

Using Dry-Run to ensure Success

Priasoft

Dry-Run?

What is a dry-run? With the Priasoft Migration Suite for Exchange, you have the option to perform dry-run migrations, as many times as you desire, prior to executing the production migration. The dry-run option performs all the same code operations as a production run, but sets up the target objects in an isolated manner so as not to interfere with normal business activities. Furthermore, no source changes are made for the same reasons.

This unique ability give you access to information that other processes (with other tools or vendors) cannot deliver. Total duration is most likely the most valuable metric gained from the dry-run. Knowing up front how long a migration will take – down to the minute – is a powerful data point.

Reasons

There are several key reasons to perform dry-run migrations. The following tasks are performed for a specific purpose and outcome and should be performed in the order presented.

1. Technical Feasibility

The first reason to execute a dry-run is to establish the technical feasibility of the process. In this task you are only attempting to validate that the migration of mailboxes will work within the environment. In most cases you will only select a handful of mailboxes, perhaps even contrived test mailboxes. The outcome of this task will tell you whether you have connectivity issues, configuration issues, and/or environment issues. In general, this task is very short, but should be horizontal enough in scope to validate all of the environment 'touch points', meaning the source and target exchange servers and domain controllers as well as the migration host.

If your environment is large or distributed in the sense that there are multiple source or target exchange servers or DCs, you need to perform this task for each of those resources. Too much risk is assumed by testing only one segment. Just because the migration works from site A does not guarantee that the migration will work from site B.

2. Performance Tuning

The next reason to use the dry-run feature is to tune the performance of the migration. The Priasoft Migration Console has the ability to migrate several mailboxes concurrently, however, the total number of concurrent mailboxes is indeterminable in the beginning of the project. Each environment is different with regards to how it responds to the efforts of a migration. Some environments perform well with many mailboxes migrating at the same time while others do not.

PHONE

602.801.2400

EMAIL

support@priasoft.com

WEB

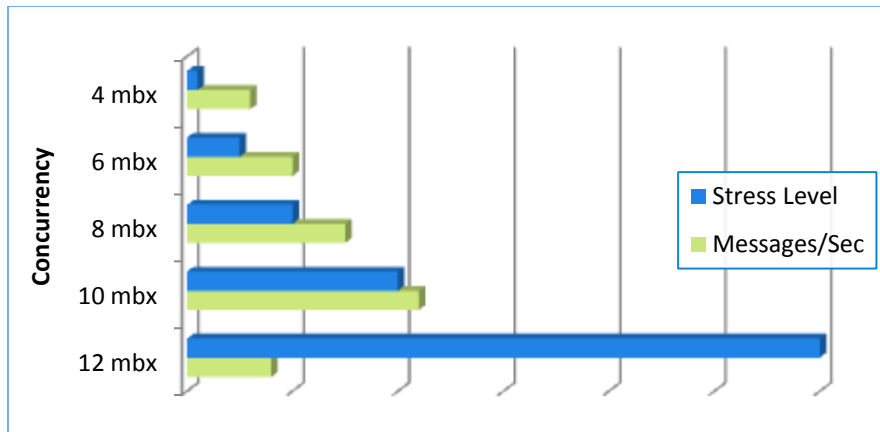
www.priasoft.com

Using the dry-run feature for performance tuning gives you the ability to ratchet up the concurrency in a controllable manner so as to determine the best level of concurrency for the environment and for the demands of the project. Also be aware that the performance tuning is per server, meaning that concurrency of one number from exchange server A may not be valid for exchange server B. Additionally, it is best to dry-run real mailboxes versus test mailboxes. The data in a real mailbox has real complexities and characteristics that a test mailbox would never have such as fragmentation and item age.

