing process cannot be reversed. Stubbing requires GroupWise 8.0HP1. Previous versions of GroupWise do not contain critical stability fixes.**

## Overview

Retain is GWAVA's long-term archiving solution and retention system designed specifically for Novell GroupWise 7.0.1 and higher. Retain uses a web-based, centralized management system to schedule and configure archive jobs. Retain can work with GroupWise retention and purge systems and settings and is designed as a cross-platform solution with full auditing built in.

## What Retain does

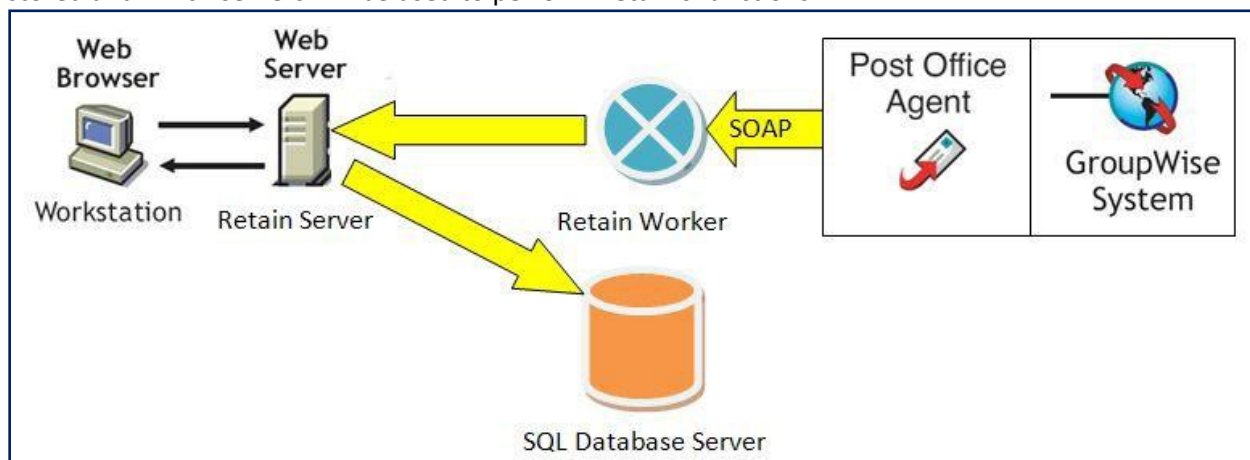
Retain provides a service of long-term storage of data as well as providing search, retrieval, and review services for retained messages. Retain is NOT a backup or emergency restoration system. (For this, we recommend GWAVA Reload.) Retain dredges messages from GroupWise post offices and stores them for long-term reference. Users may log in and review their personal archived mail and search through it.

## How Retain works

The Retain Worker process uses *SOAP* and a trusted application key to collect GroupWise data as defined in a specified profile and according to a defined schedule. This data is handed to the Retain Server which in turn stores the collected data in an administrator-defined location. Users log in to the Retain Server to search through and access their archived messages. User rights are managed by the administrator.

## Architecture

Retain consists of three main parts which can be installed on the same server or they can be spread out across different servers. This gives the administrator enormous flexibility in choosing where data gets stored and which servers will be used to perform Retain's functions.



- **Retain Server:** This is the heart and soul of Retain itself. All functions are controlled from the Retain Server. The server also manages the Retain Worker and stores its data in the database server. Usually there is only one Retain Server per system.
- **Retain Worker:** There is at least one per system, usually installed on each computer running a GroupWise Post Office Agent. (GWPOA) The worker does the actual data collection work. The collected data is handed to the server. The server stores the data in a database.

- **SQL Database Server:** This is an SQL server where Retain stores the collected data. It is not actually part of Retain itself. Retain was designed to support many different databases. The default is MySQL 5.0, but Retain also supports MySQL 5.0, Oracle 10 and 11, MS SQL server 2000, MS SQL server 2005, and PostgreSQL 8.2+

**Installation, maintenance, tuning and backup of the database is the customer's responsibility.**

### **Best Practices for Component Placement**

Retain components communicate via TCP/IP. Placing all the components in the same server will yield the best communication speed. However, this is impractical in a large scale environment. Thus, you must balance the performance of the components on the physical servers with the speed of the network links joining them.

