Loading Compositions

CloudTest version

This document applies to CloudTest build 5060 and later.

Overview of the CloudTest environment

A CloudTest environment typically contains, at a minimum, a "Main Instance" (server) that contains the Web Application and a "Maestro" service (load generator). For higher-volume situations, there may be a separate Database server that contains the Repository and the Results Service with its Result database.

Optionally, there may be additional servers that contain additional Maestros (load generators) and additional Results Services.

This general structure applies regardless of whether physical servers are being used, or a cloud-based environment is being used, or a combination.

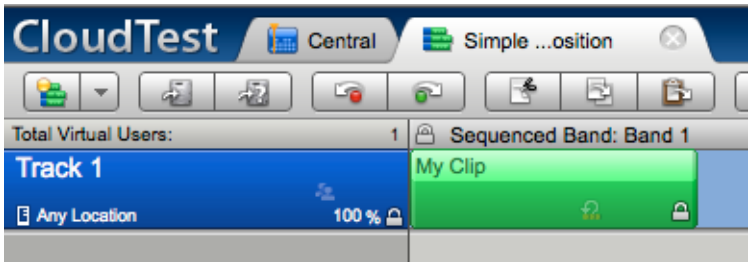
The CloudTest Lite version always consists of a single server.

The CloudTest Pro version usually consists of multiple servers, since Compositions cannot be played on the Main Instance in that environment.

Overview of Track play locations for Compositions

Every Track in a Composition may optionally specify where it is to play.

By default when a Composition is created, all of its Tracks are set to have no preference as to where they are played, as denoted by the text "Any Location":

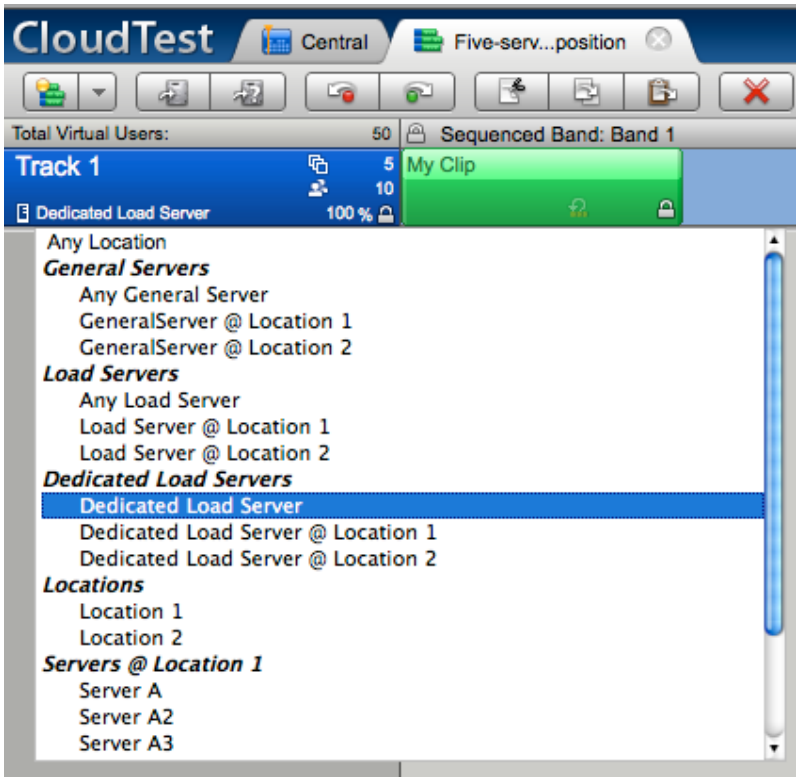


Maestro servers are configured as one of two types: "General" or "Load". A "General" Maestro server can play any number of Compositions simultaneously (except for the Lite version). A Maestro server of type "Load" can only play one Composition at time. In addition, the types of Maestro servers are used to direct Tracks to specific types of servers.

A typical multi-server CloudTest environment has one Maestro server of type "General" on the "Main instance", and any number of Maestro servers of type "Load", although it is possible for the System Administrator to configure the environment differently.

Servers are grouped into "Locations", which typically represent different geographical locations, but do not necessarily have to.

Each Track in a Composition can be directed to a different Maestro server according to the criteria in the location dropdown in the Composition Editor:

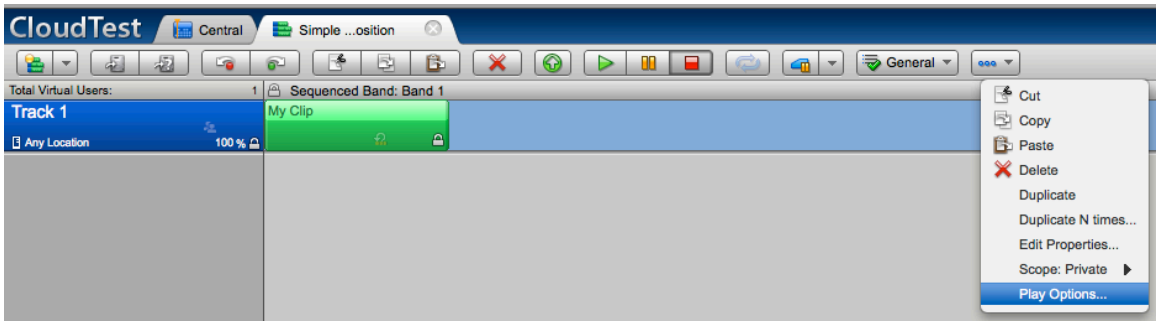


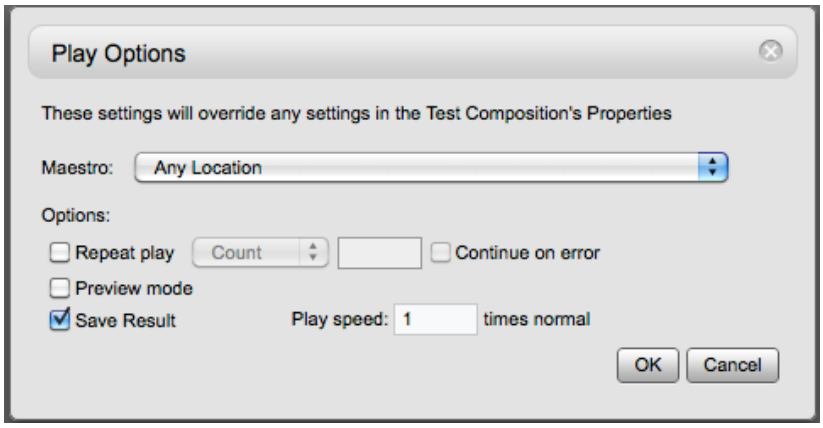
"Dedicated Load Server" means that the Track is to be played on a separate Maestro server of type "Load" all by itself and will not share that server with any other Track or Composition. This is typically used in high-volume load tests to guarantee good and consistent performance. This is different from the "Load Server" choice, in which case the Track will be played on a Maestro server of type "Load" but might share that server with other Tracks from the same Composition.

The last portions of the dropdown list each individual servers and locations separately, making it possible to assign the Track to a specific individual server or location.

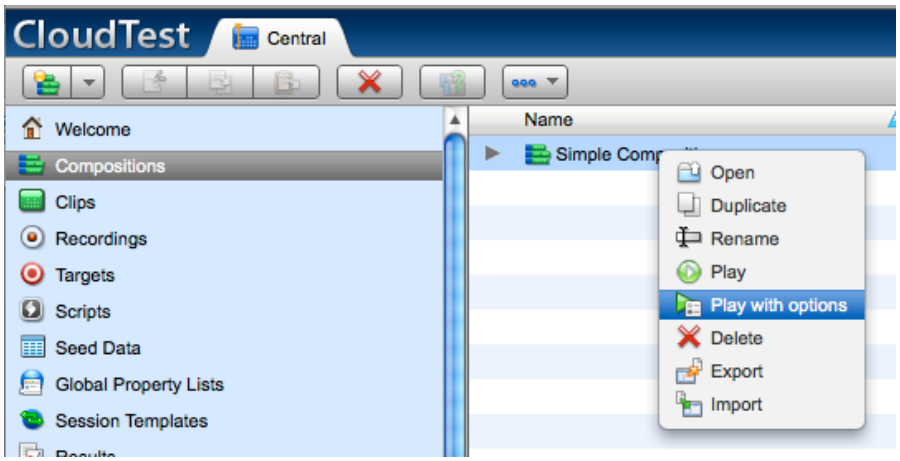
Play location for the Composition as a whole

When a Composition is started, the location on which the Composition is to be started may optionally be specified. This is done through the "Play Options" dialog in the Composition Editor:





When a Composition is played from the list of Compositions in "Central", rather than from the Composition Editor, the Play Options dialog can be accessed by choosing "Play with Options":



It is important to remember that "Play Options" are temporary settings. They are not saved to the Composition and selecting them does not change the Composition itself. In the Composition Editor the Play Options last as long as the Composition is open in the editor, and are discarded when the Composition is closed. When a Composition is played from the list of Compositions in Central, the Play Options are one-time only for that specific play of the Composition.

Overview of loading a Composition

When a Composition is loaded, the Web Application first contacts the Maestro service that resides on the same server as the Web Application itself (on the "Main Instance").

If "Play Options" was not used, the Main Maestro loads the Composition and analyzes all of the Tracks and their locations, and searches for a Maestro server that would be appropriate to play the Composition, which may be the Main Maestro itself, or may be another Maestro server. This process may require multiple different Maestros to be tried before an appropriate one is found.