The methodology for determining the max concurrency is simple and is outlined as follows:

1. Create a dry-run migration batch with several dozen mailboxes (if you have an environment of such size)
 - a. When selecting mailboxes, choose mailboxes with slightly more than average amounts of data so that they can remain active long enough to analyze the environment. Mailboxes that take 20 minutes or more are sufficient for performance tuning.
 - b. If you only have relatively small mailboxes, this is acceptable, it just means that you may need to perform the dry-run performance tests multiple times with increasing concurrency each time since the batch will complete quickly and without enough time to increase the concurrency before mailboxes start completing.
2. Start with a low maximum concurrency number, perhaps 4 (such is configured in the migration wizard)
3. While the mailboxes migrate, monitor the health of the source and target exchange servers being used over several minutes (at least 5 minutes)
 - a. NOTE: Refer to Appendix A, "What to monitor when migrating" for details on monitoring
4. If it is determined that the environment is not over-stressed by the current activity, increase the concurrency count by 2 or 3 mailboxes (this can be done in the migration console window using buttons above the list of accounts) and then monitor again for several minutes.
 - a. Be sure to wait for the additional mailboxes to start moving mail data before monitoring the environment again. It will take several seconds to a minute before the new mailboxes move into the data copy routines – if you monitor the "Active" view in the console, you will know when the new mailboxes are migrating data as you will see the progress column and item count columns populate with information.
 - b. Repeat from step 3 until you reach a point at which the migration performance degrades (step 5).
5. Your goal is to reach a concurrency number that overcommits the environment and thereby slows the migration. In simplest terms if the msgs/sec of a higher concurrent migration is markedly less than the previous concurrency value, you want to choose the previous concurrency number as the optimal value for the source-to-target path.
6. You may find that you get different optimal concurrency values depending on the source and target exchange servers involved. Furthermore, you may find that the difference is sufficient enough to warrant an additional migration host so that the concurrency settings can be "per path".

Below is a chart that expresses the idea of the performance tuning process



Note:

"StressLevel" in this chart is a pseudo-value that represents many factors. Refer to the section "What to Monitor when Migrating" to understand the details of what to look for when attempting to determine when the environment is overstressed.

Notice in the above chart that the Stress Level becomes overloaded in when 12 mailboxes migrate concurrently. In addition, the messages-per-second is considerably less than the previous level of concurrency. Based on the pattern for the given environment, 10 mailboxes concurrent is the best performing configuration and should be used for further dry-runs and the eventual production run.

The guideline for performance is to attempt to get the messages per second number up as high as possible. However, such should be weighed against the project's requirements as well. Consider a case where the project is allotted 48 hours to complete. If 4 mailboxes concurrent yield an approximate migration run time of 40 hours there may not be much need to improve the performance, unless of course you want to be over with the migration sooner than later.

Performance Influencers

When performing dry-runs to determine ideal performance, you should consider those things that impact performance. Note that the following are general ideas of things to consider. Your environment may have additional influencers.

- Virus Scanners
 - o Exchange "on-access" virus scanners bottleneck performance due to the fact that each time the migrator asks for an item to migrate, the virus scanner jumps in and scans the message first.
- Scheduled Services
 - o Any service that causes contention for the source or target exchange server can impact performance
 - o You should avoid performing migrations during backups or archive operations.
 - o You should avoid migrations if there are any actions that would cause contention with a shared storage system – for instance a complicated database operation that shares the same disks as the exchange server.
- Network Saturation
 - o Any operations that capitalize the network can impact performance

Also be aware to perform your dry-run tasks for performance with the same environment characteristics, as you'd expect for a production run. Environmental differences between dry-run and production can affect the expectations of how long the migration should take.

3. Duration & Metrics

In this task the goal is to dry-run ALL the mailboxes that are part of the project. Two sets of metrics are desired from this task: total duration and sizing metrics. Total duration is important for obvious reasons, especially if you have a defined project timeline of

completion or if you have prescribed windows within which you are allowed to migrate. Knowing the duration up front ensures the project's requirements are met.

In addition, the duration may help guide decisions on some of the options available in the Priasoft Migration Manager application. For instance, the option to use the foreground+backfill feature may be necessary if the full mailbox dry-run shows a run time that is longer than the project requirements. An additional dry-run of all mailboxes with the foreground migration only migrating the last 30 recent days of mail may take considerably less time than a full content migration. Following the foreground migration, the duration of the backfill pass can be established and both durations can be presented to the project manager or the business stakeholders for approval.