### **How Retain Stores the Archives**

Retain uses a hybrid data storage approach. The database contains all the metadata, folder structure and attachment information but does not hold the actual message text or actual attachments. These are stored on the file system. Data is hashed using an SHA-256 algorithm which can be used to detect tampering.

### **Other components that Retain depends on:**

The following items must be ready before you install Retain's core components.

- NetWare 6.5 SP 5+ or OES 1.0 SP 2
- Novell GroupWise
- Apache Web Server OR Microsoft IIS
- Sun's JDK or JRE
- Apache Tomcat Application Server
- SQL Database for storage (MySQL 5.0, Oracle 10/11, MS SQL 2000/2005, or PostgreSQL 8.2)

## System Requirements

The hardest part about installing Retain is preparing GroupWise and your server(s).

You must first decide *where* you want everything installed. Once the support and core components are installed, using Retain is the same no matter what platform it is installed on. The Retain Worker can be installed locally to the GroupWise POA on Linux, Windows, or NetWare. The Retain Server is only supported on Windows and Linux. You may spread out the components of Retain or consolidate them according to network speed and performance.

### Minimum System Requirements

- Pentium 4 (or higher) at least 1.0GHz
- 64 MB of Ram (for the Worker)
- Tomcat 4.x (Bundled with 6.5 sp1-6), OR Tomcat 5.0x (ONLY with Netware 6.5 SP7)
- Apache 2.0
- MySQL 5.0x (or alternate supported database version)
- Java JVM 1.42 (comes with NetWare)
- As much storage space as system requires. (see [Estimating storage requirements](#))

### Suggested Requirements

- 3 GHz or better Processor (Dual-Core recommended)
- 160MB Ram free for Retain Worker

### Supported NetWare Versions:

- Open Enterprise Server 1
- NetWare 6.5 SP6
- NetWare 6.5 SP7 (Recommended)

### Novell GroupWise Requirements:

- GroupWise 8, 7, (7.01 or higher, **7.03 highly recommended for bug fixes**)
- 7.0.3 Or later is recommended for bug fixes and performance improvements.
- SOAP port must be enabled on all POAs to be scanned for messages.
- SSL is ok, but costs significantly in performance.



## Server hardware, memory, and storage

At a minimum the Retain Worker component must be installed on a system with at least 100MB RAM reserved exclusively for the Retain Worker. It is recommended to leave 128-160MB of RAM for the worker. Please see the section on tuning Tomcat for details.

- ***Tomcat must be adjusted prior to running Retain. Without adjustment, Tomcat can run out of memory, causing it to crash. This will disable all web applications that depend on it. Please see the section on memory tuning for Tomcat.***

CPU should be **at least** a Pentium 4 or higher, **at least** 1.0 GHz. *The faster, the better.*

## Known Incompatibilities

***The following are systems known to be incompatible with Retain. DO NOT USE THESE on the same server as Retain.***

- **Apache 1.x** – Retain will NOT function with Apache 1.x
- **iFolder 3.x** – iFolder 3.x and Retain must NOT be used on the same server. Disruption to the Retain Worker will result.
- **GroupWise WebAccess** – only incompatible if using Microsoft IIS for running Retain
- **Retain cannot access mailboxes with either of the following set**
  - Disabled/Expired account
  - Limited client license

*Novell will be addressing the limited client license issue in GroupWise 7 SP3.*

## Design Considerations

Retain is designed to be as flexible as possible, giving you choices as to where to install its components. Here are some points to keep in mind when deciding where to put everything.