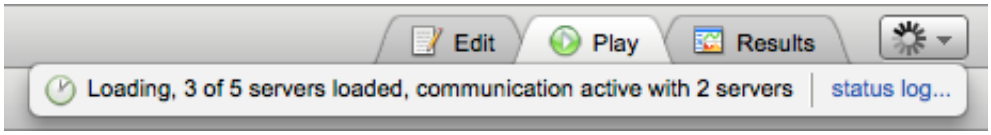
If "Play Options" was used, the Main Maestro service searches for the Maestro server that is appropriate based upon the Play Options given, and transfers the Composition load

task to that Maestro server. This process may require multiple different Maestro servers to be contacted before an appropriate one is found. The Maestro server to which the task is transferred then performs the load and analysis process described in the prior paragraph.

Note that even if a starting location was specified via "Play Options", the Composition might still be moved to a different Maestro server. For example, if "Play Options" was used to specify that the Composition play on Server A, but all of the Tracks in the Composition specify that the Tracks are to be played on Server B, then the Composition will be moved to Server B.

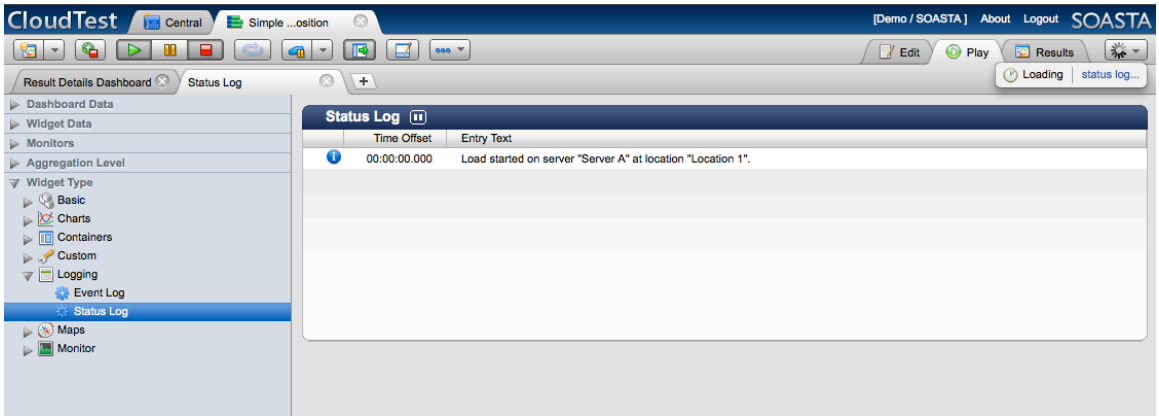
Viewing load status

Status information about a Composition load in progress is available in a short line of status text visible on the status line in the Composition Editor:



More detailed status is available in the Status Log. The Status Log is viewable in a pop-up window by clicking on the "status log…" link:

The Status Log can also be made visible by creating a Dashboard Widget in the Results Viewer:



The information displayed in the Status Log is in-memory information that is maintained only while the Composition is loaded and/or active. Once the Composition is unloaded, the information in the Status Log is no longer available. However, detailed information about the play of the Composition is saved in the Result object, and can be viewed via the "Event Log" and "Result Details" widgets in a dashboard.

The Result for a Composition play is created at the time that the Composition actually starts playing. Thus the status information shown in the Status Log from the load process appears only in the Status Log and not in the Result.

For multi-server Compositions, it can be handy to have a Dashboard added to the Composition to make viewing of the Status Log easy during the load process as well as during play and stop.

Starting the load

A Composition is first "loaded", and then it is started playing. The two operations can be performed manually as separate operations by pressing the "Load" button in the Composition Editor followed by the "Play" button. Or, the Composition can be loaded and started all in one operation by simply pressing the "Play" button.

The "Load" button is a green up arrow, and the "Play" button is green triangle:

The load operation involves reading the Composition from the Repository, analyzing it, moving it to the appropriate server, deploying portions to other servers as appropriate if it is a multi-server Composition, and reading and distributing any Seed Data as appropriate.

Once the load operation is completed, the Composition is ready to start playing. When play starts, the Result is created, Monitors are started as appropriate, and then the execution of the Tracks and Clips in the Composition begins.

The load operation can often be the point where environmental problems are first discovered, such as configuration errors, insufficient number of servers, or problem or crashed servers. Therefore, for an important load test that must be scheduled to occur at a specific time, it is often wise to do the load portion separately ahead of time, to allow time to resolve any such problems discovered. No traffic to the target(s) being tested is generated while the Composition is loaded, the load generation does not start until the Composition is played.

Here is a typical Status Log from the load of a Composition:

Status Log

Time Offset	Entry Text
00:00:00.000	Load started on server "Server A" at location "Location 1".
00:00:00.241	Moving to server "Server A2" at location "Location 1".
	Server "Server A" at location "Location 1" is not suitable.
00:00:00.000	Load started on server "Server A2" at location "Location 1".
00:00:00.302	Queuing load and distributing seed data to 4 additional servers.
00:00:00.304	Waiting for additional servers to load.
00:00:02.315	Loaded
00:00:02.315	Load completed, 5 servers.
00:00:02.315	Estimated Maximum Potential Virtual Users: 50.
00:00:02.316	Seed Data distribution summary.
	100 total rows, starting with row 1, distributed across 5 servers (20 rows per server) from: /My Seed Data

Here is some commentary on the events in the Status Log:

The first entry indicates that the load was started on server "Server A", which in this case is the Main Instance.

