Cisco AMP Threat Grid Appliance Migration Notes

Version 2.2

Last Updated: 3/3/2017

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

Contents

Introduction

The Threat Grid Appliance employs a migration process that will perform migrations in the background, thereby allowing the appliance to remain in active use while the migration is in progress.

This document describes what to expect during migration, and answers frequently asked questions about the upgrade process.

If you have any questions about your migration, please contact Threat Grid Support: support@threatgrid.com

Release 2.2 Migration

Requirement

The ElasticSearch migration in 2.1.5/2.1.6 must be completed before installing 2.2.

What to Expect

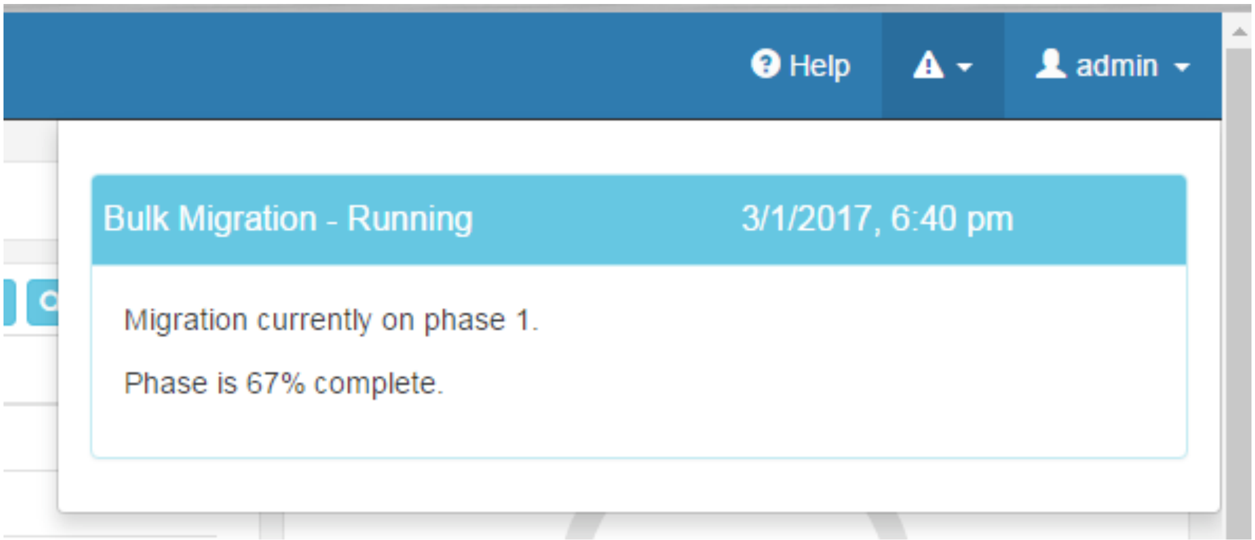
The Threat Grid Appliance 2.2 release includes a significant data migration, which is completed in four phases.

Low-entropy, non-archive-format samples, artifacts and extracted blobs created in Appliance 1.x releases are migrated in Phase 1.

Samples, artifacts and blobs that do not meet one of those above criteria are migrated in Phase 2.

Samples that were migrated in Phase 1 or Phase 2 are made available for download again in Phase 3; whereas Phase 4 converts archival-format storage of analysis results, reports, and other core content into a faster, random-access format.

The migration status is communicated via service notices, including the percentage completed for each phase:



Bulk Migration - Running 3/1/2017, 6:40 pm

Migration currently on phase 1.

Phase is 67% complete.

Time Required

The migration will take a significant amount of time. The amount of time required, particularly during Phase 1, depends entirely upon how actively the appliance was used while on 1.x.

Each appliance will vary depending on setup, usage, and storage, but some migrations have required 20 hours or more. Some have required far less.

Track the status of your migration in the service notices.

Appliance Availability During Migration

The appliance will remain usable during the migration. However, it may not be possible to retrieve the original samples from pre-update submissions while the migration is running: once the pre-update sample has been migrated in Phase 1, the sample will be unavailable for download until it is reprocessed in Phase 3.

Metadata, status, reports, and generally content other than the original samples themselves will remain continuously available.

Full functionality is again available during Phase 4.

Data Issues

The migration will include aligned blobs (disk artifacts) from 1.x. These consist of content that was carved out from a sample's analysis -- disk artifacts, network artifacts, &c.

IMPORTANT NOTE:

Release 2.2 greatly increases storage efficiency to make disk capacity available that had not been usable on systems initially installed with 1.x releases. This implements pruning - removal - of old content in the future.

Old content will be migrated in 2.2. However, older content - especially carved disk and network artifacts that are produced in extremely high volume and only rarely used - ***is no longer retained indefinitely***.

