Quest® InTrust 11.4.2

Repository Indexing for Advanced Search Capabilities

If you have any questions regarding your potential use of this material, contact:

Quest Software Inc.

Attn: LEGAL Dept

4 Polaris Way

Aliso Viejo, CA 92656

Refer to our Web site (https://www.quest.com) for regional and international office information.

Patents

Quest Software is proud of our advanced technology. Patents and pending patents may apply to this product. For the most current information about applicable patents for this product, please visit our website at https://www.quest.com/legal.

Trademarks

Quest, the Quest logo, and Join the Innovation are trademarks and registered trademarks of Quest Software Inc. For a complete list of Quest marks, visit https://www.quest.com/legal/trademark-information.aspx. All other trademarks and registered trademarks are property of their respective owners.

Legend

CAUTION: A CAUTION icon indicates potential damage to hardware or loss of data if instructions are not followed.

IMPORTANT, **NOTE**, **TIP**, **MOBILE**, or **VIDEO**: An information icon indicates supporting information.

InTrust Repository Indexing for Advanced Search Capabilities
Updated - September 2020
Version - 11.4.2

Contents

About Repository Indexing

InTrust uses indexing on repositories for fast searching and data retrieval. Indexing is optional, but it gives you the following benefits:

- You get real-time feedback as you browse audit data.
- You do not need to configure reporting for data analysis, because interactive data filtering, organization and sorting functionality is available in Repository Viewer.

Centera-based repositories currently do not support indexing.

Note that repository indexing is a resource-intensive activity. To increase indexing performance, set up a powerful dedicated computer specifically for this purpose, as described in the Dedicated Indexing topic.

When you view the contents of a repository in Repository Viewer, the most recent events are found first.

> **CAUTION: In InTrust versions prior to 10.4, events were not prioritized by how recent they are. Events gathered to an indexed repository by InTrust 10.4 or earlier will not be prioritized in the current version of Repository Viewer—they will be found in arbitrary order.**
>
> **However, in a non-indexed repository events will be prioritized correctly, even though searches will be slower.**

Estimating the Resources Required for Indexing

Performance tests conducted by InTrust quality control on repositories with typical heterogeneous data show the following results.

Conditions

- There is only one InTrust server in the organization.
- Indexing and real-time event collection are the only activities that take place in the tests.
- Every agent processes data from 6 to 10 active event logs, and the rate of event generation is about 50 events per second per agent.
- The InTrust server is a virtual machine. Two virtual machine host configurations are used in tests:
    - VMware ESXi on Dell PowerEdge 1950 with Intel® Xeon® E5310 processors, 2 sockets, 4 cores per socket
    - Hyper-V Server on Windows 2012 R2 on Dell PowerEdge R620 with Intel® Xeon® E5-2670 processors, 2 sockets, 16 cores per socket, hyperthreading is on

Results

On the E5310 configuration

Number of processors	Number of agents	Repository growth rate
4	200	36MB/sec
8	300	54MB/sec

On the E5-2670 configuration

Number of processors	Number of agents	Repository growth rate
4	500	90MB/sec
8	800	144MB/sec
16	1100	198MB/sec

Configuration-independent repository growth statistics (including index)

Number of agents	Repository growth per minute, bytes	Repository growth per hour, bytes	Repository growth per day, bytes	Repository growth per month, bytes
100	9,750,000	585,000,000	14,040,000,000	421,200,000,000
500	48,750,000	2,925,000,000	70,200,000,000	2,106,000,000,000
900	87,750,000	5,265,000,000	126,360,000,000	3,790,800,000,000

Estimating in Your Environment

The typical hardware requirements for repository indexing are outlined above. If the conditions in your environment differ—for example, if the indexing computers also perform a lot of audit data gathering or monitor a lot of sites—then you need to adjust your estimates.

To estimate the required disk space, you need to know how much the actual audit data takes up. For that, the most readily available solution is to use the **dir** command with the **/s** switch.

