# Managing Computer Viruses in a Groupware Environment

By Gregory Tetrault
Sybari Software Inc.
January 17, 1997



Sybari
Software Incorporated

# Introduction

Groupware seems finally destined to become an indispensable computing technology by providing scaleable communication, collaboration and coordination services to people. Much like the word processor and spreadsheet, groupware is becoming a staple of organizations big and small due to the productivity gains it consistently demonstrates. However, the enabling technologies behind groupware are creating a serious computer virus problem for organizations that have come to rely on it.

The purpose of this paper is to both illuminate the perils of computer viruses in groupware and examine the strategies to defend against this security threat. We will show how groupware has exacerbated the current virus problem, especially macro viruses, and provided the necessary elements to spawn entirely new and more devastating types of viruses that are native to groupware systems. Finally, we will conclude with a set of strategies to combat computer viruses in groupware environments and specifics on implementing these strategies.

## 1.1 Groupware Basics

Before entering into the details of managing computer viruses in a groupware environment, we will first cover the essential capabilities and technologies of groupware. This paper assumes the reader is familiar with managing viruses in traditional environments but not necessarily a groupware expert. However, a general awareness of groupware products is assumed.

### 1.1.1 Groupware Capabilities

The essential capabilities can be summarized into what is known as the three 'C's: Communication, Collaboration, and Coordination. Applications such as discussion threads, tracking, document sharing, calendaring, various approval processes, etc. are all well suited to groupware. In addition, groupware embraces the remote-computing concept by allowing geographically dispersed teams to work effective together.

### 1.1.2 Groupware Technologies

Groupware is made possible through the combination of advanced technologies pulled from many areas. Some of the key enabling technologies incorporated into groupware includes:

| | |
|---|---|
| Messaging | Messaging a core service that all groupware products must have. Workflow applications demand this facility. |
| Document Repository | Document oriented storage sub-system designed to handle the unstructured data and rich media types associated with groupware documents. |
| Document Replication | Replication overcomes geographically dispersed team |

| | |
|---|---|
| | collaboration by creating exact copies of documents on multiple servers that are located at each user's worksite. |
| Remote Access | Successful integration of mobile computing demands this technology be available. |
| Digital Signatures | Authenticates the originator of a document or message. |
| Strong Encryption | Provides the necessary privacy required by sensitive data. |
| Workflow Agents | Allows agents (macros) to be embedded in documents and messages for the purpose of executing workflow steps. |
| Macro Languages | Powerful macro languages are provided for creating applications and complex workflows within the groupware environment. |
| Internet Integration | Integrated internet protocol support for HTTP, FTP, NNTP, GOPHER, SMTP and POP3… and now Java. |

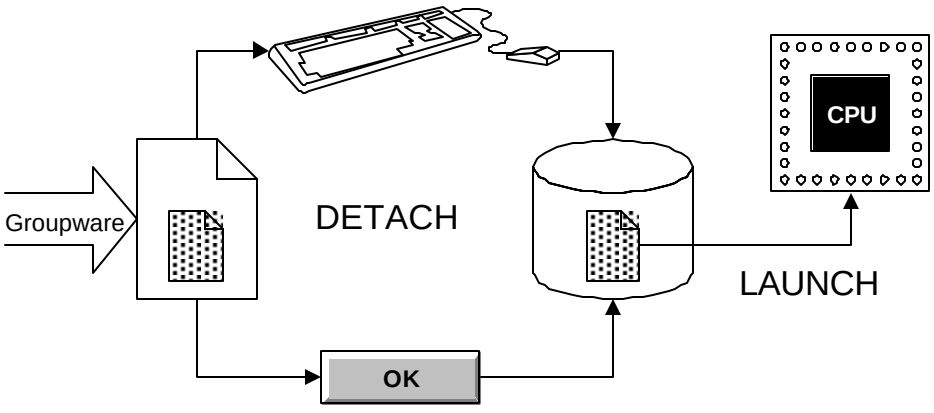
The principal groupware products available today include Lotus Notes, Microsoft Exchange, Novell Groupwise, and Netscape Collabra. Currently only Lotus Notes provides all of the technologies described, however the gap is closely quickly. We should expect that all above mentioned products will implement all of the technologies listed in the next 6 to 12 months.

# Groupware Perils

History has proven that the number and severity of security incidents increases dramatically for a computing environment that proliferates rapidly. Examples include DOS, Windows, the Internet, and Microsoft Office. Groupware implementations such as Lotus Notes are well on their way to achieving similar status. Although numerous security related issues exist, we will focus strictly on computer viruses within a groupware environment. The virus threat for groupware is essentially two-fold. First, groupware messages and documents can contain one or more file attachments that may be infected with well-known file viruses or macro viruses. Second, the technologies provided by groupware such as messaging, replication, workflow agents, and powerful macro languages can be combined to create devastating viruses native to the groupware environment.

