

# Infonic Geo-Replication and best practices for Microsoft Office SharePoint Server 2007

## Technical White Paper

Author: [Lawrence Poynter](#): [Infonic](#)

Contributors: [Ion Tobescu](#): [Microsoft](#)

Published: [July 2008](#)

Applies To: MOSS 2007, SPS 2003, Infonic

## Summary:

This technical white paper discusses the deployment of Infonic Geo-Replicator Server to Server with Microsoft Office SharePoint Server 2007 (MOSS). The paper provides a detailed overview of Infonic Server-to-Server replication process.

Geo-Replicator's Server to Server solution uses Epsilon, a patented byte-level differencing technology, to reduce transmission times of new portal content, each way, by up to 99%.

In a nutshell, this means that WAN users at remote locations have the same instant document access as LAN users at your corporate HQ – dramatically improving productivity across the enterprise by joining regional collaborating teams into one global group.

For distant office workers, or personnel dependent on remote servers, it's as if they were working on the headquarters LAN. Local server replication provides business continuity and LAN-speed access, in SharePoint or other web applications, even when connection to the central server is completely lost.

## About INFONIC:

Infonic is a publicly traded company listed on the UK's AiM Stock Exchange with head quarters and primary software development laboratories in the UK, with sales offices in the USA and Singapore, and sales agents in Hong Kong and Germany.

Infonic Geo-Replicator product improves Microsoft SharePoint access for users at the edge of military, commercial and aid organisations. Infonic's customers are supported by a team of experienced technical staff who ensure that our products are effectively integrated and supported within each client's organisation.

## Copyright

This is a preliminary document and may be changed substantially prior to final commercial release of the software described herein.

This document has been written by Infonic plc, with contributions for Microsoft Corporation. The copyright for this document is jointly owned by Infonic and Microsoft.

The information contained in this document represents the current view of Infonic and Microsoft on the issues discussed as of the date of publication. Because Infonic and Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Infonic or Microsoft, and Infonic or Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

This White Paper is for informational purposes only. INFONIC AND MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Infonic plc or Microsoft Corporation.

Infonic and Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Infonic and/or Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

Unless otherwise noted, the example companies, organisations, products, domain names, e-mail addresses, logos, people, places and events depicted herein are fictitious, and no association with any real company, organisation, product, domain name, e-mail address, logo, person, place or event is intended or should be inferred.

© 2008 Infonic plc /Microsoft Corporation. All rights reserved.

Microsoft *SharePoint Server 2007*, *Windows Server*, and *SQL* are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

The names of actual companies and products mentioned herein may be the trademarks of their respective owners.

## Table of Contents

|   |                                     |
|---|-------------------------------------|
| Infonic Geo-Replication and best practices for Microsoft Office SharePoint Server 2007..... | 1                                   |
| Technical White Paper .....   | 1                                   |
| Summary:.....   | 1                                   |
| About INFONIC:.....   | 1                                   |
| Introduction .....  | 4                                   |
| Prerequisites .....   | 4                                   |
| Deploying Prerequisite Components .....   | 5                                   |
| Infonic Publisher .....   | 5                                   |
| Infonic Client .....  | 6                                   |
| Infonic Server Extensions.....  | 6                                   |
| Infonic Web Service .....   | 6                                   |
| Tips for Deploying Infonic Server to Server Geo-Replicator .....                            | 7                                   |
| SharePoint Structural Configuration to Support Replication.....                             | <b>Error! Bookmark not defined.</b> |
| Establishing Replication Update Frequency.....  | 7                                   |
| Pre-Filtering Replication Data .....  | 7                                   |
| Monitoring and Error Handling -.....  | 7                                   |
| Automated Conflict Management .....   | 9                                   |
| Tips for Installing MOSS in a Geo-Distributed environment.....                              | 11                                  |
| Configuring Infonic Custom Components.....  | 12                                  |
| Conclusion.....   | 16                                  |
| References .....  | 16                                  |
| Acknowledgements.....   | <b>Error! Bookmark not defined.</b> |

## Introduction

Global organizations are struggling with the challenges associated with making SharePoint available to their dispersed employees and remote offices. The key challenge is to provide users with a consistent "SharePoint" experience and rapid access to critical content and collaborative applications. This must be done in a way that addresses the end users needs, while also being manageable from an IT perspective.