Metrics gained from the dry-run are valuable for many reasons. The following are some common points of value are listed below, but you may find others that are more relatable to your project. It's worth exploring the reports and summaries Priasoft provides in order to get the most out of the data.

- Mailbox sizes and counts
 - The mailbox size data provided by the results of a dry-run will be the most accurate value available.
 - If you will have database limits in the target environment, knowing the size of the source mailbox can help determine action to take for large mailboxes
 - You may need to ask the owner to reduce the mailbox size
 - You may need to place the mailbox into a different target database (or create a special database)
 - You may need to set limit overrides on the mailbox to allow it to be stored in the target database
 - You may want to separate large mailboxes into a separate batches (in the migrator) since they will likely take longer to migrate than an average sized mailbox
 - You may find mailboxes that have excessive item counts. This is an important concern; mailboxes with very large item counts [anything over 100,000 items should be considered large] will take a long time to migrate
 - Note that item count (not size) is the primary metric of concern when migration duration is a critical factor
 - An analogy is the duration it takes to load a freight truck: if you have 10,000 small boxes of goods that you must load onto the truck, such will take a considerable amount of time compared to 100 large boxes, even if the large boxes are slow to load because of their weight or dimension.
 - Conversely you may find mailboxes with only a handful of items. Such mailboxes might be worth investigating as they may be mailboxes that do not need to be migrated.
- Mailbox durations
 - You will have access to data that discloses how long a single mailbox takes to migrate. This information can be valuable when trying to setup batches of mailboxes in the sense that it could best to migrate the long running mailboxes first so that they can start earliest. Conversely the opposite may be more valuable in the sense that a greater percentage of the total mailboxes can complete first leaving only the long running mailboxes at the end.
 - Additionally, it may be important to identify long running mailboxes in order to have the owner reduce the mailbox size or item count so as to reduce the time the mailbox takes to migrate

4. Issue Identification

The final purpose, and possibly the most valuable, is the identification of mailboxes that have issues. In most other migration processes the discovery of issues come during the production migration effort. The project quickly becomes stressed as mailboxes fail to migrate or data is lost due to some environmental issue.

Performing dry-run migrations – ultimately a dry-run on all mailboxes – will expose any mailboxes that are of issue – in a non-stressful situation since the issues found do not affect any user directly. Issues found can be worked during normal business hours, which is extremely valuable as often the rectification requires input or interaction with the end user. Once the mailbox is said to be repaired, you can dry-run it again to validate. If it again fails the dry-run, then you can send it back again for further action.

The Process

During a dry run the application creates placeholder windows user accounts in the target for each source mailbox user selected in the migration batch. For each placeholder a mailbox is created and the data and attributes for the source mailbox will be migrated. This is not a simulation – data from the source mailbox will actually be migrated, but the target mailbox will be isolated from the target environment so as not to interfere with any production objects that might already exist. Changes are made to the placeholder so that it is not visible in any Exchange Address Books and all of the email addresses are modified so that no delivery can be made to them. The source objects are not modified in any way and the entire process is a read-only operation out of the source environment. The placeholder objects are safe to delete when they are no longer needed; the migration wizard will prompt you to delete existing dry-run placeholders as well for easy cleanup.

NOTE: YOU SHOULD NEVER ATTEMPT TO MODIFY A DRY RUN OBJECT IS SUCH A WAY AS TO ATTEMPT TO PUT INTO PRODUCTION.

What to monitor when migrating

During a migration, several data values can be monitored in order to evaluate the impact the migration is having on the system and to help tune the migration for optimal performance.

Storage system I/O

Using utilities provided by your storage vendor, you should be able to monitor the read and write queues on the storage system. Monitoring queues is important as it is the primary indicator of over-commitment of resources. In most cases, the read or write queue (source environment = read queue; target environment = write queue) should stay at an average value less than 3. Sustained values above 3 mean that the storage system is being asked for requests 3x as much as it can deliver, and as such it must start delaying requests for data in order to serve current requests. It's ok to see bursts of as much as 12 or more, but should only sustain for 1 or 2 seconds. If you see the average queue value stay above 3 for more than 10 seconds, such is too high and is creating a bottleneck in performance.