The size of the fully indexed repository is approximately twice the data size.

Troubleshooting

When you configure indexing, it is useful to track indexing-related events in the InTrust Server event log. For details about these events, see Events from InTrust Repository Services. The following tips will only indicate the event IDs, which you can look up in that topic.

Event IDs 13875, 13877, 13878

These warnings mean the number of files that haven't been indexed has exceeded a threshold value. This warning normally indicates a temporary state. For example, it may keep recurring for only four hours a day.

Event ID 8321

This error indicates that the InTrust server's event queue for a particular agent has overflowed. Reduce the activity of that agent.

Mind Your Disk Queue

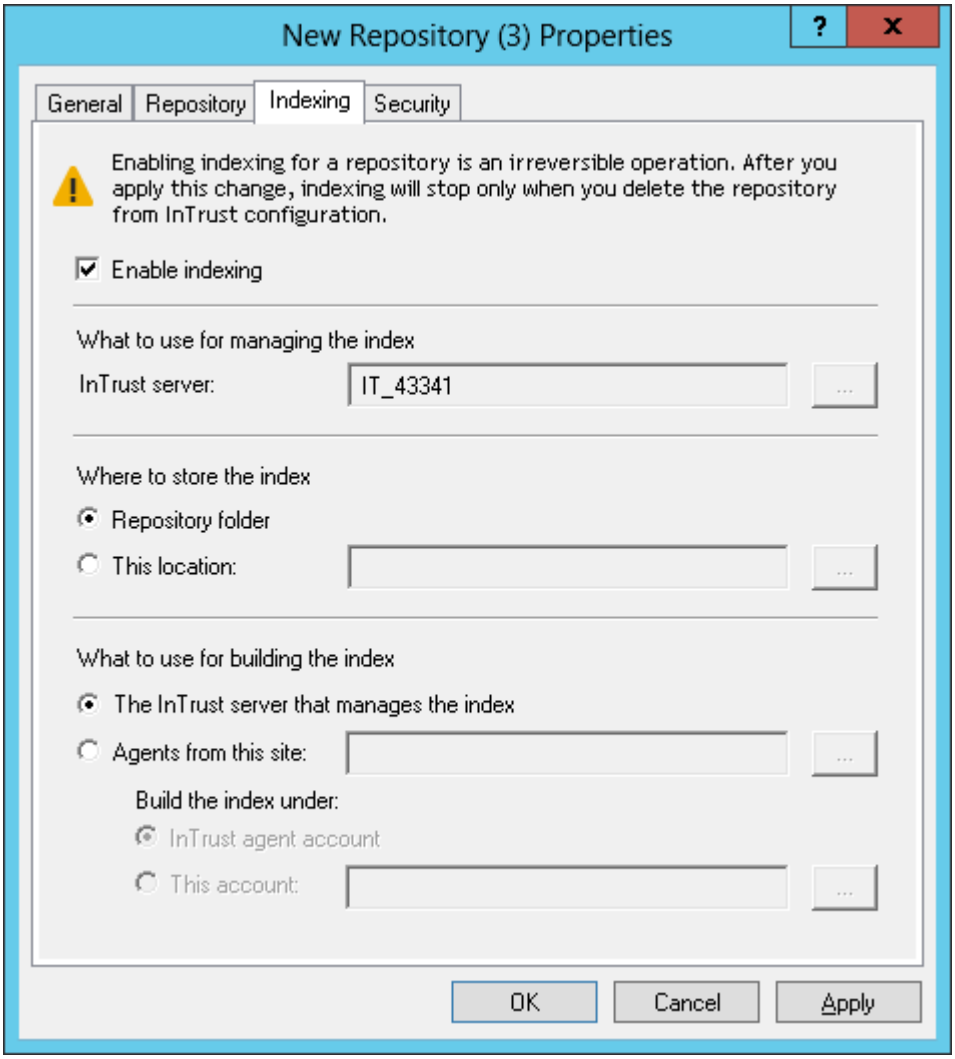
The average disk queue length on the indexing InTrust server's disk and on the disk that contains the repository should not exceed the value **2**. A higher number means the disk is a bottleneck resource.