This is followed by entries that indicate that the Composition was moved to server "Server A2". This occurred because the Composition has specified that all of the Tracks be played on a server of type "Load" and "Server A" (on the Main Instance) is a server of type "General" (as is typical for the Main Instance).

The next entries indicate that this is a multi-server Composition and additional servers are being told to start loading the Composition in parallel. At the same time, any Seed Data is being broken up and prepared for distribution to those servers.

Next, at :00.304, the system is waiting for all of the additional servers to finish loading. This can take a while, depending upon the number of servers, network speed, and whether any problems are encountered. This is a point where networking issues and problem servers often show up.

Next, an entry indicates that 5 servers have been loaded successfully.

The next entry gives an estimate of the maximum number of Virtual Users that the loaded Composition might have. This is only an estimate, since the value could be changed at runtime due to scripting and conditional logic inside of the Composition. A "Virtual User" is one repeat of a Track. In this example, there is a Track that repeats 10 times that is copied to 5 servers, so the estimate is 50. This allows you to check if the load you're about to generate is approximately what you expected, before you start the play.

The final entry gives some information about the seed data that was distributed, in this case from a Seed Data object named "My Seed Data".

Here is the same Composition loaded again, but this time before loading the "Play Options" dialog was used to specify that the Composition be loaded onto server "Server A3":

Status Log

	Time Offset	Entry Text
ℹ	00:00:00.000	Load started on server "Server A3" at location "Location 1".
ℹ	00:00:00.539	Queuing load and distributing seed data to 4 additional servers.
ℹ	00:00:00.542	Waiting for additional servers to load.
ℹ	00:00:02.577	Loaded
ℹ	00:00:02.577	Load completed, 5 servers.
ℹ	00:00:02.577	Estimated Maximum Potential Virtual Users: 50.
ℹ	00:00:02.577	Seed Data distribution summary.
		100 total rows, starting with row 1, distributed across 5 servers (20 rows per server) from: /My Seed Data

Many of the entries are the same as before, except that there is no initial move of the Composition load process to a different server. As Play Options was used to specify that the Composition be loaded onto server "Server A3", it was not loaded onto the Main

Instance, but was loaded directly onto server “Server A3”. Since server “Server A3” is a server of type “Load”, it is acceptable and no moving is required.

Unloading instead of playing

Once the Composition has been loaded, you can play it. Or, if you decide you don’t want to play it after all, you can simply unload it. After the load completes, the green up arrow “load” button changes to an orange down arrow “unload” button:

Partial loads

Sometimes the load of a Composition can end with a “partial load”, meaning not all of the needed servers according to the Composition’s definition could be found or loaded. This can happen because of some sort of error while loading certain servers, or just because there simply aren’t enough servers available. It can also happen if the load process is stopped prematurely, which will be covered in the next section.

If there is a partial load, the Composition can still be played. It will just play with fewer servers than required to implement the full Composition as defined, meaning that some Tracks and/or copies of Tracks will not be played. Events in the Status Log will indicate which Tracks could not be loaded.

Here is a Status Log from a partial load:

Status Log

Time Offset	Entry Text
00:00:00.000	Load started on server "Server A" at location "Location 1".
00:00:00.251	Moving to server "Server A2" at location "Location 1".
	Server "Server A" at location "Location 1" is not suitable.
00:00:00.000	Load started on server "Server A2" at location "Location 1".
00:00:00.292	Queuing load and distributing seed data to 4 additional servers.
00:00:00.293	Waiting for additional servers to load.
00:00:02.307	Loaded
00:00:02.307	Partial load completed, 3 of 5 servers.
	Tracks not loaded:
	Track 1 (copy 4 of 5)
	Track 1 (copy 5 of 5)
00:00:02.307	Estimated Maximum Potential Virtual Users: 30.
00:00:02.307	Seed Data distribution summary.
	60 total rows, starting with row 1, distributed across 3 servers (20 rows per server) from: /My Seed Data

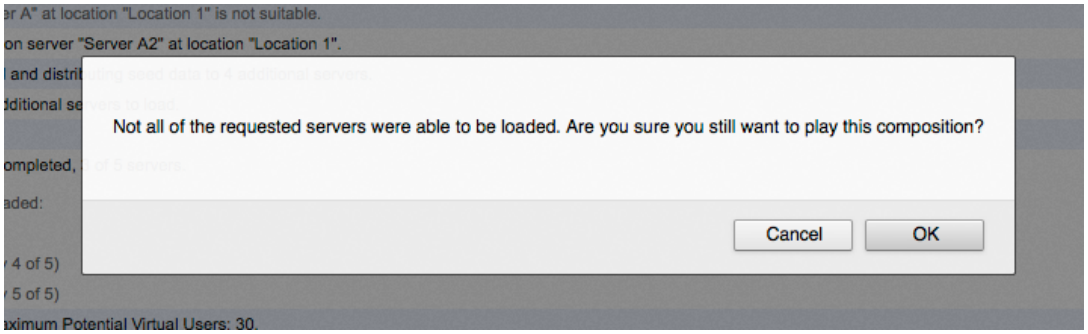
In the above example, there were no errors. There simply weren’t enough servers to satisfy the Composition’s definition. The Composition is defined to play on 5 servers, but only 3 are available.

