Migrating to Zimbra Collaboration Suite

ZCS Migration Methods & Technologies Whitepaper Version 1.0
Nov 2006

Table of Contents

1 Introduction

1.1 About the Whitepaper

This document describes the methods and technologies that would be used to manage a large-scale migration of user accounts and data for service providers, from their current messaging and collaboration solution to Zimbra™ Collaboration Suite. This whitepaper provides a framework for review and discussion at the appropriate time by the key parties that would be involved in the migration effort. This document includes insights based on Zimbra team's extensive past experience in managing large-scale (multi-million mailbox) messaging system migrations, including those completed successfully at service providers.

We welcome any feedback or questions regarding the methods, tools and timelines discussed in this document.

1.2 Service Provider Interoperability & Migration Needs

Messaging, and in particular collaboration, has established itself as the critical means of communication, particularly in service provider environments. Users (both consumers and enterprises) typically retain much of their email for future reference and safekeeping. Regardless of its real value, users consider most of their email as critical information. While storage space for email messages is often limited to the space allocated by the administrator, 100MB+ quota per user is not uncommon nowadays.

The challenge to successfully implementing a new messaging and collaboration solution is to ensure a smooth migration from the legacy system, and this includes the ability to transfer data between the two while having fallback mechanisms securely in place. Also, to preserve users' working environment, this migration cannot be simply limited to mailbox information; it must also include other objects: calendar, contacts, and user preferences, distribution lists (private and public), and aliases.

Service providers often have strict requirements for developing a smooth migration strategy. Key ones among them include:

1. **No data loss**: It is imperative that no user messages and other information be lost during or after the migration is completed.

2. **Robust fallback procedure**: The migration methods and tools must clearly specify procedures at each step to robustly revert to using the original system in case of unexpected issues.

3. **Minimal user access downtime**: System downtime for user access should be limited to regularly scheduled maintenance windows and during non-peak times.

4. **"In-place" migration**: This migration should not require modifications to any external system or network components

5. **All collaboration services simultaneously available for users**: All aspects of the migration are generally required to be completed before service is restored for the users.

With detailed upfront planning and execution using the Zimbra migration methods and tools, Zimbra can meet these requirements and transition users over to ZCS with no visible user service disruption.

Note: In addition to the migration tools described below, ZCS also provides built-in and well-tested wizards that are designed to migrate data seamlessly from Microsoft Exchange, Lotus Domino and Novell Groupwise systems as well as tools to import .pst (Outlook) files. For more information, please refer to the appropriate document source in the ZCS References section below.

1.3 ZCS References

ZCS Documentation: https://www.zimbra.com/products/documentation.html

ZCS Wiki: http://wiki.zimbra.com/

ZCS Migration Tool from Microsoft Exchange Server 5.5, 2000, 2003: https://www.zimbra.com/pdf/Zimbra%20Migration%20Wizard.pdf

ZCS Migration Tool from Lotus Domino (coming soon)

ZCS Migration Tool from other messaging systems: http://wiki.zimbra.com/index.php?title=User_Migration

ZCS Blog: http://www.zimbra.com/blog/

ZCS Bug & Enhancement Tracking (*Bugzilla*): http://bugzilla.zimbra.com

ZCS Software Download: https://www.zimbra.com/community/downloads.html

ZCS Roadmap: https://www.zimbra.com/products/roadmap.html

2 Determining What Content to Migrate

The initial requirements gathering phase of the project determines the scope of the ZCS migration. During this phase, data is collected about the current email environment and decisions must be made about what type of data will be migrated to ZCS.

During this phase, service providers need to determine what type of content to migrate to ZCS. Will all users email messages, calendar appointments, and address books be migrated? In addition, which of the users' settings and preferences should be migrated?

ZCS can migrate the following information about users:

- User Settings/Preferences

- Email messages

- Calendar appointments

- Contacts (Personal Address Book)

- Distribution Lists & Groups

- Signatures

- Vacation Messages

- Family/household account configuration (primary/secondary account hierarchy and access privileges)

- Tasks (Coming Soon)

To help gather, evaluate, and plan ZCS system requirements, the following Zimbra worksheets are available for service providers. These are available upon request:

- Workload profile worksheet - to size your deployment based on workload.

