

Terracotta Server Administration Guide

Version 10.3

April 2019

This document applies to Terracotta DB and Terracotta Ehcache Version 10.3 and to all subsequent releases.

Specifications contained herein are subject to change and these changes will be reported in subsequent release notes or new editions.

Copyright © 2010-2019 Software AG, Darmstadt, Germany and/or Software AG USA Inc., Reston, VA, USA, and/or its subsidiaries and/or its affiliates and/or their licensors.

The name Software AG and all Software AG product names are either trademarks or registered trademarks of Software AG and/or Software AG USA Inc. and/or its subsidiaries and/or its affiliates and/or their licensors. Other company and product names mentioned herein may be trademarks of their respective owners.

Detailed information on trademarks and patents owned by Software AG and/or its subsidiaries is located at <http://softwareag.com/licenses>.

Use of this software is subject to adherence to Software AG's licensing conditions and terms. These terms are part of the product documentation, located at <http://softwareag.com/licenses> and/or in the root installation directory of the licensed product(s).

This software may include portions of third-party products. For third-party copyright notices, license terms, additional rights or restrictions, please refer to "License Texts, Copyright Notices and Disclaimers of Third Party Products". For certain specific third-party license restrictions, please refer to section E of the Legal Notices available under "License Terms and Conditions for Use of Software AG Products / Copyright and Trademark Notices of Software AG Products". These documents are part of the product documentation, located at <http://softwareag.com/licenses> and/or in the root installation directory of the licensed product(s).

Use, reproduction, transfer, publication or disclosure is prohibited except as specifically provided for in your License Agreement with Software AG.

Table of Contents

About This Documentation.....	5
Online Information and Support.....	5
Data Protection.....	6
Cluster Architecture.....	7
Active and Passive Servers.....	9
Clients in a Cluster.....	13
Electing an Active Server.....	15
Failover.....	17
Starting and Stopping the Terracotta Server.....	19
Safe Cluster Shutdown and Restart Procedure.....	21
The Terracotta Configuration File.....	23
Configuring the Terracotta Server.....	27
System Recommendations for Hybrid Caching.....	31
System Recommendations for Fast Restart (FRS).....	33
Failover Tuning.....	35
Connection Leasing.....	41
Cluster Tool.....	43
Licensing.....	63
Backup, Restore and Data Migration.....	65
Overview.....	66
Data Directory Structure.....	66
Online Backup.....	67
Offline Backup.....	69
Restore.....	69
Data Migration of Ehcache data.....	70
Technical Details.....	71
Restarting a Stripe.....	73
IPv6 support in Terracotta DB.....	75

SSL / TLS Security Configuration in Terracotta.....	79
Security Core Concepts.....	80
Cluster Security.....	85
TMS Security.....	91
LDAP Properties.....	96
SSL / TLS Troubleshooting guide.....	98
Terracotta Server Migration from Terracotta BigMemory Max v4 to Terracotta DB v10.....	107
Using Command Central to Manage Terracotta.....	109

About This Documentation

Online Information and Support

Software AG Documentation Website

You can find documentation on the Software AG Documentation website at "<http://documentation.softwareag.com>". The site requires credentials for Software AG's Product Support site Empower. If you do not have Empower credentials, you must use the TECHcommunity website.

Software AG Empower Product Support Website

If you do not yet have an account for Empower, send an email to "empower@softwareag.com" with your name, company, and company email address and request an account.

Once you have an account, you can open Support Incidents online via the eService section of Empower at "<https://empower.softwareag.com/>".

You can find product information on the Software AG Empower Product Support website at "<https://empower.softwareag.com/>".

To submit feature/enhancement requests, get information about product availability, and download products, go to "[Products](#)".

To get information about fixes and to read early warnings, technical papers, and knowledge base articles, go to the "[Knowledge Center](#)".

If you have any questions, you can find a local or toll-free number for your country in our Global Support Contact Directory at "https://empower.softwareag.com/public_directory.asp" and give us a call.

Software AG TECHcommunity

You can find documentation and other technical information on the Software AG TECHcommunity website at "<http://techcommunity.softwareag.com>". You can:

- Access product documentation, if you have TECHcommunity credentials. If you do not, you will need to register and specify "Documentation" as an area of interest.
- Access articles, code samples, demos, and tutorials.
- Use the online discussion forums, moderated by Software AG professionals, to ask questions, discuss best practices, and learn how other customers are using Software AG technology.
- Link to external websites that discuss open standards and web technology.

Data Protection

Software AG products provide functionality with respect to processing of personal data according to the EU General Data Protection Regulation (GDPR). Where applicable, appropriate steps are documented in the respective administration documentation.

1 Cluster Architecture

The Terracotta cluster can be viewed topologically as a collection of *clients* communicating with a Terracotta Server Array (TSA).

The server array can be broken down into one or more logically independent stripes. The total storage capacity of the TSA can be increased with the addition of more stripes.

A stripe can be broken down into one or more servers. Each stripe contains a single *active server* and zero or more *passive servers*. These stripe members are all configured by a single Terracotta configuration file. Refer to the section [“The Terracotta Configuration File” on page 23](#) for more details.

For more information on clients, active servers and passive servers, see the sections [“Clients in a Cluster” on page 13](#) and [“Active and Passive Servers” on page 9](#).

TSA Topologies

There are multiple types of TSA topology, each offering different resource and availability capabilities.

TSA Topology	Description
Single-server cluster	<p>This is a TSA which consists of one stripe containing a single server. This server is always the active server.</p> <p>This scheme offers the least amount of both resource and availability capabilities:</p> <ul style="list-style-type: none"> ■ If this server should become unavailable, your client endpoints will fail to operate. ■ The resource services exposed to your clients are limited to those of the underlying server JVM and OS.
High-availability cluster	<p>This refers to a TSA where each stripe consists of at least two servers. In addition to the active server there will be at least one passive server. The stripe will continue operating in the event of an active server failure, as long as at least one passive server is available.</p> <p>Note that stripes do not share passive servers, so each stripe will need at least one passive to possess high-availability.</p>
Multi-stripe cluster	<p>Multi-stripe refers to a TSA that consists of multiple independent stripes. This scheme offers the ability for</p>

TSA Topology	Description
	increased storage, with each stripe contributing to the available total amount of storage. For a multi-stripe TSA to possess high-availability, each stripe must consist of more than one server. This setup offers the maximum of both resource and availability capabilities.

Client perspective

Each client is logically independent of other clients. It sees the TSA as a collection of one or more stripes. It connects to the active server of each stripe in order to issue messages to the cluster.

Stripe perspective

Each stripe is logically independent of other stripes in the TSA. Each stripe member only concerns itself with the clients connected to it and its sibling servers.

Specifically, the active server is the key point in each stripe: each stripe has exactly one active server and it is this server which interacts directly with each connected client and each passive server within the same stripe.

2 Active and Passive Servers

Introduction

Terracotta Servers exist in two modes, *active* and *passive*. The description of each mode is given below.

Active servers

Within a given stripe of a cluster, there is always an active server. A server in a single-server stripe is always the active server. A multi-server stripe will only ever have one active server at a given point in time.

The active server is the server which clients communicate with directly. The active server relays messages on to the *passive servers* independently.

How an active server is chosen

When a stripe starts up, or a failover occurs, the online servers perform an *election* to decide which one will become the active server and lead the stripe. For more information about elections, see the section [“Electing an Active Server” on page 15](#).

How clients find the active server

Clients will attempt to connect to each server in the stripe, and only the active server will accept the connection.

The client will continue to only interact with this server until the connection is broken. It then attempts the other servers if there has been a *failover*. For more information about failover, see the section [“Failover” on page 17](#).

Responsibilities of the active server

The active server differs from passive servers in that it receives all messages from the clients. It is then responsible for sending back responses to the calling clients.

Additionally, the active server is responsible for replicating the messages that it receives on the passive servers.

When a new server joins the stripe, the active server is responsible for synchronizing its internal state with the new server, before telling it to enter a standby state. This state means that the new server is now a valid candidate to become a new active server in the case of a failover.

Passive servers

Any stripe of a cluster which has more than one running server will contain passive servers. While there is only one *active* server per stripe, there can be zero, one, or several passive servers.

Passive servers go through multiple states before being available for failover:

<i>UNINITIALIZED</i>	This passive server has just joined the stripe and has no data.
<i>SYNCHRONIZING</i>	This passive server is receiving the current state from the active server. It has some of the stripe data but not yet enough to participate in <i>failover</i> .
<i>STANDBY</i>	This passive contains the stripe data and can be a candidate to become the active, in the case of a failover.

Passive servers only communicate with the active server, not with each other, and not with any clients.

How a server becomes passive

When a stripe starts up and a server fails to win the election, it becomes a passive server.

Additionally, newly-started servers which join an existing stripe which already has an active server will become passive servers.

Responsibilities of the passive server

The passive server has far fewer responsibilities than an active server. It only receives messages from the active server, not communicating directly with other passive servers or any clients interacting with the stripe.

Its key responsibility is to be ready to take over the role of the active server in the case that the active server crashes, loses power/network, or is taken offline for maintenance/upgrade activities.

All the passive server does is apply messages which come from the active server, whether the initial state synchronization messages when the passive server first joined, or the on-going replication of new messages. This means that the state of the passive server is considered consistent with that of the active server.

Lifecycle of the passive server

When a passive server first joins a stripe and determines that its role will be passive, it is in the *UNINITIALIZED* state.

If it is a *restartable* server and also discovers existing data from a previous run, it makes a backup of that data for safety reasons. Refer to the section [“Clearing Automatic Backup Data” on page 11](#) for more details.

Refer to the section *Restarting a Stripe* in the *Terracotta Server Administration Guide* for information on the proper order in which to restart a *restartable* stripe.

From here, the active server begins sending it messages to rebuild the active server's current state on the passive server. This puts the passive server into the *SYNCHRONIZING* state.

Once the entire active state has been synchronized to the passive server, the active server tells it that synchronization is complete and the passive server now enters the *STANDBY* state. In this state, it receives messages replicated from the active server and applies them locally.

If the active server goes offline, only passive servers in the *STANDBY* state can be considered candidates to become the new active server.

Clearing Automatic Backup Data

After a passive server is restarted, for safety reasons, it may retain artifacts from previous runs. This happens when the server is restartable, even in the absence of restartable cache managers. The number of copies of backups that are retained is unlimited. Over time, and with frequent restarts, these copies may consume a substantial amount of disk space, and it may be desirable to clear up that space.

Backup rationale: If, after a full shutdown, an operator inadvertently starts the stripe members in the wrong order, this could result in data loss wherein the new active server initializes itself from the, possibly, incomplete data of a previous passive server. This situation can be mitigated by (1) ensuring all servers are running, and (2) the cluster is quiesced, prior to taking the backup. This ensures that all members of the stripe contain exactly the same data.

Clearing backup data manually: The old fast restart and platform files are backed up under the server's data directories in the format `terraccotta.backup.{date&time}/ehcache/` and `backup-platform-data-{date&time}/platform-data` respectively. Simply change to the data root directory, and remove the backups.

It may be desirable to keep the latest backup copy. In that case, remove all the backup directories except the one with the latest timestamp.

3 Clients in a Cluster

Within the overall structure of the cluster, the clients represent the application end-points. They work independently but can communicate through the *active* servers of the stripes to which they are connected.

Note that a client only ever interacts with an active server, never directly communicating with a *passive* server.

In a single-stripe cluster, each client is connected to the active server of that stripe. In a multi-stripe cluster, each client is connected to the active server of *each* stripe, interacting with them quasi-independently.

Within the logical structure of the cluster, the client isn't the process making the connection, but the connection itself. This means that a single JVM opening multiple connections to the same stripe will be seen by the stripe as multiple, independent clients.

How a client finds an active server

When establishing a connection to a stripe, the client must find the active server. It does this by attempting to connect to each server in the stripe, knowing that only the active server will not reject the connection attempt.

How a client handles failover or restart

If an active server to which a client is attached goes offline, the client will attempt to reconnect to one of the other servers in the stripe, if there are any. This is similar to what happens during its initial connection.

Note that there is no default time-out on this reconnection attempt. In the case that each stripe member is unavailable, this means that it is possible for all clients to wait, blocking their progress, until a server is restarted, potentially days later.

4 Electing an Active Server

When a new stripe comes online, the first thing the servers within it need to do is elect an *active* server which will coordinate the *client* interactions and *passive* servers within the stripe.

Additionally, if the active server of an existing stripe goes offline, the remaining passive servers need to elect a replacement active server. Note that only passive servers in the *STANDBY* state are candidates for this role. For related information, see the section [“Failover” on page 17](#).

In either of these situations, the servers involved address this problem by holding an election.

A server that is started up from rest needs to get votes from all of its peer servers to get elected as an active server. In contrast, votes from a smaller set of peers are sufficient for a running *PASSIVE-STANDBY* server to become elected as an active server.

If for some reason, not all servers in a stripe can be started up, you can still forcefully get a candidate server elected as active using the cluster tool. For more information about this manual intervention using the cluster tool, see the section *The promote command* of the [“Cluster Tool” on page 43](#).

High-level process

In an election, each server will construct a "vote" which it sends to the other involved servers. The vote with the highest score can be determined statically, so each server knows it has agreement on which server won the election.

In the case of a tie, the election is re-run until consensus is achieved.

Vote construction

The vote is a list of "weights" which represent the factors which should be considered when electing the most effective active server. The list is ordered in a way that the next element is only considered if the current element is a tie. This allows the earlier elements of the vote to be based around important concepts (such as how many transactions the server has processed), followed by concrete concepts (such as server up-time), then ending in more arbitrary concepts designed to break edge-case ties (such as a randomly generated number).

5 Failover

In a high-availability stripe, the failure of a single server represents only a small disruption, but not outright failure, of the cluster and the *client* operations (for related information on high availability, see the section “[Cluster Architecture](#)” on page 7).

In the case of a failing passive server, there is no disruption at all experienced by the clients.

In the case of a failing active server, however, there is a small disruption of client progress until a new active server is elected and the client can reconnect to it. *Failover* is the name given to this scenario.

Client Reconnect Window

When a failover happens, the clients connected to the previous active server automatically switch to the new active server. However, these clients have a limited window of time called the *client reconnect window* to complete the failover (120 seconds, by default). The new active server will stop processing any client requests until all the previously known clients connect back or until this window expires. This could cause all the clients to stall even if a single client fails or takes too long to fail over to the new active server.

If clients fail to connect back to the new active server within the reconnect window, the server will consider them unreachable and will continue processing requests from the connected clients. Clients reconnecting after the reconnect window will be rejected by the server and they will rejoin the cluster as a new client by establishing a new connection.

This reconnect window can be configured in the Terracotta configuration file using the `<client-reconnect-window>` element. The following XML snippet shows how the client reconnect window can be changed to 60 seconds:

```
<tc-config>
  ...
  <servers>
    ...
    <client-reconnect-window>60</client-reconnect-window>
  </servers>
</tc-config>
```

Server-side implications

Once all clients have reconnected (or the reconnect window closes), the server will process all re-sent messages it had seen before for which the client had not been notified of completion.

After this, message processing resumes as normal.

Client-side implications

Clients will experience a slight stall while they reconnect to the new active server. This reconnection process involves re-sending any messages the client considers to be in-flight.

After this, client operations resume as normal.

6 Starting and Stopping the Terracotta Server

Starting the Terracotta Server

The command line script to start the Terracotta Server is located in the `server/bin/` directory of the server kit. UNIX users use `start-tc-server.sh` while Windows users use `start-tc-server.bat`. All arguments are the same for both.

The usage of script is as follows:

```
start-tc-server.sh [-f /path/to/tc-config.xml] [-n server_name]
```

Options to the script

While it is possible to run the script without any arguments, this will result in using empty defaults for the configuration, which is generally not useful for anything other than verifying that the Terracotta Server is able to run.

Specific arguments which should be used are:

- `[-f /path/to/tc-config.xml]` - This is the path to the `tc-config.xml` file for the stripe this server is expected to join. Note that all servers in the same stripe are expected to use the same configuration file.

The file `tc-config.xml` describes per-server details such as listening TCP port and log directory.

- `[-n server_name]` - This is the name the server should use for itself, which determines which server stanza described within the `tc-config.xml` should be used.

Environment variables read by the script

- `JAVA_HOME` - Points to the JRE installation which the server should use (the Java launcher in this JRE will be used to start the server).
- `JAVA_OPTS` - Any additional options which should be passed to the underlying JVM can be passed via this environment variable and they will be added to the Java command line.

Stopping the Terracotta Server

If your server is not running in a Terracotta cluster, you can use the standard procedure offered by your operating system to terminate the server process.

If you are running the server as part of a Terracotta cluster, you can stop all servers in the cluster by using the cluster tool. See the section [“Cluster Tool” on page 43](#) for details.

7 Safe Cluster Shutdown and Restart Procedure

Although the Terracotta Server Array is designed to be crash tolerant, like any distributed system with HA capabilities, it is important to consider the implications of shutting down and restarting servers, what sequence that is done in, and what effects that has on client applications and potential loss of some data.

The safest shutdown procedure

For the safest shutdown procedure, follow these steps:

1. Shut down all clients and ensure no critical operations such as backup are running on the cluster. The Terracotta client will shut down when you shut down your application.
2. Use the `shutdown` command of the cluster tool to shut down the cluster.

If you want to partially shut down a stripe with passive servers configured, you can use the partial shutdown commands provided by the cluster tool. See the section [“Cluster Tool” on page 43](#) for details.

The safest restart procedure

To restart a stripe for which the failover priority is *consistency*, servers can be started up in any order as it is guaranteed that the last active server is re-elected as the active server, thus preventing data loss. This is guaranteed even if there are multiple former active servers in the stripe at the time of shutdown (for example, one active server and one or more suspended active servers or former active servers that were shut down, decommissioned or had crashed).

However, if the failover priority is *availability*, restarting the servers in any random order might result in data loss. For example, if an older active server is started up before the last active server, it could win the election and become the active server with its old data. To avoid such data loss scenarios, the last known active server must be restarted first. All other servers must be started up after this last known active server becomes the active server again.

However, if you do not know the most recent active server at the time of restart and still want to restart the stripe safely without data loss, it can still be done by starting all the servers in that stripe using the `--consistency-on-start` option of the server startup script. When the servers are started up using this option, they will wait for all peer servers to come up and then elect the most recent active server as the new active server.

