

Optimizing Worldwide SharePoint Data and Content Access with DocAve



Table of Contents

Introduction	2
About AvePoint	2
Today's 21 st Century Knowledge Worker	3
When SharePoint End-User Demands Conflict with IT Reality.....	4
SharePoint Deployment Architecture Options	5
Single, Centralized Deployment.....	5
Centralized Deployment with Local Content.....	5
Fully Distributed, Replicated Deployment.....	5
Distributed Model with Centralized Backup.....	5
Vital Replication Questions.....	5
Native SharePoint Replication Features	6
Replication Considerations for Web Content Management	8
Conclusion.....	9

Speed and agility have always been two hallmarks of any successful business, but particularly in today's 21st century "always on" culture, these two traits are even more important. Organizations are spreading their presences into multiple, global offices with knowledge workers charged with working together on projects. Additionally, with the advent of the internet, businesses have a constant need to ensure its website content and information is current and truly resonating with their customers.

Many companies have turned to Microsoft SharePoint Products and Technologies in order to provide a platform for managing information and connecting knowledge workers in order to optimize their business processes. With the release of SharePoint Server 2010, many organizations are now using the platform as the basis for its web content management (WCM) initiatives.

However, once the decision on SharePoint has been made, more regarding the proper farm architecture come to the fore. For WCM, how will content approved on a testing farm quickly and automatically be pushed to the publishing farm powering the customer-facing website without any glitches? Particularly for those with a geographically distributed workforce, will there be a one centralized SharePoint server farm, or multiple, distributed SharePoint servers and farms?

For both of these scenarios, there are several technological and infrastructural barriers that must be overcome to ensure their SharePoint platform not only provides an effective, efficient means of promoting effective global collaboration, but can also quickly respond to ever-changing needs at "Web speed".

In the coming pages, we will describe the business situation facing many 21st century knowledge workers, the inherent conflict between business demands versus IT management, the different scenarios for architecting SharePoint environments, and how third-party tools such as DocAve Replicator can ensure organizations can implement a comprehensive, reliable, and efficient replication strategy to optimize collaboration and increase productivity.

About AvePoint

AvePoint is a global technology company and proven software leader headquartered in the United States. Since its founding in 2001, AvePoint has become the world's largest provider of infrastructure management software solutions for Microsoft SharePoint Products and Technologies. Propelled by the world's largest SharePoint-exclusive research & development team, AvePoint is the premier provider for EPG, SMB, Mid-Market and Government organizations demanding the most powerful and flexible infrastructure management solutions for their SharePoint environments and assets. AvePoint's award-winning DocAve Software Platform is recognized as the industry standard for comprehensive, scalable, and truly integrated SharePoint backup and recovery, administration, replication, migration, archiving, deployment management, reporting, storage optimization, and content lifecycle management.

AvePoint is headquartered and maintains its principle engineering center in Jersey City, NJ, with wholly owned sales and engineering centers in the USA - San Jose, Los Angeles, Seattle, Chicago, Washington DC, Houston, Boston; Ontario, Canada; Melbourne, Australia; London, United Kingdom; Munich, Germany; Johannesburg, South Africa; Tokyo, Japan; Singapore; and China - Beijing, Changchun, Dalian. AvePoint's global team, fortified by an expansive network of certified partners, helps more than 8,000 enterprise customers – including many Fortune 500 companies and government agencies – to protect, manage, optimize, and integrate their mission-critical SharePoint environments. AvePoint is a Depth Managed Gold Certified Microsoft Partner and GSA provider.

Today's 21st Century Knowledge Worker

The days of companies having every single one of their employees literally under one roof are essentially over. Particularly at the turn of the 21st century, business has become geographically expansive and, subsequently, a globally distributed workforce has emerged. Recent surveys find **up to 90 percent** of employees work in locales beyond headquarters. Some of these destinations aren't necessarily known for their ubiquitous internet access, either. Places of business now include remote, mobile locations such as ships and submarines. More than half the workforce finds itself in different locations than the rest of their team, or even their supervisors, but they still require access to the same information. Moreover, many knowledge workers today are also charged with managing, authoring, or publishing content to help keep customer-facing web presences up to date.