- Server and Storage layout worksheet – to plan the type and number of servers and storage subsystems to be deployed.

Data stored in auxiliary systems (e.g. user profile information stored in a centralized directory, users' messages in a voicemail system, users' pictures/albums/videos) may or may not be migrated and consolidated into ZCS depending on service providers' preferences & requirements. If data stored in auxiliary systems is not consolidated into ZCS, the auxiliary systems will be integrated "in place" within the ZCS offering, and modifications to existing applications is not usually required. Also, in such a case, some of the data can also be made available via the ZCS application, through the use of Zimlets as appropriate.

3   Developing a Migration Strategy

Migration of user data to the ZCS system can be achieved in one of the following ways:

- A simple Single-Domain Migration (also referred to as a one-phase migration)

- Shared-Domain Migration (also referred to as a multi-phase migration

Service providers must evaluate their current usage and traffic growth projections before deciding which of these strategies would be best to achieve their objectives.

3.1 Single-Domain Migration Strategy

Single-Domain Migrations is a one-phase migration that is conducted so that there is no overlap in the usage of the incumbent and new systems during the migration process. In other words, all data for all users is migrated during a short maintenance window. New messages are routed to the new ZCS system and user access is

switched over instantly for all users. Single-Domain Migrations are suitable for small deployments in which migration of all user data can be completed during the short maintenance window or when users can tolerate long periods of service inaccessibility during the migration process. Zimbra does not recommend Single-Domain Migration for large-scale migrations.

3.2 Shared-Domain Migration Strategy

Unless a migration can be accomplished in a single "big bang" event, the legacy and new messaging systems must coexist as seamlessly as is feasible within the principal domain (e.g., serviceprovider.net) for the duration of the migration. Also called multi-phase migration, Shared-Domain Migrations is a multi-phase migration that is conducted in such a manner that both incumbent system and ZCS continue to jointly operate for a short period of time. The user data is migrated in batches during a short window (e.g. 100K mailboxes during each regularly scheduled maintenance window). New messages and user access may or may not be switched to use the new system all at once.

Zimbra Collaboration Suite has been extensively tested to support a domain that is shared between two messaging systems. The necessary pre-requisites for a successful Shared-Domain Migration include routing/proxying infrastructure to handle:

- In-bound SMTP requests for delivering mail to user mailboxes

- In-bound POP requests (and IMAP if supported) from existing "fat" messaging clients accessing mail via POP (and IMAP, if supported)

- Web UI proxying for launching the appropriate Web user interface from the carrier portal (including single sign-on/SSO)

Note: see www.zimbra.com website documentation and forum discussions about ZCS migration with a domain that is shared between two messaging systems.

In most migration scenarios that service providers consider, a user mailbox is housed in the legacy system until it has been migrated to the new system. However, Zimbra's POP (current) and IMAP (future) aggregation capabilities enable more flexible migration scenarios in which users' mail is still delivered to the legacy messaging platform, but the user's stateful mailbox, web user interface, and POP requests are all provided by Zimbra. In this scenario:

- The mailbox is "migrated" *without changing the in-bound mailbox* for the purpose of receiving new emails from the Internet

- New emails that arrive in the legacy system are automatically "POP'ed" to ZCS (from the legacy system end the effect is similar to a user converting from the Web interface to a fat client like MS Outlook).

- The user now interacts exclusively with ZCS, whether via the Web UI or via POP or a combination of the two.

With POP aggregation, ZCS uses the imapsync utility (see below) to migrate a snapshot of the user's mailbox which then becomes the new mailbox "of record." At the same time ZCS continues to use the POP aggregation capabilities to leave the legacy mailbox in place. POP aggregation can also be used by users seeking to consolidate email bound to other public email domains (AOL, Earthlink, MSN, Yahoo!, Gmail, …) within the new ZCS messaging system. In particular, POP aggregation could allow service providers to defer any contractual migration trigger points during the trial period and potentially even until the later stages of the migration.

