The technologies and architecture behind iOra Geo-Replicator

Contents

01. Executive Summary

Globally distributed organizations have been successfully reducing costs by deploying web applications and consolidating branch office servers at central server farms. An increasing number of organizations who have implemented these strategies have experienced dissatisfaction from a large percentage of end users, which threatens the value of their overall IT investment.

The problem is that the ever increasing remote and mobile user populations have become frustrated with the performance and availability of enterprise web applications. The result is that these users either stop using the enterprise applications, or manually copy subsets of the information in web applications to local servers and hard drives. The consequence is that the return on investment in web applications is diminished both because of their reduced use and inefficiency in the way the information in them is accessed. Potentially even more serious is that the workarounds used to copy confidential enterprise data locally are threatening the integrity of the data itself.

Even those organizations that have invested heavily in both fast Wide Area Networks (WANs) and central server farms have also experienced unacceptable performance. In many cases these organizations have hit a fundamental limitation to performance, namely, the speed of light. As applications become more complex, and documents grow in size, the number of network requests needed to browse to a web page or open a document has increased. However fast the network and servers are, the response time cannot be reduced below the time it takes light to make the round trip to transmit and receive all the requests, and their responses, needed to deliver the web page or document. With global networks this can mean some web pages and documents cannot be delivered in less than 5 seconds, and with the addition of WAN and server peak time delays this can easily increase to tens of seconds. The only solution to this fundamental problem is to locate content closer to the user requesting it by either replicating the applications to remote servers or all the way to the hard drives of the end users. The challenge is to do this in a way which does not remove the cost advantages of web applications and central server consolidation, and continues to guarantee the security of the data.

iOra develops a new advanced generation of replication software which uses patented technologies to address the fundamental problems of speed and availability of applications. This software enables distributed organizations with large numbers of remote and mobile workers, to eliminate long response times over WANs, and provides guarantees of application availability even when no network is available.

Typically when organisations plan to deploy content closer to the end user for reasons of performance they generally derive the following set of requirements.

---

01. Executive Summary (continued)

- **Cost Reduction**
  To reduce the need for costly upgrades to servers or WANs, or investment in expensive wireless solutions, it is necessary to reduce the load on the central servers and networks from the remote and mobile users, especially during peak business hours. This implies using advanced compression technologies to reduce the amount of data being transmitted, as well as allowing the users to access distributed replicas of applications and portals

- **User Interface Compatibility**
  To avoid costly re-training of end users, or redevelopment of existing applications, all replicated content must provide a transparent interface to the end user. What they see on the live server, is what they should see on the remote replica or when offline, without any change to existing applications.

- **Content and Platform Compatibility**
  Support for any document type and web technology is key to ensuring you can choose a single replication solution for all your applications. This implies support for the widest possible range of portal, document management and web technologies such as Microsoft SharePoint, Lotus Domino, .Net and J2EE. In addition, the solution must be able to work across all your network types from low bandwidth dial up or wireless to wide area networks that use intermittent satellite connections.

- **Security**
  Without providing a secure replication solution the tendency is for remote and mobile workers to make copies of documents and maintain them on their desktop, "My Documents" or email folders. Not only does this mean that the mobile workers waste time making copies, it also means that there is no way of managing the security of these documents. The ideal solution is one which integrates replication support with the existing security infrastructure.

- **Scalability**
  With the need for organizations to operate globally the number of remote offices and users has been growing. In addition there is an exponential growth in document, web and email content. Any content replication solution must have a scalable architecture that can manage tens of thousands of users and terabytes of replicated content.

- **Network Availability**
  Irrespective of network quality, organizations need to plan for the event of network disconnection at either the remote server or remote client. When network connectivity is lost, users need to be able to access business data to continue to deliver operational services.

iOra addresses each of these challenges. At the root of this are the world's most sophisticated replication technologies called Web Virtualization and Epsilon Compression which jointly address the five primary requirements.

02. Industry Trends

2.1. **The Demand for Replication**

For the last 10 years, corporations have been increasingly focused on the adoption and implementation of corporate portals to store, search, categorise and archive their critical information assets. These portals have moved from being a competitive technical opportunity, to being an expected component of the corporate infrastructure.