Accessing any web based application over an extended WAN typically introduces performance degradation for remote users. The need for LAN speed of WAN content grows as bandwidth resources are typically constrained. Without the provision of equality of service, deployments run the risk of remote users disengaging from the SharePoint platform which would jeopardize the corporate collaboration initiative.

## Prerequisites

The Infonic Software must have the following on the Infonic servers:

- At least 1 GB RAM (2GB recommended if searching is to be enabled).

- 1.8 GHz CPU or above recommended

- 32-bit windows server platform (64-bit server platform support planned Q3 2008)

- 50 MB free disk space for the Infonic Geo-Replicator (IGR) Publisher software.

- Sufficient disk space for log files. The publisher log files can take up to 50 MB of space with an additional 30 MB for each publication (a "publication" is a unit of replication).

- Sufficient disk space to hold information about any published publications. This varies depending on how many publications are published, and how large their content is.

- Windows XP or Windows 2003 operating system.

- Microsoft .NET 2.0 Framework.

- Microsoft Internet Explorer version 6.0 (or above) installed on the computer

## Network Specific Infonic Requirements

- The SharePoint server and IGR server should be linked by a high speed LAN.

- The Infonic Publisher and client communicate with services via HTTP and use the standard URL address of the SharePoint site, specified in the Infonic publication, to determine the address of the web services

- IGR can transfer publication amendments between locations via HTTP/HTTPS, FTP or a file share.

- For HTTP/HTTPS transfer of amendments, IIS must be installed on the IGR server at each location.

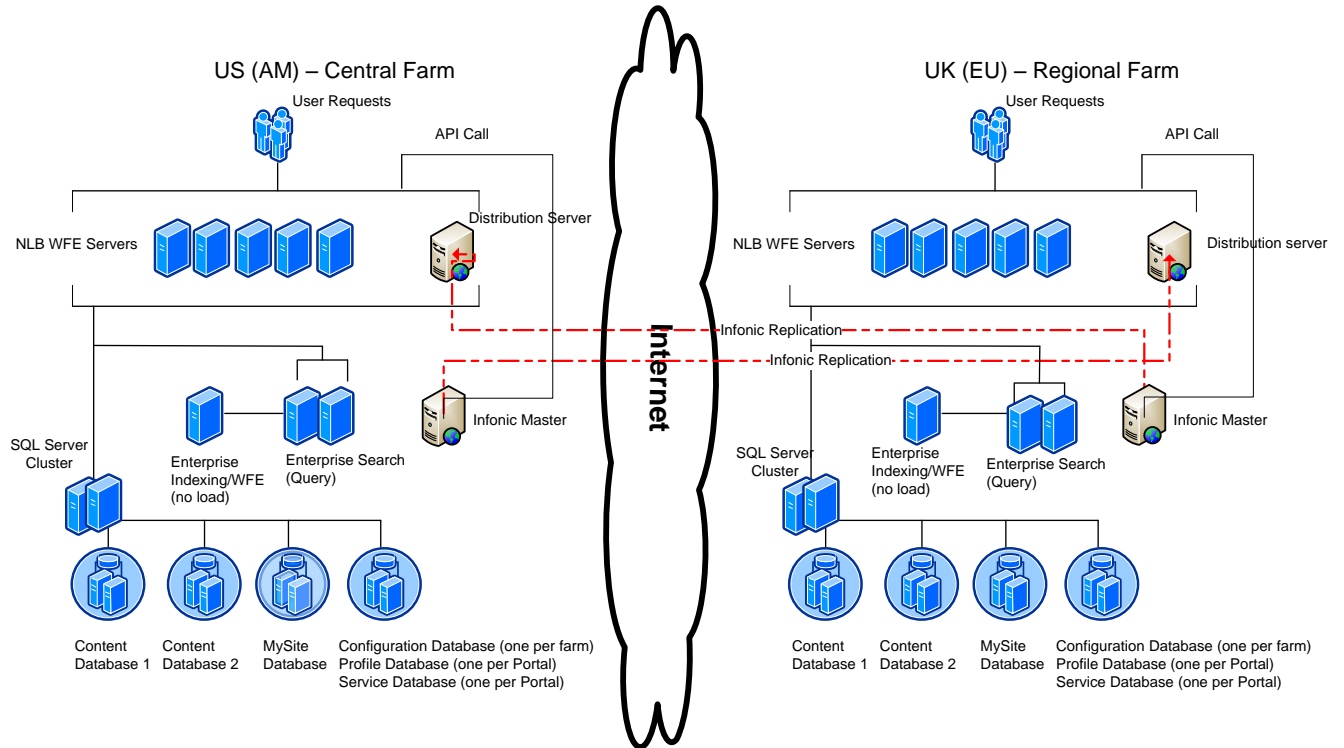
- Each replica IGR client must be able to download amendments from the master and vice versa.