For more information, please see the related document, *Threat Grid Appliance Data Retention*:

http://www.cisco.com/c/dam/en/us/td/docs/security/amp_threatgrid/amp-threat-grid-appliance-data-retention.pdf.

Release 2.1.5 Migration

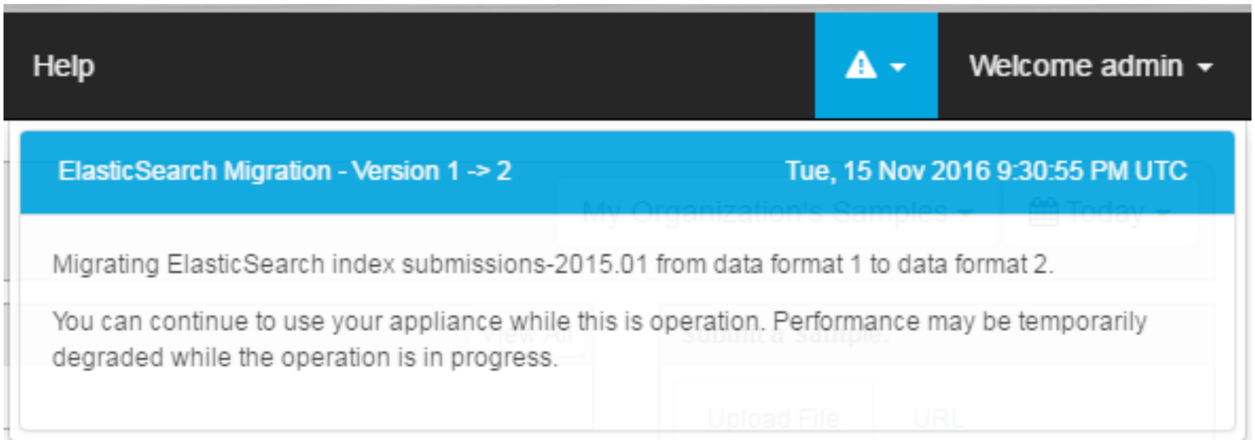
The 2.1.5 Threat Grid Appliance release updates the datastore that backs various search functionalities in the application, most notably the API calls that back appliance integration with Cisco ESA and WSA devices. This datastore could become extremely slow when the quantity of contents grew large, causing API calls to frequently fail or time out. The new version resolves these issues, retaining high performance even with high volume. However, upgrading to the new version is a time-intensive process, which requires rebuilding database indexes twice.

What to Expect

First, upgrade your appliance to version 2.1.5.

Phase 1

Immediately after upgrading your appliance, the CSA API performance will be similar to what it was before the migration. Within 10 minutes after boot, however, you should see a service notice similar to the following:



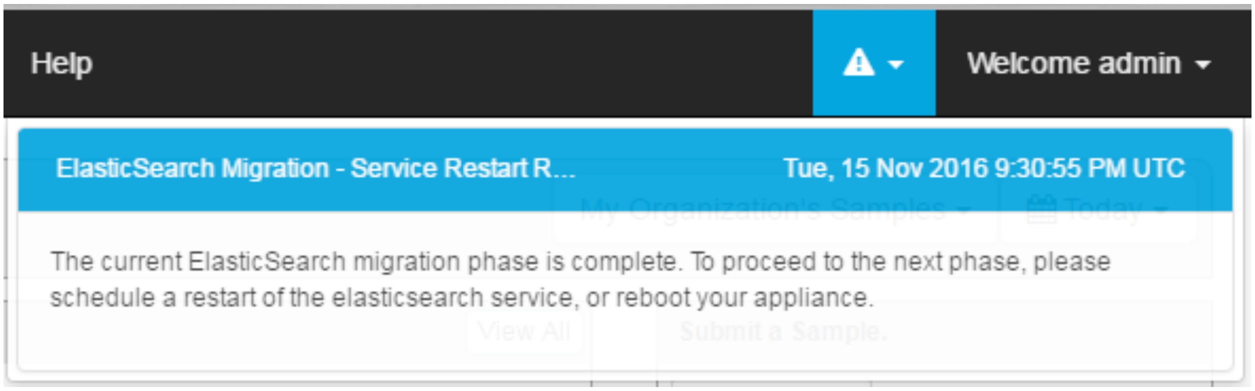
ElasticSearch Migration - Version 1 -> 2 Tue, 15 Nov 2016 9:30:55 PM UTC

Migrating ElasticSearch index submissions-2015.01 from data format 1 to data format 2.

You can continue to use your appliance while this is operation. Performance may be temporarily degraded while the operation is in progress.

This notice will be periodically updated to migrate data from different months.