Two key drivers are at the heart of the staggering rate of portal adoption; Compliance with the ever increasing regulatory framework that corporations need to address to manage risk and the archiving of corporate information as part of a fully managed repository of corporate knowledge.

Typically as organizations attempt to realize this corporate collaboration goal on a global basis they identify that they need to address the following deployment considerations:

Web Platform Standardization

A major trend that has also helped to drive down IT costs, is that organizations across all sectors have invested extensively in web applications. The web browser has become the defacto standard for the delivery of a range of enterprise applications. From collaboration and portal products like Microsoft SharePoint, to ERP and CRM solutions, to customized enterprise applications, all manner of applications are now made available through the browser. The browser interface provides a consistent user experience, simplifying access to disparate sources of information and applications, and reduces the need to roll-out and install multiple applications on remote desktops and laptops.

The benefits of web applications have been readily apparent: users are empowered, application management is simplified and collaboration across an entire extended organization is vastly improved.

There is, however, a problem – one that is limiting the potential return on investment any organization will see from deploying web applications across WANs. The problem is that remote users are experiencing unacceptable performance delays because of WAN performance. Without a solution that accelerates their access to enterprise web applications, such that WAN users access them at LAN speeds, the remote users will not use the web applications and the expected ROI from web enablement will not be achieved.

02. Industry Trends (continued)

The Distributed Enterprise

A related trend has been that most organizations are becoming more distributed. Company growth and mergers are increasing the number of branch, regional and international offices. Use of offshore services in countries such as India requires near-real time information sharing. Military and emergency operations require deployment of personnel to remote locations with limited or no networking infrastructures.

Whatever the capacity of corporate WAN's it would appear that incrementally adding global business functions generates an ever increasing strain on available bandwidth. Deploying Enterprise applications is no exception and organizations look for ways of scheduling and managing the bandwidth resource requirements.

Continuation of Operations/Failover

All major corporate system roll-outs need to address redundancy as key component of the deployment architecture. This is particularly important when the application is an Enterprise portal server being used to store operational, technical and policy based information that is accessed and applied on a frequent basis. The loss of access to a server and the inability to access this information can have significant operational consequences.

As part of a typical deployment plan, organisations are designing server based redundancy in as part of the architecture plan from day one. In the event of server outage these replica servers automatically switch into operation, providing the end user with an immediate continuation of operation.

Mobile Users

It is usual for most organizations to have a significant proportion of their users engaging in roles that take them outside of the corporate network. Knowledge workers like sales forces, field service teams, consultants, auditors, dispersed R&D teams and many others, require access to information if they are to perform to their full potential. But currently too much of that information is only available online as part of the web enabled corporate infrastructure. Even the promise of new wireless technologies will not truly solve this issue: capacity restraints and reception problems (as well as high costs) will mean that for the foreseeable future, at least, 'always-on' will remain a myth. To all intents and purposes they become disconnected from the corporate network, although they have a clear requirement to access material stored as part of the Enterprise applications.

02. Industry Trends (continued)

To maximize the effectiveness of knowledge workers, improve their job satisfaction and improve service to clients, organizations today need the ability to access enterprise information offline as well as online using the same consistent user interface.

True Global Collaboration

It would appear that given the difficulties of deploying web based applications on a global scale results in an inability to realize one of the key goals for these enterprise applications and portals. True global collaboration cannot be a viable proposition unless all of the participating users have consistent and equal access to the key collaborative data. The problem is only exacerbated if collaboration is planned between partner organizations. Successfully managed business partnerships have become a key competitive consideration for most organisations. Typically data is difficult to share in partnership environments where competing collaboration tools, security schemes and servers actively work against the collaboration goal.

In the absence of any inter-company or partner collaboration support, users are forced to collaborate using fax and email that directly impacts the ability of the organization to control and archive the storage of business communication as part of a single corporate portal.