Mind Your Bandwidth

Make sure you have enough bandwidth to accommodate the traffic generated by all the agents that talk to the InTrust server.

Configuring Indexing

If a repository is created in InTrust Deployment Manager, indexing is enabled automatically for it, but if it is created in InTrust Manager, indexing is disabled by default. To configure indexing options for an existing repository, open its properties in InTrust Manager and go to the **Indexing** tab.



This tab lets you do the following:

- Enable indexing capabilities for the repository.
  This is an irreversible operation; indexing of the repository will continue until it is deleted from InTrust

configuration.

- Select the InTrust server that will manage the index-processing operations.
- Specify the location of the index files.
  The most convenient method is to place the index in the repository folder (this is the default option).
- Configure whether the indexing work is offloaded elsewhere.
  You can make the index-managing server also build the index, or you can use dedicated computers for the purpose. For details about offloading indexing, see the Dedicated Indexing topic. If you select to set up dedicated indexing, you have the option to specify the account that will be used for it.

Make sure that Active Directory delegation is enabled for the following:

- The computer account of the InTrust server that manages the index
- The user accounts that perform indexing

Best Practice: Lean Indexed Short-Term Repository

If you use InTrust Manager, the recommended setup is to have more than one repository:

1. An indexed repository with the most current data
2. Another repository (indexed or non-indexed) for archival of data that is no longer current

A freshly-deployed default repository is indexed. In an upgraded InTrust deployment, all repositories keep their indexing settings after the upgrade, so if indexing was disabled prior to the upgrade, it is not enabled automatically.

To keep the short-term repository current, set up regular repository cleanup jobs that clear all data older than specified. To move data from the short-term repository to the archive, use regular cleanup jobs that occur straight after the consolidation.

This configuration helps achieve the following:

- Maximize indexing performance
- Keep the current repository lean, fast and at a roughly constant volume

The size of your short-term repository depends on your auditing needs. To prevent indexing from slowing down the auditing workflow, you may want to consider the suggestions in the Dedicated Indexing topic.

Adjusting the Task-Based Gathering Workflow for Repository Indexing

Getting the most out of InTrust repository indexing requires that you organize your gathering workflow in a particular way. To follow the best-practice recommendation of using a short-term repository and an archive repository, take the following steps:

1. If you already have a production repository with data, make an archive repository based on it, as described in the **To convert an existing repository to a short-term repository** procedure below. If you only have an empty default repository (in a fresh InTrust deployment), simply create an additional repository for archival purposes.
2. Prepare your production repository to be used as the short-term repository, as described in the **To convert an existing repository to a short-term repository** procedure below.

To convert an existing repository to a short-term repository

1. In InTrust Manager, open the properties of the repository and note its path. You will need this path later.
2. Change the path. The new path will be used for short-term storage.
3. Create a new repository in InTrust. In the New Repository wizard, specify the old repository location from step 1.
   The resulting repository is ready to be used for audit data archival, and the original repository is ready for regular data extraction and cleanup.

NOTE: In InTrust Manager, access to repositories is based on IDs, not names or paths. This means it does not lose track of a repository no matter how you rename or relocate it. However, if you delete a repository and recreate it with the old name and path, it will not be recognized as the old repository. Therefore:

- To repurpose a repository without disrupting the workflows associated with it, rename it.
- To relocate a repository without disrupting the workflows associated with it, change its path.
- If you don't mind losing associations with any current workflows, create a new repository. You can use the path from a previously deleted repository to populate it.

To organize archival of repository data

1. Decide on the retention period in the short-term repository—for example, 90 days. To make the decision, take into account how far back the events you view usually date.