  - For HTTP/HTTPS, the IGR Client uses the Internet Explorer settings from its local machine.

- DNS Mapping can be used at a Replica Location to allow the local Replica SharePoint site to be accessed via the same URL as the Master site.

## Deploying Prerequisite Components

### Diagram Server-to-Server Geo-replication

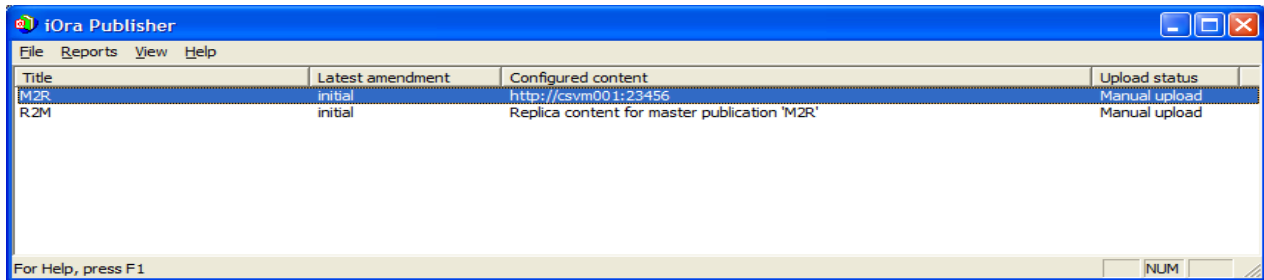


This section provides key pointers for deploying the pre-requisite components. It will not cover the detailed steps required to install these components, as the Microsoft documentation, referred to in the References section, provides all the necessary information. However, there are certain tips and tricks discussed below that you need to be aware of.

The Infonic software is composed of separate components, each performing a different role in the creation, management, and update of publications:

#### Infonic Publisher

You use the Infonic Publisher software to create publications, fill them with initial content, distribute them to the target computers, and generate and distribute amendments containing content changes. The Infonic Publisher main window lists the publications you have produced and displays status information about them.



You can create new publications and work with existing ones. Infonic Publisher provides wizards to guide you through these tasks. You use the Infonic Publisher as the central tool for creating and distributing publications. In 2-Way mode, you install Infonic Publisher to process publications on both the master server and the replica server environments. The publisher can be co-located on the SharePoint server, but typically organizations install on a dedicated server with access to the SharePoint services.

### Infonic Client

Infonic Client applies the amendments (files that detail changes on the master server) distributed by the Infonic Publisher to update publications. Typically the Infonic Client operates in the background and master and replica server managers do not need to be aware of its presence. The client is installed in an environment with network access to the SharePoint master or replica server. Installation of the client is only required where updates to the locally installed SharePoint server are required. If the SharePoint configuration is being used as the basis for information to be broadcast to replicas without information returning, only the publisher is required.

### Infonic Server Extensions

The server extensions are required to be installed to support the communication of the publisher and client components to the SharePoint platform.

### Infonic Web Service

The web service is required to communicate directly with the SharePoint server via its standard application programmer's interface. The web service is required to be installed on the SharePoint web front end.

These components need to be installed in the order listed above.

## Tips for Deploying Infonic Server to Server Geo-Replicator

### Targeting Content for Replication

Before you undertake a replication program initiate an analysis exercise that identifies the key structure and content that is required for replication. With potentially scarce bandwidth resources, you want to ensure that only data that is required in remote locations is replicated. Targeting the focus of your replication will also allow you to increase the frequency with which replication operates.

### Establishing Replication Update Frequency

Typically organizations work to ensure that server-to-server replication operates in a network efficient manner. Generally it is unacceptable to schedule synchronization of servers on a minute-by-minute basis, except were the transport of data is mission critical. As part of the configuration process, it is essential that an update frequency is defined that not only ties in with the available network resources, but also addresses the update urgency required by the business. Defining the update frequency will in effect determine the closeness to 'real-time' replication and is therefore a constant trade-off against network efficiency. Scoping of replication should therefore include a clear definition of the data volatility, and as a result different replication units will all be replicated on different schedules. In addition, specific to SharePoint replication, it can be critical to ensure that the site is self-contained and not dependent on inherited information from other sites.