In the midst of these two business challenges, enabling collaboration among global workers and distributing web content management tasks to business owners, below are some additional hurdles knowledge workers face every day:

- Over-reliance on technology with often unreliable features
- Non-technical users are forced to use the latest technologies
- Information overload, as technology has enabled knowledge workers to create enormous amounts of electronic data
- According to Forrester Research, workers spend 40 percent of their time searching for information, and 70 percent of content is typically recreated rather than reused
- In today's hesitant economic climate, workers are often being asked to do more with less

Let's use an AvePoint customer, Anglo American, as an example to describe the challenges associated with geo-distribution. Anglo American is one of the world's largest diversified mining groups, with 107,000 full-time employees in Africa, Europe, South America, North America, Australia, and Asia. Anglo deployed SharePoint as an intranet portal and collaboration platform to provide its workers with up-to-date content. The mining giant has a decentralized Microsoft SharePoint platform with three farms (totaling 14 front end web servers) in Johannesburg, South Africa; Santiago, Chile; and Brisbane, Australia consisting of 20,000 end-users. Some of these locales, due to the nature of its mining work, were remote and had inconsistent network connectivity. Also, due to the large volume of traffic and data – approximately 600 gigabytes (GB) – stored across Anglo's SharePoint farms, it was critical to ensure that there was no degradation of service.

Anglo quickly realized it needed to find a way to ensure:

- Real-time, two-way synchronization of SharePoint content across unreliable networks
- Automated scheduling of content replication according to business needs
- Intelligent network bandwidth usage in order to prevent business disruption

To better depict challenges associated with the distribution of web content management to business owners, we'll introduce another AvePoint customer, The City of Prince George. The City of Prince George (located in British Columbia, Canada), deployed SharePoint to replace its intranet and internet site management system in order to make the sites more manageable. The City's SharePoint administrators set up a two-farm SharePoint environment – one for authoring, and one for production. The original plan was to devote one farm to internal editorial management, then publish updates to the external farm nightly. However, the City faced several obstacles:

- Using SharePoint’s Content Deployment Wizard tool to push site collections, webs, lists, folders, and list items from the authoring to production farm for publication, the tool often failed.
- When Content Deployment Wizard did work, it did not preserve critical content metadata.

Now that we have highlighted some of the challenges information workers face on a daily basis, we will examine the various requirements end-users have for accessing enterprise content, as well as concerns IT has for managing such a system, and learn how geo-replication helps organizations meet these challenges. We will examine multiple SharePoint architecture options, and assess native Microsoft replication functionality. Third Party replication solutions will also be considered and evaluated.

When SharePoint End-User Demands Conflict with IT Reality

The message from knowledge workers today is clear: Empower me with the information and technology necessary so that I can do my job most effectively. Any web-based platform, especially SharePoint, will lead to end-users to have the same expectations they would have for any website they visit. They want the sites to load immediately, files and documents should download in seconds (and large files, in maybe a minute or two), and real-time collaboration should be seamless. They expect the technology they use at work to be just as fast and robust as the technology they use at home.

These expectations remain the same whether the user has local or remote access to SharePoint. However, due to native limitations in technology as well as cost concerns, remote users are often looked at as second-class citizens and, subsequently, have a poor experience compared to other users. Distance and bandwidth constraints with wide area networks (WAN) also tend to come about. The table below outlines the three different WAN-related limitations: latency, contention, and congestion.

Latency	The time it takes for a data packet to travel from one designated endpoint to another. For SharePoint, this is related to the time necessary for requested content to progress from the server that hosts SharePoint to a user’s browser at a remote location. The distance between these locations directly impacts network latency, and can result in significant delays – experienced as poor quality of service.
Contention	The phenomena of multiple relay processes competing for limited resources. This occurs when data packets originating from different processes – such as legacy applications, email, and communications – are competing for the same network bandwidth. Administrators must also measure the balance between Internet browsing for business purposes (like competitive intelligence) as well as recreational browsing (like Twitter and YouTube).
Congestion	The sheer volume of traffic attempting transmission over the channel. When congestion levels are elevated too much, prioritization no longer addresses the problem of limited bandwidth, as various types of mission-critical data are being delayed. In SharePoint this is particularly true, as much of the traffic is generated through the interaction of content during regular business hours. Large file sizes and vigorous platform activity will combine to reduce end-user experience and decelerate response times.

If action is not taken, usability will suffer. As usability suffers, user adoption will decline. If user adoption declines, it defeats the entire purpose for SharePoint’s existence as a centralized collaboration and document repository platform.