In order to make this multi-month migration period as seamless as possible for users, careful attention needs to be paid to the following infrastructure:

- *Shared LDAP directory* − a common directory for both the ZCS and legacy systems provided by the service provider. A common LDAP repository is required for preserving user security/configuration information during the migration. The migration is complete from the user's perspective when the directory is updated to indicate that this user's mailbox is now to be found in ZCS rather than within the legacy system. Our expectation is that a common directory would remain in place throughout the migration scenario described below. (As an aside, Zimbra will store ZCS user specific configuration data within OpenLDAP and access the legacy directory only as necessary for non-ZCS configuration data, like userid, password, etc. However, our expectation is that the ZCS will be given write access to the legacy directory in order to "squirrel" away certain ZCS specific configuration data within a single unified directory.)

- *Mail/SMTP routing* − A mail proxy would provide this function for the service provider (with accompanying edge security software for anti-spam/anti-virus). The proxy would support LDAP integration with the unified directory (see above) in order to determine which overall system (ZCS or legacy) and which primary server hosts a particular user's mailbox. This front-end routing infrastructure is essential for preservation of service providers' domain email addresses while the migration is underway. (We deem it a non-starter to go with the alternative of introducing a new domain as a way to define independent MX records for the two systems, since it exposes complexity to users.) If POP aggregation is used instead, then it is not necessary for MTA-routing to legacy system or ZCS to be in place for the trial period. This MTA routing must be deployed prior to the final step of an individual migration in which the user's destination mailbox is switched from the legacy system to ZCS. It also must be deployed as a prerequisite to "cap and grow" in which new-users are provisioned solely into ZCS.

- *POP/IMAP proxying infrastructure* − The ZCS distribution includes POP3/IMAP4 proxying capabilities. (IMAP proxying would only be required if a customer elects to offer that service directly to its users) These proxies read from the common directory in order to determine whether to forward

POP/IMAP protocol requests to the appropriate server within the legacy or ZCS system, based on whether the mailbox in question has been migrated. This proxying infrastructure is essential for the preservation of email client references to, for example, "pop.service-provider.net", and would make the migration completely transparent for POP users that do not rely upon a Webmail UI. POP proxying should be in place prior to any beta/trial period, unless beta/trial users would be restricted to only using the Web UI.

- ***Web portal proxying to the appropriate ZCS or legacy web application*** − This allows a user log-in (with single sign-on/SSO) to launch the appropriate application based on whether that user is (1) currently configured to use the legacy system, or (2) has been migrated to ZCS (but might still be relying upon ZCS POP aggregation to receive email originally delivered to the legacy system). This portal proxying function with single-signon should be in place prior to any beta period/trials, since giving users alternative URLs for the new email trial would likely cause confusion and generate support cases. (And, of course, this work has to be done anyway for the subsequent migrations.) ZCS SSO via the service provider portal is easily accomplished with the existing Zimbra Pre-Auth API.

4 Shared-Domain Migration Scenarios

In a Shared-Domain Migration, data is migrated as a batch process (e.g. batches of 100K users). Batches can be determined by the service provider, but we recommend household proximity and geography as the preferred batching function so support calls can be managed by region. Migration plan must aim to minimize users calling support.

After an account is provisioned in ZCS, a user is considered "switched to ZCS" when the following steps are completed:

1. New mail is routed to ZCS, either directly via the MTA tier or indirectly via POP aggregation

2. Web login takes user to ZCS, and POP access is provided via ZCS

3. Mail and contacts in the legacy system has been moved to ZCS

Depending on the duration of regularly scheduled maintenance windows and the service providers' (and consequently users') tolerance to temporary inaccessibility to stored data, the provider may choose from among the following three data migration options:

4.1 Trickle Migration

"Trickle Migration" entails migrating users' data in the background even when the user is switched over to use the features of the new ZCS system. In particular, this

methodology suggests performing steps 1 and 2 mentioned above simultaneously followed by step 3 in the background.

This methodology has the advantage that users experience almost no downtime with respect to their ability to read and send new messages. The downside is that some users may be confused by the fact that some of their legacy mail is still on its way. During the trickle migration, the Zimbra web UI displays "mail migration is in progress, please be patient" prominently to mitigate support issues.