## 2.1 File Attachments

Groupware supports embedded file attachments in both messages and documents. The file attachment feature allows users to send binary data and executable files to each other by attaching them to a mail message or document. The known risk is that an attached file may be infected with a platform specific virus. However, to activate the virus, the user needs to detach the file and then open/run it. The exception is if a Trojan horse detaches and executes the attachment for you. It is very common in the Lotus Notes environment to include button macros that perform file detaching and execution in a message. With very little effort, this technique can be converted into a virus dropper that might even include logic to disable the local anti-virus scanner.




A bigger problem however is the speed with which replication and messaging can spread an infected file attachment throughout a groupware environment. The spread

characteristics can potentially expose a large number of users to a single virus in a very short amount of time.

## 2.2 Native Macro Trojans and Viruses

Groupware also introduces a number of technologies that make the specter of native groupware viruses very real. The combination of workflow agents with powerful macro languages is an ideal environment for supporting viruses. In fact it is actually easier to create a simple virus in Lotus Notes than in Microsoft Word. Thus far, virus activity has been restricted to simple trojans being created as a result of user experimentation and workflow programming errors. To date, there has not yet been a reported native groupware virus in the wild.

One aspect of a native groupware virus is that it can spread extremely fast. By activating when the document or message is read, a native virus can quite easily mail itself to a random set of valid recipients and/or copy itself to new databases. The following table illustrates a fairly conservative model of a 1000 user organization that checks there mail on average twice a day.

| | | Total Infections | Infected Mail Sent | New Target Mail Sent | New Infections | Percent Infected |
|---|---|---|---|---|---|---|
| Day 1 | 1:00 PM | 1 | 10 | 10 | 10 | 1% |
| Day 1 | 5:00 PM | 11 | 100 | 100 | 93 | 10% |
| Day 2 | 1:00 PM | 104 | 1000 | 930 | 359 | 46% |
| Day 2 | 5:00 PM | 463 | 9300 | 3590 | 70 | 53% |
| Day 3 | 1:00 PM | 533 | 35900 | 700 | 38 | 57% |
| Day 3 | 5:00 PM | 571 | 7000 | 380 | 39 | 61% |
| Day 4 | 1:00 PM | 610 | 3800 | 390 | 30 | 64% |
| Day 4 | 5:00 PM | 640 | 3900 | 300 | 23 | 66% |
| Day 5 | 1:00 PM | 663 | 3000 | 230 | 17 | 68% |
| Day 5 | 5:00 PM | 680 | 2300 | 170 | 13 | 69% |

As one can see, almost half of the organization is infected in 36 hours. If the organization is connected to external groupware systems, the virus will very likely infect those systems as well. In addition, a mail storm occurs at the height of the infection, thus disabling the groupware environment itself.

# Anti-Virus Strategies

*"Do not use a hatchet to remove a fly from your friend's forehead" – Chinese proverb*

The simplest protection from groupware threats is to simply not use groupware. Of course if this rationale prevailed, we would have reverted back to the slide-rule years ago as an anti-virus strategy. Instead, we must move to design and implement sound anti-virus strategies as groupware environments expand and proliferate. To adequately address the complex task of protecting groupware environments, we will divide this discussion into four areas:

- General principles that will influence the various strategies discussed.
- Historically successful anti-virus strategies that should be avoided.
- Strategies geared toward electronic mail and routed workflow applications.
- Strategies geared toward document repositories and replication.

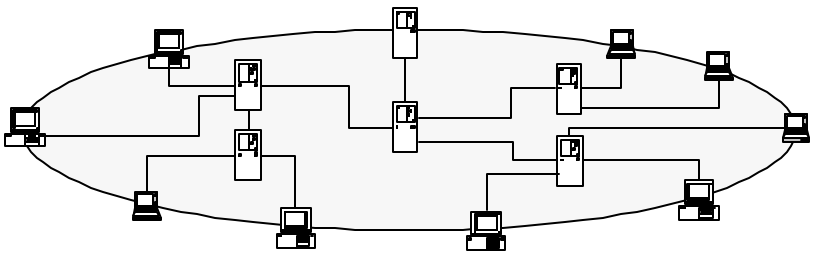
The strategies that will be discussed in this section are intended to be used as part of an overall anti-virus policy and implementation. Specific implementation details for the strategies described are found in section four.

## 3.1 Anti-virus Principles

Groupware is an inherently complex environment that combines client/server technology, mobile computing, heterogeneous networking, electronic messaging, data replication, cryptography, and a host of other technologies. To create workable anti-virus strategies that can be understood and implemented demands some basic principles be defined and followed. Two principles are central to creating effective anti-virus strategies for groupware environments. The first is perimeter protection of the environment and the second is incident containment within the environment.