After all of the ElasticSearch version 1 content has been upgraded to version 2, the system will then prompt asking for a reboot:



ElasticSearch Migration - Service Restart R... Tue, 15 Nov 2016 9:30:55 PM UTC

The current ElasticSearch migration phase is complete. To proceed to the next phase, please schedule a restart of the elasticsearch service, or reboot your appliance.

Phase 2

After the reboot has completed, the system will then migrate content in the background from version 2 to version 5, with a notice similar to that used for the migration from version 1 to version 2.

After this migration is complete, the notice will be closed, and no future service notices will be generated.

Frequently Asked Questions

What kind of downtime window(s) do I need to schedule?

Installing the 2.1.5 release will take no longer than any other typical appliance upgrade, and should be less than 30 minutes. The system should then be immediately usable. After a period of time which varies by the amount of data on the system (see *How long will this take?* below), the system will announce that it needs to be rebooted to start the second phase of the migration process. This reboot will take only the usual amount of time required for an appliance reboot, typically less than 10 minutes. No further downtime is required.

How long will this take?

Phase 1 of the migration runs at approximately 150,000 samples/hour when otherwise idle, or 75,000 samples/hour when heavily loaded.

Thus, for a TG5500 appliance continually using its entire 5000/day sample rate, one hour of migration time in Phase 1 is to be expected per month of operation if unloaded, or two hours per month of operation if heavily loaded during the migration process.

For a TG5000 appliance continually using its entire 1500/day sample rate, Phase 1 should require roughly 20 minutes per month of operation is expected if unloaded, or 40 minutes per month of operation if heavily loaded.

Phase 2 should require slightly less time than Phase 1; however, since performance is already improved at the start of Phase 2 (and no further reboot or other maintenance operation is required at its end), exact timing for its completion is less critical.

When will I see performance improvements?

After the reboot separating Phase 1 from Phase 2, performance will be significantly improved from what was seen with prior releases, though it will not yet be at the rates expected after the completion of Phase 2.

After Phase 2 completes, CSA API performance issues should be completely and entirely resolved.

Are there any functional caveats?

There is a very brief period near the end of the migration of each index where database contents added to that index during the migration process itself can be duplicated twice in search results.

Does the migration process modify my appliance's storage requirements?

The storage format changes allowing faster search do increase disk storage requirements. However, we've also enabled newly available options to use improved compression, largely negating this effect.

In our QA lab, we've seen the net result typically be an increase in storage requirements for the data/elasticsearch filesystem on the scale of 13%.

How much disk space needs to be available for the migration to succeed?

Roughly three times the size of the largest index should be available to allow for transient storage requirements.

The largest index sizes seen in practice tend to be roughly 15gb; thus, customers with less than 45gb of disk space available should contact customer support before attempting this upgrade.

How can I monitor the progress of the operation?

If email notifications are configured for the appliance, status regarding the upgrade will be sent to email. Regardless of whether email notifications are configured, logging into the Threat Grid application and clicking on the triangle-and-exclamation-point icon at the top-right side of the screen (if present) will display any service notices which are pending.

What was the cause of the performance problem?

With prior versions of ElasticSearch, database queries used to respond to /report/list and /report/latest API calls generated an inordinate amount of I/O load. This would also cause other API calls referring to the same store -- such as disposition queries -- to be greatly slowed even if they would otherwise perform well on their own.

Why does this process require two stages?

ElasticSearch 5 cannot directly load or migrate the on-disk data format used by ElasticSearch 1. Consequently, it is necessary to migrate data to ElasticSearch 2 as an interim step.

I'd prefer to make my maintenance window between Phase 1 and Phase 2 as short as possible; can I avoid rebooting?

It's possible to restart only the ElasticSearch service between Phase 1 and Phase 2, rather than rebooting the entire system. There are some caveats:

- If using the `tgsh` interface to perform this service restart yourself, note that the first attempt to start ElasticSearch after an upgrade may report failure. This is normal and expected: During service startup, we perform a check as to whether a later version of the service which is available and installed on the system will be successfully able to operate this datastore; if so, we then shut down the just-started version of the service and report a failure, which will cause an automatic restart to take place later. Telling the ElasticSearch service to attempt to start a second time should succeed.

- API calls from application components may fail for up to 30 seconds due to attempts to use ElasticSearch queries built for a prior version of the database. When the 30-second timeout is met, the application will query the datastore for its version information and begin using version-appropriate queries, so this is a strictly transient error.

I just rebooted my appliance, but the message saying I need to reboot is still there.

The appliance checks its migration status, and starts any required or pending operations -- or clears migration-related service notices which are no longer relevant -- up to ten minutes after it was booted. Prior to that time, a message from before it was booted may be displayed.