The event at :02.307 indicates which Tracks could not be loaded. This particular Composition has a single Track named "Track 1" that is specified to be copied to 5 "Dedicated Load Servers". In that case, 5 copies of "Track 1" are made, one for each server. In this case, copies number 4 and 5 could not be placed anywhere and thus will not be played if the Composition is played.

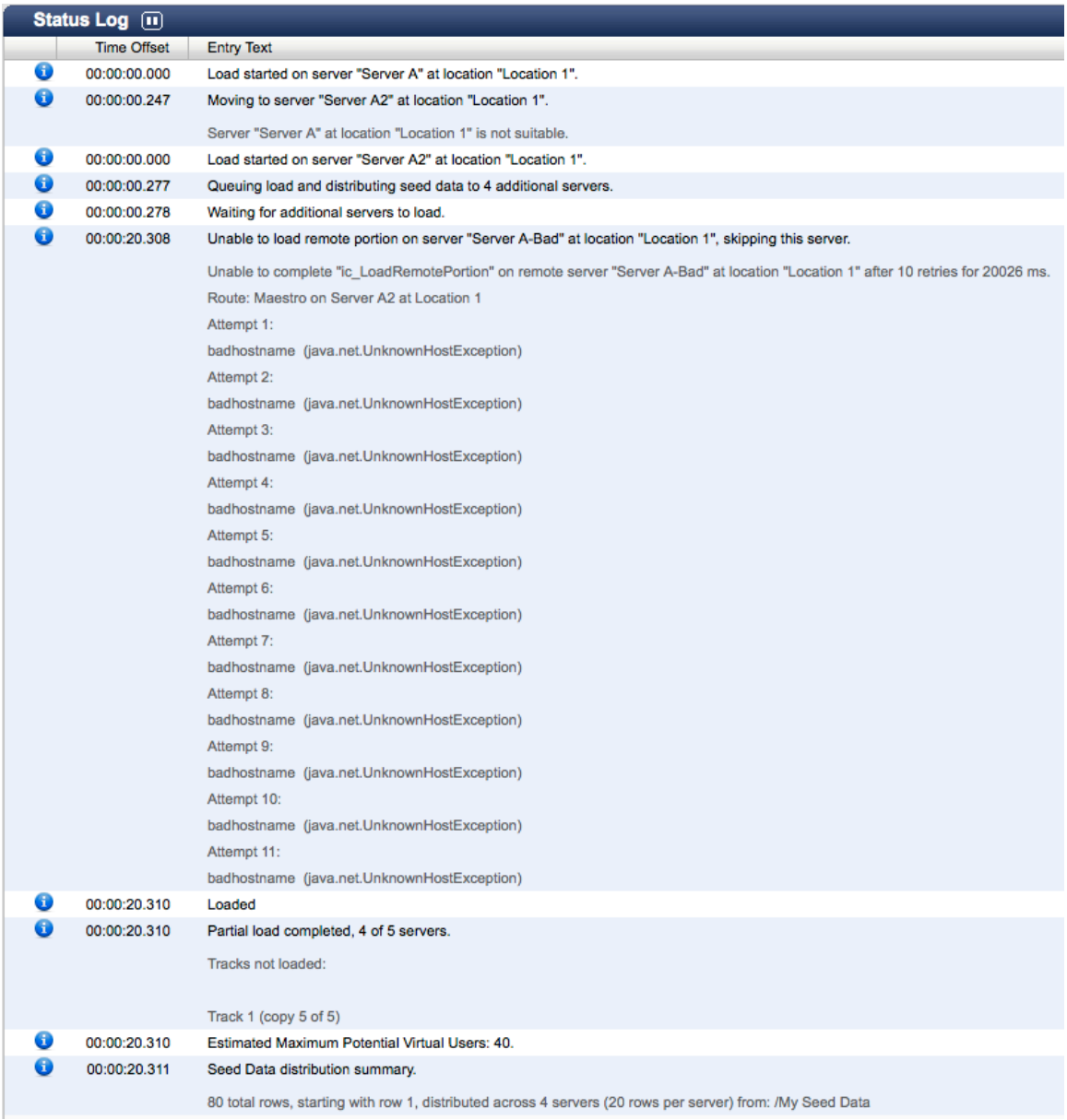
Notice that the "Estimated Maximum Potential Virtual Users" is reduced to reflect the reduced number of Tracks that will play, and the amount of Seed Data distributed is also less.