### Retain Worker

- The piece that actually grabs data on a scheduled basis from the POAs is the Retain Worker. It then passes this data to Retain Server.
- Requires a reliable, speedy network connection between Retain Server and Retain Worker.
- Requires a reliable, speedy network connection to Post Office Agents being accessed.
- How many and where? Local to the SOAP POA has its privileges and is recommended. That way you have maximum network throughput and don't need SSL at the SOAP POA. We then compress what we send to the Retain Server. So often one per POA and locally placed is better. If you run more than one POA on the same server, you will encounter a long-standing Novell bug. You will have to manually edit the PO redirection tables. Novell doesn't recommend more than one POA on a server for a variety of reasons. You can have more than one Worker on a server but one worker will handle only one job at a time.
- DO NOT put iFolder 3 on a Retain Worker machine.
- To reiterate, it's recommended to put the Worker on the same server as the POA it dredges for performance reasons.

### GroupWise Address Book

Retain gathers mail from known users. Users known to GroupWise are stored in GroupWise's **System Address Book**. Retain caches this information locally.

Often, the address book needs to be updated as users are added. The Retain Server component NEVER deletes a user from the cached address book.

Over time, Retain will know about all GroupWise users on the system, both current and past.

Because Retain stores past users in its cached address book, it can distinguish between two users of the same name. For example, "John Smith" will be recognized as a different user from "John Smith" who worked at the company six months ago.

### GroupWise

- SOAP MUST be enabled at all Post Office Agents being contacted by Retain Worker.
- SSL must be universally on/off at all SOAP Ports. One or the other, but not a mixed state. We can always change from plaintext to SSL, but not the reverse downgrade. However, it costs in performance. SSL in general is VERY performance intensive. Therefore we recommend leaving SSL off, and having a Retain Worker installed local to each POA in question.
- During the setup wizard, you'll need to generate a **Trusted Application Key** on a Win32 workstation with full rights to the primary domain.
- (optional) Retention and/or Smart Purge flags enabled.
- It's recommended that access by Retain Worker occurs at off hours, to avoid overloading the Post Office Agent.
- Address Book caching is recommended every 24 hours.

## Tomcat

- **Tomcat memory usage must be fine tuned.** This cannot be stressed enough. Default values are insufficient for normal operation of Retain. See [tuning Tomcat section](#), below.
- Support for Tomcat 5.0x is available, but only using NetWare 6.5.7 or above, which installs a newer version of Tomcat. The Tomcat 5.0 available prior to NetWare 6.5.7 will never be supported.

## Installation Preparation

### Preparing Your Server(s)

The general steps to preparing your server are:

- 1) Java (JDK)
- 2) Apache
- 3) Tomcat
- 5) You're ready to install the Retain Worker.

### NetWare

NetWare 6.5 SP5 and future versions. OES 1.0 NetWare SP2 and future versions

### Java

Java 1.42x is included with every NetWare installation. Nothing else needs to be done.

### Apache

Apache Web Server 2.0 is included in NetWare. Nothing else needs to be done.

### Tomcat

Apache Tomcat 4.1x is included in NetWare. The optional "Tomcat 4 Servlet Container" choice needs to have been chosen using the NetWare Installer. Please remember that support for Tomcat 5.0x is available, but only using NetWare 6.5.7 or above, which installs a newer version of Tomcat. The Tomcat 5.0 available prior to NetWare 6.5.7 will never be supported.

## Retain Installation

**\*\*\* NOTE: BE SURE YOUR SERVER'S CLOCK IS ACCURATE BEFORE YOU BEGIN! USING NTP IS RECOMMENDED TO KEEP YOUR CLOCK ACCURATE.** Now that your server is ready, you are ready to install Retain for GroupWise's core components. Retain for GroupWise is installed as follows:

1. Download the Retain for GroupWise package from gwava.com.
2. Unzip the package to a work directory. Almost any name will do, as long as you remember it.
3. Copy all files and folders to the server(s) you want to install to, into another work directory. This will be your "distribution" directory. After installation, you may delete this directory.

Once you've downloaded the Retain Installation package and unzipped it, the files in the working directory should look like this:

readme.txt	RetainServer.zip (core Retain Server)
license.txt	<common>
RetainInstall.sh (Linux install script)	<linux>
RetainInstall.ips (NetWare install script)	<NetWare>
RetainInstall.exe (Windows installer)	<doc>
RetainWorker.zip (core Retain Worker)	<tools>
Setup Tomcat for Windows.exe	<win32>
tomcatinstall.sh	<win64>

4. Go to NWCONFIG at your server console
5. Choose Product Options
6. Choose Install a product not listed
7. Type `volume:path`, where volume, path refer to your "distribution" directory where you've copied the installation files to.