If there are multiple active servers at the time of shutdown, which can happen if the failover priority of the cluster is *availability*, one of them will be chosen automatically on restart. This choice is made based on factors like the number of clients connected to those servers at the time of shutdown, the server that was started up first, etc.

Considerations and implications of not following the above procedure

Facts to understand:

- Servers that are in "active" status have the "master" or "source of full truth" copy of data for the stripe they belong to. They also have state information about in-progress client transactions, and client-held locks.
- Mirror servers (in "passive standby" state) have a "nearly" up to date copy of the data and state information. (Any information that they don't have is redundant between the active server and the client.)
- If the active server fails (or is shut down purposely), not only does the standby server need to reach active state, but the clients also need to reconnect to it and complete their open transactions, or data may be lost.
- A Terracotta Server Array, or "Cluster" instance has an identity, and the stripes within the TSA have a "stripe ID". In order to protect data integrity, running clients ensure that they only "fail over" to servers with matching IDs to the ones they were last connected to. If cluster or stripe is completely "wiped" of data (by purposely clearing persisted data, or having persistence disabled and having all stripe members stopped at the same time), that will reset the stripe ID.

What happens if clients are not shut down

If clients are not shut down:

- Client applications will continue sending transactions (data writes) to the active server(s) as normal, right up until the active server is stopped. This may leave some successful transactions unacknowledged, or falsely reported as failed to the client, possibly resulting in some data loss.
- Clients will continue to try and connect and when the server is restarted, the clients will fail the current operation and enter a reconnect path to try and complete the operation. When clients enter a reconnect path, it is left to the client to ensure idempotency of the ongoing operation as the operation might either have been made durable just before shutdown or it may have been missed during shutdown.

What happens if the active server is shut down explicitly

If the active server is shut down first:

- Before shutting down any other servers, or restarting the server, ensure that you wait until any other servers in the stripe (that were in 'standby' status) have reached *active* state, and that any running clients have reconnected and re-sent their partially completed transactions. Otherwise there may be some data loss.

8 The Terracotta Configuration File

This document describes the elements of the Terracotta configuration file, which is an XML-formatted file named `tc-config.xml` by default.

This file serves to configure all the members of a single Terracotta Server Array (TSA) stripe. Refer to the section [“Cluster Architecture” on page 7](#) for details on various different TSA topologies.

You can use a sample configuration file provided in the kit as the basis for your Terracotta configuration. Some samples have inline comments describing the configuration elements. Be sure to start with a clean file for your configuration.

Explanation of Configuration Sections

The Terracotta configuration file is divided into several main, unordered sections. For general configuration purposes, the most relevant sections are:

Section	Description
Servers	<p>The <code>servers</code> section defines all the servers which will make up this stripe of a cluster. High-availability is enabled by configuring and running at least 2 servers in this stripe. Note that there is no explicit configuration of which server takes an active or passive role, as that may change over the lifetime of the cluster.</p> <p>Each <code>server</code> element in the <code>servers</code> section is identified by a name given by the <code>name</code> attribute:</p> <pre><server ... name="ServerName"></pre> <p>To start the server with the name "ServerName", you pass the option <code>-n ServerName</code> to the start script. Refer to the section “Starting and Stopping the Terracotta Server” on page 19 for more details.</p>
Plugins	<p>The <code>plugins</code> section extends the capabilities of the listed servers by (1) registering and configuring additional <code>services</code> made available to those servers, and (2) providing general configuration information made available to all services.</p> <p>The use of extension points is required for the server to work with your application end-points.</p> <p>Refer to the section “Configuring the Terracotta Server” on page 27 for concrete examples.</p>

Section	Description
Properties	The <code>tc-properties</code> section exposes a list of key-value pairs to further customize the behavior of the cluster. Note that this section is normally empty.

Simple Configuration Sample

This is an example of a very simple server configuration file.

```
<tc-config xmlns="http://www.terracotta.org/config">
  <servers>
    <server host="localhost" name="testServer0">
      <logs>terracotta-kit-test/testServer0/logs</logs>
      <tsa-port>26270</tsa-port>
      <tsa-group-port>26271</tsa-group-port>
    </server>
  </servers>
</tc-config>
```

This shows the key components of a standard configuration but describes only a single server with no extension points used.

Key points:

- Configuration namespace: ["http://www.terracotta.org/config"](http://www.terracotta.org/config).
 - Only a single server with name `testServer0`.
 - Server name is important as it is used when starting a server so it knows which server it is.
 - `localhost` should be replaced with the actual fully-qualified hostname or IP address of the server, in a real deployment.
- Note:** Using `localhost` as hostname will not work correctly if clients and operator console are connecting from other hosts, since the configured hostname is sent to clients and other members of the stripe as the authoritative hostname to reach this server.
- A relative path to a logs directory is given.
 - All relative paths are with respect to the location of the containing configuration file.
 - The `tsa-port` is the port that clients will use when connecting to the server (default: 9410).
 - The `tsa-group-port` is for inter-server communication among stripe members, even though there are no other servers in this case (default: 9530).
 - Neither *restartability* nor *failover* would be possible with the above sample configuration, as restart support requires an extension point which provides that capability, and failover requires at least 2 servers in the cluster.

Parameter Substitution

Parameter substitution provides a way to substitute variables with pre-defined system properties in the Terracotta Server configuration file. Thus, a significant number of fields can be intelligently populated based on machine specific properties. Parameter substitution is most commonly done for hostname, IP address and directory path substitutions.

The following predefined substitutions are available for use:

Parameter	Description
%h	the fully-qualified host name
%i	the IP address
%H	the user's home directory
%n	the username
%o	the operating system name
%a	the processor architecture
%v	the operating system version
%t	the temporary directory (on Linux or Solaris, e.g., /tmp)
%(property)	the Java system property of the JVM (e.g. %(java.home), %(logs.path))
%D	the time stamp (yyyyMMddHHmmssSSS)

These parameters can be used where appropriate, including for elements or attributes that expect strings or paths for values:

- the name, host and bind attributes of the <server> element
- the logs child element of the <server> element
- data-roots

Important: Be careful when you specify substitution parameters in the host attribute of the <server> element in tc-config. If multiple servers from the same stripe containing substitution parameters in host are started on different hosts, they will not be able to communicate with each other since each host

substitutes the parameter based on its own host name. Similarly, tools like the cluster tool will not be able to connect to the servers, as they don't know which host the server is running on.

Note: The variable `%i` is expanded into a value determined by the host's networking setup. In many cases that setup is in a `hosts` file containing mappings that may influence the value of `%i`. Test this variable in your production environment to check the value it interpolates.

9 Configuring the Terracotta Server

Overview

For your application end-points to be useful they must be able to utilize storage resources configured in your Terracotta Servers. The services offered make use of your server's underlying JVM and OS resources, including direct-memory (offheap) and disk persistence.

These server resources are configured in the `plugins` section of the Terracotta configuration file. For related information, see the section [“The Terracotta Configuration File” on page 23](#).

A Note on System Requirements and Considerations

We recommend to familiarize yourself with the following recommendations concerning the Terracotta Server:

Planning a Successful Deployment (*Installation Guide*)

System Recommendations for Hybrid Caching (*in Terracotta Server Administration Guide > Configuring the Terracotta Server*)

System Recommendations for Fast Restart (FRS) (*in Terracotta Server Administration Guide > Configuring the Terracotta Server*)

Offheap Resources

The use of JVM Direct-Memory (offheap) is a central part of the operation of a Terracotta Server. In effect, you **must** allocate and make available to your server enough offheap memory for the proper operation of your application.

In your configuration file you define one or more named *offheap resources* of a fixed size. These named resources are then referred to in the configuration of your application end-points to allow for their usage.

Refer to the section *Clustered Caches* in the *Ehcache API Developer Guide* for more details about the use of offheap resources.

Example Offheap Resource Configuration

```
<plugins> <!--1-->
  <config>
    <ohr:offheap-resources
      xmlns:ohr="http://www.terracotta.org/config/offheap-resource">
      <ohr:resource name="primary-server-resource"
        unit="MB">384</ohr:resource> <!--2-->
      <ohr:resource name="secondary-server-resource"
        unit="MB">256</ohr:resource> <!--3-->
    </ohr:offheap-resources>
  </config>
</plugins>
```

1	The <code>plugins</code> element is a direct child of the <code>tc-config</code> element
2	Defines <code>primary-server-resource</code> with size 384MB
3	Defines <code>secondary-server-resource</code> with size 256MB

Data Directories

A data directory is a location on disk, identified by a name, and mapped to a disk location, where a Terracotta Server's data resides.

Data directories are commonly configured by server administrators and specified in the Terracotta Server configuration. Data directory names can be used by products that need durable storage for persistence and fast restart from crashes. For example, restartable cache managers need to be supplied with a data directory name to persist the restartable CacheManager specific data.

For information on restartable servers, see the section [“Platform Persistence” on page 29](#) below. See also the sections *Fast Restartability* and *Creating a Restartable Cache Manager* in the *Ehcache API Developer Guide*.

Sample Data Directories Configuration

```
<config xmlns:data="http://www.terracottatech.com/config/data-roots">
  <data:data-directories>
    <data:directory name="someData" > <!-- 1 -->
      /mnt1/data > <!-- 2 -->
    </data:directory>
    <data:directory name="otherData" > <!-- 3 -->
      %(logs.path)/data > <!-- 4 -->
    </data:directory>
  </data:data-directories>
</config>
```

1	<code>someData</code> is a data directory name,
2	<code>/mnt1/data</code> is the disk location to which <code>someData</code> is mapped to.
3	<code>otherData</code> is another data directory name mapped to a different disk location,
4	<code>%(logs.path)/data</code> is the data directory which <code>otherData</code> is mapped to. Note the use of <code>%(logs.path)</code> , which gets substituted with the logs path property

See the description of parameter substitution in the section [“The Terracotta Configuration File” on page 23](#) to check the complete list of available parameter substitutions.

General Notes on Configuring Data Directories

- A data directory specified in a stripe configuration file must be specified in all the configurations of all stripes of the cluster.
- Each data directory must be given a unique mount point (or disk location).
- The data directories are created if they do not exist already.
- Changing the disk location of the data directory between server restarts, without copying the data, is equivalent to erasing that data. It will cause unpredictable runtime errors that depend on the exact data lost.

Platform Persistence

The Terracotta server saves its internal state on a disk which enables server restarts without losing data. Platform persistence leverages data directories to store required data, so at least one data directory must be configured in the server configuration.

Note: Platform persistence is mandatory and a Terracotta server will refuse to start if there are no data directories defined.

Care must be taken to avoid losing data when restarting the stripe. Refer to the section [“Restarting a Stripe” on page 73](#) for more details. Passive restartable servers automatically back up their data at restart for safety reasons. Refer to the topic *Passive servers* in the section [“Active and Passive Servers” on page 9](#) for more details.

Changing the disk location of the data directory used for platform persistence before restarting a server and not copying the data will result in the server starting as if it was a new server without any data.

Platform Persistence Configuration

By default, if a single data directory is defined, it will be used for platform persistence. If more than one data directory is defined, one of them must have the `use-for-platform` attribute set to `true`.

If the server cannot resolve a data directory to be used for platform persistence, it will fail to start.

Sample Server Configuration with Platform Persistence

```
<tc-config xmlns="http://www.terracotta.org/config"
  xmlns:data="http://www.terracottatech.com/config/data-roots"
  xmlns:persistence="http://www.terracottatech.com/config/platform-persistence">
  <plugins>
    <config>
      <data:data-directories>
        <data:directory name="platform"
          use-for-platform="true"/>/mnt/platform</data:directory> <!-- 1 -->
      </data:data-directories>
      <data:data-directories>
        <data:directory name="data"/>/mnt/data</data:directory>
      </data:data-directories>
    </config>
  </plugins>
</tc-config>
```

```
</plugins>
<servers>
  <server host="localhost" name="server1">
    <tso-port>9410</tso-port>
  </server>
</servers>
</tc-config>
```

- | | |
|---|--|
| 1 | Indicates that the <code>platform</code> data directory is to be used for platform persistence |
|---|--|

Relation to Fast Restartability

The Ehcache *Fast Restartability* feature depends on, and makes use of, platform persistence.

Refer to the section *Fast Restartability* in the *Ehcache API Developer Guide* for more information.

10 System Recommendations for Hybrid Caching

Hybrid Caching supports writing to one single mount, so all of the Hybrid capacity must be presented to the Terracotta process as one continuous region, which can be a single device or a RAID.

The mount should be used exclusively for the Terracotta server process. The software was designed for usage on local drives (SSD/Flash in particular) - SAN/NAS storage is not recommended. If you utilize SAN/NAS storage you will experience notably reduced and inconsistent performance - any support requests related to performance or stability on such deployments will require the user to reproduce the issue with local disks.

Note: System utilization is higher when using Hybrid Caching, and it is not recommended to run multiple servers on the same machine. Doing so could result in health checkers timing out, and killing or restarting servers. Therefore, it is important to provision sufficient hardware, and it is highly recommended to deploy servers on different machines.

Hybrid Caching is described in detail in the Developer Guide.

11 System Recommendations for Fast Restart (FRS)

Fast Restart (FRS) supports writing to one single mount, which can be a single device or a RAID.

The mount should be used exclusively for the Terracotta server process. The software was designed for usage on local drives (SSD/Flash in particular) - SAN/NAS storage is not recommended. If you utilize SAN/NAS storage you will experience notably reduced and inconsistent performance - any support requests related to performance or stability on such deployments will require the user to reproduce the issue with local disks.

Fast Restartability is described in detail in the Developer Guide.

12 Failover Tuning

Overview

In a clustered environment, network, hardware, or other failures can cause an active server to get disconnected from the rest of the servers in its stripe. When your cluster needs to remain tolerant to such failures, you have a choice to make: choose either *consistency* or *availability* but not both (CAP theorem). If consistency is chosen over availability, then the cluster will halt processing client requests as consistent reads/writes can't be guaranteed when the cluster is partitioned. But when availability is chosen over consistency, the cluster will respond to client requests even when the cluster is partitioned but the response is not guaranteed to be consistent. In the absence of such failures, the cluster can provide both consistency and availability.

If the cluster is tuned to favor availability over consistency, then when such failures happen, the behavior of a stripe is that the remaining passive servers will then run an election and, if not able to find the old active server, the passive server that wins the election becomes the new active server. While this configuration ensures high availability of the data, risks of experiencing a so-called split-brain situation during such elections are increased. In the case of a TSA, split-brain would be a situation in which multiple servers in a stripe are acting as active servers. For example, if an active server gets partitioned from its peers in that stripe, the active server will remain active and the passive servers on the other side of the partition would elect a new active server as well. Any further operations performed on the data are likely to result in inconsistencies.

When tuned for consistency, a stripe would need at least a majority of servers connected with each other to elect an active server. Thus, even if the stripe gets partitioned into two sets of servers due to some network failure, the set with the majority of servers will elect an active server among them and proceed. In the absence of a majority, an active server will not be elected and hence the clients will be prevented from performing any operations, thereby preserving data consistency by sacrificing availability.

Server configuration

When configuring the stripe, the user needs to choose between availability and consistency as the failover priority of the stripe. To prevent split-brain scenarios and thereby preserve data consistency, failover priority must be set to consistency. However, if availability is preferred, failover-priority can be set to availability at the risk of running into split-brain scenarios.

The following xml snippet shows how to configure a stripe for *consistency*:

```
<tc-config>
  ...
  <servers>
    ...
  </servers>
  <failover-priority> <!-- 1 -->
    <consistency/> <!-- 2 -->
  </failover-priority>
```

```
</tc-config>
```

1	Failover priority is tuned to favor...
2	Consistency over availability for this stripe

Similarly, the stripe can be tuned for *availability* as follows:

```
<failover-priority>
  <availability/>
</failover-priority>
```

Note: Even though the availability vs. consistency configuration is done at the stripe level, it must to be consistent across all the stripes in the cluster.

This is a mandatory configuration and if not provided in the tc-config, the server will refuse to start.

Choosing Consistency versus Availability

A Terracotta Server Array (TSA), being a distributed system, is subject to the constraints of the *CAP Theorem*. The CAP Theorem states that it is impossible for a distributed system to simultaneously provide guarantees for Consistency, Availability, and Partition tolerance. A TSA always seeks to be tolerant of network partitions so a choice must be made between data consistency and service availability.

For a TSA supporting only Ehcache users, choosing availability over consistency can be an effective and proper choice. If a network partition occurs leaving two servers unable to communicate with each other (to coordinate data) but each able to serve clients, choosing availability allows continuing operations using both servers running independently. Data residing on the separated servers will no longer be coordinated (consistency is abandoned) and will drift apart. Once the network partition is resolved and the servers are able to communicate with each other again, consistency rules are re-applied and one server is chosen as the holder of the *current* data - *the data on the other servers is discarded*. For a caching application, this may result in some cache misses and delays in some processing but, since a cache is not the system of record for the data being processed, no data is lost.

When using a TSA configured for availability over consistency to support TCStore users, data loss is a real possibility - if a network partition occurs and clients are actively updating TSA servers on both sides of the partition, data loss *will* occur once the network partition is resolved. If the applications using TCStore are tolerant of missing/inconsistent data - not an easy task - then configuring for availability over consistency is appropriate for the TSA. However, if a TCStore dataset is used as the system of record, this data loss or other inconsistency is generally undesirable if not catastrophic. If the TSA is used for a TCStore dataset which is either a system of record or for which loss/inconsistency is an undesirable outcome, then the TSA **must** be configured for consistency over availability.

You may use a single TSA supporting both Ehcache and TCStore but the choice of availability versus consistency is a cluster-level configuration - both Ehcache and

TCStore users in a TSA are subject to that configuration. If your Ehcache users need high availability and your TCStore users need data consistency, you must use a separate TSA for each user base.

External voter for two-server stripes

Mandating a majority for active server election in certain topologies introduces additional availability issues. For example, in a two-server stripe the majority quorum is two as well. This means that if these servers get disconnected from each other due to a network partition or because of a server failure, the surviving server would not promote itself as the active server as it requires 2 votes to win the election. But since the other voting server is not reachable, it will not be able to get that second vote and hence will not promote itself. In the absence of an active server, the stripe is not available.

