
Liferay Portal EE Reference Architecture & Hardware Requirements

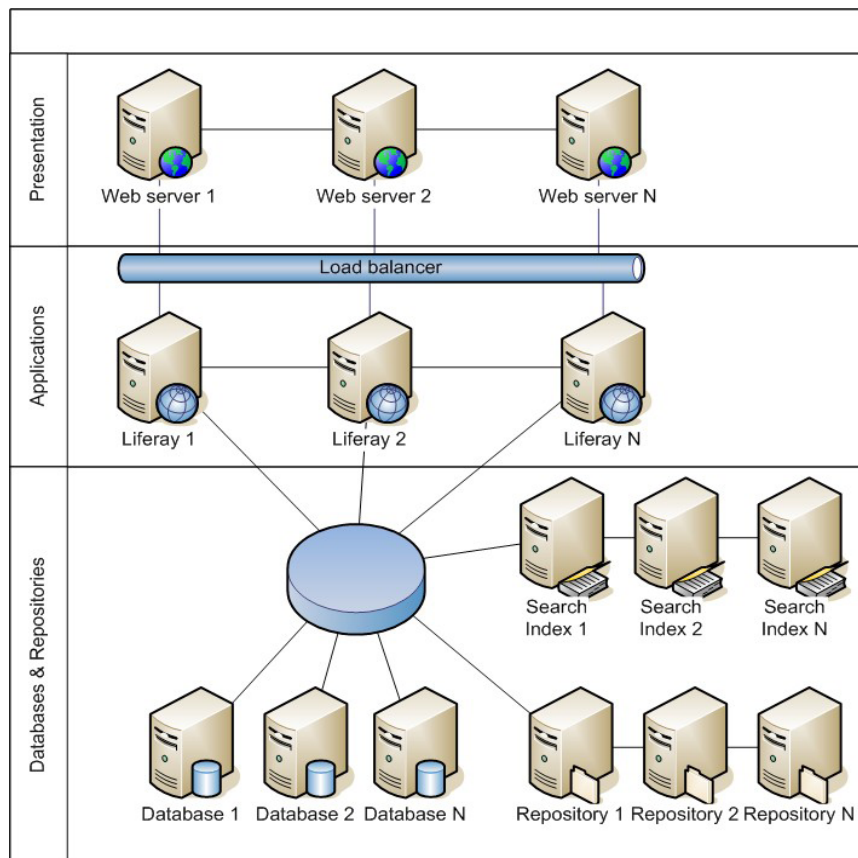
Table of Contents

Reference Architecture	2
Overview	2
Reference Hardware	3
Example Software/Hardware Configurations	4
Apache + mod_jk + Tomcat + Liferay:.....	4
Apache + mod_proxy + Tomcat + Liferay:	5
Apache + mod_proxy_balancer + Tomcat + Liferay:.....	6
Liferay Portal Performance Metrics.....	7
Sizing Approach	7
Scenarios	7
Anonymous Content	7
Isolated Login	8
Authenticated Content	10
Other Scenarios	11
Sizing Your System.....	12
Anonymous Content Browsing Hardware Requirements	12
Isolated Login Hardware Requirements	13
Authenticated Content Browsing Hardware Requirements	13
Realistic Scenarios	13
In Conclusion.....	14
Moving Forward	16
Liferay Enterprise Edition Support.....	16
Liferay Professional Services	16