For example, if you copied this entire distribution to `sys:\inst`, type `sys:\inst`. The installer shuts down Java and Apache. So if you're installing on a server using things like iManager or WebAccess, expect these services to go down.

8. You will be prompted for an installation location. The default shown is `SYS:\Retain`.

**The installation for ONLY the Retain Worker is supported in NetWare. The Retain 1.6 Server cannot be installed on NetWare.**

9. The Retain Worker for GroupWise will be installed and Tomcat will be configured.
10. You will be asked if you want to modify **AUTOEXEC.NCF**. You should normally say yes, to guarantee that Apache and Tomcat are loaded on startup.

**You are now ready to go on to the general administration and use guide.**

## Appendix A

### Tomcat Memory Tuning

#### General Recommendations

The following are our general recommendations for the maximum memory allocation parameter. The initial memory allocation parameter may be anywhere from 50%-100% of these values. These parameters are discussed in more detail in the next section.

#### Retain Worker Component

The Retain Worker component(s) have been designed to be light weight. Nonetheless, a minimum of 64 MB RAM is recommended, and 160 MB is strongly recommended, to give room for future functionality. Remember, if this RAM is in use by Tomcat, it is unavailable to other server processes – leave sufficient physical RAM for these. Also, remember this memory is shared with other web applications.

#### Tomcat Configuration Parameters

There are three significant parameters:

##### 1) Memory allocated upon Tomcat startup

- Indicates how much memory is immediately allocated and reserved to Tomcat upon startup. This memory will be in usage for the entire lifetime of Tomcat and never available to the other server processes.
- Typically this is set to 50%-100% of the maximum memory parameter discusses below. It can be lower, but pre-allocating a sizeable percentage of memory *enhances performance* and reduces memory fragmentation.

##### 2) Maximum Memory available to Tomcat

- If the memory usage grows beyond the startup allocation, Tomcat will allocate additional blocks of memory in chunks as needed up to this limit. It will never return this memory to the general server memory pool. (Although memory internally will be freed and reused for Tomcat applications).
- Tomcat guarantees that it will never exceed this memory allocation parameter. If Tomcat runs out of memory, it will try to reclaim unused memory via garbage collection. If this is insufficient, the web application will be denied the memory allocation. Unpredictable (but invariably unpleasant) results will then occur.
- **This is the most critical parameter to tune.**

### 3) Stack Size

- For each thread (which includes each and every concurrent user request), Tomcat will allocate stack space.
- This value is typically measured in KB, and defaults to 512KB.
- Generally we recommend reducing this number to 128k-160k. Increase conservatively, in 64 KB chunks, if you see out-of-stack-space errors in the logs.

#### **Netware/OES NetWare Tomcat tuning:**

- 1) Make a backup copy of **tomcat4.ncf**.
- 2) Edit **tomcat4.ncf**. You'll find these command line switches, that will look something like this:

- -Xmx256m 256MB maximum memory by default.
- -Xss160k 160KB stack size
- -Xms256m 256MB initial memory allocated.

Netware cannot allocate more than 512m of memory without modifying the autoexec.bat.

To allocate more than 512m of memory to Tomcat, place the following switch in the autoexec.bat startup file.

Server -u 3000000000

**Without the Server -u switch, Netware cannot allocate more than 512m of memory to java and Tomcat.**

## Appendix B

### Estimating Storage Requirements

How much storage space you will need in your Retain system is dependent on many factors. While there is no hard-and-fast rule, we have provided the following guide to help you evaluate your own needs. If you have GWAVA Redline, you could use its reports to help you with your estimates. In all cases, you are making an estimate for how much disk space to provide to Retain. When rounding your figures, round up to provide more disk space than you estimate you need. **Calculation guide:** Use the table below to help you calculate your storage requirements for Retain. Do this for each post office you plan to store in Retain and add up the results to get a final figure.