Adding a third server is the best option, so that even if one fails, there is a majority (2 out of 3) surviving to elect an active. A three-server stripe can provide data redundancy and high availability at the same time even when one server fails. If adding a third server is not feasible, the alternate option is to get high availability without risking data consistency (via split-brain scenarios) using an external voter. But this configuration cannot offer data redundancy (like a three-server stripe) if a server fails.

An external voter is a client that is allowed to cast a vote in the election of a new active server, in cases where a majority of servers in a stripe are unable to reach a consensus on electing a new active server.

External voter configuration

The number of external voters needs to be described in the server configuration. It is recommended that the total number of servers and external voters be kept as an odd number.

External voters need to get registered with the servers to get added as voting members in their elections. If there are n voters configured in the server, then the first n voting clients requesting to get registered will be added as voters. Registration requests of other clients will be declined and put on hold until one of the registered voters gets de-registered.

Voters can de-register themselves from the cluster so that the voting rights can be transferred to other clients waiting to get registered, if there are any. A voting client can de-register itself by using APIs or by getting disconnected from the cluster.

When a voting client gets disconnected from the server, it will automatically get de-registered by the server. When the client reconnects, it will only get registered again as a voter if another voter has not taken its place while this client was disconnected.

Server configuration

A maximum count for the number of external voters allowed can optionally be added to the `failover-priority` configuration if the stripe is tuned for `consistency`, as follows:

```
<failover-priority>
  <consistency>
    <voter count="3"/> <!-- 1 -->
  </consistency>
```

```
</failover-priority>
```

1	Here you are restricting the total number of voting clients to three.
---	---

The failover priority setting and the specified maximum number of external voters across the stripes must be consistent and will be validated during the cluster configuration step. For more information on how to configure a cluster, see the section [“Cluster Tool” on page 43](#).

Client configuration

External voters can be of two variants:

1. Standalone voter
2. Clients using the voter library (client voter)

Standalone voter

An external voter can be run as a standalone process using a script provided with the kit. The script takes the tc-config files of the stripes in the cluster as arguments. A variant that takes the `<host>:<port>` combinations instead of the server configuration files is also supported. Each `-s` option argument must be a comma separated list of `<host>:<port>` combinations of servers in a single stripe. To register a multi-stripe cluster, multiple `-f` or `-s` options can be provided for each stripe.

Usage:

```
start-tc-voter.(sh|bat) -f TC-CONFIG [-f TC-CONFIG]...
```

or

```
start-tc-voter.(sh|bat) -s HOST:PORT[,HOST:PORT]... [-s HOST:PORT[,HOST:PORT]...]...
```

To connect the voter to a secure cluster, the path to the security root directory will also have to be provided using the `-srd` option. For more details on setting up security in a Terracotta cluster, see [“SSL / TLS Security Configuration in Terracotta” on page 79](#).

Client voter

Any TCStore or Ehcache client can act as an external voter as well by using a voter library distributed with the kit. A client can join the cluster as a voter by creating a `TCVoter` instance and registering itself with the cluster.

Note: A cluster must be configured using the cluster tool before a client voter can be registered with it.

When the voter is no longer required, it can be de-registered from the cluster either by disconnecting that client, or by using the `deregister` API.

```
TCVoter voter = new EnterpriseTCVoterImpl(); // 1
voter.register("my-cluster-0" // 2
              "<host>:<port>", "<host>:<port>"); // 3
...
voter.deregister("my-cluster-0") // 4
```

1	Instantiate a <code>TCVoter</code> instance
2	Register the voter with a cluster by providing a cluster name ...
3	and host port combinations of all servers in the cluster.
4	De-register from the cluster using the same cluster name that was used to register it.

To connect to a secure cluster, the voter must be instantiated using the overloaded constructor of `EnterpriseTCVoterImpl` that takes in the security root directory path.

Manual promotion with override voter

Since an external voter is just another process, there is no guarantee that it will always be up and available. Especially in the form of client voters, the moment the client leaves, the external voter leaves too. In the rare event of a failure happening (partition splitting the active and passive servers or the active server crashing) and the external voter not being around either, none of the surviving servers will be acting as an active server. The servers will be stuck in a *suspended* state where operations from the regular clients are all stalled. A manual intervention will be required to get the cluster out of this state by fixing the cause of the partition or by restarting the crashed server. If neither is feasible, then the third option is to get a server manually promoted using an override vote from an external voter.

The voter process can be started in an *override* mode to promote a single server stuck in that intermediate state to be an active server. When the voter process is started in this special mode, it will connect to the server that you want to promote, give it an override vote and exit. The voter process can be started in *override* mode as follows:

```
start-tc-voter.(sh|bat) -o HOST:PORT
```

Running this command will forcibly promote the server at `HOST:PORT` to be an active server, if it is stuck in that intermediate state.

Note: This override voting will work even if external voters are not configured in the server configuration.

Be cautious not to start two different override voters on both sides of the partition separately so that both sides win and cause a split-brain.

Server startup

When the failover priority of the stripes is tuned for consistency, it has an impact on server startup as well. In a multi-server stripe, when the servers are started up fresh, a server will not get elected as an active server until it gets votes from all of its peers. This will require all the servers of that stripe to be brought up. Bringing up regular voters is not going to help as they need to communicate with all the active servers in the cluster

to get registered. But if bringing up the other servers is not feasible for some reason, then an override voter can be used to forcibly promote that server.

13 Connection Leasing

Why Leasing

When a client carries out a write with IMMEDIATE or STRONG consistency, the server ensures that every client that could be caching the old value is informed, and the write will not complete until the server can ensure that clients will no longer serve a stale value.

Where network disruptions prevent the server communicating with a client in a timely manner, the server will close that client's connection to allow the write to progress.

To achieve this, each client maintains a lease on its connections to the cluster. If a client's lease expires, the server may decide to close that client's connection. A client may also close the connection if it realises that its lease has expired.

When TCStore serves data from its client-side cache, the client checks its lease. If it detects that the lease has expired, it will not use potentially stale data held in the cache.

Lease Length

When selecting the length of lease, consider the range of possible client to server roundtrip latencies over a network connection that can be considered as functional. The lease should be longer than the largest possible such latency.

On a server that is heavily loaded, there may be some additional delay in processing a client's request for a lease to be extended. Such a delay should be added into the roundtrip network latency.

In addition, leases are not renewed as soon as they are issued, instead the client waits until some portion of the lease has passed before renewing. A guideline suitable for the current implementation is that leases should be approximately 50% longer to allow for this.

Setting long leases, however, has the downside that, when clients are unreachable by a server, IMMEDIATE writes could block for up to the length of a lease.

The default value of leases is currently two and a half minutes.

Lease Configuration

To configure the lease length, add a connection leasing service plugin configuration to the tc-config file. For example:

```
<tc-config xmlns="http://www.terracotta.org/config"
  xmlns:lease="http://www.terracotta.org/service/lease">
  <plugins>
    <service>
      <lease:connection-leasing>
        <lease:lease-length unit="seconds">60</lease:lease-length>
      </lease:connection-leasing>
    </service>
  </plugins>
</tc-config>
```

```
</plugins>  
</tc-config>
```

which configures a lease length of sixty seconds.

Valid values for the `unit` attribute are: `milliseconds`, `seconds`, `minutes` and `hours`.

Any positive integer may be used for the value within the `lease-length` element as long as the length of time configured for the lease length is not more than `Long.MAX_VALUE` nanoseconds, which is approximately 290 years.

14 Cluster Tool

The cluster tool is a command-line utility that allows administrators of the Terracotta Server Array to perform a variety of cluster management tasks. For example, the cluster tool can be used to:

- Configure or re-configure a cluster
- Obtain the status and configuration information of running servers
- Dump the state of running servers
- Take backups of running servers
- Promote a suspended server on startup or failover
- Shut down an entire cluster
- Perform a conditional partial shutdown of a cluster having one or more passive servers configured for high availability (for upgrades etc.)

The cluster tool script is located in `tools/cluster-tool/bin` under the product installation directory as `cluster-tool.bat` for Windows platforms, and as `cluster-tool.sh` for Unix/Linux.

Usage Flow

The following is a typical flow in a cluster setup and usage:

1. Create Terracotta configuration files for each stripe in the deployment. See the section [“The Terracotta Configuration File” on page 23](#) for details.
2. Start up the servers in each stripe. See the section [“Starting and Stopping the Terracotta Server” on page 19](#) for details.
3. Make sure the stripes are online and ready.
4. Configure the cluster using the `configure` command of the cluster tool. See the section [“The “configure” Command”](#) below for details.
5. Check the current status of the cluster or specific servers in the cluster using the `status` command. See the section [“The “status” Command” on page 49](#) below for details.

Cluster Tool commands

The cluster tool provides several commands. To list them and their respective options, run `cluster-tool.sh` (or `cluster-tool.bat` on Windows) without any arguments, or use the option `-h` (long option `--help`).

The following section provides a list of options common to all commands, and thus need to be specified before the command name:

Precursor options

1. `-v` (long option `--verbose`)

This option gives you a verbose output, and is useful to debug error conditions.

2. `-srd` (long option `--security-root-directory`)

This option can be used to communicate with a server which has TLS/SSL-based security configured. For more details on setting up security in a Terracotta cluster, see the section [“Security Core Concepts” on page 80](#).

Note: If this option is not specified while trying to connect to a secure cluster, the command will fail with a `SECURITY_CONFLICT` error.

3. `-t` (long option `--timeout`)

This option lets you specify a custom timeout value (in milliseconds) for connections to be established in cluster tool commands.

Note: If this option is not specified, the default value of 30,000 ms (or 30 seconds) is used.

Each command has the option `-h` (long option `--help`), which can be used to display the usage for the command.

The following is a comprehensive explanation of the available commands:

The "configure" Command

The `configure` command creates a cluster from the otherwise independent Terracotta stripes, taking as input a mandatory license key. No functionality is available on the server until a valid license is installed. See the section [“Licensing” on page 63](#) for details.

Note: All servers in any given stripe should be started with the same configuration file. The `configure` command configures the cluster based on the configuration(s) of the currently known active server(s) only. If there is a configuration mismatch among the active and passive server(s) within the same stripe, this command can configure the cluster while taking down any passive server(s) with configuration mismatches. This validation also happens upon server restart and changes will prevent the server from starting. See the section on the `reconfigure` command for more information on how to update server configurations.

The command will fail if any of the following checks do not pass:

1. License checks
 - a. The license is valid.
 - b. The provided configuration files do not violate the license.

2. Configuration checks

- The provided configuration files are consistent across all the stripes.

The following configuration items are validated in the configuration files:

1. `config`:

a. `offheap-resource`

Offheap resources present in one configuration file must be present in all the files with the same sizes.

b. `data-directories`

Data directory identifiers present in one configuration file must be present in all the files. However, the data directories they map to can differ.

2. `service`

a. `security`

Security configuration settings present in one configuration file must match the settings in all the files.

For more details on setting up security in a Terracotta cluster, see the section [“Security Core Concepts” on page 80](#).

b. `backup-restore`

If this element is present in one configuration file, it must be present in all the files.

3. `failover-priority`

The failover priority setting present in one configuration file must match the setting in all the files.

Refer to the section [“The Terracotta Configuration File” on page 23](#) for more information on these elements.

The `servers` section of the configuration files is also validated. Note that it is not validated between stripes but rather against the configuration used to start the servers themselves.

- `server`

- `host`

- It must be a strict match

- `name`

- It must be a strict match

- `tlsa-port`

- It must be a strict match

Note: Once a cluster is configured, a similar validation will take place upon server restart. It will cause the server to fail to start if there are differences.

Usage:

```
configure -n CLUSTER-NAME [-l LICENSE-FILE] TC-CONFIG [TC-CONFIG...]
configure -n CLUSTER-NAME [-l LICENSE-FILE] -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

A name that is to be assigned to the cluster.

- `-l LICENSE-FILE`

The path to the license file. If you omit this option, the cluster tool looks for a license file named `license.xml` in the location `tools/cluster-tool/conf` under the product installation directory.

- `TC-CONFIG [TC-CONFIG ...]`

A whitespace-separated list of configuration files (minimum 1) that describes the stripes to be added to the cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option. Any one server from each stripe can be provided. However, multiple servers from the same stripe will work as well. The cluster will be configured with the configurations which were originally used to start the servers.

Note: The command configures the cluster only once. To update the configuration of an already configured cluster, the `reconfigure` command should be used.

Examples

- The example below shows a successful execution for a two stripe configuration and a valid license.

```
./cluster-tool.sh configure -l ~/license.xml -n tc-cluster
~/tc-config-stripe-1.xml ~/tc-config-stripe-2.xml
Configuration successful
License installation successful
Command completed successfully
```

- The example below shows a failed execution because of an invalid license.

```
./cluster-tool.sh configure -l ~/license.xml
-n tc-cluster ~/tc-config-stripe-1.xml ~/tc-config-stripe-2.xml
Error (BAD_REQUEST): com.terracottatech.LicenseException: Invalid license
```

- The example below shows a failed execution with two stripe configurations mismatching in their offheap resource sizes.

```
./cluster-tool.sh configure -n tc-cluster -l
~/license.xml ~/tc-config-stripe-1.xml ~/tc-config-stripe-2.xml
Error (BAD_REQUEST): Mismatched off-heap resources in provided config files:
[[primary-server-resource: 51200M], [primary-server-resource: 25600M]]
```

The "reconfigure" Command

The `reconfigure` command updates the configuration of a cluster which was configured using the `configure` command. With `reconfigure`, it is possible to:

1. Update the license on the cluster.
2. Add new offheap resources, or grow existing ones.
3. Add new data directories.
4. Add new configuration element types.

The command will fail if any of the following checks do not pass:

1. License checks
 - a. The new license is valid.
 - b. The new configuration files do not violate the license.
2. Stripe checks
 - a. The new configuration files have all the previously configured servers.
 - b. The order of the configuration files provided in the `reconfigure` command is the same as the order of stripes in the previously configured cluster.
3. Configuration checks
 - a. Stripe consistency checks

The new configuration files are consistent across all the stripes. For the list of configuration items validated in the configuration files, refer to the section *The "configure" Command* above for details.
 - b. Offheap checks

The new configuration has all the previously configured offheap resources, and the new sizes are not smaller than the old sizes.
 - c. Data directories checks

The new configuration has all the previously configured data directory names.
 - d. Configuration type checks

The new configuration has all the previously configured configuration types.

Usage:

```
reconfigure -n CLUSTER-NAME TC-CONFIG [TC-CONFIG...]
reconfigure -n CLUSTER-NAME -l LICENSE-FILE -s HOST[:PORT] [-s HOST[:PORT]]...
reconfigure -n CLUSTER-NAME -l LICENSE-FILE TC-CONFIG [TC-CONFIG...]
```

Parameters:

- `-n CLUSTER-NAME`
The name of the configured cluster.

- TC-CONFIG [TC-CONFIG ...]

A whitespace-separated list of configuration files (minimum 1) that describe the new configurations for the stripes.

- -l LICENSE-FILE

The path to the new license file.

- -s HOST[:PORT] [-s HOST[:PORT]]...

The `host:port(s)` or `host(s)` (default port being 9410) of servers, each specified using the `-s` option.

Servers in the provided list will be sequentially contacted for connectivity, and the command will be executed on the first reachable server.

`reconfigure` **command usage scenarios:**

1. License update

When it is required to update the license, most likely because the existing license has expired, the following `reconfigure` command syntax should be used:

```
reconfigure -n CLUSTER-NAME -l LICENSE-FILE -s HOST[:PORT] [-s HOST[:PORT]]...
```

Note: A license update does not require the servers to be restarted.

2. Configuration update

When it is required to update the cluster configuration, the following `reconfigure` command syntax should be used:

```
reconfigure -n CLUSTER-NAME TC-CONFIG [TC-CONFIG...]
```

The steps below should be followed in order:

- a. Update the Terracotta configuration files with the new configuration, ensuring that it meets the reconfiguration criteria mentioned above.
- b. Run the `reconfigure` command with the new configuration files.
- c. Restart the servers with the new configuration files for the new configuration to take effect.

3. License and configuration update at once

In the rare event that it is desirable to update the license and the cluster configuration in one go, the following `reconfigure` command syntax should be used:

```
cluster-tool.sh reconfigure -n  
CLUSTER-NAME -l LICENSE-FILE TC-CONFIG [TC-CONFIG...]
```

The steps to be followed here are the same as those mentioned in the *Configuration update* section above.

Examples

- The example below shows a successful re-configuration of a two stripe cluster `tc-cluster` with new stripe configurations.

```
./cluster-tool.sh reconfigure -n tc-cluster
~/tc-config-stripe-1.xml ~/tc-config-stripe-2.xml
License not updated (Reason: Identical to previously installed license)
Configuration successful
Command completed successfully.
```

- The example below shows a failed re-configuration because of a license violation.

```
./cluster-tool.sh reconfigure -n tc-cluster
-l ~/license.xml -s localhost:9410
Error (BAD_REQUEST): Cluster offheap resource is not within the limit of the license.
Provided: 409600 MB, but license allows: 102400 MB only
```

- The example below shows a failed re-configuration of a two stripe cluster with new stripe configurations having fewer data directories than existing configuration.

```
./cluster-tool.sh reconfigure -n tc-cluster
~/tc-config-stripe-1.xml ~/tc-config-stripe-2.xml
License not updated (Reason: Identical to previously installed license)
Error (CONFLICT): org.terracotta.exception.EntityConfigurationException:
Entity: com.terracottatech.tools.client.TopologyEntity:topology-entity
lifecycle exception:
Entity: com.terracottatech.tools.client.TopologyEntity:topology-entity
lifecycle exception:
Entity: com.terracottatech.tools.client.TopologyEntity:topology-entity
lifecycle exception: org.terracotta.entity.ConfigurationException:
Mismatched data directories. Provided: [use-for-platform, data],
but previously known: [use-for-platform, data, myData]
```

The "status" Command

The `status` command displays the status of a cluster, or particular server(s) in the same or different clusters..

Usage:

```
status -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
status -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

The name of the configured cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

When provided with option `-n`, servers in the provided list will be sequentially contacted for connectivity, and the command will be executed on the first reachable server. Otherwise, the command will be individually executed on each server in the list.