Reference Architecture

OVERVIEW

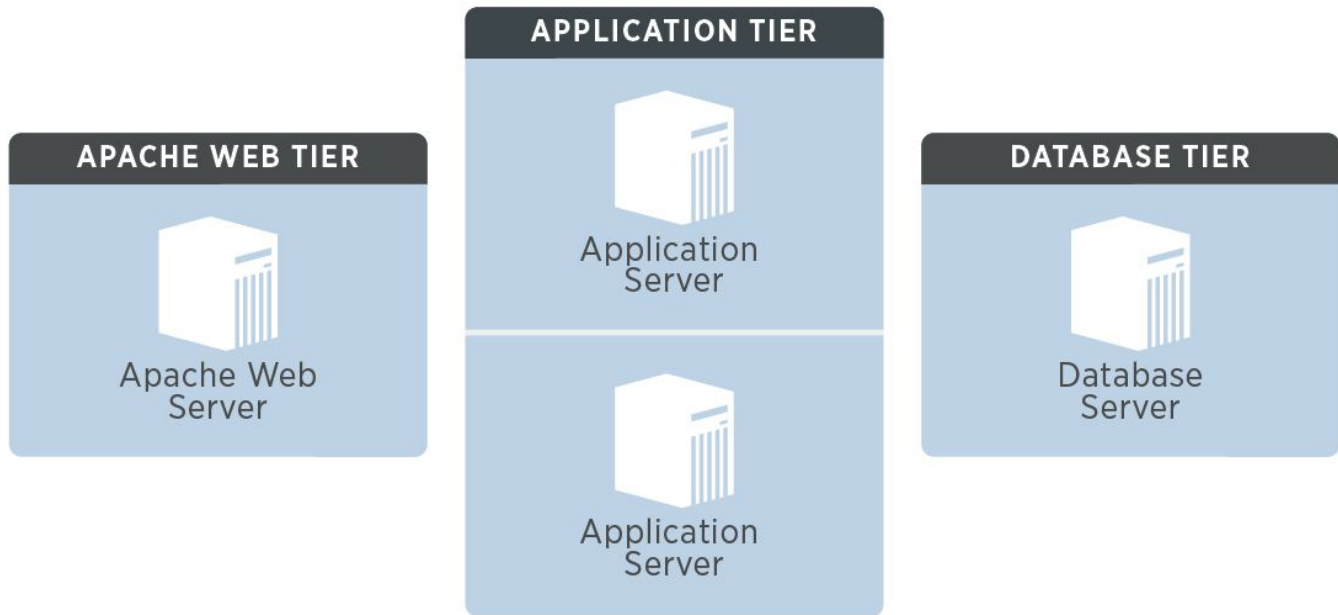
The production environment for a typical Liferay deployment consists of a number of HTTP/web servers, application servers or servlet containers, and database servers playing specific roles. For example, the HTTP server's main role is to load balance the entire system. It also serves as a proxy and a single entry point to the system shielding it from the Internet. The web server is the front-end of the system and its main role is to fulfill web requests. All the business components will reside in application servers. The database server is responsible for storing all data. An example configuration of these servers is captured in the diagram below.



There are a number of other non-production servers to consider. The QA/Staging server acts as a testing environment to assure that newly developed features are working as designed. This server can also serve as the code repository for storing and versioning custom enhancements. The data backup server will backup all essential data on a nightly basis. This includes the database, code repository, and all other repositories.

REFERENCE HARDWARE

For the purposes of this document we will use a very simple reference hardware configuration illustrated here:



Hardware platforms:

1. Web Server
 - 1 x Intel Core 2 Duo E6405 2.13GHz CPU, 2MB L2 cache (2 cores total)
 - 4GB memory
 - 1 x 146GB 7.2k RPM IDE
2. Application Server
 - 2 x Intel Core 2 Quad E5430 2.66GHz CPU, 12MB L2 cache (8 cores total)
 - 8GB memory
 - 2 x 146GB 10k RPM SCSI
3. Database Tier
 - 1 x Intel Core 2 Quad E5430 2.66GHz CPU, 12MB L2 cache (8 cores total)
 - 16GB memory
 - 4 x 146GB 15k RPM SCSI