IT administrators have a completely different set of requirements and concerns when using a platform like SharePoint to power geographically distributed offices, as well as maintain public-facing websites. They

seek a platform simple to manage and centralized in nature. They are being asked to do more work with less resources, just as other knowledge workers are. There is also an issue regarding a lack of skilled technical staff in remote locations. Even if an organization wanted to deploy to other locations, no one would be there to manage it. Consequently, this would fall into the laps of the central IT administrators. This becomes another hurdle because when issues occur, it would be difficult for administrators to understand them as they cannot experience it first-hand.

SharePoint Deployment Architecture Options

Now that we've gone over the details on WAN limitations, let's review options administrators have for designing and architecting their SharePoint deployments with an eye toward its distributed user base. All but the first of the topology options outlined below – single, centralized deployment – assume bandwidth limitations of the related WAN, and offers a different approach to managing network traffic associated with delivery of SharePoint to a geographically disperse audience.

Single, Centralized Deployment

This is the simplest, most efficient deployment architecture – the entire SharePoint farm is deployed and managed from one location. For example, a company maintains its one farm at central headquarters in New York City, and has remote users in Los Angeles, the Middle East, and the United Kingdom. Those remote users – for practical purposes – have unlimited bandwidth with which to remote-access the central SharePoint server. This is a recommended approach when bandwidth or latency does not pose a barrier, and how Microsoft originally envisioned SharePoint being deployed.

Centralized Deployment with Local Content

In this model, there is one main farm in New York City as well as satellite server farms deployed in Los Angeles, the Middle East, and the United Kingdom. Each regional farm will only host content and configurations specific to its regional business activity in order to reduce transmission requirements of such data from the central farm. This model will enable remote users to have faster access to SharePoint in their regional farms but also increase the complexity of the infrastructure, complicating organizational IT Governance plans.

Fully Distributed, Replicated Deployment

The next option for delivering SharePoint access to regional locations is a completely distributed and replicated SharePoint architecture. The regional offices in Los Angeles, the Middle East, and the United Kingdom will maintain their own local SharePoint farms, but – directly through their respective SharePoint local farms – can access content available through their local farm as well as content replicated from the other farms throughout the deployment. With this architecture, WAN limitations can be overcome and regional users can enjoy fast, reliable access to enterprise-wide information through their local farms.

Distributed Model with Centralized Backup

The final option we will discuss utilizes the same architecture as the distributed, replicated deployment above. However, this model allows each regional farm to decide to backup their data and content locally, to an alternate site, or even to cloud storage.

Vital Replication Questions

Synchronizing all regional farms is optimal for ensuring true collaboration and end-user adoption of SharePoint. However, it does not come without its challenges and obstacles. Questions organizations must ask themselves prior to selecting a global deployment architecture include:

- How do you make sure users have access to the same information when necessary?
- How do you share documents and list items, and let users work on them in each regional farm?
- How do you make sure information is kept up to date?
- How do you deal with low-bandwidth connections?

Native SharePoint Replication Features

Since the release of Microsoft SQL Server 2005, SharePoint has offered SQL Mirroring and SQL Log Shipping as native replication options for organizations wishing to utilize geo-distributed SharePoint environments.

SQL Mirroring	The replication of one database (the principal) to another (the mirror).
SQL Log Shipping	Transaction log file of a principal database is transmitted to a satellite database. SQL queries are executed by the satellite database in the log file to execute changes, so it reflects changes made in the principal database.

For replication purposes, these two options have not always delivered the appropriate synchronization functionality in three critical ways:

1. They provide no means to granularly select which content gets replicated, and little control over replicated contents' destination.
2. They require robust network connectivity. Constant connectivity is required for SQL Mirroring, while intermittent connections result in a transmission cue that delays execution of changes to the satellite server for Log Shipping.
3. They do not allow for bi-directional replication or conflict resolution.

To help enhance SQL Mirroring and SQL Log Shipping, the release of SharePoint Server 2010 – and also those of Windows 7 and Windows Server 2008 R2 operating system – introduces **BranchCache**, which enables content from file and web servers on a WAN to be cached on computers at a local branch office. This has the potential to improve application response time as well as reduce WAN traffic, and can be distributed across peer client computers or centrally hosted on a server.