The partial load is noticeable on the status line as well:

At this point the decision can be made whether it is acceptable to play the Composition with the reduced number of servers, or whether the Composition should be unloaded, more servers added, and the load tried again. If you don't wish to play it, simply unload it. If you do wish to play it, press the "play" button. To be sure that you have noticed that a partial load occurred, a warning pop-up is displayed:



Here is an example in which there were errors loading a server:



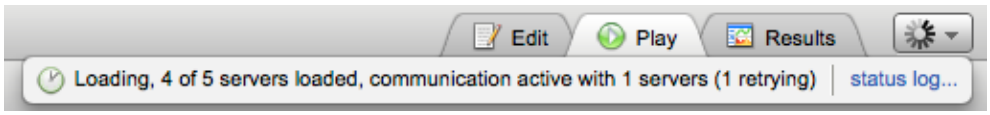
Status Log

	Time Offset	Entry Text
ℹ	00:00:00.000	Load started on server "Server A" at location "Location 1".
ℹ	00:00:00.247	Moving to server "Server A2" at location "Location 1". Server "Server A" at location "Location 1" is not suitable.
ℹ	00:00:00.000	Load started on server "Server A2" at location "Location 1".
ℹ	00:00:00.277	Queuing load and distributing seed data to 4 additional servers.
ℹ	00:00:00.278	Waiting for additional servers to load.
ℹ	00:00:20.308	Unable to load remote portion on server "Server A-Bad" at location "Location 1", skipping this server. Unable to complete "ic_LoadRemotePortion" on remote server "Server A-Bad" at location "Location 1" after 10 retries for 20026 ms. Route: Maestro on Server A2 at Location 1 Attempt 1: badhostname (java.net.UnknownHostException) Attempt 2: badhostname (java.net.UnknownHostException) Attempt 3: badhostname (java.net.UnknownHostException) Attempt 4: badhostname (java.net.UnknownHostException) Attempt 5: badhostname (java.net.UnknownHostException) Attempt 6: badhostname (java.net.UnknownHostException) Attempt 7: badhostname (java.net.UnknownHostException) Attempt 8: badhostname (java.net.UnknownHostException) Attempt 9: badhostname (java.net.UnknownHostException) Attempt 10: badhostname (java.net.UnknownHostException) Attempt 11: badhostname (java.net.UnknownHostException)
ℹ	00:00:20.310	Loaded
ℹ	00:00:20.310	Partial load completed, 4 of 5 servers. Tracks not loaded: Track 1 (copy 5 of 5)
ℹ	00:00:20.310	Estimated Maximum Potential Virtual Users: 40.
ℹ	00:00:20.311	Seed Data distribution summary. 80 total rows, starting with row 1, distributed across 4 servers (20 rows per server) from: /My Seed Data

In this case, the system was unable to contact the server named, appropriately, “Server A-Bad”. The event at :20.308 indicates that it tried 11 times to contact the server before giving up. In this specific case, the problem was a configuration error in which the URL that the system has for that server is incorrect (containing the invalid host name “badhostname”).

If the system had been able to find another suitable server to replace “Server A-Bad”, then the load would have completed successfully. In this case, since there were no other servers available, it resulted in a partial load.

While the load was occurring, another hint that maybe there was a problem was visible on the status line:



Retrying of communication with a server is not always an indication of an unrecoverable error, sometimes retries do, in fact, work. It is not unusual to see a retry every now and then, especially when using a large number of servers, and usually the system is able to recover.

Stopping the load

Sometimes loading a Composition can take a significant amount of time, for example if there are a large number of servers involved, or there are communication errors or problem servers.

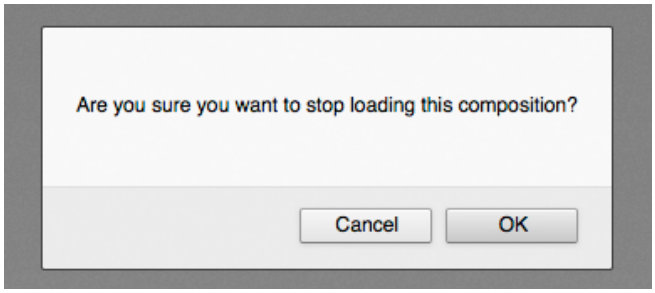
While the Composition is loading, the green up arrow "Load" button changes to have a red rectangle on it, indicating that pressing the button again will cause the load to be stopped:

You can press the "stop load" button at any time to stop the load. Depending upon how far along the load is at the time the button is pressed, the load will either be aborted entirely or stopped with a partial load.

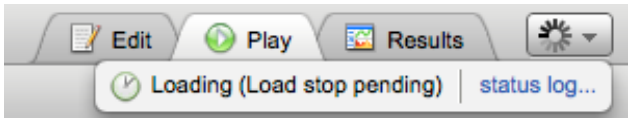
Typical reasons for stopping the load include:

1. You changed your mind about playing the Composition, or
2. There are server or network problems discovered during the load and you want to fix them and try again, or
3. The load is almost complete but there are one, two, or a handful of servers that are having problems, and you don't want to wait for those issues to be resolved because you're willing to go ahead with the number of servers that have been successfully loaded so far.