### Pre-Filtering Replication Data

To make regular automated replication work you want to leave open the opportunity for portal users to add, modify and delete content to either the replica or the master server. That said, you need to protect the automated replication cycle from the additional of significantly large structural or content inclusions that have the effect of stalling the replication cycle because their sheer size cannot be processed as part of the scheduled replication cycle. Typically, organizations filter the content added to either the master or replica to exclude specific file formats e.g. .mov, large structural additions, or files exceeding a certain file size e.g. 1GB.

### Monitoring and Error Handling

Usually organizations deploy replication so that it becomes an automated function that runs on a defined and regular schedule. Given that users rely on either the data that has been replicated from the master to the replica or vice versa, it is critical that any failure to replicate in turn automatically generates a warning detailing that for some reason replication has failed, and requires manual intervention. Two approaches are adopted:

*Server-based* - The failing server automatically generates an email that is distributed to interested parties

*User-Based* - An update time stamp is included as part of the replicated content and it is the user that generates the warning when they determine that the replication cycle has not been executed within the standard cycle.

### Mastering for Control

Any replication deployment should include a clear definition of master and replica that has a direct impact on what can and cannot be performed.

A typical deployment of SharePoint replication software is to replicate content that is mastered on one data server to one or more replica servers which are geographically dispersed. For most organizations, this centralized content ownership is highly beneficial in that it enables IT to maintain centralized control. Typically the central site is acting as the primary generator of content and portal structure and the replicas perform more passive roles.

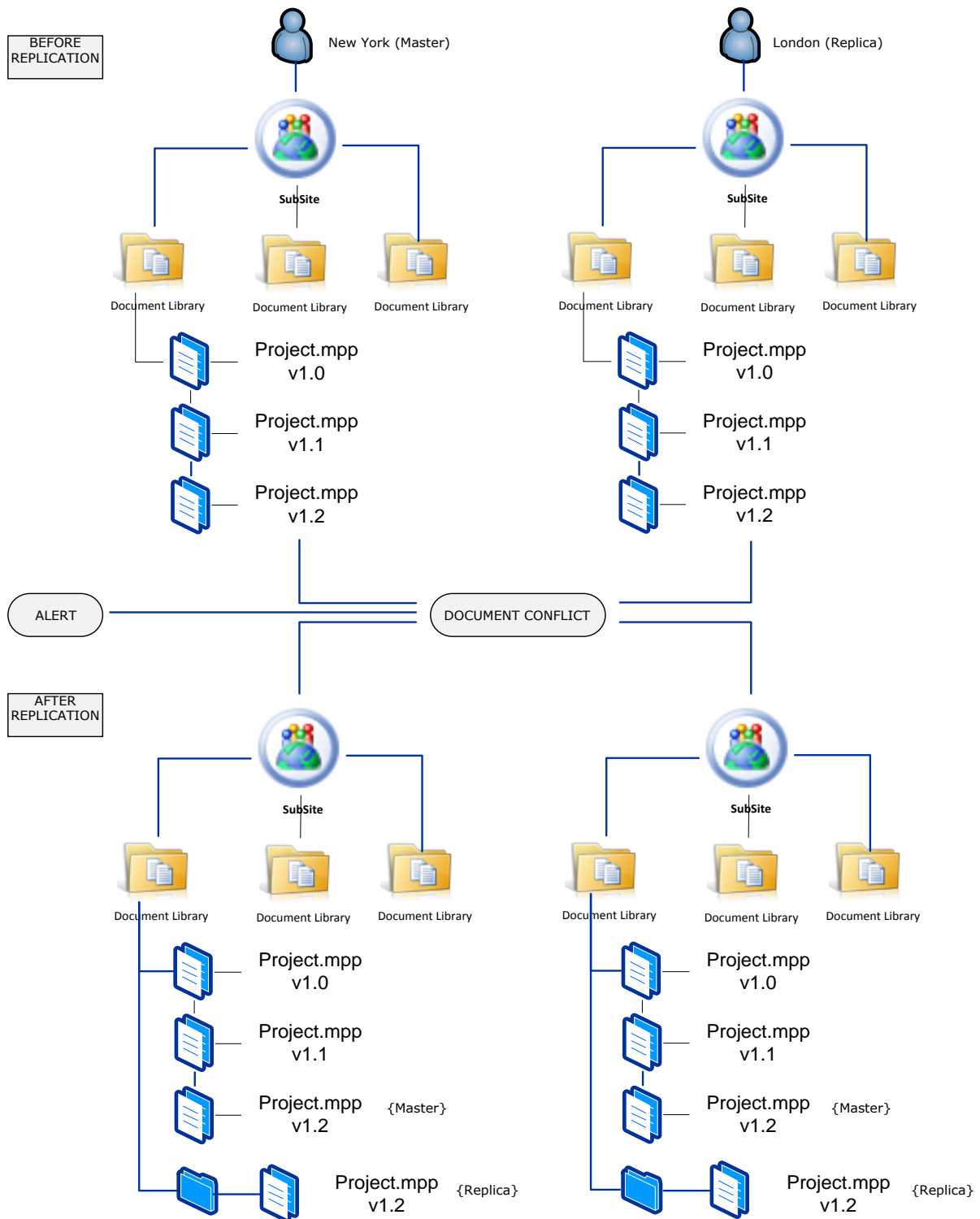
The implications of configuring the master server is that structural and access rights changes will only be replicated in a single direction – from the master to the replica. This should be the default and recommended setup as it defends against the potential problem of replica users modifying and/or deleting structural components that are then replicated up to the master environment and then in turn are propagated to the rest of the connected replica network. In this way for a configuration of a single master and 20 replicas, the deletion of a document library by a user at one of the replicas would have the impact of deleting the related document libraries at the master and 19 other connected replicas.