When organizations enable BranchCache, a copy of data accessed from intranet Web and file servers is cached locally within the branch office. Then, when an individual on the same network – but in a different physical location – requests the file, BranchCache helps download the content from the local cache without downloading the same content across the WAN. Essentially, this is to avoid duplication of efforts – and subsequent traffic – over precious network bandwidth.

BranchCache operates in one of two modes:

1. **Distributed Cache** – Using a peer-to-peer architecture, organizations can cache copies of files and send them directly to other users as necessary. Distributed Cache is geared toward branch offices that do not have a local server.
2. **Hosted Cache** – Using a client/server architecture, users can cache content to a computer on the local network – called the Hosted Cache. Then, other users who need the same content can retrieve it directly from the Hosted Cache.

BranchCache is passive, meaning that it will only retrieve data from a server when requested and subsequently will not increase WAN utilization. By only caching read requests, it will not interfere with users attempting to save files.

While this is a definite step forward, the table on the next page outlines the capabilities of BranchCache versus **DocAve Replicator for SharePoint**, AvePoint's tool for full-fidelity synchronization of SharePoint data and content – one-way, two-way, or one-to-many – across multiple SharePoint farms in real-time or according to a fully customizable schedule.

Replication Functionality	BranchCache	DocAve Replicator
Highly scalable, standalone distributed architecture	×	✓
Ability to replicate content	×	✓
Ability to replicate securities	×	✓
Improve access times for items larger than 64 KB	×	✓
Improve access times for items less than 64 KB	✓	✓
Improve access times for dynamic content such as video and audio files	×	✓
Works with any SharePoint-supported operating system	×	✓
Works with any SharePoint-supported Windows Server	×	✓
Same user experience for remote location and central location	×	✓
Ability to replicate configurations	×	✓
Ability to independently replicate content, securities, or configurations	×	✓
Ability to replicate look-and-feel of content or customizations	×	✓
Ability to replicate over HTTP/HTTPS	×	✓
Granularly select content for replication	×	✓
Byte-level replication	×	✓
Real-time replication	×	✓
Event-triggered replication	×	✓
One-to-many farm replication	×	✓
Bi-directional replication with conflict resolution	×	✓
Schedule-based replication	×	✓
Full-fidelity replication	×	✓

Graphical, user, and domain mapping	×	✓
Drag-and-drop interface	×	✓

All of these features make DocAve a scalable, robust replication solution for SharePoint. With DocAve, organizations can deliver reliable SharePoint access to global workforces, drastically improving user experience and, ultimately, SharePoint's return on investment. DocAve is able to help organizations overcome the previously discussed WAN-related limitations – latency, contention, and congestion – in the following ways:

Latency	Because DocAve enables organizations to replicate SharePoint content to multiple, global locations – remote or otherwise – the distance between users' browsers and SharePoint servers will be drastically decreased. Additionally, all users will be provided with fast, local access as well as a higher quality of service.
Contention	Administrators can limit the amount of bandwidth available to replicate SharePoint content by throttling bandwidth usage with DocAve. The amount of SharePoint data traveling over networks is limited by the throttle that the SharePoint administrator designates.
Congestion	DocAve offers byte-level differencing, meaning only the changes that occur in a document are replicated over the WAN, instead of entire documents or sites. By enabling byte-level differencing, organizations can minimize the volume of data that SharePoint is sending over a WAN.

Now, the next section will take a look at the native features in SharePoint Server 2010 firmly positioning it as an ideal platform upon which to not only launch geo-distributed offices, but also web content management initiatives for stewarding nimble, robust customer-facing websites by replicating content from authoring farms to publishing farms in real-time or as business needs dictate.

Replication Considerations for Web Content Management

As organizations strive to either deepen or expand their competitive footprint amidst an ever-growing field of competitors, it is important they maintain up-to-date, fresh content on its external-facing website. With constantly updated content and access to the website from hundreds, if not thousands or millions of customers daily, the WCM platform must be able to scale to these peaks in volume and demand, have the reliability necessary to not crash or fail when customers need it most, and provide an extensible management and publishing framework to keep up with demand.