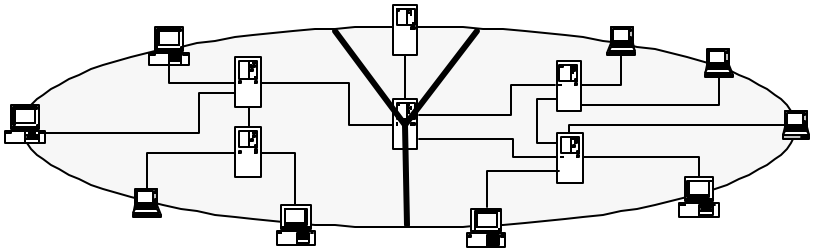
### 3.1.1 Perimeter Protection

The essence of perimeter protection is to simply never allow viruses to enter the groupware environment. This requires that every entry point into the system be assessed and suitably protected using anti-virus tools. For groupware, entry points generally include all clients and all external gateways.



### 3.1.2 Incident Containment

The essence of incident containment is to minimize the affected area when perimeter protection fails. If perimeter protection is guaranteed, incident containment is not required at all. However, few if any entry points into a system can be 100% protected at all times. Incident containment requires then that every path a virus can take through the system be mapped. For groupware, virus distribution paths generally follow from the perimeter entry points (clients and gateways) to the defined mail routing and data replication topologies that connect servers to servers and to other entry points defined in the perimeter. Therefore incident containment measures will be primarily focused on the groupware server infrastructure.



For example, suppose anti-virus tools are deployed on just the central mail routing server. As the diagram above shows, virus incidents would be contained to one half of the environment while the other half would remain virus free. If perimeter protection is in place, then incident containment acts as a barrier to the spread of a virus in the event that the perimeter protection of the environment fails. As a rule, incident containment is much cheaper to deploy since it involves far fewer systems and does not require any end-user interaction.

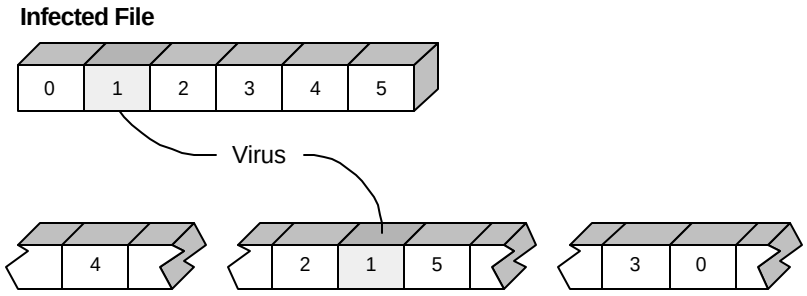
## 3.2 Strategies to Avoid

Equally important to describing effective strategies for protecting groupware environments is to explain anti-virus strategies that should be avoided. This is especially significant since the strategies that should be avoided are in fact quite effective for other types of environments.

### 3.2.1 File-based Scanning

The vast majority of anti-virus tools available today are oriented to file-based scanning. This is reasonable given that the vast majority of viruses infects files and/or boot sectors. However groupware provides a haven where file infecting computer viruses can exist undetected by these same tools. The principle reason for this is that groupware implements a proprietary storage sub-system within a file or files where all messages, documents, and associated file attachments are maintained. Because the organization of a file attachment within the groupware storage is utterly different than if it were stored normally on the file system, virus scanners have little chance of detecting viruses and no chance of cleaning viruses that infect file attachments. To illustrate this, assume an infected file is attached to a groupware document and saved

to the groupware storage sub-system. The virus would now exist within the groupware storage sub-system file, as shown below.




**Infected File within Groupware Storage**

When the groupware storage is now scanned for computer viruses at the file level, one of the following situations will occur.

- If the scanner has been configured to scan only executable files and Word documents, the groupware file is skipped.
- If the scanner contains heuristics that search only particular locations within the file, the scanner will usually miss the infected portion of the file.
- If the scanner is configured to scan the entire file, the virus will be detected assuming the underlying groupware storage stored the virus contiguously. At this point, any attempt to clean the virus will usually corrupt the internal storage structure of the groupware file and result in partial or complete data loss.
- If the file was compressed and/or encrypted when it was attached, then the virus will be impossible to detect at the file level.

File-based scanning, therefore, provides little chance of virus detection and a strong chance of data loss in the event the a virus is found and cleaned.

Thus far, the discussion of file-based scanning of groupware storage has been focused on the "on-demand" variety of scanner. The "on-access" variety of file-based scanner does in fact afford a certain amount of protection against file-based viruses embedded in groupware storage. Since the virus within groupware is dormant until detached to a file, scanning every file as it is detached will succeed in protecting the user against infection of the local file system. The "on-access" scanner is capable of scanning and cleaning files as they are detached but it does NOT clean, remove, or disable the virus as it exists within groupware storage. This restriction limits the overall effectiveness of "on-access" scanners within the groupware environment.