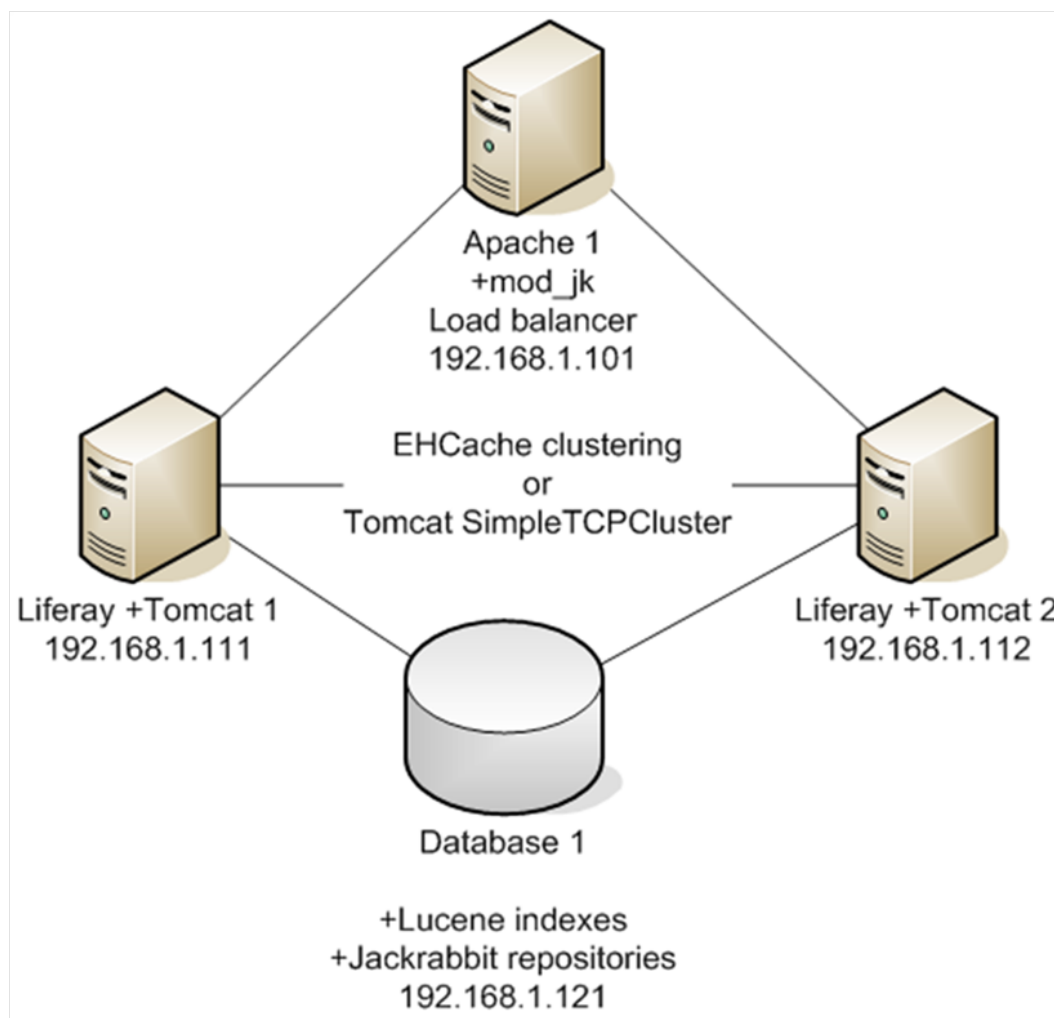
Network:

- Gigabit network between all servers and test clients

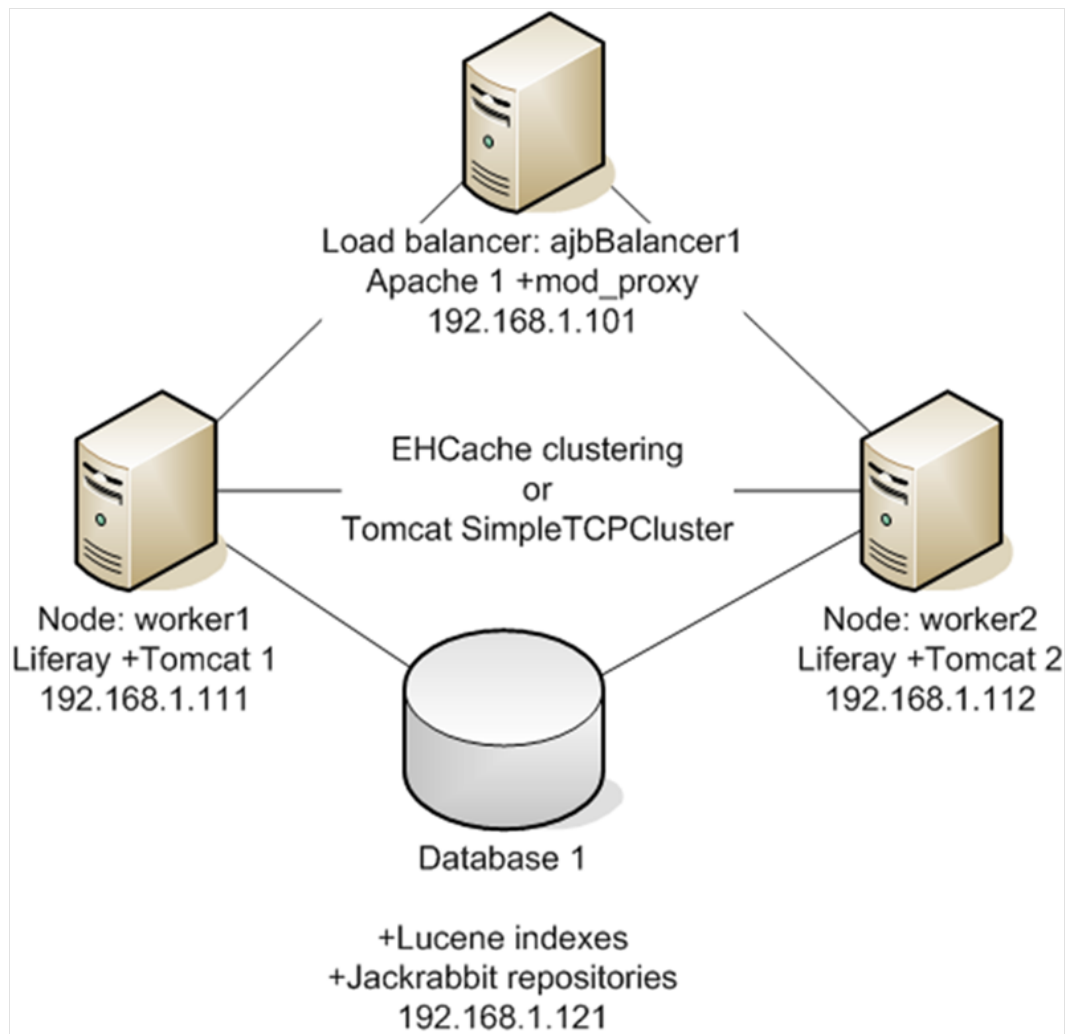
Example Software/Hardware Configurations

The following are some sample software/hardware configurations of the reference hardware that you may want to consider. You may or may not have these exact configurations, or a derivative or permutation of one of them