The issue with these trends is that they are in conflict with each other. As the applications and data is moved from remote offices to central servers, the users are effectively moving further geographically from these central servers. Centrally located users are getting the benefits of LAN speeds of 100 megabits per second and above with near zero latency. In contrast the ever growing numbers of remote and mobile users are now experiencing unacceptable levels of response times due to the use of WAN connections that are often hundreds of times slower.

03. iOra - Key Technologies

iOra has a long history of developing and patenting some of the most advanced replication technologies used in the computer industry. To address the challenges of providing complete solutions in the areas of data replication requires a range of technologies. In this section we describe two key components of iOra's replication technology, which are integrated as part of the core product set.

3.1. **Epsilon Compression**

Epsilon Compression is a patented, industry-leading compression technology. It provides highly bandwidth-efficient replication of large volumes of data, no matter what type of content, by ensuring that equivalent sequences of bytes of data are never transmitted on the network twice. Epsilon compression has been shown to out perform other industry technologies, such as delta level differencing, by an order of magnitude because delta differencing is limited to comparing pairs of files with the same name.

Differencing with a Difference

Epsilon provides byte level differencing, supported by pattern matching. Where it differs from its nearest competitor – delta-based differencing – is in the fact that it doesn't simply work within a single file, web page or document, but instead compresses across an entire iOra Publication. A publication is a operational construct used by iOra as a standard container for replication to either a remote server or remote laptop. For instance, if a new presentation is published in a portal, traditional differencing techniques would have nothing to difference the presentation against. With Epsilon Compression, the presentation is differenced against all the other content already held in the iOra Publication, regardless of file or data type. Typically this would mean that Epsilon would find that over 95% of the content of a new presentation already exists in other presentations or documents, and would reduce, for example, a 10 megabyte presentation to under 500 kilobytes Then, using Epsilon's byte level compression algorithms, the data is compressed further, reducing the 500 kilobytes in the example to less than 200 kilobytes, giving an overall compression ratio of over 98%. On a slow network this translates to over 50 times faster transfers.

3.2. **Web Virtualization**

iOra's Web Virtualization technology provides transparent access to replicated web-based content independently of the web server software used to create this content. This includes the dynamic content that plays such a vital role in web-based enterprise applications. What makes this technology so valuable is that it allows the web source of an application to be replicated so that remote and mobile users have a transparent view of the application, without requiring the supporting server software to be installed on the remote servers or laptops.

03. iOra - Key Technologies (continued)

Defining Web Virtualization

Web Virtualization technology recreates web-based enterprise applications on a remote server or laptop, enabling users to work with them in exactly the same way as they would on the central server. These virtualized replica's can be installed on a single users laptop for personal use, or on a server for group replica access as part of a network resource delivered using a proxy configuration. So, for example, a salesperson working with an offline version of his or her company's sales portal can still expect to be able to search customer account details even when he or she is not connected to the network. Virtualized replica server content is often viewed as a quick and light method for providing disconnected read access to corporate information for groups of users, without the over-head of having to install and maintain the underlying server infrastructure.

The role of Web Virtualization technology is two-fold:

Firstly, it processes the pages selected for offline delivery, identifying all the content on each page as well as all the embedded hyperlinks. It effectively follows every one of those links, like a spidering technology would, so that when the site is replicated offline, all the interactive components, graphics and text elements are visible and all the links are active.