There are those who would argue that "on-access" scanning is sufficient in that it does prevent infection of the client machine by file-based viruses within groupware. This position ignores the inherent risks of maintaining active viruses within the overall-computing environment. For instance:

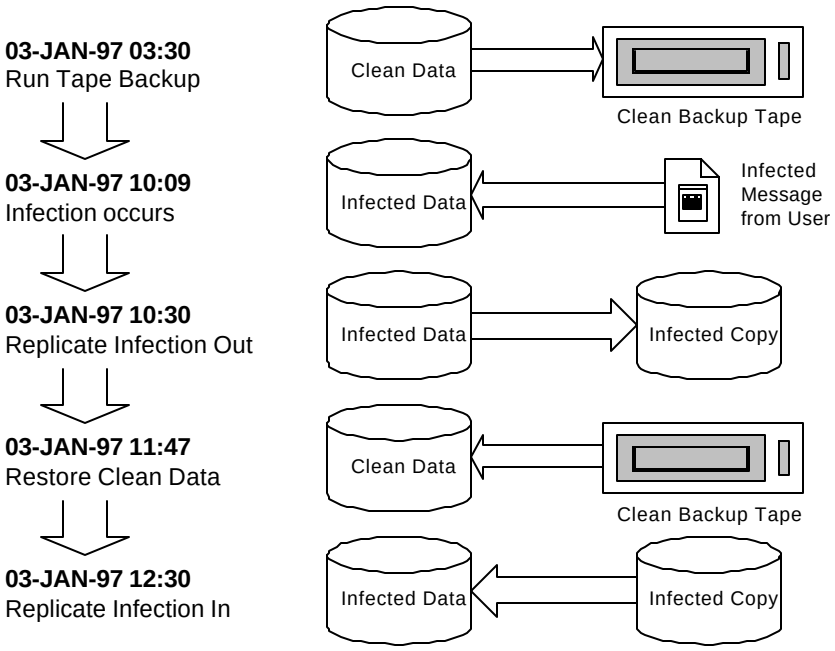
- Users may temporarily disable on-access scanning during software upgrades or when software incompatibilities arise. It is common practice for software vendors to suggest disabling virus scanning during an upgrade or installation.

- Users could mail, replicate, or copy documents with infected file attachments. Possible destinations might be home computers, colleagues or consultants that do not work on-site, or perhaps customers and vendors.

The basic flaw in this argument is that it assumes all people and systems with which the users share documents contained in groupware are similarly protected by "on-access" virus scanning. This of course is rarely true.

### 3.2.2 File Backup and Recovery

Establishing regular backups of groupware data files is always a prudent measure for preventing data loss. However, groupware environments complicate and often negate the effect of data recovery as a means of virus removal. Specifically, groupware applications that replicate data to multiple servers are problematic in that infections occurring after the backup date will be replicated into the restored database. That is, if a virus has infected a document or repository after the backup was made, then the infected documents will replicate into the restored database and overwrite the clean versions of the document, as illustrated in the diagram below.




*An example illustrating how replication re-infects data after restoration from a clean backup.*

Since this phenomenon is dependent on the replication schedule and topology of the groupware application, re-infection will be difficult to prevent unless one of the following procedures is followed:

- If the database is not replicated, then restoration from a clean backup will succeed in removing the infected document(s). However, data loss will occur since all documents created between the backup time and the restore time will be lost.

- Restore every replica copy with the clean backup. Be sure to disable replication on the application until all restores are complete. This method is only realistic if all replica copies of the data exist at the same site. Furthermore, this method will also result in data loss for documents created between the backup time and the last restore time.

Remember that not even the above procedures are foolproof since mobile users may harbor infected data for days or weeks before replicating the infections back to the server. Therefore, one should not rely on data backup and recovery procedures for virus removal in groupware environments.

More generally, data backup and restore procedures are quite ineffective as a means of removing computer viruses from data files (as opposed to executable files). There are two distinct problems that any procedure based on this scheme suffers from. First, there will be a high probability that data loss will occur as a result of a recovery operation. Specifically, edits to any data file after the backup will be overwritten and lost. The amount of data loss will be inversely proportional to the frequency of backups.

Second, there may exist a significant probability that the clean backup that was made is not clean after all. Let's suppose that a user performs daily incremental backups to tape and that a full virus scan is performed prior to the backup. No infections are found but a few weeks later, the user installs a new version of the scanner (replacing a 4-month-old version). The following day, the virus scanner reports that a new macro virus has been detected in a Word document. Restoring the document from tape does not solve the problem since the updated scanner detects a virus in the backup copy as well.