Examples

- The example below shows the execution of a cluster-level `status` command.

```

./cluster-tool.sh status -n tc-cluster -s localhost
Cluster name: tc-cluster
Stripes in the cluster: 2
Servers in the cluster: 4
  Server{name='server-1', host='localhost', port=9410},
  Server{name='server-2', host='localhost', port=9610} (stripe 1)
  Server{name='server-3', host='localhost', port=9710},
  Server{name='server-4', host='localhost', port=9910} (stripe 2)
Total configured offheap: 102400M
Backup configured: true
SSL/TLS configured: false
IP whitelist configured: false
Data directories configured: data, myData
| STRIPE: 1 |
+-----+-----+-----+
| Server Name | Host:Port | Status |
+-----+-----+-----+
| server-1    | localhost:9410 | ACTIVE |
| server-2    | localhost:9610 | PASSIVE |
+-----+-----+-----+
| STRIPE: 2 |
+-----+-----+-----+
| Server Name | Host:Port | Status |
+-----+-----+-----+
| server-3    | localhost:9710 | ACTIVE |
| server-4    | localhost:9910 | PASSIVE |
+-----+-----+-----+

```

- The example below shows the execution of a server-level status command. No server is running at localhost:9510, hence the UNREACHABLE status.

```

./cluster-tool.sh status -s localhost:9410 -s localhost:9510 -s localhost:9910
+-----+-----+-----+
| Host:Port | Status | Member of Cluster |
+-----+-----+-----+
| localhost:9410 | ACTIVE | tc-cluster |
| localhost:9910 | PASSIVE | tc-cluster |
| localhost:9510 | UNREACHABLE | - |
+-----+-----+-----+
Error (PARTIAL_FAILURE): Command completed with errors.

```

Server States

STARTING	server is starting
UNINITIALIZED	server has started and is ready for election
SYNCHRONIZING	server is synchronizing its data with the current active server
PASSIVE	server is passive and ready for replication
ACTIVE	server is active and ready to accept clients

ACTIVE_RECONNECTING	server is active but waits for previously known clients to rejoin before accepting new clients
START_SUSPENDED	server startup is suspended for all of its peers to come up
ACTIVE_SUSPENDED	server is active but blocked in the election process (consistency mode)
PASSIVE_SUSPENDED	server is passive but blocked in the election process (consistency mode)
UNREACHABLE	server is unreachable from cluster tool
UNKNOWN	server state is unknown

The "promote" command

The `promote` command can be used to promote a server stuck in a *suspended* state. For more information about suspended states, refer to the topics *Server startup* and *Manual promotion with override voter* in the section [“Failover Tuning” on page 35](#).

Usage:

```
promote -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-s HOST[:PORT]`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option. The command will be individually executed on each server in the list.

Note: There is no cluster-wide equivalent (with the `-n` option) for this command.

Examples

- The example below shows the execution of the `promote` command on a server stuck in suspended state at `localhost:9510`.

```
./cluster-tool.sh promote -s localhost:9510
Command completed successfully.
```

- The example below shows the erroneous execution of a server-level `promote` command. The server running at `localhost:9510` is not in a suspended state to be promoted, hence the failure.

```
./cluster-tool.sh promote -s localhost:9510
localhost:9510: Promote failed as the server at localhost:9510 is not
in any suspended state
```

```
Error (FAILURE): Command failed.
```

The "dump" Command

The `dump` command dumps the state of a cluster, or particular server(s) in the same or different clusters. The dump of each server can be found in its logs.

Usage:

```
dump -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
dump -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

The name of the configured cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

When provided with option `-n`, servers in the provided list will be sequentially contacted for connectivity, and the command will be executed on the first reachable server. Otherwise, the command will be individually executed on each server in the list.

Examples

- The example below shows the execution of a cluster-level dump command.

```
./cluster-tool.sh dump -n tc-cluster -s localhost:9910
Command completed successfully.
```

- The example below shows the execution of a server-level dump command. No server is running at `localhost:9510`, hence the dump failure.

```
./cluster-tool.sh dump -s localhost:9410 -s localhost:9510 -s localhost:9910
Dump successful for server at: localhost:9410
Connection refused from server at: localhost:9510
Dump successful for server at: localhost:9910
Error (PARTIAL_FAILURE): Command completed with errors.
```

The "stop" Command

The `stop` command stops the cluster, or particular server(s) in the same or different clusters.

Usage:

```
stop -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
stop -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

The name of the configured cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

When provided with the option `-n`, servers in the provided list will be contacted for connectivity, and the command will be executed on all reachable servers. An attempt will be made to shut down the entire cluster in the correct sequence by shutting down all the passive servers first followed by the active servers. The `stop` command with the `-n` option is similar to the `shutdown` command with the `--force` option.

NOTE: This command is deprecated in favor of the `shutdown` command. Refer to the description of the `shutdown` command for more details.

Examples

- The example below shows the execution of a cluster-level `stop` command.

```
./cluster-tool.sh stop -n tc-cluster -s localhost
Command completed successfully.
```

- The example below shows the execution of a server-level `stop` command. No server is running at `localhost:9510`, hence the stop failure.

```
./cluster-tool.sh stop -s localhost:9410 -s localhost:9510 -s localhost:9910
Stop successful for server at: localhost:9410
Connection refused from server at: localhost:9510
Stop successful for server at: localhost:9910
Error (PARTIAL_FAILURE): Command completed with errors.
```

The "ipwhitelist-reload" Command

The `ipwhitelist-reload` command reloads the IP whitelist on a cluster, or particular server(s) in the same or different clusters. See the section [“IP Whitelisting” on page 87](#) for details.

Usage:

```
ipwhitelist-reload -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
ipwhitelist-reload -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME` The name of the configured cluster.
- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

When provided with the option `-n`, servers in the provided list will be sequentially contacted for connectivity, and the command will be executed on the first reachable server. Otherwise, the command will be individually executed on each server in the list.

Examples

- The example below shows the execution of a cluster-level `ipwhitelist-reload` command.

```
./cluster-tool.sh ipwhitelist-reload -n tc-cluster -s localhost
```

```
IP whitelist reload successful for server at: localhost:9410
IP whitelist reload successful for server at: localhost:9610
IP whitelist reload successful for server at: localhost:9710
IP whitelist reload successful for server at: localhost:9910
Command completed successfully.
```

- The example below shows the execution of a server-level `ipwhitelist-reload` command. No server is running at `localhost:9510`, hence the IP whitelist reload failure.

```
./cluster-tool.sh ipwhitelist-reload -s localhost:9410
-s localhost:9510 -s localhost:9910
IP whitelist reload successful for server at: localhost:9410
Connection refused from server at: localhost:9510
IP whitelist reload successful for server at: localhost:9910
Error (PARTIAL_FAILURE): Command completed with errors.
```

The "backup" Command

The `backup` command takes a backup of the running Terracotta cluster.

Usage:

```
backup -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

The name of the configured cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

When provided with the option `-n`, servers in the provided list will be sequentially contacted for connectivity, and the command will be executed on the first reachable server. Otherwise, the command will be individually executed on each server in the list.

Examples

- The example below shows the execution of a cluster-level successful `backup` command. Note that the server at `localhost:9610` was unreachable.

```
./cluster-tool.sh backup -n tc-cluster -s localhost:9610 -s localhost:9410
PHASE 0: SETTING BACKUP NAME TO : 996e7e7a-5c67-49d0-905e-645365c5fe28
localhost:9610: TIMEOUT
localhost:9410: SUCCESS
localhost:9710: SUCCESS
localhost:9910: SUCCESS
PHASE (1/4): PREPARE_FOR_BACKUP
localhost:9610: TIMEOUT
localhost:9910: NOOP
localhost:9410: SUCCESS
localhost:9710: SUCCESS
PHASE (2/4): ENTER_ONLINE_BACKUP_MODE
localhost:9710: SUCCESS
localhost:9410: SUCCESS
PHASE (3/4): START_BACKUP
localhost:9710: SUCCESS
```

```
localhost:9410: SUCCESS
PHASE (4/4): EXIT_ONLINE_BACKUP_MODE
localhost:9710: SUCCESS
localhost:9410: SUCCESS
Command completed successfully.
```

- The example below shows the execution of a cluster-level failed backup command.

```
./cluster-tool.sh backup -n tc-cluster -s localhost:9610
PHASE 0: SETTING BACKUP NAME TO : 93cdb93d-ad7c-42aa-9479-6efbdd452302
localhost:9610: SUCCESS
localhost:9410: SUCCESS
localhost:9710: SUCCESS
localhost:9910: SUCCESS
PHASE (1/4): PREPARE_FOR_BACKUP
localhost:9610: NOOP
localhost:9410: SUCCESS
localhost:9710: SUCCESS
localhost:9910: NOOP
PHASE (2/4): ENTER_ONLINE_BACKUP_MODE
localhost:9410: BACKUP_FAILURE
localhost:9710: SUCCESS
PHASE (CLEANUP): ABORT_BACKUP
localhost:9410: SUCCESS
localhost:9710: SUCCESS
Backup failed as some servers '[Server{name='server-1', host='localhost', port=9410},
[Server{name='server-2', host='localhost', port=9710}]]',
failed to enter online backup mode.
```

The "shutdown" Command

The `shutdown` command shuts down a running Terracotta cluster. During the course of the shutdown process, it ensures that:

- Shutdown safety checks are performed on all the servers. Exactly what safety checks are performed will depend on the specified options and is explained in detail later in this section.
- All data is persisted to eliminate data loss.
- All passive servers are shut down first before shutting down the active servers.

The `shutdown` command follows a multi-phase process as follows:

1. Check with all servers whether they are OK to shut down. Whether or not a server is OK to shut down will depend on the specified shutdown options and the state of server in question.
2. If all servers agree to the shutdown request, all of them will be asked to prepare for the shutdown. Preparing for shutdown may include the following:
 - a. Persist all data.
 - b. Block new incoming requests. This ensures that the persisted data will be cluster-wide consistent after shutdown.
3. If all servers successfully prepare for the shutdown, a shutdown call will be issued to all the servers.

The first two steps above ensure an atomic shutdown to the extent possible as the system can be rolled back to its original state if there are any errors. In such cases, client-request processing will resume as usual after unblocking any blocked servers.

In the unlikely event of a failure in the third step above, the error message will clearly specify the servers that failed to shut down. In this case, use the `--force` option to forcefully terminate the remaining servers. If there is a network connectivity issue, the forceful shutdown may fail, and the remaining servers will have to be terminated using operating system commands.

Note: The shutdown sequence also ensures that the data is stripe-wide consistent. Although, it is recommended that clients are shut down before attempting to shut down the Terracotta cluster.

Usage:

```
shutdown -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`
The name of the configured cluster.
- `-f | --force`
Forcefully shut down the cluster, even if the cluster is only partially reachable.
- `-i | --immediate`
Do an immediate shutdown of the cluster, even if clients are connected.
- `-s HOST[:PORT] [-s HOST[:PORT]]...`
The host:port(s) or host(s) (default port being 9410) of running servers, each specified using the `-s` option.

If the `-n` option is not specified, this command forcefully shuts down only the servers specified in the list. For clusters having stripes configured for high availability (with at least one passive server per stripe), it is recommended that you use the partial cluster shutdown commands explained in the section below, as they allow conditional shutdown, instead of using the `shutdown` variant without the `-n` option.

If the `-n` option is specified (i.e. a full cluster shutdown), this command shuts down the entire cluster. Servers in the provided list will be contacted for connectivity, and the command will then verify the cluster configuration with the given cluster name by obtaining the cluster configuration from the first reachable server. If all servers are reachable, this command checks if all servers in all the stripes are safe to shut down before proceeding with the command.

A cluster is considered to be safe to shut down provided the following are true:

- No critical operations such as backup and restore are going on.
- No Ehcache or TCStore clients are connected.

- All servers in all the stripes are reachable.

If either the `-f` or `-i` option is specified, this command works differently than above as follows:

- If the `-i` option is specified, this command proceeds with the shutdown even if clients are connected.
- If the `-f` option is specified, this command proceeds with the shutdown even if none of the conditions specified for safe shutdown above are met.

For all cases, the shutdown sequence is performed as follows:

1. Flush all data to persistent store for datasets or caches that have persistence configured.
2. Shut down all the passive servers, if any, in the cluster for all stripes.
3. Once the passive servers are shut down, issue a shutdown request to all the active servers in the cluster.

The above shutdown sequence is the cleanest way to shut down a cluster.

Examples

- The example below shows the execution of a cluster-level successful `shutdown` command.

```
./cluster-tool.sh shutdown -n primary -s localhost:9610 -s localhost:9410
Shutting down cluster: primary
STEP (1/3): Preparing to shutdown
STEP (2/3): Stopping all passives first
STEP (3/3): Stopping all actives
Command completed successfully.
```

- The example below shows the execution of a cluster-level successful `shutdown` command that fails as one of the servers in the cluster was not reachable.

```
./cluster-tool.sh shutdown -n primary -s localhost:11104
Error (FAILURE): Shutdown invocation timed out
Detailed Error Status for Cluster `primary` :
  ServerError{host='localhost:25493', Error='Shutdown invocation timed out'}.
Unable to process safe shutdown request.
Command failed..
```

- The example below shows the execution of a cluster-level successful `shutdown` command with the force option. Note that one of the servers in the cluster was already down.

```
./cluster-tool.sh shutdown -f -n primary -s localhost:11104
Shutting down cluster: primary
STEP (1/3): Preparing to shutdown
Shutdown invocation timed out
Detailed Error Status :
  ServerError{host='localhost:25493', Error='Shutdown invocation timed out'}.
Continuing Forced Shutdown.
STEP (2/3): Stopping all passives first
STEP (3/3): Stopping all actives
Command completed successfully.
```

Partial Cluster Shutdown Commands

Partial cluster shutdown commands can be used to partially shut down nodes in the cluster without sacrificing the availability of the cluster. These commands can be used only on a cluster that is configured for redundancy with one or more passive servers per stripe. The purpose of these commands is to allow administrators to perform routine and planned administrative tasks, such as rolling upgrades, with high availability.

The following flavors of partial cluster shutdown commands are available:

- `shutdown-if-passive`
- `shutdown-if-active`
- `shutdown-all-passives`
- `shutdown-all-actives`

As a general rule, if these commands are successful, the specified servers will be shut down. If there are any errors due to which these commands abort, the state of the servers will be left intact.

From the table of server states described in [“The `status` Command” on page 49](#), the following are the different active states that a server may find itself in:

- `ACTIVE`
- `ACTIVE_RECONNECTING`
- `ACTIVE_SUSPENDED`

Note: In the following sections, the term 'active servers' means servers in any of the active states mentioned above, unless explicitly stated otherwise.

Similarly, the following are the passive states for a server:

- `PASSIVE_SUSPENDED`
- `SYNCHRONIZING`
- `PASSIVE`

Note: In the following sections, the term 'passive servers' means servers in any of the passive states mentioned above, unless explicitly stated otherwise.

The `shutdown-if-passive` Command

The `shutdown-if-passive` command shuts down the specified servers in the cluster, provided the following conditions are met:

- All the stripes in the cluster are functional and there is one healthy active server with no suspended active servers per stripe.
- All the servers specified in the list are passive servers.

Usage:

```
shutdown-if-passive -s HOST[:PORT] [-s
HOST[:PORT]]...
```

Parameters:

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

Examples

- The example below shows the execution of a successful `shutdown-if-passive` command.

```
./cluster-tool.sh shutdown-if-passive -s localhost:23006
Stopping Passive Server(s): `[Server{name='testServer1', host='localhost',
port=23006}]` of Cluster: primary
STEP (1/2): Preparing to shutdown
STEP (2/2): Stopping if Passive
Command completed successfully.
```

- The example below shows the execution of a failed `shutdown-if-passive` command, as it tried to shut down a server which is not a passive server.

```
./cluster-tool.sh shutdown-if-passive -s localhost:23004
Error (FAILURE): Unable to process the partial shutdown request.
One or more of the specified server(s) are not in passive state or
may not be in the same cluster
Discovered state of all servers are as follows:
Reachable Servers : 5
Stripe #: 1
Server : {Server{name='testServer1', host='localhost', port=23006}}
State : PASSIVE
Server : {Server{name='testServer0', host='localhost', port=23004}}
State : ACTIVE
Stripe #: 2
Server : {Server{name='testServer101', host='localhost', port=2537}}
State : ACTIVE
Server : {Server{name='testServer100', host='localhost', port=2535}}
State : PASSIVE
Server : {Server{name='testServer102', host='localhost', port=2539}}
State : PASSIVE
Please check server logs for more details.
Command failed.
```

The "shutdown-if-active" Command

The `shutdown-if-active` command shuts down the specified servers in the cluster, provided the following conditions are met:

- All the servers specified in the list are active servers.
- All the stripes corresponding to the given servers have at least one server in 'PASSIVE' state.

Usage:

```
shutdown-if-active -s HOST[:PORT] [-s
HOST[:PORT]]...
```

Parameters:

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option.

Examples

- The example below shows the execution of a successful `shutdown-if-active` command:

```
./cluster-tool.sh shutdown-if-active -s localhost:23726 -s localhost:26963
Stopping Active Server(s): `[Server{name='testServer0', host='localhost', port=23726}, Server
STEP (1/2): Preparing For Shutdown
STEP (2/2): Shutdown If Active
Command completed successfully.
```

- The example below shows the execution of a failed `shutdown-if-active` command as the specified server was not an active server.

```
./cluster-tool.sh shutdown-if-active -s localhost:23726 -s localhost:23730
Error (FAILURE): Unable to process the partial shutdown request.
One or more of the specified server(s) are not in active state or
may not be in the same cluster.
Reachable Servers : 5
Stripe #: 1
Server : {Server{name='testServer2', host='localhost', port=23730}}
State : PASSIVE
Server : {Server{name='testServer0', host='localhost', port=23726}}
State : ACTIVE
Stripe #: 2
Server : {Server{name='testServer100', host='localhost', port=26961}}
State : PASSIVE
Server : {Server{name='testServer101', host='localhost', port=26963}}
State : ACTIVE
Server : {Server{name='testServer102', host='localhost', port=26965}}
State : PASSIVE
Please check server logs for more details
Command failed.
```

The "shutdown-all-passives" Command

The `shutdown-all-passives` command shuts down all the passive servers in the specified cluster, provided the following is true:

- All the stripes in the cluster are functional and there is one active server in 'ACTIVE' state with no suspended active servers per stripe.

All passive servers in all the stripes of the cluster will be shut down when this command is run.