2. In InTrust Manager, create a task with the following jobs:
    - A consolidation job that copies the entire Windows network-related contents of the short-term repository to the archive repository; make this job the successor of the repository cleanup job
    - If applicable, a consolidation job that copies the entire Unix network-related contents of the short-term repository to the archive repository; make this job the successor of the repository cleanup job
    - A repository cleanup job that clears from the short-term repository all data older than your preferred retention period
3. Schedule the task to run at intervals equal to your retention period.

Handling the Archive Repository

The archive repository has a lower priority than the short-term repository. It is up to you whether you want the archive repository to be indexed. If you need it indexed, you can speed up the process by setting up dedicated indexing, as described in the Dedicated Indexing topic.

Remember to provide ample hard disk space, because it is going to be consumed by the gradual repository growth and, if indexing is enabled, by the index.

Dedicated Indexing

Tests have shown that using a dedicated index-processing computer makes indexing take roughly half as long, compared to using an InTrust server that is also involved in other activity. It also helps avoid load spikes on the production InTrust server at certain stages of the indexing process.

To decide whether dedicated indexing is really needed, answer the following questions:

- Can your existing InTrust servers handle the indexing management, or do you need one or more dedicated computers just for the indexing?
- What other work will the InTrust servers do besides indexing?

The Estimating the Resources Required for Indexing topic can help you find the answers.

If you choose dedicated indexing, ensure the following:

- An InTrust agent is deployed on the computer or computers you want to use for indexing.

> **CAUTION: If the index-managing InTrust server and the indexing computer are separate machines and the indexing computer is also an InTrust server, make sure the indexing computer does not list the index-managing server as an agent. Otherwise, all activity where the two servers involve each other (gathering, indexing and so on) will fail due to circular server–agent dependencies.**
>
> **To see the list of an InTrust server's agents, select the Quest InTrust Manager | Configuration | InTrust Servers | <*server_name*> | Agents node.**

- The following ports are open for incoming traffic on the indexing computers:
  - RPC ports: 135, 445
  - Dynamic ports: 1024–65535
- The index-managing InTrust server and the indexing computers must be accessible to each other by DNS name.

Enabling Dedicated Indexing

After you have planned dedicated indexing and prepared the computers, perform the configuration in InTrust Manager.

First, create a dedicated site with a meaningful name such as “Indexing helpers” and include your dedicated indexing workers in the site.

Next, configure indexing for the repository:

1. Open the repository properties, and select the **Indexing** tab.
2. Specify the index-managing InTrust server.

3. Select the **On agents from this site** option.
4. Specify the site you have created.
5. Specify the account to use for indexing activity. You can use the InTrust server account, which has all the necessary privileges. If you would prefer a less powerful account for indexing, make sure it has the permissions listed for "Perform indexing of a production repository" in Minimal Rights and Permissions Required for InTrust Operations.

Indexing Idle Repositories Without InTrust

The workflow described above cannot apply to non-indexed idle repositories that are not managed by any InTrust servers. To create an index for such a repository, use a standalone instance of the **IndexingTool.exe** command-line utility.

This is the utility that InTrust servers and agents use automatically to perform indexing activity, but it can also be installed separately using the **INDEXING_TOOL.*.*.*.msi** package or as part of the Repository Viewer setup using **IT_RV.*.*.*.msi**.

The syntax for indexing of idle repositories is as follows:

- Create an index:
  `indexingtool.exe -create -local <index_path> <rep_path> [-threads [how_many]]`
- Synchronize the index:
  `indexingtool.exe -sync -local <index_path> <rep_path> [-threads [how_many]]`
- Delete the index:
  `indexingtool.exe -delete -local <index_path> <rep_path> [-threads [how_many]]`

The **-threads** option sets how many threads indexing will use. If this option is omitted, the number of threads will be the same as the number of CPU cores. The following values can be used:

- -threads
  Use as many threads as there are CPU cores.
- -threads -1
  Use one less threads than there are CPU cores, but no less than one thread.
- -threads N
  Use N threads.