In order to determine if the high queue value is related to the migration, pause the migration console and allow the currently active mailboxes complete. Once the active mailboxes start to complete, monitor the queue and if the number drops in response to the mailboxes completing, then you know that the migration concurrency is too high, but only at the moment. Once all the active migrations complete, continue to monitor the queue value. If you find that with no migration activity that the queue value is still elevated or even at 3 or 4, this can mean that there is some other consumer of the

storage platform executing at the same time (perhaps SQL server, for instance). It would be wise at this point to determine what is elevating the queue and to see if it's possible to migrate exclusive of the competing process.

Exchange Server I/O

In addition to the storage system I/O, Microsoft Windows also has performance counters for read and write queues for physical disks. It is advisable to monitor Microsoft's counters on the source and target exchange servers in addition to the values from the storage system. Note that if the storage system is a direct attached array (like a SCSI disk array versus a SAN), the Microsoft Windows counter will be the only value you can monitor.

The Windows counters will show you disk requests that are being generated by services and applications that reside on the Exchange server. If you notice that the Windows counter is very low yet the storage system's counter is considerable higher, this is often an indicator of an external process (again, like SQL server) that is elevating the storage system's I/O.

However, if you find the reverse – the Windows counter is high, but the storage counter is low, this can mean an issue on the exchange server with the storage driver, or possible caching issues in windows. Such a discrepancy should be flagged and evaluated to determine if it is an indicator of a lower level issue, possibly with hardware such as the array controller or SAN controller card.

Migration Console

Another place to monitor activity is in the Priasoft Migration Console application. While data is migrating, you will see activity in the 2 histograms on the right-hand side of the console. During a properly tuned migration, the activity will be fairly consistent and will rise and fall between a fairly static range that is unique to each environment.

When excessive disk queuing occurs, the pattern changes to a noticeable pattern: activity will be shown in both the msgs/sec and the data/sec graphs for several seconds and then both will flat-line to zero for several seconds. The activity then no-activity pattern will repeat with each being a few seconds long (few being typically 5-15 seconds). This pattern is indicative of disk queuing with the period where there is zero activity being the time when the disk system is not sending data due to a have request queue. Once the disk queue is serviced sufficiently, activity resumes.

This pattern has an overall negative effect and can increase the time to migrate by a factor of 3 or more – that means that a migration in one design that takes 4 hours to complete can jump to 12 or more hours just because of an over-committal of resources!

Final Considerations

The above information provides a detailed scope of how to consider and integrate the dry-run into your migration planning. One of the most important concepts to adhere to is NOT to schedule a production migration until you have had a successful dry-run of EVERY mailbox. There may be temptation or pressure from others to dry-run less than everything – this carries a lot of unnecessary risk. Consider that just because 30% (or even 90% for that matter) of the mailboxes successfully dry-run does not guarantee that the rest of the mailboxes will not have issue, especially if those mailboxes that were not sent through the dry-run process are VIPs of any kind.

Best practices

- You should create an OU in the target environment to contain the dry-run objects that are created as part of the process.
 - NOTE: You cannot match a source account to any existing target account when in dry-run mode. You are required to specify a location to create the placeholder accounts.
- You should create mailbox databases in the target that are exclusively made for dry run purposes
 - This prevents white space from being created in existing production databases once the dry run objects are removed
 - This make cleanup easier since the only effort is to delete the database and database file from disk. No need to purge mailboxes.
- All dry run databases should be enabled for circular logging to avoid any database log space issues during the dry run.
 - If you have clustering or Database Availability Groups (Ex2010), it is also ideal to suspend replication during a dry-run (and production) migration. This reduces load on the exchange servers since data is not replicating at the same time as a bulk import of data.
- Dry-run databases should also be set to zero-day retention for deleted mailboxes so that when you delete dry-run mailboxes such are not left behind unnecessarily.