Usage:

```
shutdown-all-passives -n CLUSTER-NAME -s HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

The name of the configured cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option. These `host(s)` need not be passive servers.

The `-n` option is mandatory, as this command can be used only to shut down all the passive servers in the entire cluster. Servers in the provided list will be contacted for connectivity, and the command will then verify the cluster configuration with the given cluster name by obtaining the cluster configuration from the first reachable server.

After the necessary verifications, it proceeds to shut down all the passive servers in a multi-phase manner as explained below:

1. Check with all servers whether it is safe to shut down as a passive server.
2. Flush any data that needs to be made persistent across all servers that are going down and block any further changes.
3. Issue a shutdown request to all passive servers if all passive servers succeed in step 2.
4. If any servers fail in step 2 or above, the shutdown request will fail and the state of the servers will remain intact.

Examples

- The example below shows the execution of a successful `shutdown-all-passives` command.

```
./cluster-tool.sh shutdown-all-passives -n primary -s localhost:5252
Stopping Passive Server(s): `[Server{name='testServer0', host='localhost',
port=5252},
Server{name='testServer100', host='localhost', port=15361},
Server{name='testServer102', host='localhost', port=15365}]`
of Cluster: primary
STEP (1/2): Preparing to shutdown
STEP (2/2): Stopping if Passive
Command completed successfully.
```

The "shutdown-all-actives" Command

The `shutdown-all-actives` command shuts down the active server of all stripes in the cluster, provided the following are true:

- There are no suspended active servers in the cluster.
- There is at least one passive server in 'PASSIVE' state in every stripe in the cluster.

The active server of all stripes of the cluster will be shut down when this command returns success. If the command reports an error, the state of the servers will be left intact.

Usage:

```
shutdown-all-actives -n CLUSTER-NAME -s
HOST[:PORT] [-s HOST[:PORT]]...
```

Parameters:

- `-n CLUSTER-NAME`

The name of the configured cluster.

- `-s HOST[:PORT] [-s HOST[:PORT]]...`

The `host:port(s)` or `host(s)` (default port being 9410) of running servers, each specified using the `-s` option. These `host(s)` need not be active servers.

The `-n` option is mandatory, as this command can be used only to shut down all the active servers in the entire cluster. Servers in the provided list will be contacted for connectivity, and the command will then verify the cluster configuration with the given cluster name by obtaining the cluster configuration from the first reachable server.

After the necessary verifications, it proceeds to shut down all the active servers in a multi-phase manner as explained below:

1. Check with all servers whether they are safe to be shut down as active servers.
2. Flush any data that needs to be made persistent across all servers that are going down and block any further changes.
3. Issue a shutdown request to all active servers if they succeed in step 2.
4. If any servers fail in step 2 or above, the shutdown request will fail and the state of the servers will remain as before.

Examples

- The example below shows the execution of a successful `shutdown-all-actives` command. Note that the specified host was a passive server in this example. As the specified host is used only to connect to the cluster and obtain the correct state of all the servers in the cluster, the command successfully shuts down all the active servers in the cluster, leaving the passive servers intact.

```
./cluster-tool.bat shutdown-all-actives -n primary -s localhost:31445
Stopping Active Server(s): `[Server{name='testServer2', host='localhost',
port=31449},
Server{name='testServer100', host='localhost', port=27579}]`
of cluster: primary
STEP (1/2): Preparing For Shutdown
STEP (2/2): Shutdown If Active
Command completed successfully.
```

15 Licensing

This document describes the installation and update procedures for Terracotta Ehcache and Terracotta DB licenses.

Installing a license

A Terracotta license is installed on a Terracotta cluster using the cluster tool `configure` command, thereby enabling cluster configuration and license installation in one go. The command ensures that:

- The license is a valid Software AG license.
- The license has not expired already.
- The Terracotta configuration files do not violate the license.

The following example configures a Terracotta cluster using the license file `license.xml`, the name `tc-cluster`, and the configuration file `tc-config.xml`.

```
cluster-tool.sh configure -l license.xml -n tc-cluster tc-config.xml
Command completed successfully.
```

See the section [“Cluster Tool” on page 43](#) for a detailed explanation of the command usage.

License expiration

License expiry checks are done every midnight (UTC time) to ensure that the license in use did not expire. Midnight here is the time at the start of the day, i.e. '00:00' hours. As an example, for a license which is valid till December 31, the midnight check on December 31 will pass, but the check on January 1 midnight will fail, and license will be deemed as expired. When a license expires, a warning message like the following will be logged every 30 minutes in the server logs:

```
ATTENTION!! LICENSE expired. Time since expiry 1 day(s)
```

The license must be renewed within 7 days of expiry. If it is not done, the cluster will be shut down with the following message in the server logs:

```
Shutting down the server as a new license is not installed within 7 days.
```

License renewal

If your license expires, a new license can be obtained by contacting Software AG support. The new license can then be installed using the cluster tool `reconfigure` command as follows:

```
cluster-tool.sh reconfigure -l license.xml -s localhost:9410
Command completed successfully.
```

See the section [“Cluster Tool” on page 43](#) for a detailed explanation of the command usage.

16 Backup, Restore and Data Migration

- Overview 66
- Data Directory Structure 66
- Online Backup 67
- Offline Backup 69
- Restore 69
- Data Migration of Ehcache data 70
- Technical Details 71

Overview

The Backup and Restore feature enables you as an administrator of a Terracotta cluster to take a backup of the cluster and restore it from the backed up data when required.

Terracotta supports two ways of taking a backup:

1. Online backup using the cluster-tool. This is the recommended method.
2. Manual offline backup

Restore and Ehcache data migration are manual offline processes.

Note: Migration of TCStore data is currently not supported.

When a passive server starts and discovers it has data, the data is automatically backed up for safety reasons. However, this data is not cluster-wide consistent, and **must not** be used for restoration. Refer to the topic *Passive servers* in the section [“Active and Passive Servers”](#) on page 9 for more information.

Terms

Backup and Restore : Taking a snapshot of the cluster data such that it can later be installed back on the same cluster, bringing it back to the initial state.

Data Migration : Taking a snapshot of the cluster data, but installing it on a *different* cluster, bringing it to the state of the original cluster. Data Migration is also desirable in cases when only Ehcache data is needed, and not the platform data.

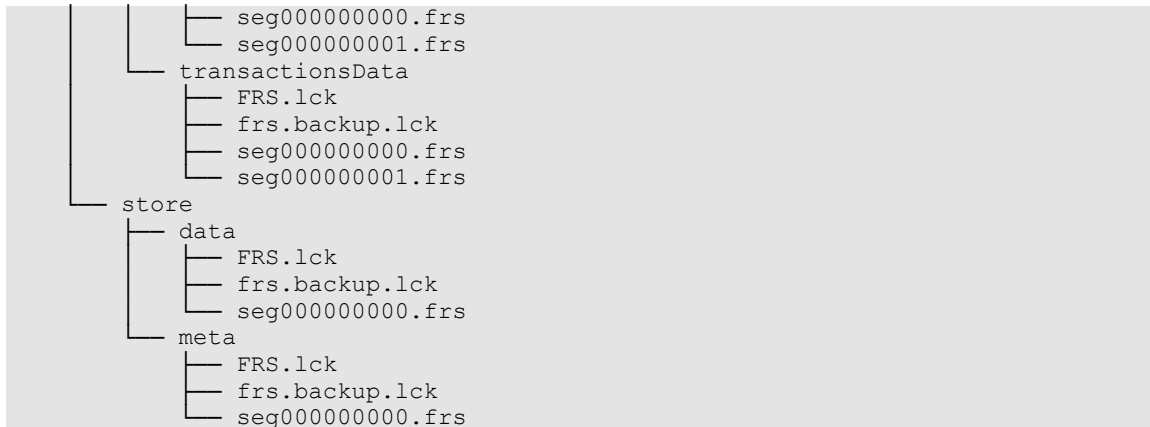
Data Directory Structure

Following is a sample data directory structure of a server containing Ehcache and TCStore data:

```

/tmp/data1/
├── server-1
│   ├── ehcache
│   │   └── frs
│   │       ├── default-frs-container
│   │       │   ├── default-cachedata
│   │       │   │   ├── FRS.lck
│   │       │   │   ├── frs.backup.lck
│   │       │   │   └── seg000000000.frs
│   │       │   └── metadata
│   │       │       ├── FRS.lck
│   │       │       ├── frs.backup.lck
│   │       │       └── seg000000000.frs
│   └── platform-data
│       ├── entityData
│       │   ├── FRS.lck
│       │   └── frs.backup.lck

```



where:

1. /tmp/data1 is the data directory path (for a given data directory) defined in the server configuration file
2. server-1 is the server name defined in the server configuration file
3. ehcache is the directory containing Ehcache data
4. platform-data is the directory containing platform specific logs
5. store is the directory containing TCStore data

Online Backup

Online backup of a Terracotta cluster is performed by the cluster-tool, and is the recommended method to take a backup. The following section describes the online backup feature and the process:

Configuring the Backup feature

To allow the server to use the Backup feature, ensure that you have set up the <backup-restore> and <backup-location> elements in the server configuration file, as shown in the following snippet:

```

<tc-config>
  <plugins>
    ...
    <service>
      <backup-restore xmlns="http://www.terracottatech.com/config/backup-restore">
        <backup-location path="/path/to/backup/dir" />
      </backup-restore>
    </service>
    ...
  </plugins>
  ...
</tc-config>

```

The path in the <backup-location> element can be absolute, or relative to the directory where the configuration file is located. If the directory specified by the backup location path is not present, it will be created during backup.

Prerequisites

Before proceeding with the online backup, ensure that:

1. At least one server in each stripe is up and running.
2. The servers have read and write permissions to the backup location.
3. Backup location has enough space available to store the backup data.
4. cluster-tool has fast connectivity to all the servers and the cluster is not heavily loaded with application requests.

Taking an online Backup

A backup is taken using the cluster-tool. Visit [“Cluster Tool” on page 43](#) for details on the `backup` command. If the backup fails for some reason, you can check the server logs for failure messages. Additionally, running the `backup` command using the `-v` (verbose) option might help.

Backup directory structure

The following diagram shows an example of the directory structure that results from a backup:

```

/tmp/backup1/
├── 7c868f83-5075-4b32-bef5-56f29fdcc6f0
│   ├── stripe0
│   │   ├── server-1
│   │   │   └── datadir
│   │   │       ├── ehcache
│   │   │       │   └── frs
│   │   │       │       └── default-frs-container
│   │   │       │           ├── default-cachedata
│   │   │       │           │   └── seg000000000.frs
│   │   │       │           └── metadata
│   │   │       │               └── seg000000000.frs
│   │   │       ├── platform-data
│   │   │       │   ├── entityData
│   │   │       │   │   ├── seg000000000.frs
│   │   │       │   │   └── seg000000001.frs
│   │   │       │   └── transactionsData
│   │   │       │       ├── seg000000000.frs
│   │   │       │       └── seg000000001.frs
│   │   │       └── store
│   │   │           ├── data
│   │   │           │   └── seg000000000.frs
│   │   │           └── meta
│   │   │               └── seg000000000.frs

```

where:

1. `/tmp/backup1/` is the backup location defined in the configuration file
2. `7c868f83-5075-4b32-bef5-56f29fdcc6f0` is an ID created by the backup command to uniquely identify a backup instance
3. `stripe0` is the stripe sequence in the cluster

4. `server-1` is the server name defined in the server configuration file
5. `datadir` is the data directory name (for a given data directory) defined in the server configuration file

Offline Backup

In the rare scenario when an online backup cannot be taken, an offline backup can be taken. The process is described as follows:

Taking an offline Backup

Follow the steps in the specified order to back up cluster data:

1. Shut down the cluster, while taking a note of the current active servers.
2. Copy the contents of the required data directories of all the servers which were actives prior to the shutdown to a desired location.
3. Name the directories in the manner described in the [“Backup directory structure” on page 68](#) section above. Although this step is optional, it helps identify different instances of backup, and keeps the restore steps consistent for online and offline backup procedures.
4. Save the configuration files as well. These files will be used to start the stripes after a restore is performed.

Restore

The restore operation is a manual operation. During the Restore operation, you use standard operating system mechanisms to copy the complete structure (directories, subdirectories and files) of the backup into the original location. Some small structural and/or naming changes are required in the restored directories after the copy, as described in the sections below.

Note: Restoring cache data will bring back cache entries which might have become stale by the time a restore is finished.

Performing a Restore

Before you start the Restore operation, ensure that all activity has stopped on the cluster and that the cluster is not running.

If you compare the structure of the backup under `/tmp/backup1` with the original structure under `/tmp/data1` (see both structural diagrams above), you will see some differences. You will also see that this is a single stripe cluster. Therefore, when you copy the `/tmp/backup1/<backup-name>` directory structure back to `/tmp/data1`, you need to make the following changes:

1. First choose a server as the active server for your stripe.
2. Note down the `name` attribute of that server in the configuration file. If there is no `name` attribute, skip step 4 mentioned below.
3. Create an empty directory for each path specified by the data directory in your configuration file. This will be the target directory for your restored data. Repeat this step for every data directory path specified in your configuration file.
4. If the `name` attribute for this server is specified, create a sub-directory with the name of the server under the data directories created above. For example, if the `name` attribute is `server-2` for the chosen active server for this stripe and the location specified for the data directory `datadir` is `/tmp/data1`, your target directory should look like `/tmp/data1/server-2`.
5. From the backup, copy the contents of `<stripe-id>/<server-name>/<data-directory>` to this newly created directory. For example, in the example given above, copy from `/tmp/backup1/<backup-name>/stripe0/server-1/data` to `/tmp/data1/server-2`.
6. Start the server with the newly created data directory with the configuration file which was backed up from the original cluster.
7. You can now bring up the passive servers in the stripe. Please note that you don't need to copy the backup data to the passive servers as they will automatically receive the data when they synchronize with the active server. It is advisable to remove any old data on the passive servers before you bring up the passive servers.
8. Repeat the above steps for other stripes in the cluster.

Data Migration of Ehcache data

Note: As noted above, data migration is currently not available for TCStore data.

Data migration can be performed to move Ehcache data to a new cluster without moving the platform data. Please note that only restartable caches contained in a restartable cache manager can be recovered. Since the data migration works at the data directory level, all the data of all restartable cache managers that use the same data directory will be recovered together.

How to perform an offline data migration

Follow the steps in the specified order to perform a migration of cluster data:

1. Shut down the source cluster and copy the contents of all ehcache directories from all required data directories of **all** active servers in the cluster. You can skip copying data directories containing restartable cache managers that you do not wish to migrate.
2. Start the target cluster (you can just start the active servers at this time) with the same number of stripes as the source cluster. Create the desired cache manager

configuration using a client. The cluster URI (including the cluster tier manager name for the cache manager) can be different in the new cluster. If the name part of the URI is different, specify the old name as the restart identifier when using the cache manager configuration API, so that the system can map the data corresponding to a given cache manager correctly. If there are more than one cache managers under the same data directory, use the configuration API to create all the cache managers in the target cluster.

For related information, see the section *Fast Restartability* of the *Ehcache API Developer Guide*.

3. Shut down the target cluster and copy the data to the matching data directories. The data directory paths can be different on the target cluster, but must have sufficient space to contain the data being copied over.
4. Once the data is available in all the stripes, you can start the target cluster. It now loads all the cache data that was moved from the source cluster.

Technical Details

Causal and Sequential Consistency across stripes

Since TCStore and Ehcache support only causal consistency (per key) and sequential consistency (across keys for a single thread of execution), the backup image across the cluster (be it single stripe or multi-stripe) must be consistent cluster wide for the point-in-time when the backup was taken.

For instance, suppose a single thread of execution from a single client synchronously made changes to keys A, then B, then C and then D in that order. Now if the backup was captured when the client had made changes to C, intuitively the backup **MUST** have all the previous changes made to A and B, regardless of the stripe where those mutations occurred. Thus on a restore of this point-in-time backup, if the restored data has C, then it **MUST** contain the changes made to A and B. Of course, it is to be expected that such a restoration may have permanently lost D, due to the point-in-time nature of restoring from backups.

As another example, say a system had long keys from 1 to 1000 and mutated them one by one exactly in that order. If the backup had 888 as the largest key, then all keys from 1 to 887 **MUST** also exist in the backup.

Causal consistency (per key) is always implied, as a key is always within a stripe. The backup taken must be consistent for a point in time snapshot, which implies that when a snapshot is taken, all mutations/changes that happen in the system **AFTER** the snapshot is taken **MUST** not reflect in the backup.

Consistency of multiple FRS logs within a stripe

Since platform data is also backed up, there are at least two FRS logs that needs to be backed up in a consistent fashion even within a single stripe.

17

Restarting a Stripe

Restart behavior is closely related to *failover*, but the difference is that the interruption period is typically much longer. On a restart, every server waits for the last active server to return instead of electing a different server as active. If a server other than the last active is elected as the active on a restart, it could cause data loss. To avoid such a data loss, every restarted server will wait in a *suspended* state until all of its peers are also started up so that the last active can be found and elected as the leader.

Unless a timeout is set, the time the clients will wait for the server to return is indefinite.

Note that a stripe can be both restartable and possess high-availability, if it is configured for restart support but also contains multiple servers. In this case, failover will progress as normal unless the entire stripe is taken offline.

Comparison with failover

The process of a *client* reconnecting to a restarted server is very similar to a newly-promoted *active* server after a fail-over. Both scenarios involve the clients reconnecting to re-send their in-flight transactions. Also, both will progress as normal once all clients have reconnected or the reconnect window closes.

The primary difference is that restart only requires one server, whereas high-availability requires at least two.

18 IPv6 support in Terracotta DB

This document describes the changes needed in configuration files, tooling, and client side APIs toward enabling IPv6 connections between clients and Terracotta cluster.

Terracotta DB supports IPv6 addresses as defined in the RFC 5952. Thus, all of following are acceptable:

- Full IPv6 address, e.g. `2001:db8:a0b:12f0:0:0:0:1`.
- Full IPv6 address enclosed in square brackets, e.g. `[2001:db8:a0b:12f0:0:0:0:1]`.
- Full IPv6 address enclosed in square brackets along with port, e.g. `[2001:db8:a0b:12f0:0:0:0:1]:9410`.
- Shortened IPv6 address, e.g. `2001:db8:a0b:12f0::1`.
- Shortened IPv6 address enclosed in square brackets, e.g. `[2001:db8:a0b:12f0::1]`.
- Shortened IPv6 address enclosed in square brackets along with port, e.g. `[2001:db8:a0b:12f0::1]:9410`.

Note that enclosing an IPv6 address in square brackets is mandatory only if a port is to be specified along with it.

Terracotta Server

If an IPv6 address of a machine hosting the Terracotta server instance is to be specified, the `host` attribute of the `<server>` element in `tc-config` would become the IPv6 address of the said machine, as exemplified below:

```
<servers>
  <server host="2001:db8:a0b:12f0::1" name="server-1">  <!-- 1 -->
    <logs>logs1</logs>
  </server>
  <server host="[2001:db8:a0b:12f0:0:0:0:2]" name="server-2">  <!-- 2 -->
    <logs>logs2</logs>
    <tsa-port>9510</tsa-port>
    <tsa-group-port>9530</tsa-group-port>
  </server>
  <server host="terracotta-host" name="server-2">  <!-- 3 -->
    <logs>logs3</logs>
    <tsa-port>9610</tsa-port>
    <tsa-group-port>9630</tsa-group-port>
  </server>
</servers>
```

1	Specifies a server host IP <code>2001:db8:a0b:12f0::1</code> . Since the <code><tsa-port></code> has not been specified, it will default to 9410.
2	Specifies a server host IP <code>2001:db8:a0b:12f0:0:0:0:2</code> enclosed in square brackets.

3	Specifies a server DNS host name <code>terracotta-host</code> which resolves to an IPv6 address.
---	--

IPv6 server sockets bind to `::0` by default, which can be overridden using the `bind` attribute in the `<server>` element.

Cluster Tool

Cluster tool commands which accept server addresses need to provide the IPv6 addresses of Terracotta servers as shown in the examples below:

- This example below shows the execution of `status` command for two IPv6 addresses - one enclosed in square brackets, and the other not enclosed in square brackets (and using the default port).