Note that the comments in the previous paragraph should only apply to the replication of structure and access rights. The actual content that is included as part of the SharePoint document libraries e.g. Word/PowerPoint files, and list items should be replicated in both directions from the master to the replica and the replica to the master. If the same scenario as detailed before were to be applied to the actual content, a file deleted at a replica would be deleted at the master and in turn at each of the replicas. This is perceived as the desirable and expected behavior and does not have the potential implications of structural deletions that could have the impact of deleting libraries of lists of gigabytes of content.

It is important to note that the designation of “master” and “replica” are publication-specific, not server or location-specific terms. This means that Geo-Replicator provides the flexibility to allow Publication “A” made up of content at Location “1” to be the master and sent to the replica server at Location “2”. While at the same time separate content and structure locally managed at Location “2” can be mastered in Publication “B” to be replicated back to Location “1” and also to Location “3”.

### **Automated Conflict Management**

Document conflicts are inevitable as part of a server to server replication scheme. These conflicts need to be managed by the replication engine to ensure that they can later be manually resolved by the contributing users. In structured deployments it may also be possible to automatically resolve the conflicting data by auto-merging content.



Generally conflict management should exhibit three main behaviors:

*Complete automation* - ensuring that any conflicts that are detected by the replication engine are managed automatically and filed in a standard location so that they can be resolved at a time that suits the interested parties. It is absolutely critical that these conflicts do not have

the affect of halting the replication process for further updates. In the diagram, note that two users based in different SharePoint environments make a simultaneous update to what is in effect the same document that is then replicated between the master and the replica. The replication process identifies that a document conflict or collision has occurred and initiates a standard process. This process regards the document created in the master environment as the one that is to be retained and replicated at the top level of the document library. The document created at the replica is shifted in a sub folder(entitled Document conflicts) and both the master and replica user can then be alerted using the SharePoint mechanism to the document conflict so that they can negotiate how to manually merge the changes.

*Zero data loss* - All versions of the conflicting data file constitute potentially valid modifications of the data. For this reason when a conflict is detected all variants of the document either created at the master or at one of the replicas have to be stored within the portal for later resolution. The replication engine does not have the capacity to delete content.

*Automated control* - Document collisions in large deployments and collaborative environments can occur with alarming frequency. Any conflict resolution program should be fully automated and should not have the effect of stalling the replication process.

### **Replication Tracking to Continue Service in the Event of Primary Replicator Failure**

Where replication constitutes a mission critical synchronization of data between operational servers, the replication engine needs to be able to be configured in a way that allows for continuation of operation in the event of the primary replication engine failing. For a replication solution to be able to hand over service to secondary server the record of what has and has not been replicated to and from the master to the replicas becomes essential.