The problem is that the older scanner did not detect the new virus. The backup is clean relative to the release date (approx.) of the scanner used to verify the backup and NOT to the time when the backup was actually performed. Reducing the time between the scanner release date and the backup date increases the probability that the backup is clean, but practically speaking, this risk can never be eliminated.

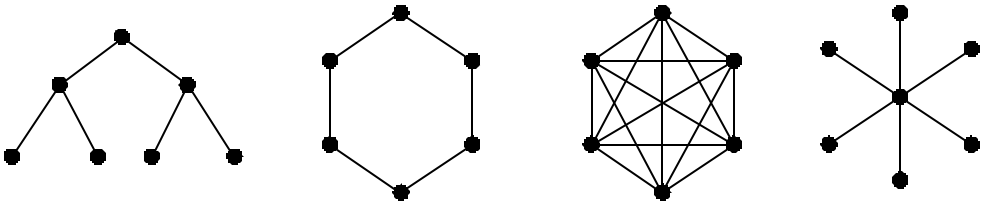
Whether from a flood, or from an activated virus payload, backup procedures play a vital role in disaster recovery. Just keep in mind that if there was a virus problem before the disaster, there will likely be a virus problem after recovery.

## 3.3 AV Strategies for Messaging

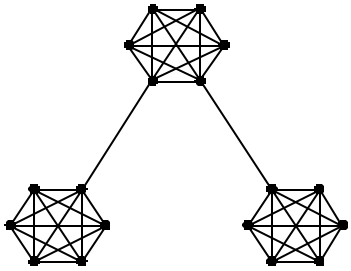
The first step in developing an anti-virus strategy for messaging within groupware is to understand the overall message routing topology. The topology defines the paths used to deliver messages from one user to another. It serves to identify all routing servers that can be used for incident containment as well as any messaging gateways. The messaging gateways along with the client workstations that use messaging define the perimeter that needs to be protected. Once the perimeter is defined, the final step is then to identify the routing servers that will be used for incident containment.

### 3.3.1 Message Routing Topology

Determining the message routing topology is a process of identifying the routing servers and the connections between each server. The connections between routing servers are bi-directional for all popular groupware platforms but may be limited to a single direction in the future. For now, we will assume bi-directional routing paths. Routing topologies are derived from one of four basic types, which are shown below.




Routing topologies can combine different types to satisfy the overall requirements of the groupware environment. For instance, a common topology is to use clusters for each site and then connect each cluster in a tree topology as show below.



Remember that the above diagram is only showing the routing paths between routing servers and not the connections from client workstations to the routing server.

### 3.3.2 Perimeter Protection

There is one simple rule to perimeter protection: use it on every client workstation and gateway that exists. If that is not possible, or if perimeter protection is suspect in certain areas such as home computers dialing into a routing server, incident containment must be deployed.

### 3.3.3 Incident Containment

With the message routing topology in hand, the next step is to define routing nodes that will provide incident containment. First, decide what the acceptable user population exposed to a virus incident should be. Remember that this is not how many users will have their workstations infected but rather how many workstations may become infected. Generally, perimeter protection deployed on the workstations will prevent viruses from entering and infecting workstations. There is no simple method for choosing incident containment nodes since it is a function of how secure the perimeter protection is, how large the routing topology is, and the acceptable risk level to the environment. There is however a few rules of thumb that should be observed.

- Use a divide and conquer approach that creates equal size containment areas within the topology.

- The simplest method is to simple deploy containment on every routing server. It affords the highest level of containment while remaining unaffected by any changes that may occur to the routing topology.
- Another simple method is to deploy containment only on routing servers that interact with client workstations. However, if the site has a cluster routing topology (this is usually the case) then deployment on all servers is necessary since every server interacts with client workstations.

### 3.3.4 Encrypted Messaging

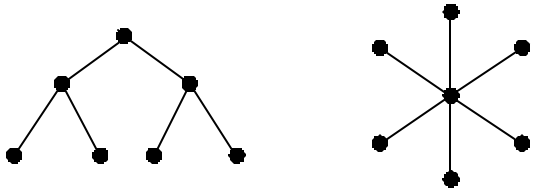
Most groupware environments provide very substantial encryption capabilities to guarantee privacy of messages. However, this capability defeats the use of incident containment on routing servers since the routing server does not have sufficient access rights to decrypt messages for scanning. If we look at the extreme case where all messages are encrypted, then incident containment should be ignored and all focus should be placed on the perimeter protection. It is possible in some groupware environments to segment the use of encryption such that messages crossing a domain boundary cannot be encrypted. This can be an effective technique to guarantee messages that cross the boundary are always scanned. However this must be weighed against the security requirements of the organization.