There are several new and enhanced features in SharePoint Server 2010 which firmly position it as a scalable, reliable WCM solution for today's organization:

- **Content Deployment** uses database snapshots to collect content, and then packages exported content for movement to the destination box. This improves the scalability and performance of the SharePoint platform to handle a consistent lifecycle of content creation and publication.
- With the evolution of the *type* of content hosted on websites, not just text and pictures but also rich media including audio and video, it is important page load times and the overall performance of the website are not affected. SharePoint has **improved its caching for Binary Large Object (BLOB) content** – also known as unstructured data – on the web front ends. The content from SQL is read in smaller chunks and will start sending file to the client immediately so the user doesn't need to wait for the whole file to download.

- With the influx of content that must be created and published on today’s dynamic websites, SharePoint has taken great strides to improve its capability for organizing and supporting voluminous content. The introduction of **Large Pages Library** supports thousands of pages in a given library.
- **Pages Library** introduces the ability to organize pages in folders in order to perpetuate a streamlined hierarchy for governance purposes.
- **Content Organizer** helps end-users organize web content with customizable rules that will decide where given pages should go. This lets authors concentrate on creating more compelling content, rather than sweating out whether or not the page will go to the proper location.
- The **Content Query Web Part (CQWP)** enables site designers to aggregate interesting and relevant slices of information on web pages by offering the creation of queries through an easy-to-use UI, and then display that content in unique, configurable ways. For example, showing “Top Rated News Stories” on a web page. The CQWP is designed to return roll-up content over several different scopes – from a single list and library, to all lists and libraries across an entire Site Collection

While SharePoint Server 2010 has fantastic scalability features to aid in its being an ideal WCM platform, there are some concerns that remain. Below is a table showing how **DocAve Replicator** enables a seamless synchronization of web content prepared in testing or development SharePoint farms to the production farm which powers a company’s internet site.

Challenge	Opportunity with DocAve
Synchronize SharePoint content across multiple, globally distributed farms to ensure optimal global access.	DocAve Replicator for SharePoint enables one-way, two-way, and one-to-many replication of content and data among multiple SharePoint environments.
Granularly select content for publication to various sites and farms for enhanced flexibility and bandwidth optimization.	DocAve Replicator can also move content from staging and testing farms through to external-facing production farms on-demand or according to a pre-defined, customizable schedule.

Conclusion

Today’s rapidly evolving business landscape requires that organizations have a truly global reach, and are always abreast of the latest developments and trends. As more companies have knowledge workers dispersed worldwide, it is important they are all still able to access the most up-to-date content quickly in order to collaborate on business-critical projects. Due to branch offices, remote locations, and WAN limitations, this team collaboration can oftentimes be inhibited.

Today’s business environment also requires organizations maintain an active web presence that can quickly be augmented and changed with content to meet customers’ ever-changing needs. Many companies are turning to Microsoft SharePoint in order to provide one central home for their enterprises, as well as power their web content management initiatives.

However, for a devoted, nimble web presence – as well as a geo-distributed workforce – to actually succeed, there must be support for on-demand and schedulable replication of data and customizations. Natively, SharePoint has limited support for replication. While the platform is meant to be managed centrally and accessed globally, there is no server replication functionality. Consider AvePoint’s DocAve Replicator to deliver granular, flexible, and comprehensive synchronization of multiple SharePoint farms in order to meet geo-distribution and WCM initiatives.

AvePoint Global Headquarters
3 Second Street
Jersey City, NJ 07311, USA

Phone: 1.201.793.1111
Fax: 1.201.217.8709

For more information, visit www.AvePoint.com.

Copyright

2011 AvePoint, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written consent of AvePoint, 3 Second Street, Suite 803, Jersey City, NJ 07311, USA

Trademarks

AvePoint DocAve®, AvePoint logo, and AvePoint, Inc. are trademarks of AvePoint, Inc. Microsoft, MS-DOS, Internet Explorer, Microsoft SharePoint Server 2010, Microsoft Office SharePoint Servers 2007, SharePoint Portal Server 2003, Windows SharePoint Services, Windows SQL server, and Windows are either registered trademarks or trademarks of Microsoft Corporation. Adobe Acrobat and Acrobat Reader are trademarks of Adobe Systems, Inc. All other trademarks are property of their respective owners.

Changes

The material in this document is for information only and is subject to change without notice. While reasonable efforts have been made in the preparation of this document to assure its accuracy, AvePoint assumes no liability resulting from errors or omissions in this document or from the use of the information contained herein. AvePoint reserves the right to make changes in the product design without reservation and without notification to its users.