### **Maintaining replication state for continuation of operation following unexpected service disruption**

Akin to all automated operations, synchronizing data replication must provide a trusted method of reconstructing the replication scheme after an unexpected system failure. Unexpected system failures are defined as the point where the network or operational environment becomes unavailable e.g. natural disasters, hardware failures, power interruptions, malicious intent etc. Key to restarting the replication process is the detailed record of the status of replication at the point of failure. Any ambiguity regarding the status of replication will call into question the validity of the replica as a copy of the master. Consequently the only method of reestablishing validity is to rebuild both master and replica from scratch or backup – a potentially disastrous option for a commercial enterprise.

In a geographically-distributed MOSS architecture, failover to another farm is a viable option using a replication strategy. As Geo-Replicator mimics the path of all replicated content beyond the server name of the web application, utilizing a single URL alias for all employees globally permits seamless operational continuity. Regardless of which MOSS farm the end-user is accessing, they use the same URL for their SharePoint application. This URL is resolved by the DNS server in their local environment to redirect them based on their IP address back to the WFE or NLB address for MOSS. In the event of a local outage of their MOSS farm, the DNS server would have a secondary path to a replicated MOSS farm in another region where the users would be re-directed. The application would be slower because of latency and available bandwidth, but would at least be available for use until the local MOSS farm was restored to full operation.

## Tips for Installing MOSS in a Geo-Distributed environment

Many organizations support employees working in multiple sites either within a region or around the world. Microsoft currently supports three different deployment configurations to accommodate geographically dispersed sites with Microsoft Office SharePoint Server 2007 and Windows SharePoint Services 3.0. This planning content includes the following articles:

- [Supported global solutions for Office SharePoint Server](#)
- [Extending Office SharePoint Server global solutions with Office Outlook 2007 and Office Groove software](#)
- [Design global information architecture and governance](#)
- [Plan for global enterprise search](#)
- [Plan for bandwidth requirements](#)
- [Optimizing Office SharePoint Server for WAN environments](#)
- [Optimizing custom Web parts for the WAN](#)

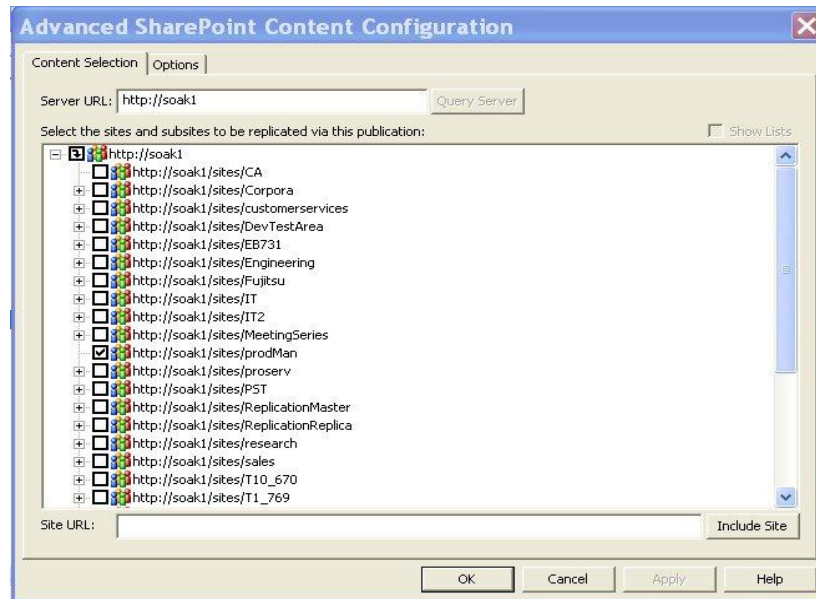
### Configuring Infonic Custom Components

*Deploying Infonic Servers* - To operate in 2-Way mode, you need to take care in setting up your system and in the sequence of building publications on the master then the replica and ensuring that they are logically connected.

In the steps below, you must use accounts with site collection administrator rights on the respective SharePoint sites when capturing or updating their content. This applies to the account you use when running Infonic Client as a service and when running Infonic Publisher automatically through the scheduler.