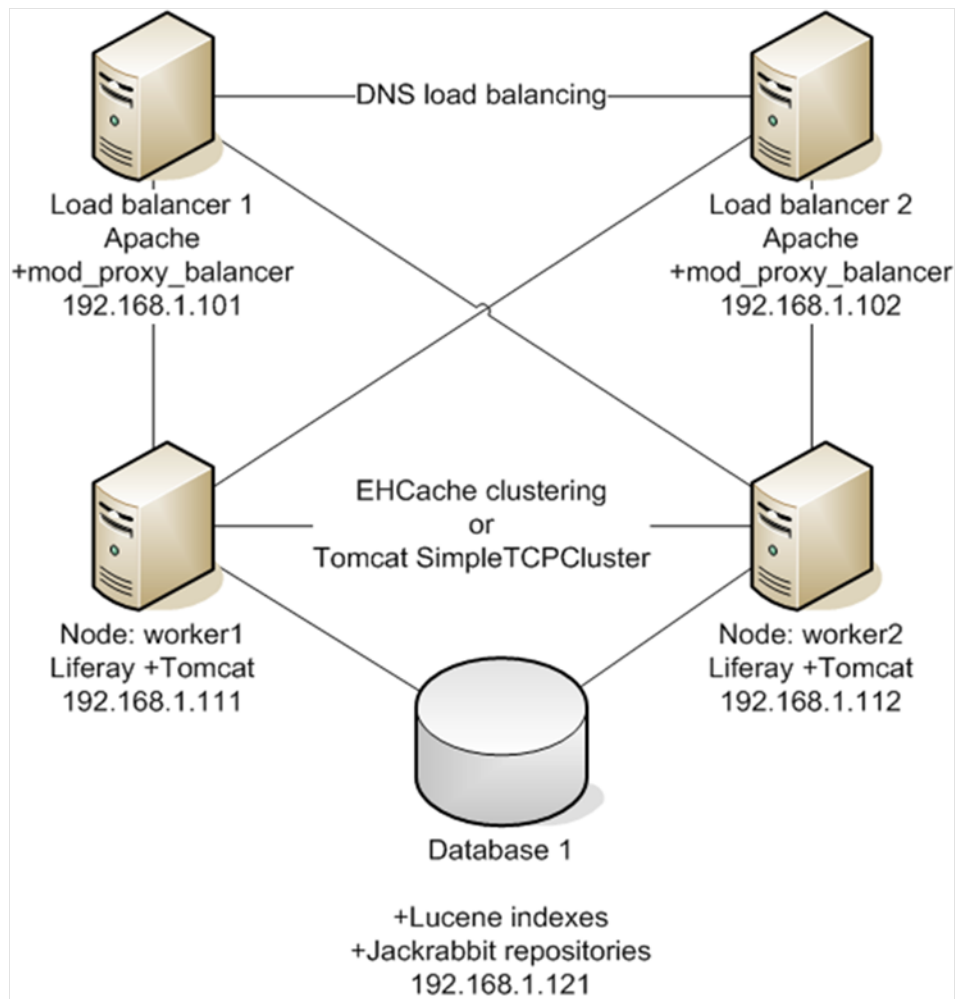
APACHE + MOD_JK + TOMCAT + LIFERAY:



APACHE + MOD_PROXY + TOMCAT + LIFERAY:



APACHE + MOD_PROXY_BALANCER + TOMCAT + LIFERAY:



Liferay Portal Performance Metrics

SIZING APPROACH

To properly size your eventual deployment of Liferay, you need to understand the number of concurrent users your system needs to service, as well as the utilization of your users. **In other words, what is the max number of users you will have and what will they be doing exactly?** There is no way to predict with 100% accuracy, what your users will be doing once you deploy to production. The only thing you can do is to prepare as best as you possibly can for certain situations and scenarios. This is where you use this document in conjunction with the whitepaper, “Liferay Portal Performance: Benchmark Study of Liferay Enterprise Edition” available here:

<http://www.liferay.com/documentation/additional-resources/whitepapers>

SCENARIOS

Let us examine a few scenarios from the benchmark whitepaper. You can see the rest of the scenarios in more detail in the whitepaper above, but for the purposes of this sizing discussion, we will use three scenarios to demonstrate how to go about sizing your system. We will look at anonymous content browsing, isolated login, and authenticated content browsing.

Anonymous Content

During tests, Liferay Portal was able to service over 150,000 users browsing for content anonymously. The tests show the WCM handling the load with ease, using only 35% of a single server's resources. At 150,000 users, the 95% of the transactions (2σ) remains comfortably below 50ms while the system only consumes 34% of total processing power. At the 150,000 user mark, the benchmark environment's network becomes a bottleneck as the gigabit network strains to keep up with the volume of requests. Total throughput appears to remain constant at roughly 3,900 transactions per second. Based on this statistic, it appears we have reached the maximum throughput of a single JVM hosting Liferay Portal. Since we have sufficient CPU resources, to achieve higher throughput, we should instantiate a second JVM on the same physical server.

