Performance Management Advisors Hardware Sizing Guide

FA Deployment Architecture and Recommendations for Optimal Performance
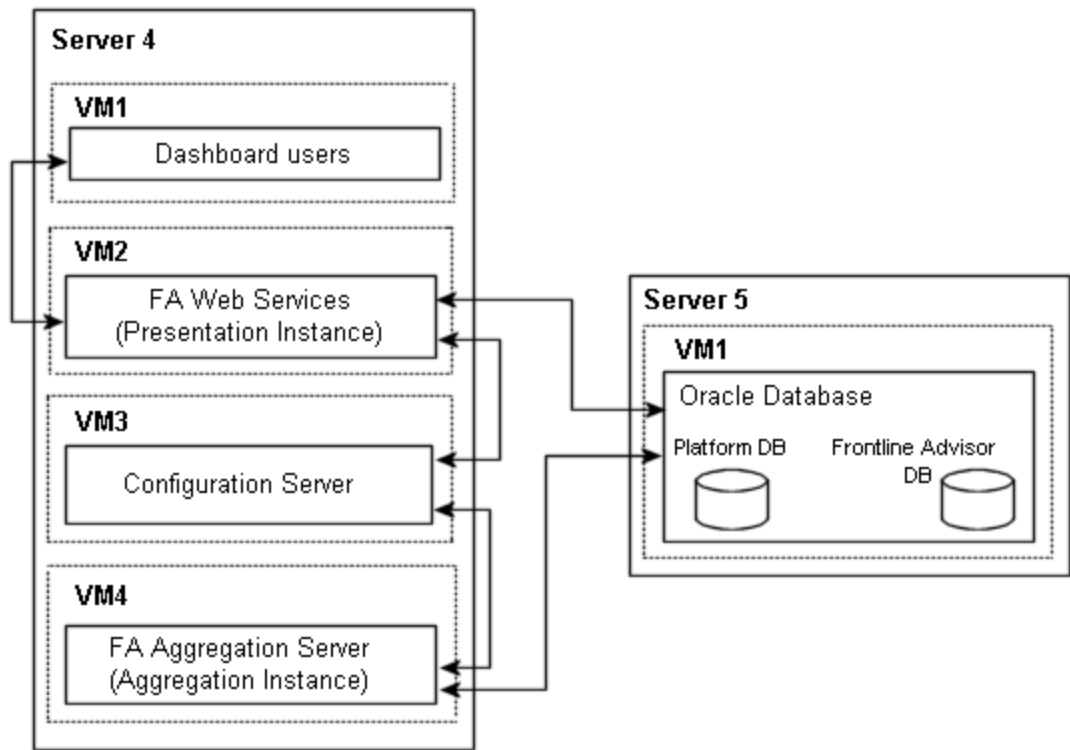
4/16/2025

Contents

FA Deployment Architecture and Recommendations for Optimal Performance

The following table describes the high-level dimensions controlling the environment used to achieve the results described in Load-Carrying Capacity.

Frontline Advisor Presentation Object Configuration

Object	Count
Agents	30 000
Depth (levels)	6
Multiplicity (refers to the average number of teams to which an agent belongs)	1
Agent Groups	5000 (with agents)
Time Profiles	3

The following diagram shows the environment topology used to successfully achieve 1500 concurrent users of Frontline Advisor (FA). "VM" is a virtual machine. The Aggregation Instance performs data aggregation; the Presentation Instance services the dashboard.

Frontline Advisor Environment Topology

Frontline Advisor Presentation Machine and VM Information

The following table describes the characteristics of the hardware and virtualization environment shown in the figure, Frontline Advisor Environment Topology, which supports 1500 concurrent users of FA.