Secondly, it reconstructs the pages on request at the remote server or on a laptop. It does this without needing any server applications to be stored on the replica device: literally, what the user sees and interacts with is a virtual copy of the content.
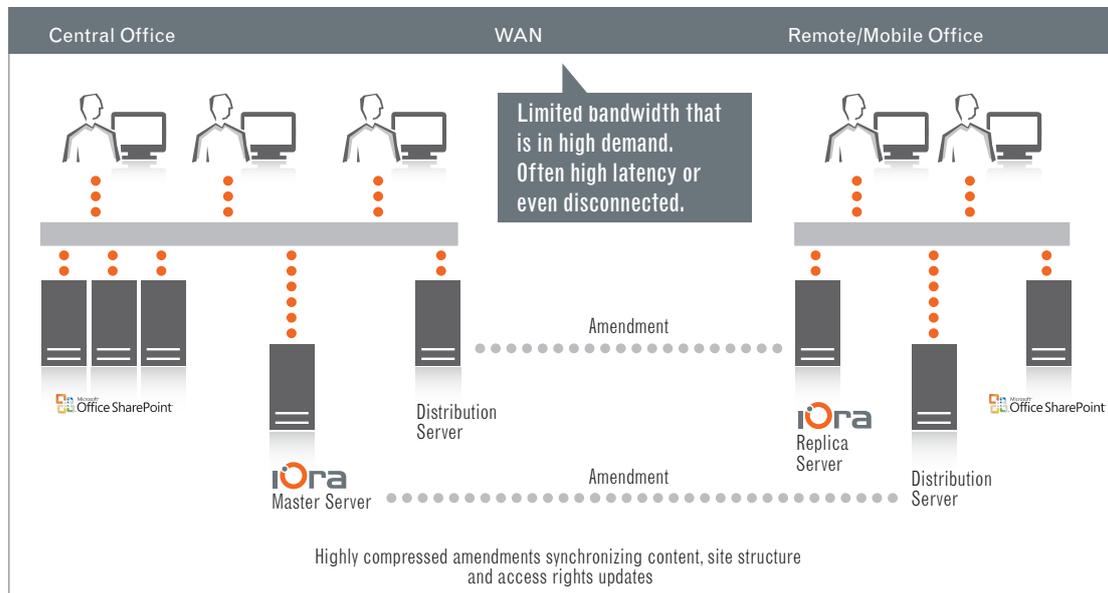
3.2.   File Virtualization

In addition to Web Virtualization, iOra also supports File Virtualization. This is a similar technology, in that it creates a virtual copy of all the content within a particular file server, which can then be replicated for use offline. This, combined with Epsilon compression, provides a highly effective system for replicating large volumes of purely file-based content to mobile users or remote sites.

04. Architecture

4.1. Server to Server Replication

To deploy a server to server replication solution two primary components are required to be installed. The iOra master server component is installed in the master content environment and is the tool that is used to define the scope of the replication. Scope can be as broad as the complete content and structure of the master environment, or as small as a site.

The second installed component is the iOra replica server, that is installed within the replica environment and acts to receive the updates from the master environment and either append these changes to its local repository for delivery using its web application virtualization capability, or pass the updates into the local content server e.g. with Microsoft SharePoint via the standard update API.

Central Office | WAN | Remote/Mobile Office

Limited bandwidth that is in high demand. Often high latency or even disconnected.

Microsoft Office SharePoint

Distribution Server

iOra Master Server

Amendment

Amendment

iOra Replica Server

Distribution Server

Microsoft Office SharePoint

Highly compressed amendments synchronizing content, site structure and access rights updates

Once configured the master and replica servers check against their local content repositories on a update schedule that is dictated by the volatility of the data. Scheduled operation ensures that server peak usage can be efficiently managed and avoids potential system overflow that can occur when triggering incremental replication on each and every server update event.

Specifically the Microsoft SharePoint implementation uses the SharePoint platform API to read and write the structure, content and access rights changes on the master and replica.

04. Architecture (continued)

The update amendments that are passed between the master and replica server benefit from the high levels of compression that can only be achieved by the Epsilon technology. Hence the network connection that is required to connect master and replica can be anything from 1Kbps to 100Mbps. In the same regard the bandwidth used by the replication phase is dramatically reduced, which has direct implications on cost.