If required, the fallback procedure to revert to using the original mail system during or after migration would be to simply reverse steps 1 and 2 above and run the IMAP sync utility to synchronize new messages delivered to ZCS with the original mailbox. Updated contacts would need to be synchronized to the legacy mailbox via a third-party tool or by utilizing the legacy systems' address book import capabilities.

4.2   One-Shot Migration

"One-Shot Migration" entails migrating all of the data of a group/batch of users prior to granting them access to the new ZCS system. This methodology suggests disabling user/s login, performing step 1 first (mail routing) followed by step 3 (mail migration). Finally, once all data has been migrated, perform step 2 (enable web and POP logins) for the users in question.

The advantage of one-shot migration is that users are not confused by seeing only some of the messages (even if you display the best of warnings, some subset of users may be confused). The disadvantage with one-shot is that there is user downtime during migration—they are unable to login and send or access received mail. However, that downside can be mitigated by doing the migration during the night, and providing an appropriate status message for user's attempting to log-in. Moreover, for most accounts, the downtime to migrate their entire mailbox need not be long.

If IMAP is available in the legacy system, one additional optimization available for "one-shot" migration is that the accounts can be pre-provisioned on Zimbra and a preliminary imapsync can be performed prior to migrating the mailbox "of record." Then when the final migration step is actually done, step 2 takes less time because most data has already been copied.

If required, the procedure to fallback to the legacy system during or after migration would be to reverse steps 1 and 3, using imapsync to sync messages delivered to ZCS—before switching user access (2) to web and POP logins back to the original system. Address book is handled similarly as within Scenario 4.1.

4.3   User-led Migration

User-led migration is the same as one-shot, but allows the user head-of-household to initiate the migration (Compare with: "do you wish to switch your default maps to Yahoo Maps Beta?"). In this scenario, the service provider identifies that a certain

set of users are ready to be migrated. When users from that group login, they are offered the choice to migrate to the new mail system. The advantage here is that users initiate their own downtime, for which they can receive appropriate warning (e.g., "This migration is expected to take 20 minutes."). The disadvantage to this plan is that not all users login via the Web client (such as POP users), so alternative policies must be determined for the migration of their accounts.

Reverting all (or self-selected) households to the original mail system would simply require pointing the proxy to redirect new messages for those users to the original mail system (step 1) and reverse synchronizing user mailboxes from ZCS to the legacy (step 3) prior to switching back access to web and POP logins for those users to the original system (step 2).

5 Suggested Migration Timeline & Coexistence of Legacy System and the ZCS

When the new ZCS and the legacy messaging systems are expected to co-exist for a predetermined period of time, Zimbra suggests the following process to minimize risk and to ensure a smooth migration experience:

1. Begin the migration with a trial period during which the service provider's migration strategy and the new ZCS messaging solution will be validated. During the trial period, the service provider and Zimbra can determine if any critical enhancements need to be incorporated. Trial users can be selected by the service provider or self-selected by giving some number of Web users a trial option at log-in.

2. Following the trial period and the incorporation of any critical enhancements, the service provider would start provisioning new users into ZCS ― the "cap-and-grow" strategy. By this time, the MTA routing tier must be in place.

3. Full migration could begin once the cap-and-grow phase had been completed. The full migration can be expected to take several weeks or months to complete depending on the number of users, bandwidth, and the available maintenance windows to migrate users. In addition to the overhead of copying and indexing all user mailboxes into the Zimbra data stores, there are inherent constraints in bandwidth to move such large amounts of data. For example, to move 100K mailboxes with an average of 20Mb of data each over a ten hour period would require (100K accounts * 20Mbytes/account * 8bits/byte = 16K Gbits) divided by (10 hours * 60 min./hour * 60 sec./min. = 36K sec.) equals <.5Gbits/sec. Sufficient bandwidth will be required to accomplish this level of daily mailbox migrations. Irrespective, Zimbra's migration implementation is sufficiently hardened that lower network bandwidths or other quality of service issues do not introduce any data corruption or integrity problems. Of course, any issues with connecting into the legacy system could extend the overall time required for migration. It would also be safe to validate the IMAP performance and scalability of the legacy system ― it is possible that this can become a more significant gate than bandwidth between the data centers.