```
./cluster-tool.sh status -s [2001:db8:a0b:12f0::2]:9510 -s 2001:db8:a0b:12f0::1
```

Host:Port	Status	Member of Cluster
[2001:db8:a0b:12f0::2]:9510	ACTIVE	tc-cluster
2001:db8:a0b:12f0::1	PASSIVE	tc-cluster

- The example below shows the execution of a `stop` command that stops two servers. One server is identified by its host name and the other by its IPv6 address.

```
./cluster-tool.sh stop -s terracotta-host:9610 -s [2001:db8:a0b:12f0::1]:9410
Stop successful for server at: terracotta-host:9610
Stop successful for server at: [2001:db8:a0b:12f0::1]:9410
Command completed successfully.
```

Commands which don't accept server addresses directly (e.g. [“The “configure” Command” on page 44](#) of the Cluster Tool using `tc-configs`) do not need any change to work with IPv6.

Ehcache Client

An Ehcache client can specify the IPv6 addresses of the servers it wants to connect to either through Java APIs or XML configuration, as shown below:

API Example

```
InetSocketAddress firstServer = InetSocketAddress.createUnresolved("2001:db8:a0b:12f0::1", 0);
InetSocketAddress secondServer = InetSocketAddress.createUnresolved("2001:db8:a0b:12f0:0:0:0:2",
List<InetSocketAddress> servers = Arrays.asList(firstServer, secondServer);
String cacheManagerIdentifier = "cacheManager-1";
PersistentCacheManager cacheManager = CacheManagerBuilder
    .newCacheManagerBuilder()
    .with(EnterpriseClusteringServiceConfigurationBuilder.enterpriseCluster(servers,
        cacheManagerIdentifier) // 1
        .autoCreate())
    .build(true);
```

1	<code>EnterpriseClusteringServiceConfigurationBuilder.enterpriseCluster(Iterable, String)</code> lets you create a <code>CacheManager</code> by specifying IPv6 addresses of the servers. The first argument is the <code>Iterable<InetSocketAddress></code> of
---	---

	the servers in the cluster, while the second argument is the cache manager identifier.
--	--

Like other Ehcache APIs, the above API has a secure variant as well with the signature `EnterpriseClusteringServiceConfigurationBuilder.enterpriseSecureCluster(Iterable, String, Path)`, where the last argument is the path to the client's security root directory.

XML Example

```
<ehcache:config
  xmlns:ehcache="http://www.ehcache.org/v3"
  xmlns:tc="http://www.terracottatech.com/v3/terracotta/ehcache">
  <ehcache:service>
    <tc:cluster>
      <tc:cluster-connection cluster-tier-manager="cacheManager-1"> <!-- 1 -->
        <tc:server host="[2001:db8:a0b:12f0::1]"/> <!-- 2 -->
        <tc:server host="2001:db8:a0b:12f0:0:0:0:2" port="9510" /> <!-- 3 -->
      </tc:cluster-connection>
    </tc:cluster>
  </ehcache:service>
</ehcache:config>
```

1	Cache manager identifier.
2	Terracotta server IP [2001:db8:a0b:12f0::1]. Since the port is not specified, it will default to 9410.
3	Terracotta server IP 2001:db8:a0b:12f0:0:0:0:2 and port 9510.

TCStore Client

```
InetSocketAddress firstServer = InetSocketAddress.createUnresolved("2001:db8:a0b:12f0::1", 0);
InetSocketAddress secondServer = InetSocketAddress.createUnresolved("2001:db8:a0b:12f0:0:0:0:2",
List<InetSocketAddress> servers = Arrays.asList(firstServer, secondServer);
DatasetManager datasetManager = DatasetManager.clustered(servers) // 1
    .build();
```

1	<code>DatasetManager.clustered(Iterable)</code> lets you create a <code>DatasetManager</code> by specifying IPv6 addresses of servers.
---	--

Like other TCStore APIs, the above API has a secure variant as well with the signature `DatasetManager.clustered(Iterable, Path)`, where the last argument is the path to the client's security root directory.

XML Example

```
<clustered xmlns="http://www.terracottatech.com/v3/store/clustered"> <!-- 1 -->
  <cluster-connection>
    <server host="[2001:db8:a0b:12f0::1]"/> <!-- 2 -->
    <server host="2001:db8:a0b:12f0:0:0:0:2" port="9510" /> <!-- 3 -->
  </cluster-connection>
</clustered>
```

1	Declares a clustered <code>DatasetManager</code> configuration.
2	Terracotta server IP <code>[2001:db8:a0b:12f0::1]</code> . Since the port is not specified, it will default to 9410.
3	Terracotta server IP <code>2001:db8:a0b:12f0:0:0:0:2</code> and port 9510.

Terracotta Management Console

To connect to a Terracotta cluster running on IPv6, you can specify the server IPv6 addresses in the **Connection URL** as shown in the snapshot:

Screenshot: Terracotta Management Console, field for Connection URL



The screenshot shows the Terracotta Management Console interface. At the top, there is a blue header with the text "TERRACOTTA Management Console" and a hamburger menu icon. Below the header, the "Connection URL:" label is followed by a text input field containing the URL `terracotta://2001:db8:a0b:12f0::1,[2001:db8:a0b:12f0:0:0:0:2]:9510`. To the right of the input field are two buttons: "Next" and "Cancel".

19 SSL / TLS Security Configuration in Terracotta

■ Security Core Concepts	80
■ Cluster Security	85
■ TMS Security	91
■ LDAP Properties	96
■ SSL / TLS Troubleshooting guide	98

Security Core Concepts

Security features in Terracotta DB

Terracotta DB provides the following security features:

1. **Connection encryption** - encrypts connections between clients and servers, and also between servers, using the SSL/TLS protocol.
2. **Cluster authentication** - validates the identity of processes initiating connections to a server in a Terracotta cluster.
3. **TMS authentication** - validates the identity of users attempting to use the TMS.
4. **TMS authorization** - determines if an authenticated user has access to perform a given operation in TMS.
5. **Auditing** - writes security-relevant events to audit logs.
6. **IP whitelisting** - restricts access to only allow clients from known IP addresses.

Security Root Directory

To configure security features in Terracotta DB, each server and, in most cases, each client, must have a security root directory. The security root directory is a filesystem location for certificates, passwords and other security-related files.

The security root directory, and the files and subdirectories contained in it should be readable by the respective client or server process.

Note: Ensure that the security root directory is accessible only by users who are permitted to run the respective Terracotta DB client or server.

When a security root directory is specified, it can be given as absolute or relative. If specified in a `tc-config.xml` file, this is relative to the directory where the `tc-config` file is located.

Structure of the Security Root Directory

The directory structure below lists all possible subdirectories and files that can be present in a security root directory. Note that any one security root directory will have only a subset of these.

```
<security-root-directory>
├── access-control
│   ├── ldap.properties
│   └── users.xml
├── identity
│   ├── credentials.properties
│   └── <common-name>-<timestamp>.jks
├── trusted-authority
│   └── trusted-authority-<timestamp>.jks
└── whitelist.txt
```


Note: It is recommended to keep only the required files and directories under the security root directory to prevent configuration errors. If any unidentifiable files or directories are found, an appropriate warning message will be logged.

Access control subdirectory

The `access-control` subdirectory should only be specified on the server-side. It contains files to configure authentication and authorization:

`ldap.properties` - configuration properties for using an LDAP server for authentication or authorization

`users.xml` - used for file-based authentication or authorization

Identity subdirectory

The `identity` subdirectory can be specified on both the client-side and the server-side. It contains certificates or password-based credentials to prove the identity of the process.

`credentials.properties` - a standard Java properties file containing two properties: `username` and `password`. These credentials are sent to the server when the connection is established, in order to authenticate. This file should be present when file-based authentication or LDAP-based authentication is configured.

`<common-name>-<timestamp>.jks` - a keystore containing a certificate for proving identity. Note that the `<common-name>` part of the filename must match the Common Name specified in the certificate and the `<timestamp>` represents the time that the certificate was created.

Trusted authority subdirectory

The `trusted-authority` subdirectory can be specified both client-side and server-side. It contains trusted root certificates.

`trusted-authority-<timestamp>.jks` - a truststore containing a trusted root certificate for validating identity certificates. Note that the `<timestamp>` represents the time that the certificate was created.

The `whitelist.txt` file

The `whitelist.txt` file should only be specified on the server-side. It contains details of client IP addresses permitted to establish connections with the cluster.

Certificates

This section assumes a good understanding of SSL/TLS fundamentals.

Certificate creation

Note: Certificates must be created using the RSA algorithm, preferably with a key size of 4096.

The keystores and truststores, that are deployed to the `identity` and `trusted-authority` directories respectively, can be created in any way desirable as long as the following rules are followed:

Keystore creation

- Keystore must be of type `jks`.
- Keystore filename must be in `${common name}-${yyyyMMddThhmmss}.jks` format (e.g. `com.organization.host-20180223T102319.jks`). `yyyyMMddThhmmss` should represent the time of creation (timestamp) of the file. When multiple keystores are present, the keystore with the latest timestamp is used.
- Keystore must have only one `terracotta_security_alias` entry, and it should contain the identity certificate and the corresponding private key.
- Common Name field in the Distinguished Name in the certificate must match the common name fragment in the keystore filename. For a server the common name must match the host name.
- The identity certificate must be within its period of validity.
- The password for the keystore and the `terracotta_security_alias` store entry must be `terracotta_security_password`.
- The certificate must be created using the RSA algorithm, preferably with a key size of 4096.

Truststore creation

- Truststore must be of type `"jks"`.
- Truststore filename must be in `trusted-authority-${yyyyMMddThhmmss}.jks` format (e.g. `trusted-authority-20180223T102319.jks`). When multiple truststores are present, all of them are used irrespective of their timestamps.
- Truststore must have only one `terracotta_security_alias` entry, and it should contain a trusted certificate.
- The trusted certificate must be within its period of validity.
- The password for the truststore must be `terracotta_security_password`.
- The certificate must be created using the RSA algorithm, preferably with a key size of 4096.

Certificate rotation

If the certificates expire or get compromised, they must be rotated. Following are the different ways of rotating them:

Certificate rotation with cluster shutdown

The followed steps need to be performed *in order*:

1. Generate new keystore and truststore files following the rules mentioned in the *Keystore creation* and *Truststore creation* sections above.
2. Shut down all the servers and clients.
3. Replace the old keystores and truststores with the new keystores and truststores in the corresponding `identity` and `trusted-authority` directories of each client and server.
4. Start all the servers and the clients for the new certificates to take effect.

The above sequence has the advantage that it is simple to perform, with the drawback of requiring the entire cluster and the clients to be restarted.

Certificate rotation with rolling restarts

The followed steps need to be performed *in order*:

1. Generate new keystore and truststore files following the rules mentioned in the *Keystore creation* and *Truststore creation* sections above.
2. Deploy new truststores in corresponding `trusted-authority` directories of each client and server.
3. Restart all the passive servers and clients. Once the passive servers reach `PASSIVE-STANDBY` status, restart all the active servers.
4. Replace old keystores with the new keystores in corresponding `identity` directories of each client and server.
5. Restart all the passive servers and clients. Once the passive servers reach `PASSIVE-STANDBY` status, restart all the active servers.
6. Delete old truststores from corresponding `trusted-authority` directories of each client and server.
7. Restart all the passive servers and clients. Once the passive servers reach `PASSIVE-STANDBY` status, restart all the active servers.

The above sequence has the advantage that it does not require cluster downtime, with the drawback of having a significant number of steps.

Auditing

Terracotta DB can audit when security-relevant events occur. You can configure this by specifying an audit directory.

Note: Ensure that the audit directory is accessible only by users who are permitted to run the respective Terracotta DB server or by users who are allowed to read the audit log.

Audit directory structure

The auditing process creates a hierarchical filesystem structure under the audit directory:

```
<audit-directory>
├── server
│   ├── <hostname>-<port>
│   │   └── <yyyy-MM-dd>
│   │       ├── terracotta-server-audit-<yyyy-MM-dd>.0.log
│   │       └── terracotta-server-audit-<yyyy-MM-dd>.1.log
├── tmc
│   ├── <hostname>-<port>
│   │   └── <yyyy-MM-dd>
│   │       ├── tmc-audit-<yyyy-MM-dd>.0.log
│   │       └── tmc-audit-<yyyy-MM-dd>.1.log
```

When an auditable event occurs, an appropriate message is written to the audit log file. Each line in the audit log corresponds to a single event.

Audit log files generated by a server in a Terracotta cluster are stored under the `server` directory. Audit log files generated by the TMS are stored under the `tmc` directory.

Each server has a separate directory named after the hostname of the server and the port number on which the server process is listening. This allows multiple servers to use the same audit directory.

Under the server-specific directory, logs are organised by date. If any audit log file exceeds 10 MB, a new audit log file is created.

Audit entry sanitization

Any dynamic information, such as the username used in a login attempt, is sanitized before being logged.

If the sanitization process finds no unacceptable characters or character sequences, then the information is logged surrounded with backticks.

For example, if the user `alex` successfully authenticated, that is audited as:

```
LDAP authentication success:- IP: `203.0.113.1`, User: `alex`
```

However, if the sanitization process finds unacceptable characters or character sequences, then the information is logged surrounded with tildes and prefixed with `sanitized`. Characters replaced in the sanitization process appear in the format `|U+xxxx|` where the `xxxx` is the unicode codepoint in hexadecimal.

Note that there may be more than four digits in the hex codepoint due to supplementary unicode characters.

For example, if the user `jürgen` successfully authenticated, that is audited as:

```
LDAP authentication success:- IP: `203.0.113.1`, User: ~sanitized j|U+00fc|rge~
```

If a string exceeds 2000 characters, then it will be truncated and a `*` added to the end to indicate the truncation.

~, |, *, ` and + are all treated as unacceptable characters and sanitized. Other strings sequences are also sanitized, particularly any characters outside the standard, printable ASCII range.

Cluster Security

Connection encryption

You can enable connection encryption by adding an `<ssl-tls>` tag to the `<security>` section of the `tc-config.xml` file. When an `<ssl-tls>` tag is specified, you must also supply:

- a `<security-root-directory>` tag, with the path to the security root directory.
- an `<authentication>` tag, with the authentication scheme to use. See “[Authentication](#)” on page 85 for more information.

With SSL/TLS configured in the `tc-config.xml` file, the specified security root directory must contain a valid truststore in the `trusted-authority` subdirectory and a valid keystore in the `identity` subdirectory. The identity certificate is required even if file-based or LDAP-based authentication is chosen. The choice of authentication scheme may mean other files are also required in the security root directory.

On the client, the security root directory must contain a trusted authority certificate in the `trusted-authority` subdirectory. Again, the choice of authentication scheme may mean other files, including an identity certificate, are also required in the security root directory.

Cluster Authentication

To configure authentication in the cluster, add an `<authentication>` tag to the `<security>` section of the `tc-config.xml` file. Inside that tag you should configure the authentication scheme required.

When authentication is configured, a `<security-root-directory>` tag, with the path to the security root directory, must also be added to the `tc-config.xml` file.

Note: It is highly recommended that if you configure an authentication scheme, you also configure encrypted connections using an `<ssl-tls>` tag, otherwise an attacker could acquire credentials by eavesdropping on the unencrypted connection.

Certificate-based authentication

To configure certificate-based authentication, add an `<authentication>` tag containing a `<certificate/>` tag to the `<security>` section of the `tc-config.xml` file.

If certificate-based authentication is configured, an `<ssl-tls>` tag must also appear in the `<security>` section of the `tc-config.xml` file. Also, clients must have an appropriate identity certificate keystore in their security root directory.

File-based authentication

To configure file-based authentication, add an `<authentication>` tag containing a `<file/>` tag to the `<security>` section of the `tc-config.xml` file.

If file-based authentication is configured, the server's security root directory must contain a `users.xml` file. The `users.xml` file must contain the list of valid users with a password hash for each user. The only password hashing algorithm currently supported is `bcrypt`.

Example of a `users.xml` file for authentication:

```
<users>
  <user>
    <username>alex</username>
    <password>
      <algorithm>bcrypt</algorithm>
      <hash>$2a$10$UoM85/5I4SnIbrOQuFZ43ekffuQKSxZmL93bR9VMcdr2URmPyjyX2</hash>
    </password>
  </user>
  <user>
    <username>beth</username>
    <password>
      <algorithm>bcrypt</algorithm>
      <hash>$2a$10$6D6c791E0k/0SxrEtnfhGe2Yr.ygG0rFP1QzeyD9qshIMRrpUMOAS</hash>
    </password>
  </user>
</users>
```

Note: Hashes should start with `bcrypt` version `$2a$`.

Generating bcrypt password hashes

There is a command line utility for generating `bcrypt` password hashes. The `bcrypt` script is located in `tools/security/bin` under the product installation directory as `bcrypt.bat` for Windows platforms, and as `bcrypt.sh` for Unix/Linux.

When running the `bcrypt` script, you must specify the number of rounds. The number to choose depends on both the performance of the servers on which authentication takes place and also how you decide to trade off security against speed. A higher number creates hashes that are harder for attackers to crack, but are also harder for your servers to verify. Increasing the number of rounds by 1 doubles the difficulty. You may find that a value between 10 and 13 is suitable.

The `bcrypt` script can read the password from `stdin`:

```
bcrypt.sh -n 10
```

or directly from the the command line:

```
bcrypt.sh -n 10
pa$w0rd
```

The `stdin` form of the command is preferred because when a password is used on the command line, it can be written into files such as the command history and also readable using commands such as `ps`.

LDAP-based authentication

To configure LDAP-based authentication, add an `<authentication>` tag containing an `<ldap/>` tag to the `<security>` section of the `tc-config.xml` file.

If LDAP-based authentication is configured, the server's security root directory must contain a `ldap.properties` file. The properties in the `ldap.properties` file are the same properties that are used to configure LDAP integration in other Software AG products. See “[LDAP Properties](#)” on page 96 for a full list of supported properties.

Example of an `ldap.properties` file for authentication:

```
url=ldap://ldapsrvr.example.com:389
userrootdn=ou=People,dc=example,dc=com
uidprop=uid
personobjclass=person
```

Auditing

To configure auditing in the cluster, add an `<audit-directory>` tag to the `<security>` section of the `tc-config.xml` file.

IP Whitelisting

Introduction

The IP whitelisting feature enables you as the cluster administrator to ensure that only clients from known IP addresses can access the TSA. You can use this feature to prevent malicious clients from establishing connections to the TSA. The term "clients" here refers to clients communicating with the Terracotta servers using the TSA wire protocol.

Note: It should be understood that usage of this feature in itself does not provide a strong level of security for the TSA. The ideal way to enforce connection restrictions based on IP addresses would be to use host-level firewalls.

The whitelist file

A whitelist file is a plain-text file containing a list of IPs. Only clients running on these IPs are allowed to access the TSA. The server IPs specified in the `tc-config`, and the localhost IPs of the server are always whitelisted. An empty whitelist file has the semantics of blacklisting all the IPs, except the IPs fetched from the `tc-config`, and those corresponding to `localhost`.

The following rules need to be followed for a whitelist file to be considered valid, and the entries in it to be parsed properly:

1. The whitelist file must be named `whitelist.txt` and placed in the security root directory.
2. The entries can be IP addresses, or CIDR notations (to represent IP ranges). Any entry that is not a valid IP address or a valid CIDR is ignored.

3. Each line in the file can contain either a single IP address or a comma-separated list of IP addresses.
4. Lines beginning with # are considered as comments, and are ignored during parsing.
5. Blank lines are ignored.

The following is an example of a valid whitelist file:

```
# Whitelist for Terracotta cluster
# Caching clients
192.168.5.28, 192.168.5.29, 192.168.5.30
10.60.98.0/28
# Other clients
192.168.10.0/24
```

Usage

To configure IP whitelisting, add a `<whitelist>` tag to the `<security>` section of the `tc-config.xml` file.

When IP whitelisting is configured, a `<security-root-directory>` tag, with the path to the security root directory, must also be added to the `tc-config.xml` file.

If the `whitelist.txt` file is not found in the security root directory, or there is an error reading the file, the server startup will fail with an appropriate error message.

If hostnames are used in the `tc-config.xml` file instead of IPs, the server attempts to resolve these hostnames to IPs. If the resolution fails, the server startup fails with an appropriate error message. Note that hostname resolution is done for the `tc-config.xml` file only, and any hostnames present inside the `whitelist.txt` file are ignored.

A multi-stripe cluster should be started with the same `whitelist.txt` file, or identical copies of the file. Similarly, when updates to the file are desired, they should be performed on all the stripes, as described in the following section.

Dynamic updates

After a cluster is started with whitelisting enabled, entries can be dynamically added to or removed from the whitelist without the need for server restarts. To perform a dynamic update, edit the `whitelist.txt` file as needed, and run the `ipwhitelist-reload cluster-tool` command to notify the servers in the cluster to reload the `whitelist.txt` file. Refer to the [“The “ipwhitelist-reload” Command” on page 53](#) section for details of the usage.

Errors during whitelist reload, if any, are logged in individual Terracotta Server logs. Thus, after every update operation, server logs should be checked to verify that the updates took effect in all the servers.

If a cluster is not yet configured and the whitelist file needs to be reloaded on the servers, the `server-levelipwhitelist-reload` command can be used. It may also be helpful when the machine from where the cluster tool is to be used is itself not whitelisted initially. In this scenario, adding this machine's IP to the whitelist, and running the server-level `ipwhitelist-reload` command ensures that cluster tool can configure the cluster later.

If any failures happen while reading the `whitelist.txt` or the `tc-config.xml` files during a dynamic update, the updates are ignored and the server continues with the current whitelist. No partial updates are applied. The update is not be retried, and the `ipwhitelist-reload` command will have to be run again after resolving the problems.

Connection Behaviour

When a client connects to a server on the `tsa-port`, the server accepts the socket connection, and verifies the IP of the incoming client connection against the whitelist. If it finds that the client IP is not whitelisted, it closes the socket connection.

If a whitelisted client is removed from the whitelist via a dynamic update, it remains connected to the cluster as long as there is no network disconnection or explicit connection closure from the client. Subsequent connection attempts from the client to cluster will fail.

Full example of secure server XML configuration

Here is an example of a `tc-config.xml` security configuration section that configures auditing, encrypted connections, LDAP authentication and IP whitelisting:

```
<service>
  <security xmlns="http://www.terracottatech.com/config/security">
    <audit-directory>/path/to/audit-directory</audit-directory>
    <security-root-directory>/path/to/security-root-directory</security-root-directory>
    <ssl-tls/>
    <authentication>
      <ldap/>
    </authentication>
    <whitelist/>
  </security>
</service>
```

With the configuration in the above example, the server's security root directory would contain a truststore in the `trusted-authority` subdirectory, a keystore and a `credentials.properties` in the `identity` directory, an `ldap.properties` in the `access-control` subdirectory and a `whitelist.txt` file.

The keystore is required so that the server can prove to clients that it is the server they are expecting. The `credentials.properties` file is required so that the server can connect to other servers in the stripe. If there are no other servers in the stripe, the `credentials.properties` file should not be present.

Client configuration

To configure a client to connect to a secured cluster, you need to give the client a path to the client's security root directory. This should contain, for example, the credentials that the client needs to connect to the cluster.

Cluster tool

To enable the cluster tool to connect to a secure cluster, each cluster tool command must be prefixed with `-srd` (or long option `--security-root-directory`) as in the following example:

```
./cluster-tool.sh -srd /path/to/security-root-directory configure -l ~/license.xml -n tc-cluster
Configuration successful
License installation successful
Command completed successfully.
```

Attempting to connect to a secure cluster without the `-srd` option will fail. Commands without this option retain their behavior. For more information on cluster tool commands, see the section “[Cluster Tool](#)” on page 43.

Ehcache Client

An Ehcache client can define either an XML or a programmatic configuration, both of which support security configuration. The following are the examples of usages of each:

1. API Example

```
PersistentCacheManager cacheManager = CacheManagerBuilder
    .newCacheManagerBuilder()
    .with(EnterpriseClusteringServiceConfigurationBuilder
        .enterpriseSecureCluster(connectionURI,
            securityRootDirectoryPath) // 1
        .autoCreate())
    .build(true);
```

1	<p><code>EnterpriseClusteringServiceConfigurationBuilder enterpriseSecureCluster (URI, Path)</code> lets you create a <code>CacheManager</code> using a secure connection. The first argument is the URI of the Terracotta cluster, appended with the <code>CacheManager</code> name. The second argument is the Path to the client's security root directory. <code>EnterpriseClusteringServiceConfigurationBuilder enterpriseSecureCluster (Iterable<InetSocketAddress>, String, Path)</code> serves the same function, with the added support for IPv6 addresses.</p>
---	--

Note: If the URI or `Iterable<InetSocketAddress>` contains host names, make sure that they match the host names specified in the server certificates.

The `EnterpriseClusteringServiceConfigurationBuilder enterpriseCluster (URI)` API continues to provide unsecured connections to unsecured clusters.

2. XML Example

```
<ehcache:config
  xmlns:ehcache="http://www.ehcache.org/v3"
  xmlns:tc="http://www.terracottatech.com/v3/terracotta/ehcache">
  <ehcache:service>
    <tc:cluster>
      <tc:connection url="${cluster-uri}/CM" security-root-directory="${security-root-director
      <tc:server-side-config auto-create="true"/>
    </tc:cluster>
  </ehcache:service>
</ehcache:config>
```

1	<code>security-root-directory</code> lets you specify the Path to the client's security root directory. Not passing this option retains the behavior of communicating with an unsecured cluster.
---	--

TC Store Client

1. API Example

```
DatasetManager datasetManager = DatasetManager.secureClustered(
    connectionURI, securityRootDirectoryPath) // 1
    .build();
```

1	<code>DatasetManager.secureClustered(URI, Path)</code> lets you create a <code>DatasetManager</code> using a secure connection. The first argument is the URI of the Terracotta cluster. The second argument is the Path to the security root directory which is to be used for the connection. <code>DatasetManager.secureClustered(Iterable<InetSocketAddress>, Path)</code> serves the same function, with the added support for IPv6 addresses.
---	---

Note: If the URI or `Iterable<InetSocketAddress>` contains host names, make sure that they match the host names specified in the server certificates.

The `DatasetManager.clustered(URI)` API continues to provide unsecured connections to unsecured clusters.

2. XML Example

```
<clustered xmlns=
    "http://www.terracottatech.com/v3/store/clustered"> <!--1-->
    <cluster-connection>
        <server host="localhost" port="9410"/>
        <security-root-directory>/path/to/security-root-directory
        </security-root-directory> <!--2-->
    </cluster-connection>
</clustered>
```

1	Declares a clustered <code>DatasetManager</code> configuration.
2	<code>security-root-directory</code> lets you specify the Path to the client's security root directory. Not specifying this element retains the behavior of communicating with an unsecured cluster.

TMS Security

TMS security is configured using the `tmc.properties` file.

Setting the security root directory

If you need to set the security root directory to use for connections between a browser and the TMS, then you should set the `tms.security.root.directory` property to the path for the security root directory.

Windows platforms

Note that Windows paths often contain backslashes. The Java properties format requires backslashes to be escaped, so `C:\tcd\security-root-directory` would be configured using:

```
tms.security.root.directory=C:\\tcd\\security-root-directory
```

Alternatively, you can use forward slashes:

```
tms.security.root.directory=C:/tcd/security-root-directory
```

Encrypted connections between the browser and the TMS

You can enable HTTPS by setting:

```
tms.security.https.enabled=true
```

When HTTPS is enabled, you must also set the `tms.security.root.directory` property to the path for the security root directory. See [“Setting the security root directory” on page 92](#).

When HTTPS is configured, the specified security root directory must contain a valid truststore in the `trusted-authority` subdirectory and a valid keystore in the `identity` subdirectory.

When HTTPS is configured, the TMS is only accessible over HTTPS. Thus, any URL used to access the TMS must start with `https://`.

Browser warnings

If you are using self-signed certificates, you may see a security warning in your browser when connecting to the TMS. This is because the certificate was not signed by a certificate authority registered with the browser as a trusted certificate authority. We suggest that you check the certificate and verify its authenticity.

Furthermore, you should add the certificate to your browser's list of trusted root certificate authorities.

Note: Each cookie entry is associated with a certain domain (not including port), and some browsers may remember the protocol under which the cookie has been set. So if you switch between `http` and `https`, or between `localhost` and `127.0.0.1`, you should clear related cookies before reloading TMC in the browser.

Authentication

To configure TMS to require users to log in, you should set the `tms.security.authentication.scheme` property. There are two options: `file` and `ldap`, for file-based authentication and LDAP-based authentication respectively.

When authentication is configured, you must also set the following properties:

- `tms.security.root.directory` - see [“Setting the security root directory” on page 92](#).
- `tms.security.authorization.scheme` - see [“Authorization” on page 93](#).

File-based authentication requires a `users.xml` file to be added to the security root directory; LDAP-based authentication requires an `ldap.properties` file. See [“Cluster Authentication” on page 85](#) for more details.

Authorization

To control which access to TMS operations, you should set the `tms.security.authorization.scheme` property. There are three options: `authenticated`, `file` and `ldap`. The `authenticated` option allows access to all operations to all authenticated users. The `file` and `ldap` options correspond to file-based authorization and LDAP-based authorization respectively.

When authorization is configured, you must also set the following properties:

- `tms.security.root.directory` - see [“Setting the security root directory” on page 92](#).
- `tms.security.authentication.scheme` - see [“Authentication” on page 93](#).

TMS supports one role: `admin`. Users with no role can use most of the TMS functionality, but some operations, such as shutting down a server, require the `admin` role.

File-based authorization

File-based authorization requires a `users.xml` file to be added to the security root directory. See [“Cluster Authentication” on page 85](#) for more details.

For file-based authorization, the `users.xml` file must contain role information.

Example of a `users.xml` file for authorization:

```
<users>
  <user>
    <username>alex</username>
    <roles/>
  </user>
  <user>
    <username>beth</username>
    <roles>
      <role>admin</role>
    </roles>
  </user>
</users>
```

In this example `users.xml` file, `beth` has the `admin` role whereas `alex` does not.

Note that you can use a mixture of authentication and authorization schemes. For example, you could use LDAP-based authentication with file-based authorization. As in the example above, the `users.xml` file need not contain password hash information, if it is not being used for authentication. However, if you use `file` for both the authentication and authorization schemes, then you must specify both password hashes and role information.

Example of a `users.xml` file for both authentication and authorization:

```
<users>
  <user>
    <username>alex</username>
    <password>
      <algorithm>bcrypt</algorithm>
      <hash>$2a$10$UoM85/5I4SnIbrOQuFZ43ekffuQKSxZmL93bR9VMcdr2URmPyjyX2</hash>
    </password>
    <roles/>
  </user>
  <user>
    <username>beth</username>
    <password>
      <algorithm>bcrypt</algorithm>
      <hash>$2a$10$6D6c791E0k/0SxrEtnfhGe2Yr.ygG0rFP1QzeyD9qshIMRrpUMOAS</hash>
    </password>
    <roles>
      <role>admin</role>
    </roles>
  </user>
</users>
```

LDAP-based authorization

LDAP-based authorization requires an `ldap.properties` file. See [“Cluster Authentication” on page 85](#) for more details.

The properties in the `ldap.properties` file are the same properties that are used to configure LDAP integration in other Software AG products. See [“LDAP Properties” on page 96](#) for a full list of supported properties.

Example of an `ldap.properties` file for authorization:

```
url=ldap://ldapserver.example.com:389
userrootdn=ou=People,dc=example,dc=com
uidprop=uid
personobjclass=person
memberinfoingroups=true
mattr=uniqueMember
grouprootdn=ou=Group,dc=example,dc=com
gidprop=gid
groupobjclass=group
```

You may wish to map the role defined in the LDAP server to the TMS `admin` role, in which case, you can add the `tcdm.roleMap` property to the `ldap.properties` file. For example:

```
tcdm.roleMap=terracottaTmsAdmin=admin
```

would map the `terraccottaTmsAdmin` group defined in the LDAP server onto the TMS `admin` role.

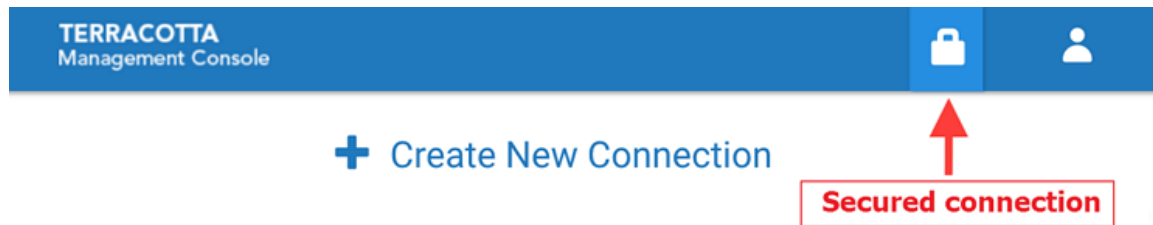
Auditing

To configure auditing for the TMS, set the `tms.security.audit.directory` property to the audit directory.

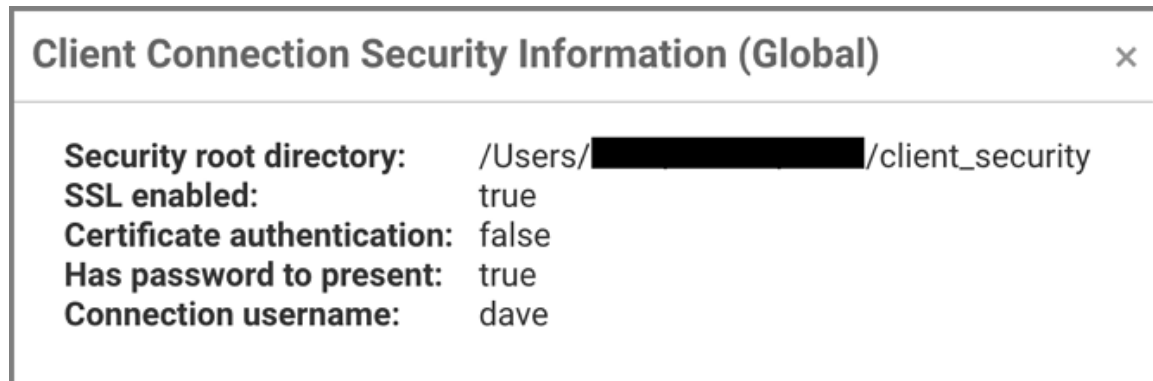
Connecting the TMS to secured clusters

To allow the TMS to connect to a secure cluster, you should set the `tms.security.root.directory.connection.default` property to the path for the security root directory containing, for example, credentials to connect to the secured cluster.

After configuring a secured cluster and setting up the TMS client security root directory, you should see a padlock icon in TMC:



When you click on the padlock, a pop-up window shows some security related information:



Note: This security root directory and its configuration relates to the connection between the TMS server and the secured cluster. It is independent of the connection security between browser and TMS server.

SSL enabled	Connection is SSL enabled - i.e. the security root directory has a trusted authority certificate.
-------------	---

Certificate authentication	Connection is 2-way SSL enabled - i.e. the security root directory has an identity certificate.
Has password to present	The TMS/TMC client has a <code>credentials.properties</code> file.

Windows platforms

Note that Windows paths often contain backslashes. The Java properties format requires backslashes to be escaped, so `C:\tcd\audit-directory` would be configured using:

```
tms.security.audit.directory=C:\\tcd\\audit-directory
```

Alternatively, you can use forward slashes:

```
tms.security.audit.directory=C:/tcd/audit-directory
```

LDAP Properties

The minimum set of properties to specify is:

```
url
userrootdn
uidprop
personobjclass
```

If you are using LDAP for authorization, you must also specify:

```
memberinfoingroups
mattr
```

and, if you set `memberinfoingroups` to `true`, you must also specify:

```
grouprootdn
gidprop
groupobjclass
```

Most other properties need only be used if you have specific requirements.

Connection related properties

`url` - The URL of the LDAP server (e.g. `ldap://ldapserver:389` or `ldaps://ldapserver:636`).

`keystoreUrl` - The URL from which a keystore can be retrieved (e.g. `file:///usr/local/ldap/keystore.jks`) - used to authenticate to the LDAP server.

`keystoreType` - The store type of the keystore (e.g. `JKS`).

`keystorePassword` - The password to verify the integrity of the keystore.

`keyAlias` - The alias in the keystore where the certificate and key are stored.

`keyPassword` - The password to allow access to the specified alias. Defaults to `store password`.

`truststoreUrl` - The URL from which a truststore can be retrieved. Used to validate the certificate presented by the LDAP server during an SSL/TLS handshake.

`truststoreType` - The store type of the truststore (e.g. `JKS`).

`truststorePassword` - The password to verify the integrity of the truststore.

`noPrinIsAnonymous` - Set to true for LDAP servers that allow anonymous connections.

`prin` - The username to use to authenticate to the LDAP server.

`cred` - The password to use to authenticate to the LDAP server.

Note: If `prin` and `cred` are not specified and `noPrinIsAnonymous` is not set to true, then the username and password of the user attempting to authenticate to the cluster / TMS will be used to authenticate to the LDAP server.

`watt.server.ldap.ignore.serverCertificateValidity` - If true, then invalid certificates presented by the LDAP server are ignored.

`watt.server.ldap.extendedProps` - Extra properties to add to the LDAP context.
Format: `key1=value1;key2=value2`

`watt.server.ldap.retryCount` - How many times to retry a connection if it fails.

`watt.server.ldap.retryWait` - How many milliseconds to wait between connection retries.

Properties related to how to interact with the LDAP server

`timeout` - LDAP query timeout in milliseconds.

`watt.server.ldap.DNescapeChars` - A list of characters that should be escaped.

`watt.server.ldap.DNescapePairs` - A list of characters that should not be re-escaped.

`watt.server.ldap.DNstripQuotes` - If false, then quotes that get added when escaping are not striped from DN's.

`watt.server.ldap.DNescapeURL` - If true, then the start of a DN is escaped. This is useful for referrals when DN's can start with a URL.

`watt.server.jndi.searchresult.maxlimit` - The maximum number of results to return from an LDAP search. Zero means unlimited.

Properties related to the schema of a user

`userrootdn` - The DN under which users can be found (e.g. `ou=People,dc=example,dc=com`).

`uidprop` - The attribute on a user which contains the primary ID of the user (e.g. `uid`).

`personobjclass` - The LDAP schema class for users (e.g. `person`).

`useaf` - If true, then the `dnprefix` and `dnsuffix` properties should be used.

`dnprefix` - A string added to the beginning of a username for the LDAP lookup.

`dnsuffix` - A string added to the end of a username for the LDAP lookup.

Properties related to the schema of a group

`grouprootdn` - The DN under which groups can be found (e.g. `ou=Group,dc=example,dc=com`).

`gidprop` - The attribute on a group which contains the primary ID of the group (e.g. `gid`).

`groupobjclass` - The LDAP schema class for groups (e.g. `group`).

Properties related to how the schema connects users and groups

`group` - A role automatically given to every user.

`memberinfoingroups` - If true, then group membership is in the group definitions under the `grouprootdn`. If false, then group membership is in the user definitions under the `userrootdn`.

`matr` - The attribute on a user that specifies a group to which the user belongs OR the attribute on a group that specifies a user is a member. The semantics depends on the choice of `memberinfoingroups`.

`recursiveSearchDepth` - How deep to search for groups that are members of other groups.

Properties that Terracotta DB supports in addition to other Software AG products

`tcdb.roleMap` - A mapping from group names on the LDAP server to roles used in Terracotta DB. Format: `group1=tcdbRole1;group2=tcdbRole2`

Note: Multiple LDAP groups can map to the same role.

SSL / TLS Troubleshooting guide

This document provides a list of the most commonly seen problems related to [“Cluster Security” on page 85](#), and their solutions:

Problem category: Host fails to start

This section describes the most commonly seen problems related to a host (server or a client) startup.

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
configured security-root-directory /path/to/security-root-directory
does not exist
```

Diagnosis

The specified security root directory does not exist.

Action

Make sure that the directory exists and contains `identity` and `trusted-authority` directories with valid keystores and truststores in them.

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:  
com.terracottatech.security.common.exception.SecurityConfigurationException:  
identity directory doesn't exist in configured  
security-root-directory /path/to/security-root-directory
```

Diagnosis

The specified security root directory does not contain an `identity` directory.

Action

Make sure that the directory exists inside the security root directory and contains valid keystores.

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:  
com.terracottatech.security.common.exception.SecurityConfigurationException:  
trusted-authority directory doesn't exist in configured  
security-root-directory /path/to/security-root-directory
```

Diagnosis

The specified security root directory does not contain a `trusted-authority` directory.

Action

Make sure that the directory exists inside the security root directory and contains valid truststores.

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:  
com.terracottatech.security.common.exception.SecurityConfigurationException:  
No acceptable keystore files found in identity directory  
/path/to/security-root-directory/identity
```

Diagnosis

Either of:

- `identity` directory does not contain any keystores.

- `identity` directory contains keystores, but their file names are not in the format `${common name}-${yyyyMMddThhmmss}.jks` (e.g. `com.organization.host-20180223T102319.jks`).

Action

Make sure that `identity` directory contains keystores which follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
No acceptable truststore files found in trusted-authority directory
/path/to/security-root-directory/trusted-authority
```

Diagnosis

Either of:

- `trusted-authority` directory does not contain any truststores.
- `trusted-authority` directory contains truststores, but their file names are not in the format `${common name}-${yyyyMMddThhmmss}.jks` (e.g. `trusted-authority-20180223T102319.jks`).

Action

Make sure that `trusted-authority` directory contains truststores which follow the truststore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
Tried to use the password terracotta_security_password to load the
keystore file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
but that failed
```

Diagnosis

Latest keystore file does not have `terracotta_security_password` as its password, where *latest keystore file* is the keystore file with the latest timestamp string in the filename (e.g., `host-20180131T120830.jks` is considered newer than both `host-20170131T120830.jks` and `host-20180131T120822.jks`).

Action

Make sure the keystores follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
```

```
com.terracottatech.security.common.exception.SecurityConfigurationException:
Tried to use the password terracotta_security_password to read the
keystore entry with alias terracotta_security_alias in the keystore
file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
but that failed
```

Diagnosis

Latest keystore file does not have `terracotta_security_password` as `terracotta_security_alias` entry password, where *latest keystore file* is the keystore file with the latest timestamp string in the filename (e.g., `host-20180131T120830.jks` is considered newer than both `host-20170131T120830.jks` and `host-20180131T120822.jks`).

Action

Make sure the keystores follow the follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
Unable to find required private key/certificate chain entry using
alias terracotta_security_alias in keystore file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
```

Diagnosis

Latest keystore file does not have `terracotta_security_alias` as certificate alias, where *latest keystore file* is the keystore file with the latest timestamp string in the filename (e.g., `host-20180131T120830.jks` is considered newer than both `host-20170131T120830.jks` and `host-20180131T120822.jks`).

Action

Make sure the keystores follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
Certificate in keystore file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
is expired
```

Diagnosis

Latest keystore file contains an expired certificate, where *latest keystore file* is the keystore file with the latest timestamp string in the filename (e.g., `host-20180131T120830.jks` is considered newer than both `host-20170131T120830.jks` and `host-20180131T120822.jks`).

Action

Make sure the keystores follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
Certificate in keystore file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
is not valid yet
```

Diagnosis

Latest keystore file contains a certificate with a future start date, where *latest keystore file* is the keystore file with the latest timestamp string in the filename (e.g., `host-20180131T120830.jks` is considered newer than both `host-20170131T120830.jks` and `host-20180131T120822.jks`).

Action

Make sure the keystores follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
The common name org.host of the certificate that was loaded from
keystore file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
doesn't match the common name com.organization.host in the filename
com.organization.host-20180131T120830.jks
```

Diagnosis

Common Name field in the Distinguished Name in the host's certificate does not match the common name fragment in the latest keystore filename.

Action

Make sure the keystores follow the keystore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
No valid trusted certificates found in trusted-authority directory
/path/to/security-root-directory/trusted-authority; Unable to find
required trusted certificate entry using alias
terracotta_security_alias in truststore file
/path/to/security-root-directory/trusted-authority/trusted-authority-20180131T120832.jks
```

Diagnosis

No valid truststores were found. The specific truststore reported in the Exception message contains a certificate which uses an alias other than `terraccotta_security_alias`. Note that this Exception can be followed by one or more Suppressed Exceptions that can indicate why other truststores could not be used.

Action

Make sure the truststores follow the truststore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terraccottatech.security.common.exception.SecurityConfigurationException:
No valid trusted certificates found in trusted-authority directory
/path/to/security-root-directory/trusted-authority; Certificate in
truststore file
/path/to/security-root-directory/trusted-authority/trusted-authority-20180131T120834.jks
is expired
```

Diagnosis

No valid truststores were found. The specific truststore reported in the Exception message contains an expired certificate. Note that this Exception can be followed by one or more Suppressed Exceptions that can indicate why other truststores could not be used.

Action

Make sure the truststores follow the truststore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terraccottatech.security.common.exception.SecurityConfigurationException:
No valid trusted certificates found in trusted-authority directory
/path/to/security-root-directory/trusted-authority; Certificate in
truststore file
/path/to/security-root-directory/trusted-authority/trusted-authority-20180131T120834.jks
is not valid yet
```

Diagnosis

No valid truststores were found. The specific truststore reported in the Exception message contains a certificate with a future start date. Note that this Exception can be followed by one or more Suppressed Exceptions that can indicate why other truststores could not be used.

Action

Make sure the truststores follow the truststore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
No valid trusted certificates found in trusted-authority directory
security-root-client\trusted-authority; Tried to use the password
terracotta_security_password to load the truststore file
security-root-client\trusted-authority\trusted-authority-20180131T120832.jks
but that failed
```

Diagnosis

No valid truststores were found. The specific truststore reported in the Exception message does not have `terracotta_security_password` as its password. Note that this Exception can be followed by one or more Suppressed Exceptions that can indicate why other truststores could not be used.

Action

Make sure the truststores follow the truststore creation rules as described in the section [“Certificates” on page 81](#).

Symptom

Host fails to start with an Exception message similar to:

```
java.lang.RuntimeException:
com.terracottatech.security.common.exception.SecurityConfigurationException:
Unable to validate certificate chain with alias
terracotta_security_alias in keystore file
/path/to/security-root-directory/identity/com.organization.host-20180131T120830.jks
using truststore file(s)
```

Diagnosis

Host certificate in latest keystore file is not signed by any of the known trusted authorities, and thus cannot be validated by any of the truststore files. *Latest keystore file* here is the keystore file with the latest timestamp string in the filename (e.g., `host-20180131T120830.jks` is considered newer than both `host-20170131T120830.jks` and `host-20180131T120822.jks`).

Action

Make sure that the latest keystore file contained in the identity directory is signed by the truststores in the `trusted-authority` directory.

Problem category: Connection fails to establish

Symptom

```
org.terracotta.connection.ConnectionException:
com.terracottatech.connection.ProbableSecurityConfigurationException:
Handshake with server failed when this client tried to initiate a
non-secure connection. Possible reason: server is running with
security enabled.
```

Diagnosis

The client which tried to establish a connection to a server running with SSL/TLS configuration is not using an SSL/TLS configuration.

Action

Make sure the client uses a correct SSL/TLS configuration (via a secure API, an XML config, command parameters etc.) so that it can establish a secure connection with an SSL/TLS security-enabled server.

Symptom

```
org.terracotta.connection.ConnectionException:  
com.terracottatech.connection.ProbableSecurityConfigurationException:  
Handshake with server failed when this client tried to initiate a  
secure connection. Possible reasons: client security configuration  
is not valid, or server is not running with security  
enabled.
```

Diagnosis

Either of:

1. The client is using an SSL/TLS security configuration but the server is not.
2. The client cannot validate the server because the server's CA certificate is not present in the client's trusted certificates.
3. The server cannot validate the client because the client's CA certificate is not present in the server's trusted certificates.

Action

Make sure that:

1. The client uses an unsecured configuration (via an unsecured API, an XML config, command parameters etc.) if the server is running with an unsecured configuration.
2. The client and the server certificates are signed by the same CA and their `trusted-authority` directories contain the same truststores.

20 Terracotta Server Migration from Terracotta BigMemory Max v4 to Terracotta DB v10

Terracotta DB 10.x is significantly different from Terracotta BigMemory Max 4.x and Terracotta 3.x in terms of handling of cluster topology, data storage formats, and other functionality. Because of this, you cannot migrate data and configuration for a Terracotta BigMemory Max server to a Terracotta DB server. If you install Terracotta DB on the same machines that host Terracotta BigMemory Max, however, you can find the host names, addresses, and so on that you used for the Terracotta DB BigMemory Max installation in the server configuration files.

21 Using Command Central to Manage Terracotta

Software AG Command Central is a tool that release managers, infrastructure engineers, system administrators, and operators can use to perform administrative tasks from a centralized location. It assists with configuration, management, and monitoring tasks in a simple and flexible manner.

Terracotta server instances can be managed from Command Central like other Software AG products. Both the Command Line and Web Interfaces of Command Central are supported.

Disk location

The Terracotta related files can be found under `{installation_root}/TerracottaDB/server/SPM`. This directory contains the following:

1. **bin**: Contains scripts to start and shut down the server.
2. **conf**: Contains Terracotta specific configuration files, including `tc-config.xml`. If any changes to the configuration are required, such as increasing the offheap or changing the server name, the `tc-config.xml` file needs to be updated manually.
3. **logs**: Contains Terracotta server and Terracotta DB Platform Manager Plug-in logs.

Supported Commands

Terracotta supports the following Command Central CLI (Command Line Interface) commands:

1. Inventory

- `sagcc list inventory components` : Lists information about run-time components.
- `sagcc get inventory components` : Retrieves information about a specified run-time component.

2. Lifecycle

- `sagcc exec lifecycle` : Executes a lifecycle action against run-time components. See “[Lifecycle Actions for Terracotta](#)” on page 110 for Terracotta-specific information about Lifecycle Actions.

3. Monitoring

- `sagcc get monitoring state` : Retrieves the run-time status and run-time state of a run-time component.
- `sagcc get monitoring alerts` : Lists the alerts for a specified run-time component.

- `sagcc get monitoring runtimestatus`: Retrieves the run-time status of a run-time component.

4. Configuration

- `sagcc get configuration data`: Retrieves data for a specified configuration instance that belongs to a specified run-time component.
- `sagcc list configuration types`: Lists information about configuration types for the specified run-time component. See “[Supported Configuration Types](#)” on page 110 for Terracotta-specific information about configuration types.
- `sagcc list configuration instances`: Retrieves information about a specific configuration instance that belongs to a specified run-time component.

5. Diagnostics

- `sagcc list diagnostics logs`: Lists the log files that a specified run-time component supports.
- `sagcc get diagnostics logs`: Retrieves log entries from a log file. Log information includes the date, time, and description of events that occurred with a specified run-time component.

For information about Command Central CLI commands, see the Command Central Help.

Supported Configuration Types

Terracotta supports creating instances of the following configuration types:

- `JVM-OPTIONS`: The JVM memory settings for the Terracotta Server instance in `JAVA_OPTS` environment variable format.

Changes to this configuration will be effective upon a server restart.

- `TC-SERVER-NAME`: The name for the Terracotta Server instance. The default is the hostname of the machine the server is running on. This name should match the server name in the Terracotta configuration file. Therefore, after changing it from Command Central, proceed to the `conf` directory mentioned in the “[Disk location](#)” on page 109 section above, and update the `tc-config.xml` with the same name.

Changes to this configuration will be effective upon a server restart.

Lifecycle Actions for Terracotta

Terracotta supports the following lifecycle actions with the `sagcc exec lifecycle` CLI command and the Command Central Web Interface:

- `Start`: Start a server instance.
- `Restart`: Restart a running server instance.
- `Stop`: Stop a running server instance.

Runtime Monitoring Statuses for Terracotta

Terracotta can return the following statuses from `sagcc get monitoring runtimestatus` and `sagcc get monitoring state` CLI commands and the Command Central Web Interface:

- **ONLINE_MASTER**: The server instance is running and is the master (Active) in its stripe.
- **ONLINE_SLAVE**: The server instance is running and is a slave (Passive) in its stripe.
- **STARTING**: The server instance is starting. This is usually shown when the startup takes longer than expected because either it is a slave (Passive) synchronizing with its master (Active), or it is recovering from an error condition.
- **STOPPING**: The server instance is stopping.
- **STOPPED**: The server instance is not running.
- **FAILED**: The server instance was running, but crashed. Check the server logs to find out the reason.
- **UNRESPONSIVE**: The server instance is running, but is not responding.
- **UNKNOWN**: The state of the server instance is not known. This is most likely because of an unexpected exception or error that occurred while trying to fetch the server status.