Server	# of Processors	Processor Type	# of Cores	Total Cores	RAM	OS	Application
Server 4: VM-Host	2	Intel Xeon X5675 @ 3.07GHz	6	12	32 GB	VMWare ESXi 5.0 Standard 64bit	
Server 4: VM-1	2	Same as host	4	8	8 GB	Windows Server 2008 Standard SP1 64bit	Dashboard simulator
Server 4: VM-2	2	Same as host	4	8	8 GB	Windows Server 2003 Standard	Geronimo (FA Presentation instance

Server	# of Processors	Processor Type	# of Cores	Total Cores	RAM	OS	Application
						SP1 64bit **NOTE:** At the time of testing, Genesys supported Windows Server 2003. In your enterprise, be sure to use a Genesys-supported operating system.	on Advisors Platform)
Server 4: VM-3	2	Same as host	1	2	4 GB	Windows Server 2003 32bit **NOTE:** At the time of testing, Genesys supported Windows Server 2003. In your enterprise, be sure to use a Genesys-supported operating system.	Configuration Server
Server 4: VM-4	2	Same as host	4	8	8 GB	Windows Server 2008 R2 Standard SP1 64bit	FA Aggregation instance on Advisors Platform
Server 5: VM-Host	2	AMD Opteron 2439SE @ 2.8GHz	6	12	32 GB	VMWare ESXi 5.0 Standard 64bit	
Server-5: VM-1	2	Same as host	1	2	8 GB	RHEL Server 5.7 64bit	Oracle database

Recommendations for Frontline Advisor Performance Improvement

The following table describes settings you can change to improve the performance of Frontline

Advisor.

Location	Sub-directory or File, where applicable	Settings

On each FA node	`<FA home>\geronimo-tomcat6-minimal-2.2.1\bin\setenv.bat`
	
	Tip
	Starting with release 8.5.2, the path to the `setenv.bat` file changes to `<FA home>\apache-tomcat-<version>\bin\setenv.bat`.
		Change
		
		`GERONIMO_OPTS=-ms128m -mx1024m`
		`  -XX:MaxPermSize=128m`
		
		To
		
		`GERONIMO_OPTS=-Xms4g -Xmx8g`
		`  -XX:MaxPermSize=512m`
		
		Starting with release 8.5.2, `GERONIMO_OPTS` becomes `CATALINA_OPTS`. In addition, the `XX:MaxPermSize` setting no longer exists. So, starting with release 8.5.2, you make the following change in the `setenv.bat` file:
		
		Change
		
		`CATALINA_OPTS=-ms1024m -mx2048m`
		
		To
		
		`CATALINA_OPTS=-Xms6g  -Xmx6g`

	`<FA home>\geronimo\var\catalina\server.xml`	Under the `<Connectorname="TomcatAJPConnector">` section, add `maxThreads="2000"` **Tip** Starting with release 8.5.2, you do not change the `TomcatAJPConnector` setting because it does not exist.

On AGA	**Releases 8.1.5 and 8.5.0:**
	`<AGA home>\conf\wrapper.conf`
	
	Release 8.5.1 and later:
	conf\run.bat (Windows)
	bin\setenv.sh (Linux)
		Releases 8.1.5 and 8.5.0:
		
		Change
		
		`Wrapper.java.initmemory=128`
		`Wrapper.java.maxmemory=1024`
		
		To
		
		`Wrapper.java.initmemory=4096`
		
		`Wrapper.java.maxmemory=14336`
		
		Release 8.5.1 and later:
		In the JAVA_OPTS parameter, change
		
		`-ms128m`
		`-mx1024m`

Location	Sub-directory or File, where applicable	Settings

		To
		
		```-ms4096m```
		```-mx14336m```

	<AGA home>\conf\ ```inf_genesys_adapter.properties```	Change
		
		```informiam.genesys_connector.stat```
		``` Server.addp.clienttimeout```
		```= 120```
		
		```informiam.genesys_connector.timing.```
		
		```messagerate.numberofmessages.batch```
		```= 500```
		
		To
		
		```informiam.genesys_connector.stat```
		``` Server.addp.clienttimeout```
		```= 360```
		
		```informiam.genesys_connector.timing.```
		
		```messagerate.numberofmessages.batch```
		```= 100000```

On each Apache HTTP proxy	```httpd.conf```	• Uncomment or add the following modules:
		    ```LoadModule```
		    ```deflate_module```
		    ```modules/```
		    ```mod_deflate.so```
		
		    ```LoadModule```
		    ```headers_module```
		    ```modules/```
		    ```mod_headers.so```
		