4.2.   **Server to Laptop Replication**

To deploy a server to laptop replication solution two primary components are required to be installed. The iOra publisher component is installed in the master content environment and, similar to the server to server implementation, is the tool that is used to define the scope of content to be deployed to mobile users. The second installation component is the iOra Client that is installed on each mobile device and controls the local virtualization of content through the standard Browser, as well as automatically managing the scheduling and updating of content amendments.
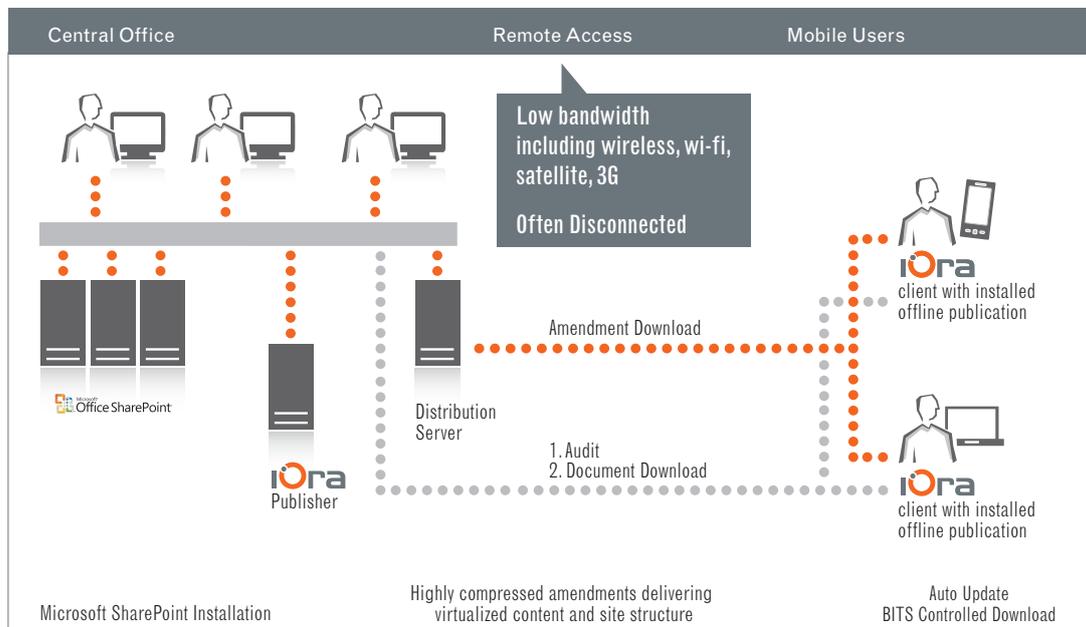
The process of defining a publication for distribution is initiated by the publisher defining the scope of what is included and what is excluded from replication. Once this initial publication has been created it can be downloaded by the mobile clients and at that point they have access to a virtualized copy of the master server application, with the same look and feel – but without the requirement for a network connection.

04. Architecture (continued)

The iOra Publisher is then configured to run on a schedule to check the master content server for changes. When a set of changes has been detected the Publisher creates an amendment file for distribution to the mobile clients as the primary means of update. Once more this amendment benefits from the Epsilon compression technology resulting in a dramatic reduction on the bandwidth that is required for transport. The typical method used as distribution is via an HTTP distribution server, that is not an iOra component, but just an available server on the existing corporate network. Once downloaded to the mobile users machine, the amendment is unpacked and applied as an update to the virtualized web application that is stored on the device. All updates are delivered automatically without the requirement for user intervention.

In addition to the broadcast of content, the server to laptop replication architecture can also be configured to consolidate updates e.g. new documents or forms, from users that were submitted to the virtualized web application while operating disconnected from the corporate network.

05. Conclusion

To support enterprise replication in complex globally distributed organizations, any 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.

- **Network Availability**
  A platform that is able to provide access within networks that have potential for disconnection

iOra meets all of these requirements. Thanks to Web Virtualization technology, iOra can transparently replicate web applications without the need to install these applications on remote servers or laptops. This in turn allows remote and mobile users to access local content, thereby turning WAN response times into LAN response times.

Using Epsilon Compression enables iOra to reduce network bandwidth by factors of 10 to 100 times and be scalable to large numbers of users and terabytes of content.

iOra's industry leading 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.

Disclaimer

Data provided in this material is furnished for information only and is subject to change without notice, and does not represent a commitment on the part of iOra. iOra assumes no responsibility or liability for any errors or inaccuracies that may appear in this material.

For more information on how iOra Geo-Replicator can help your organization, get in-touch via **support@iora.com**

iOra