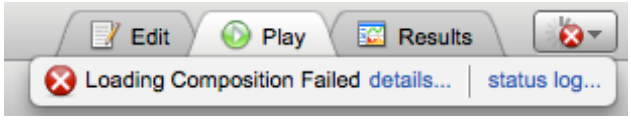
When the “load stop” button is pressed, a confirmation pop-up appears:

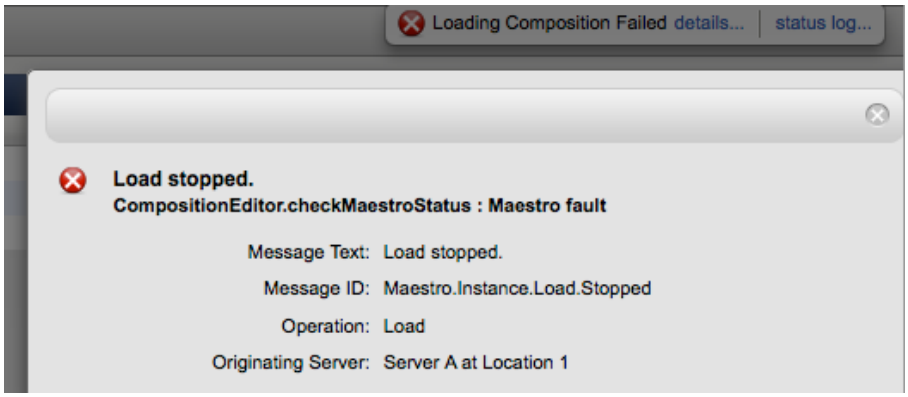


If you choose to go ahead and stop the load, the status line will change to indicate that the load is being stopped:

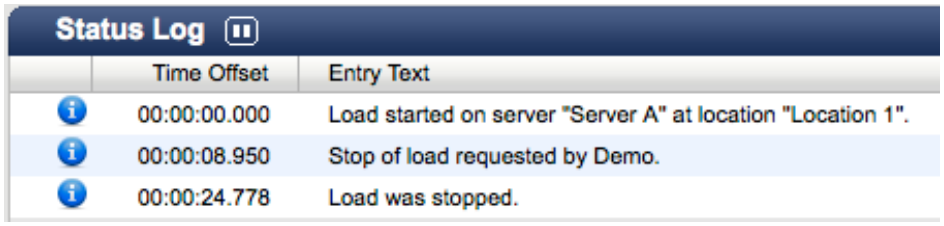


If the load is stopped so early in the load process that nothing was loaded successfully, then the load will end in an error:



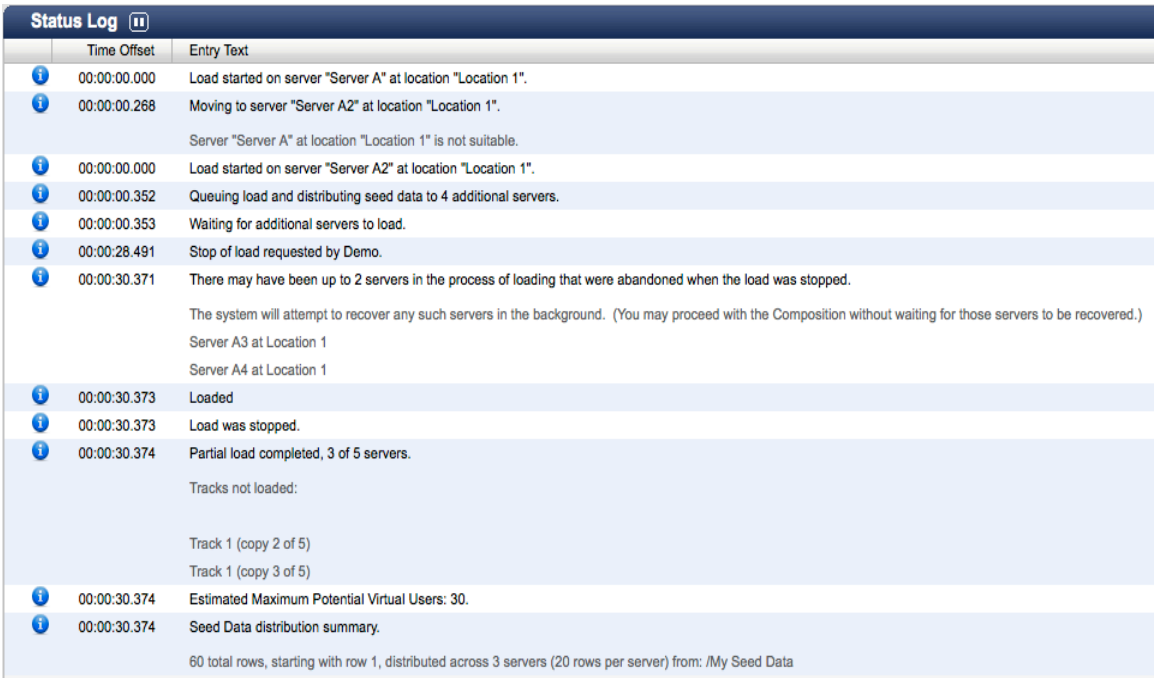


The fact that the load was stopped will also be reflected in the Status Log, if it was visible at the time of the stop:



In the above example, the event indicates that the load was stopped by the user named “Demo”.