## 3.4 AV Strategies for Replication

Developing anti-virus strategies for replication within groupware is similar to routing in that a replication topology must be identified. The topology for replication defines the paths taken by documents as they are copied from one replicating server to another. With the exception of gateway servers that replicate with external sources, replicating servers within the topology can be used for incident containment. The gateway servers and client workstations that interact with replicating servers define the perimeter that needs to be protected.

### 3.4.1 Document Replication Topology

Determining the document routing topology is a process of identifying the replicating servers and the connections between each server. The connections between routing servers are bi-directional but may be restricted to a single direction. For now, we will assume bi-directional replication paths. Replication topologies are derived from one of two basic types, which are shown below.



Tree Hub and Spoke

### 3.4.2 Perimeter Protection

There is one simple rule to perimeter protection: use it on every client workstation and gateway that exists. If that is not possible, or if perimeter protection is suspect in certain areas such as home computers dialing into a replicating server, incident containment must be deployed.

### 3.4.3 Incident Containment

With the document replication topology in hand, the next step is to define replicating servers that will provide incident containment. First, decide what the acceptable user population exposed to a virus incident should be. Remember that this is not how many users will have their workstations infected but rather how many workstations may become infected. Generally, perimeter protection deployed on the workstations will prevent viruses from entering and infecting workstations. There is no simple method for choosing incident containment nodes since it is a function of how secure the perimeter protection is, how large the routing topology is, and the acceptable risk level to the environment. There is however a few rules of thumb that should be observed.

- Use a divide and conquer approach that creates equal size containment areas within the topology.
- The simplest method is to simple deploy containment on every replication server. It affords the highest level of containment while remaining unaffected by any changes that may occur to the replication topology.
- Another simple method is to deploy containment on just the hub servers. This increases the size of the containment areas but reduces the number of servers that need to have anti-virus tools installed.

# Implementation Guide

With solid strategies now in hand, it is time to describe implementation as part of an overall anti-virus plan. This begins by identifying and classifying every computer (node) that directly interacts with the groupware environment. Then each node needs to be assessed in terms of risk and appropriate anti-virus tools selected and deployed. The result is a manageable and secure groupware environment that does not limit productivity.

## 4.1 Classification

The first step in implementing an anti-virus plan for groupware is to identify and classify all of the computers that directly interact with the environment. Every computer must be classified into at least one of the following categories:

| | | |
|---|---|---|
| Workstation Client | The workstation client is a computer that allows the user to interact with the groupware environment. | PERIMETER |
| Routing Server | The routing server transfers messages from a client or routing server to a client or routing server. | INTERIOR |
| Replicating Server | The replication server exchanges documents from a client or replication server to a client or replication server. The replicating server also stores shared document repositories. | INTERIOR |
| Gateway Server | The gateway server interacts interior routing and/or replication servers as well as one or more external (not trusted) servers. Generally, the gateway server is used to connect environments to each other. | PERIMETER |

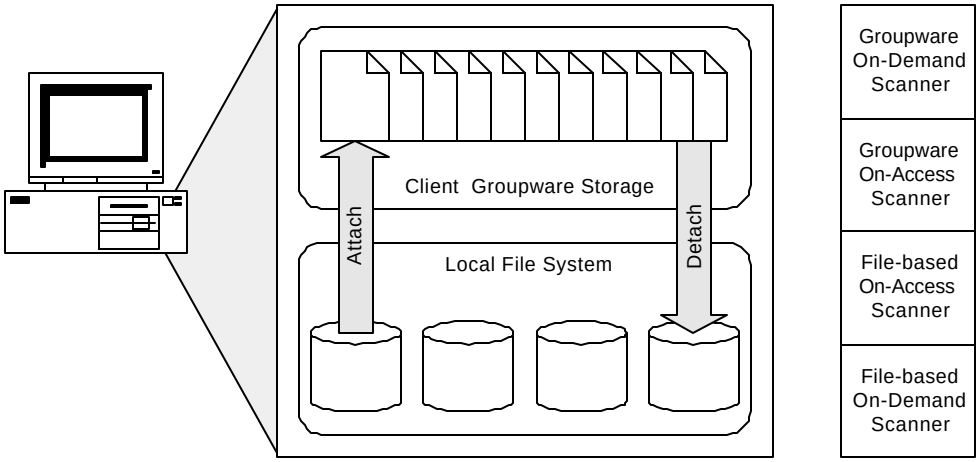
Often, a single computer will assume multiple roles. For instance, a server may perform both routing and replication. A server might also be used as a workstation client. Failure to identify all such situations will increase the risk of virus infection to the groupware environment. For instance, if dial-up access is provided to users, then be sure to check if home computers are being used to connect to the environment. A home computer represents a high risk, unmanaged node on the perimeter of the groupware environment. Also ensure that new computers are suitably protected before they are allowed to interact with the groupware environment.