Follow these steps to configure 2-Way mode set up:

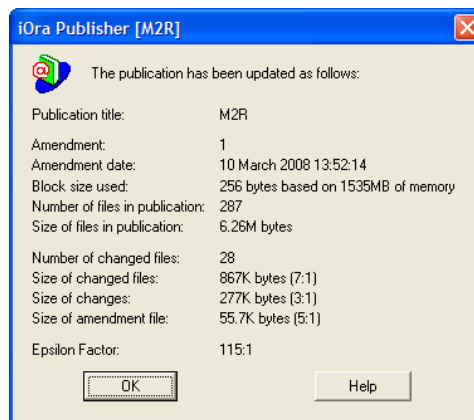
1. Install the Infonic software components and other system software required on both the master server and the replica server as defined in the previous section
2. On the master server, create a master publication containing the master content from your central content server. For SharePoint content, the account you use to create the publication must have SharePoint site collection administrator privileges.



3. Make an installer executable of the master publication.
4. For SharePoint content, configure an account with the same privileges on the replica server as you used to capture the content on the central content server.
5. Execute the master publication installer on the replica server. This installs the master content on the replica server by directly interfacing to the locally installed SharePoint environment to create sites and structure, add content and define user rights and permissions.
6. On the replica server, create a replica publication containing the replica content in exactly the same way that you created the master content as part of step 2. This definition will need to include the name of the master publication so that conflict resolution can be managed efficiently. The publisher makes this a simple process by presenting the names of the already created master publications for selection.
7. Make an installer executable of the replica publication.
8. Execute the replica publication installer on the master server. This establishes the 2-Way relationship between the central content master server and the remote replica server.

Take these steps to configure 2-Way mode in day-to-day operation:

1. Schedule amendments to the master publication on the master server using the standard Windows scheduler and the Infonic Publisher command line. Detailed below is a diagram illustrating the effectiveness of the embedded Infonic Epsilon compression technology. The dialog reports that without compression 28 changed files would result in an update packet of just under 900KB. With Epsilon data redundancy extraction and compression the replication requirement has been reduced to 56KB. A key factor in successful replication is ensuring that replication is as efficient and bandwidth friendly as possible.



2. Configure the Infonic Client on the replica server to automatically check for updates on the master server. Choose a period between checks appropriate to the schedule you have set for amendments.
3. Enable the Infonic Client on the replica server as a service.
4. Schedule amendments to the replica publication on the replica server using the standard Windows scheduler and the Infonic Publisher command line
5. Configure the Infonic Client on the master server to automatically check for updates on the replica server. Choose a period between checks appropriate to the schedule you have set for amendments.

After you have followed these steps, the Infonic software will automatically:

- Record changes in the master content and amend the replica content accordingly.
- Record changes in the replica content and amend the master content accordingly.

The Infonic software also:

- Ensures that all changes eventually converge to consistent content on all servers.
- Records any conflicts between files (that is, where the same file is manipulated on two or more servers within the same replication cycle).

Note: No amendment will be created at the regularly scheduled cycle time if no changes have been made on the source SharePoint environment since the last amendment.

## Conclusion

This paper has provided a guide to setup a complete MOSS and Infonic Server to Server Geo-Replication solution in a worldwide farms deployment. This paper also covered the key scenarios encountered by end-users when looking for rapid access to critical content and collaborative applications. The Microsoft platform can be leveraged by partners such as Infonic to deliver compelling applications to today's marketplace.

To support enterprise replication in complex globally distributed organizations, any replication solution must be able to meet the following demanding requirements:

- Performance: Replace WAN response times with LAN response times e.g. reduce response times in access of 30 seconds for opening and saving large documents to less than 5 seconds.
- Cost reduction: Minimize WAN traffic with advanced compression technology and enable web applications and file content to be replicated during off peak hours.
- Transparency: To avoid costly re-training of end users, or redevelopment of existing applications, all solutions must provide the same interface to the remote or mobile user.
- Compatibility: Support all file types, commonly used web technologies, network and security infrastructures.
- Security: Integrate replication support with the existing security infrastructure, and extend to support remote and offline users.
- Scalability: A scalable architecture that can manage tens of thousands of users and terabytes of replicated content.

Server to server replication technology is creating a new opportunity for companies to resolve existing connectivity issues, and is also enabling the deployment of systems which would not have otherwise been considered.

## References

This section lists sites on the Web where you can find more information about the topics discussed in this paper.

Infonic Technical White Paper

- [http://www.infonic.com/whitepaper\\_enquiry\\_gr.php](http://www.infonic.com/whitepaper_enquiry_gr.php)

Infonic Product Information

- [http://www.infonic.com/product\\_geo\\_replicator\\_server\\_to\\_server.php](http://www.infonic.com/product_geo_replicator_server_to_server.php)