If the load progressed far enough before the stop that at least some servers were successfully loaded, a partial load will result. See the prior discussions about partial loads. Here is the Status Log from such a partial load:



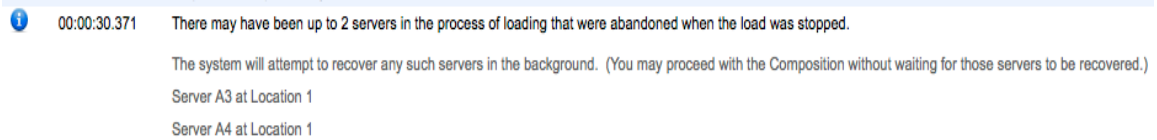
Note that this Status Log is very similar to the partial load examples from earlier. The one difference is the event at :30:371 regarding "abandoned servers", which will be discussed in the next section.

"Abandoned" servers

When a load is stopped by pressing the "load stop" button, often there may be a number of servers in the process of being loaded at that time.

Many times the load is being stopped because there are network problems and/or problem servers, and these problems can often result in long delays while communication is timing out and/or being retried. Since the load is often being stopped precisely because you don't wish to wait for further attempts at resolving those problems, the load stopping processing does not wait for those issues to be resolved.

When the load is stopped, the Composition load is considered to be complete, and any servers that were still in the middle of being loaded are excluded from the Composition's partial load, so that the stop can occur quickly. Any such servers are noted as "abandoned" servers in the Status Log, in entries such as this:

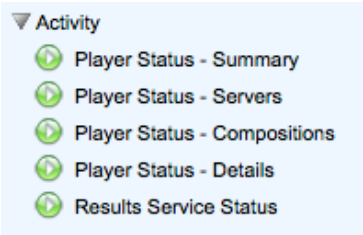


As the text of the event indicates, the system doesn't totally abandon those servers. They are "abandoned" from the point of view of the loaded Composition, which will no longer

consider them to be a part of the Composition. However, in the background the system will continue to attempt to resolve whatever was happening with those servers, in order to get them back into an idle state and ready for use again in the future.

For the purposes of playing the partially loaded Composition, you can continue without concern about the "abandoned" servers.

The background system processes that are cleaning up any such servers can be seen in the various Player Status displays. The Player Status displays are accessed from the left pane of the "Central" page, under the "Activity" header:



Here is an example from "Player Status – Servers":

Location	Name	Type / Status	Items
Location 1	Server A	General / Idle	
Location 1	Server A2	General	Composition-Main: Load completed, 3 of 5 servers: Draft of Five-server Composition cr
			StoppedLoadCleanup: Cleaning up 2 servers after stopped load.:
Location 1	Server A3	Load	Composition-Remote: Loaded: Draft of Five-server Composition created on September
Location 1	Server A4	Load	Composition-Remote: Loaded: Draft of Five-server Composition created on September
Location 1	Server A5	Load / Idle	
Location 2	Server B	Load / Idle	
Information available from 6 of 6 servers.			

You can see the partially loaded Composition, but also the "StoppedLoadCleanup" process shown in the status display. This process is cleaning up after the "abandoned" servers in the background.

Here is the same view from "Player Status – Compositions":

Name Playlist name, if any Status message	User name Start time Elapsed time	Result name Status Summary text	Primary server Server count
Cleaning up 2 servers after stopped load.	Demo – –	CleanupWait	Server A2 at Location 1 1
Draft of Five-server Composition created on September 28, 2011 4:0… Load completed, 3 of 5 servers	Demo – –	Loaded	Server A2 at Location 1 3 (partial load)
Information available from 6 of 6 servers.			

And from “Player Status – Summary”:

Player Status Summary	
Total number of servers	6
Servers registered	6
Servers reporting status	6
Servers with registration pending	0
Status communication failures	0
Idle General servers	1
Idle Load servers	2
Loaded items	4
Items of type "Composition-Main"	1
Items of type "StoppedLoadCleanup"	1
Items of type "Composition-Remote"	2
Items in state "Loaded"	3
Items in state "CleanupWait"	1
Most recent status received	1.94 seconds (Server A at Location 1)
Oldest status	27.55 seconds (Server A5 at Location 1)
Average status age	10.43 seconds

The background “clean up” processes will go away on their own when they have finished. Depending upon the states of the various servers at the time that the load was stopped, some of the “abandoned” servers could be considered as “busy” and unavailable for use as Load Servers until the cleanup process has completed.