## 4.2 Client Workstations

The client workstation is positioned at the perimeter of the groupware environment where the user interacts with the messages and documents it contains. The client workstation is the primary consumer and primary producer of information within a groupware system. That means virtually all virus paths through the groupware system will involve client workstations. Effective perimeter protection of client workstations demands that both the local file system (including boot sectors) and local groupware storage remain virus free. It is essential that active viruses are not allowed to exist in either medium.

### 4.2.1 Anti-Virus Tools & Techniques

Since the local file system and the local groupware storage need to be virus free, both file-based anti-virus scanners and native groupware anti-virus scanners need to be implemented. It is unacceptable to use just on-demand scanning techniques for the client since a virus could easily find its way into or out of the groupware system between scheduled scans. Therefore, both on-demand and on-access scanning for files and groupware need to be implemented. The on-demand scanners are especially effective at verifying the client workstation virus free before it is allowed to interact with the rest of the groupware system. Furthermore, the on-demand scanners should be used to re-scan the file system and groupware storage each time the scanners are updated. This ensures that a newly detectable virus does not already exist on the client workstation.




### 4.2.2 Viruses Encrypted By Groupware

Viruses that are contained in documents, messages, or file attachments that have been encrypted by the groupware environment pose an added challenge to protecting client workstations. First, server-based groupware storage that contains encrypted data owned by the user needs to be scanned by the client workstation. Scanning this same storage from the server fails because the encryption cannot be penetrated. For native groupware viruses that are encrypted, the on-demand groupware scanner is the only viable incident containment mechanism. All other points of containment (i.e. the groupware servers) along the virus path are not capable of detecting the virus since the encryption is impenetrable. Only the client workstation has access to messages and documents encrypted for a particular user.

### 4.2.3 Anti-Virus Recommendations

- Deploy on-demand file scanning.
- Deploy on-access file scanning.
- Run the on-demand file scanner after each upgrade of the scanner.
- Run the on-access file scanner continuously. Monitor reads and writes to all removable media. Monitor just writes to fixed media (hard disks).
- Deploy on-demand groupware scanning.
- Deploy on-access groupware scanning.
- Run the on-demand groupware scanner after each upgrade of the scanner and optionally to scan unread messages and/or documents (especially those that have been encrypted).
- Run the on-access groupware scanner continuously. Monitor reads only. Monitor writes if redundant scanning is desired in the case of file attachments.

## 4.3 Routing Servers

Routing servers exist in the interior of the groupware environment. As such they do not interact directly with users but instead provide electronic mail routing and workflow services to client workstations and gateway servers. Because all groupware messages pass through one or more routing servers, the routing server makes an excellent candidate for incident containment of mail-bourne viruses. Incident containment dramatically reduces the impact of virus infections where perimeter protection measures fail or cannot be deployed or enforced.

### 4.3.1 Anti-Virus Tools & Techniques

The routing server requires native groupware virus scanning of all messages that pass through the routing server. The scanner implementation must also guarantee that every message is scanned in the message stream. Polling techniques are not acceptable since they inherently provide an opportunity for infected messages to pass through the server un-scanned (the routing occurs faster then the polling frequency). Acceptable techniques include detecting viruses in real-time for each message as it is routed (on-access) or by modifying the message stream and queuing messages for scanning before final routing (queued). Both the on-access and inline techniques are equally effective in terms of detection, and provide a trade-off between server overhead and delivery time. Practically speaking, queued scanning is more effective because the delay in delivery is imperceptible to users (just a few seconds typically) while real-time scanning may have a significant impact on server performance.

### 4.3.2 Encryption and Digital Signatures

The encryption of messages within the groupware environment provides a one-to-one type of access to the message. That is, each encrypted mail message has one and only one user that can decrypt and read the message. Since the routing server is a transfer point and not the recipient of mail messages, scanning encrypted messages is not possible at the server level. Any suggestion that a vendor can do such scanning is

implying that they have defeated the encryption technology used by the groupware product.

Another important feature of a server-based message scanner is the ability to properly handle messages with embedded digital signatures. These signatures serve to identify the sender of the message as authentic and to verify the message has not been altered. Therefore, the scanner must ensure that the digital signature survives the scanning process. Failure to do so undermines the security of the groupware system by delivery unauthenticated messages to the user.

### 4.3.3 Trusted Message Routing

In most situations, message scanning is deployed on multiple routing servers within an organization. If a message is sent that traverses two or more routing servers, the message will be scanned multiple times. This behavior could put potentially serious drain on the overall message routing system within the groupware environment. The scanner must be capable of generating and detecting "trusted" messages that do not require additional scanning if they pass the first scan. Additionally, if multiple versions of the scanner exist in the routing path, newer versions of the scanner should not trust scans performed by older versions of the scanner.