Values less that **-1** will cause errors. The default value is **-1** (one less threads than there are CPU cores).

CAUTION:

- **When the IndexingTool.exe utility is used in standalone mode, a repository can be processed by only one instance of the utility at a time.**
- **You cannot open an idle repository in Repository Viewer while it is being indexed.**

Tracking Indexing Progress

Indexing progress is recorded to the InTrust Server event log. For a list of specific events, see Events from InTrust Repository Services.

To track the progress, gather the InTrust Server log and view the events in Repository Viewer.

Recreating the Index

In some situations, the index for a repository needs to be rebuilt from scratch—for example, if it becomes damaged. To recreate the index for a repository, take the following steps:

1. If you have multiple InTrust servers, find out which server manages the repository. To look up the server, open the properties of the repository and go to the **Indexing** tab.
2. Stop the **Quest InTrust Server**, **Quest InTrust Real-Time Monitoring Server** and **Quest InTrust Agent** services on that server.
3. Delete the **IndexingRoot$** folder, which stores the indexing data. This folder can be in the following locations:
    - If the **Store index in** option is set to **Repository folder** (in the repository properties on the **Indexing** tab), **IndexingRoot$** is in the folder specified on the **Repository** tab.
    - If the **Store index in option** is set to **This location**, then the location of **IndexingRoot$** is specified there. In addition, you also need to delete **IndexingRoot$** in the repository folder (look it up on the **Repository** tab).
4. Start the services from step 2 again.

After you have done this, the index will be recreated automatically, but it can take a very long time.

Notable Index Implementation Specifics

Where the Index Data Is Located

InTrust stores indexing information for a repository in a set of files located either in a folder on the InTrust server or in a network share on a remote computer.

To specify the location of the index data, open the properties of the repository, go to the **Indexing** tab and use the **Store index in this location** field.

Where the Processing Occurs

In the simplest case, the index of a repository is processed by a single InTrust server. This server is specified as the **Index manager** server in the repository properties.

However, if multiple repositories refer to the same InTrust server for indexing, this affects performance considerably. For better scalability, use a dedicated indexing worker, as described in the Dedicated Indexing topic.

What Happens to Index Data for Missing Events

When audit data is removed from a repository (for example, by a repository cleanup job), index data related to the removed events is not purged immediately; this cleanup operation starts automatically within 24 hours. If a lot of such orphaned index data stays behind, this may slightly slow down repository search.

Index data is made consistent automatically the next time the indexing utility runs—it starts every minute to process the newest events.

The Index Is Not Transferable

If you transfer the contents of a repository to another repository through a consolidation job, the index data from the source repository is not integrated with the target repository. All new data coming into the target repository needs to be indexed anew.

Therefore, indexing is not necessary in such situations, unless the source repository and its index are in frequent use.

Quest creates software solutions that make the benefits of new technology real in an increasingly complex IT landscape. From database and systems management, to Active Directory and Office 365 management, and cyber security resilience, Quest helps customers solve their next IT challenge now. Around the globe, more than 130,000 companies and 95% of the Fortune 500 count on Quest to deliver proactive management and monitoring for the next enterprise initiative, find the next solution for complex Microsoft challenges and stay ahead of the next threat. Quest Software. Where next meets now. For more information, visit www.quest.com.

Contacting Quest

For sales or other inquiries, visit www.quest.com/contact.

Technical support resources

Technical support is available to Quest customers with a valid maintenance contract and customers who have trial versions. You can access the Quest Support Portal at https://support.quest.com.

The Support Portal provides self-help tools you can use to solve problems quickly and independently, 24 hours a day, 365 days a year. The Support Portal enables you to:

- Submit and manage a Service Request
- View Knowledge Base articles
- Sign up for product notifications
- Download software and technical documentation
- View how-to-videos
- Engage in community discussions
- Chat with support engineers online
- View services to assist you with your product