6   ZCS Migration Tools

Among the key ZCS migration tools for service providers include IMAP synchronization for migration of messages and REST bindings for migration of address book, calendars, and other user data that has an analog within Zimbra.

6.1   Email

"*imapsync"* is used to copy the email messages from the legacy mailboxes to ZCS. Once this operation is successfully completed, the mailbox of record is moved to ZCS and all future Web mail and POP requests are redirected to ZCS. Careful attention must be paid to optimizing imapsync settings: The syncinternaldates switch is essential for preserving the original delivery/sending dates of the messages; and because Zimbra's IMAP implementation does not include access control lists (ACLs), setting the nosyncacls switch improves efficiency of the migration.

imapsync is an idempotent operation in that if there is any question that it has not been completed successfully (e.g., sanity check of the number of messages in each mailbox), then it can simply be repeated (without recopying any data that has been successfully migrated), although there is clearly a computational cost for doing so. Special consideration must be paid to ensure that messages that arrive during a imapsync, but before the "mailbox of record" has been moved, will be reflected in the new mailbox.

6.2   REST

Zimbra's web application server includes support for Representational State Transfer, or REST. This means that the Zimbra server supports URLs which translate data to and from non-html data types. The following Blog posting includes an explanation and examples of how Zimbra's REST interface works:

http://www.zimbra.com/blog/archives/2006/05/updating_via_re_1.html

Because REST is flexible and very easy to script, it offers a method of automating migration of a wide variety of legacy data. Another advantage is that the Zimbra REST interface allows user data updates by authenticating as an administrator. This means that data can be injected into Zimbra via REST without having to know user passwords.

6.3   Address book

To migrate user contacts into Zimbra using REST, the contact data must first be exported from the legacy server in, for example, vCard (*.vcf) or Comma Separated Value (*.csv) format. Please note that for a CSV based export/import to work properly, the individual field names must match. If the legacy system's CSV formats are different, then the fields will be renamed via scripts or code before injecting into ZCS.

REST migrations are also idempotent in that they can be repeated in the event of failures, but unlike repeated imapsync's, they do require all of the data to be recopied from the legacy to new system.

7 The Zimbra Team's Deployment & Migration Experience

The Zimbra team has extensive experience managing deployments for multi-million-mailbox service-provider deployments, including several migrations of 1M+ mailboxes. A representative list of such migration projects that Zimbra employees worked on at companies including Openwave Systems, Phone.com and Software.com include:

- Rogers: 1+ million mailboxes, @Home to OPWV
- ZDNET: 1+ million mailboxes, legacy IMAP in-house system to Onebox
- Nextel: 1+ million mailboxes, new deployment
- Bell Canada Mobility: from OPWV Mx4 to Mx5
- Bell South: from OPWV Mx4 to Mx6
- Cox: from OPWV Mx4 to Mx5
- Sprint: new deployment 3G messaging
- Telstra: from Qmail to OPWV
- Verizon: from OPWV to iPlanet to Exchange
- TIMnet Brazil: from Sendmail to OPWV

Following are some of the key Zimbra employees who have extensive experience in large-scale migrations. These individuals have helped to build Zimbra's body of knowledge related to migrations and would be involved in various supporting roles with service provider team members who are on the project. This could include providing training and transfer of knowledge the service provider team members.

- **Ari Rubenstein, Sr. Systems Architect:** 2+ years at Zimbra, 4 years at Openwave Systems, 2 years at Onebox, 2 years at Digex (data center ops)

- **Anup Patwardhan, Sr. Systems Architect;** 4 years at Openwave, 2 years at Onebox, 3 years at Sun (including Sun Messaging)

- **Sean Gallagher, Technical Support:** 1 year at Zimbra, 4 years at Openwave, 2+ years at Software.com

- **Anand Palaniswamy, Dir. Server Engineering:** 3+ years at Zimbra; 1 year at BEA Systems; 3 years at Openwave; 3 years at Sun JavaSoft

- **Abinash Tripathy, General Manager, India:** 6 years in Openwave technical product management, 2 years at Software.com, 3 years at Oracle. Abinash has experience managing multi-million-user migrations, and has been building our India team for a variety of migration, deployment, customization, and support services.

Copyright