### 4.3.4 Anti-Virus Recommendations

- Deploy on-access or queued message scanning on one or more routing servers. At a minimum, select all servers that interact directly with client workstations and any servers that route between sites.
- Prevent client workstations from accessing multiple routing servers where possible. In other words, provide access to a single drop-off and pick-up. This will improve the efficiency of incident containment within the messaging subsystem of groupware.

## 4.4 Replicating Servers

Replicating servers exist in the interior of the groupware environment where they provide shared access to document repositories of all types. Client workstations interact with replicating servers to read and write documents. Replicating servers also interact with each other by comparing and transferring documents to create exact "replica" copies of each document on both replicating servers. The process of replication occurs at prescribed times and between prescribed servers as defined by the replication schedule and topology. Because of the shared nature of documents that exist on the replicating servers, there is a strong need to contain viruses before they are distributed via replication to a larger audience.

### 4.4.1 Anti-Virus Tools & Techniques

The replicating server requires native groupware virus scanning of all documents that exist on the server. There are essentially two methods for scanning that are available to replicating servers. The first method is to use a native groupware on-access scanner that checks each document as it is accessed or updated by a user's client workstation or by another replicating server. This is effective but at the price of significantly increased server overhead. In many situations, the burden of on-access scanning degrades performance below acceptable levels.

The second method uses scheduled on-demand incremental scanning that only checks new or updated documents for viruses. The scanning schedule is interleaved with the replication schedule so that documents are always scanned before they are replicated. This introduces some risk in that users who share documents on the same replicating server may pass a virus between a shared document before the periodic scanning is performed. Generally, the window for this is quite small (less than two hours) and it assumes a breakdown of perimeter protection for at least two client workstations.

### 4.4.2 Encryption and Digital Signatures

The encryption of shared documents within the groupware environment provides a many-to-many type of access to the message. That is, each encrypted document has many users that can decrypt and read the message using a shared encryption key. To provide scanning of shared document databases, simply provide the server with access to the appropriate encryption key. If security policy prevents granting servers access to the encryption key, then a user who has the necessary encryption key from a client workstation must perform the scanning.

As with message scanners, an important feature of a server-based document scanner is the ability to properly handle messages with embedded digital signatures. These signatures serve to identify the sender of the message as authentic and to verify the message has not been altered. Therefore, the scanner must ensure that the digital signature survives the scanning process. Failure to do so undermines the security of the groupware system by delivery unauthenticated messages to the user.

### 4.4.3 Scanning Collisions

When a scanner detects a virus within a document on a replicating server, the scanner usually will allow a choice of actions to be taken such as cleaning or deleting. These actions generally modify the infected document in some way. The modifications are then replicated to other servers. A potential problem arises when the same document is scanned on multiple replicating servers. In this case, the changes to the document will collide. There is also the added possibility that users change the document on one server while the scan occurs on another. The selected scanner must adequately deal with these scenarios or the scanning will need to be partitioned. Partitioned scanning essentially limits the scope of the scan to only a subset of documents on each replicating server. The partitions are created so that there is no overlap and hence no possibility for scanning collisions. This of course adds an additional level of complexity to the implementation that would be better handled by the scanner itself.

### 4.4.4 Anti-Virus Recommendations

- Deploy on-access or scheduled on-demand scanning on one or more replication servers. Use partitioned scanning where scanning collisions need to be avoided or when load balancing is desired.
- Reduce the number of accessible replicating servers to client workstations to a minimum. The reduces the possibility of a virus incident from appearing on multiple paths to the replicating server and thereby reducing its overall effectiveness.

## 4.5 Gateway Servers

Gateway servers in the groupware environment exist on its perimeter and are characterized as providing routing and/or replicating services to external sources. The external source may be another groupware environment or foreign environments such as Internet SMTP mail routing or World Wide Web document (via HTTP) access. A careful examination of the source is required to determine if in fact perimeter protection against viruses is required. For instance, a gateway to a relational database containing personnel information would not require protection unless binary data such as file attachments or embedded programs were being stored (not likely). However most gateway servers do interact with external environments that can sustain and transmit viruses. Due to the normally high volume of message and/or document flow through the gateway server and its perimeter location, the gateway server will usually be the first node in the groupware environment to be protected.

### 4.5.1 Anti-Virus Tools and Techniques

The most obvious, and often easiest form of protection for gateway servers are native groupware scanners for messages and/or documents. For mail routing gateways, the tools and techniques described for routing servers are appropriate. Likewise, for replication/document transfer gateways, the tools and techniques for replicating servers should be applied.

A less obvious solution involves protecting the external source itself rather than the gateway server. For instance, lets suppose that a gateway server provides SMTP mail routing services to the Internet. If a native SMTP message scanner is implemented between the gateway server and the Internet, then all messages will be scanned regardless of the anti-virus tools implemented on the gateway server. This allows us to re-classify the gateway server as belonging to the interior and not the perimeter of the groupware environment.