- 1) Get the disk space used by the post office, usually measurable in Gigabytes (GB). If you use GroupWise Document Management and its data is stored within the post office, run a library report using **GWCHECK** and subtract these values. Alternately, get the size of your post office directory and subtract the size of the DMS directory within it.
- 2) Estimate how much your post offices will grow in size over the upcoming years. There are many ways to estimate this. See below for suggestions. Remember that Retain is cumulative.
- 3) Decide how long you want to keep your data (retention time) in years.
1. 4) Calculate your disk space requirements by multiplying your estimated growth per year by the number of years and adding the current size of your post office.

Parameter	Answer	Value #
Post Office Space Used (GB): (not including document management)		1
Yearly Additional Disk Space (GB): (estimate)		2
How long will you keep your data (Years)		3
Disk Space Requirements (GB): Calculate: [value 2] x [value 3] + [value 1]		4
Add 35% for indexes to get a Grand Total (GB) Calculate [value 4] x 1.35		5

Remember, this is only an ESTIMATE, assuming you will be storing EVERYTHING. As you restrict what you will actually store, this figure will go down. For example, do you want to store 50MB attachments? It is your choice.

### **Suggestions on methods to estimate your post office growth per year**

Post office growth varies wildly from company to company and even from post office to post office. There is no hard-and-fast rule which works for everybody. However, some historical data from your own company can help you make a reasonable estimate to determine your ongoing disk space requirements. It is very strongly recommended that the file system you use to store your Retain data be either LVM, NSS or some other system which will allow you to simply add to the storage pool as your needs dictate. Retain tends to need less disk space than your GroupWise post office. It's a small amount less but a conservative estimate would be to simply say "the same amount". Naturally, this only applies if you are storing "everything". Not every organization will store "everything". If, for instance, you're only storing data for certain specific users, you can run GWCheck to determine approximately how much of the post office is used by their data. Redline has a report on post office usage which will tell you right away how much space these users are taking. Take this figure as your baseline "post office size today". Everything else will be calculated based on this figure.  $\text{Current Size} = \text{Past Size} + \text{Additions} - \text{Deletions}$  Retain Needs = Past Size + Additions So the real challenge is making a reasonable estimate of what was or will be deleted. Calculate these figures for each post office you plan to store in your Retain archive and add the results together to get a final total. Suggested Method

- 1) Take your post office size today. (Present Size)
- 2) Take the size of the post office at some point in the past – as close to a year as you can get. (Past Size)
- 3) Calculate the size difference between them (Size Increase) and the time difference (Time Difference)
- 4) Extrapolate the size difference over the year. You'll get an estimate of NET growth. (mail was deleted too).
- 5) Estimate how much mail was deleted from the system. This will be a wild guess at best but only you can make such an estimate.
- 6) Your annual growth will be Net Growth + Deletions.

Even if you can measure your actual growth from last year to this year, it is a NET growth. In addition, rates of growth tend to go up as bandwidth increases because people are simply sending larger and larger items. Some people store things in GroupWise like it was a filing cabinet. So it's a good idea to overestimate your needs and plan accordingly.



## Appendix C

### Uninstalling Retain

Should you wish to uninstall Retain, use the following general procedure. Actual steps will vary according to your host Operating System.

1. Shut Down Tomcat
2. Shut Down Apache
3. Remove the data storage path (delete the files)
  - a. Can be: sys:retain, \retain, /var/opt/beginfinite/retain ... etc.
  - b. See the Retain Administration Guide section on storage to see where your data actually is placed.
4. Remove the program files
  - a. Stored under storage path
5. Remove the RetainWorker.xml and RetainServer.xml from Tomcat.
  - a. Netware Tomcat4 sys:\tomcat\4\webapps
6. Remove the directories under tomcat/work (but NOT work itself)
  - a. Linux: `rm -rf /usr/share/tomcat5/work`
7. Remove references to Retain for GroupWise in httpd.conf
  - a. Can be: /etc/apache2, sys:\apache2/conf, or <InstallPath>\conf).
8. (optional) remove MySQL, Apache, Tomcat
9. If Apache and Tomcat were not removed, restart them both to verify that they load without problems.