		    ```LoadModule```
		    ```proxy_module modules/```
		    ```mod_proxy.so```
		
		    ```LoadModule```
		    ```proxy_ajp_module```
		    ```modules/```
		    ```mod_proxy_ajp.so```
		
		    ```LoadModule```
		    ```proxy_balancer_module```
		    ```modules/```
		    ```mod_proxy_balancer.so```

Location	Sub-directory or File, where applicable	Settings

		LoadModule proxy_http_module modules/ mod_proxy_http.so
		
		• Add the following block to increase the number of Apache worker threads (note that this is for a Windows-based server):
		
		`<IfModule mpm_winnt_module>`
		`ThreadsPerChild 512`
		`MaxConnectionsPerChild 0`
		`</IfModule>`
		
		    • If you use a Linux server, add the following block:
		
		`<IfModule mpm_event_module>`
		`StartServer 6`
		`ServerLimit 32`
		`MinSpareThreads 150`
		`MaxSpareThreads 250`
		`ThreadsPerChild 25`
		`MaxRequestWorkers 800`
		`MaxConnectionsPerChild 0`
		`</IfModule>`
		
		• Add the following to enable a request response proxy:
		    `ProxyPass /fa/ ajp://localhost:8009/ fa/`
		
		    `ProxyPass /am/ ajp://localhost:8009/ am/`
		
		    `ProxyPass /admin/ ajp://localhost:8009/ admin/`
		
		    `ProxyPass /am-admin/ ajp://localhost:8009/ am-admin/`
		
		    `ProxyPass /ca/ ajp://localhost:8009/ ca/`
		
		    `ProxyPass /ca-ws/ ajp://localhost:8009/ ca-ws/`

Location	Sub-directory or File, where applicable	Settings
| | | ```
ProxyPass /ea-ws/
ajp://localhost:8009/
ea-ws/

ProxyPass /base-ws/
ajp://localhost:8009/
base-ws/

ProxyPass /dashboard/
ajp://localhost:8009/
dashboard/

ProxyPass /nav-
service/
ajp://localhost:8009/
nav-service/

ProxyPass /prefs-
service/
ajp://localhost:8009/
prefs-service/

ProxyPass /wu/
ajp://localhost:8009/
wu/

ProxyPass /rmc/
ajp://localhost:8009/
rmc/

ProxyPass /gc-admin/
ajp://localhost:8009/
gc-admin/

ProxyPass /ca-xml/
ajp://localhost:8009/
ca-xml/
``` |

Best Practices for Frontline Advisor Sizing

Use the following notes and best practices for optimizing FA performance:

- Use Gigabit connectivity between the FA aggregation node and FA presentation node(s).
- Enable an AJP connection between the Apache HTTP proxy and FA presentation node(s).
- Allocate as much CPU resource to FA as possible; FA performance is improved if you provide multiple CPU cores and faster clock speeds.
- Allocate sufficient memory for FA components (Genesys recommends 6GB).
- You may require multiple AGAs (up to two).

- Regarding Stat Server performance:

  - Stat Server is a single threaded process. Carefully monitor the CPU usage of your Stat Server(s).

  - Consider adding more Stat Server pairs if a Stat Server is saturating a CPU. You may require up to six pairs of Stat Servers for best performance.

Frontline Advisor Dashboard Age

Dashboard age is a performance measure used in assessing the state and performance/rule processing cycles. It represents the age of the statistics on the dashboard, and includes the following:

1. Pre-Rollup Delay: The time from the end of the last rollup until the scheduled start of the next rollup.

2. Rollup Duration: The duration of the rollup + the time to publish to the distributed cache.

3. Request Response Time (RRT): The 95th percentile of response time for a dashboard request. That is, 95% of the time, a response is returned to a dashboard request after X number of seconds, where X is a constant.

The following table shows results from Genesys' performance testing. *State processing* refers to the state metric rollup cycle and *performance/rule processing* is the performance metric rollup cycle (state and performance metric cycles run independently).

Dashboard Age Results from FA Performance Testing

Measure (seconds)	State Processing			Performance/Rule Processing		
95th Percentile	Median	Average	95th Percentile	Median	Average	
Dashboard age	18 seconds	12 seconds	12 seconds	70 seconds	62 seconds	62 seconds