CONCURRENT USERS	DURATION (MIN)	μ (MS)	σ (MS)	2σ (MS)	THROUGHPUT	CPU (%)
25,000	30	.82	2.97	6.76	3,960	22
50,000	30	1.95	8.96	19.87	3,880	25
100,000	30	2.47	12.4	27.27	3,930	28
150,000	30	3.98	19.2	42.38	3,970	32

Table 3 – Browsing for Content Anonymously

Due to the test reaching maximum network bandwidth, Liferay deemed it unnecessary to proceed with allocating a second physical server to test linear scalability.

Isolated Login

The first of two transaction centric scenarios focuses on the login process of Liferay Portal. The login and permission retrieval process is one of the most resource intensive processes within the portal. At login, the portal must retrieve user and security information from the database and calculate authorizations.

We first examine Liferay's performance with simple content portlets on the page. These portlets are extremely fast, lending average rendering times of less than 10ms.

Table 1 illustrates the performance observed during this test. The mean time for login remains less than 200ms as we approach the performance inflection point. At 3,000 concurrent users, we have a mean time (μ) of 162ms and 95% of the logins (2σ) below 600ms. The optimal performance point for this test scenario occurs somewhere between 3,000 to 3,200 concurrent users. At 3,200 users, we see 2σ increasing to over 1s, moving above the acceptable threshold. At this inflection point, we see CPU utilization at roughly 71% on the application server.

CONCURRENT USERS	DURATION (MIN)	μ (MS)	σ (MS)	2σ (MS)	THROUGHPUT (TPS)	CPU UTILIZATION (%)
1,800	30	97.9	44.2	186.3	51	36
2,100	30	119	199	517	52.6	44
2,400	30	121	197	515	59.6	50
2,700	30	139	199	537	67.4	58
2,800	30	141	177	495	70	63
3,000	30	162	206	574	74.2	71
3,200	30	310	367	1,044	76	78
3,300	30	430	452	1,334	79.1	81

Table 1 – Isolated Login

Figure 2 graphically illustrates the test scenario as we approach the optimum performance point. Although the portal itself is at peak performance, overall CPU utilization remains at 78%. This indicates we have excess capacity on the CPU. Thus, we could allocate a second JVM instance and delay the inflection point to roughly 3,500 users.

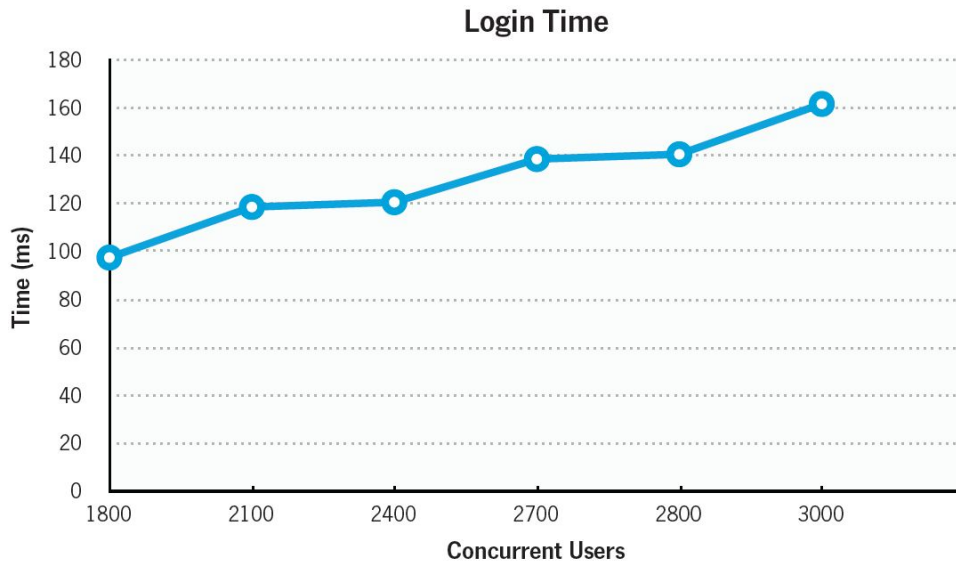


Figure 2 - Mean Login Time

In terms of throughput, the portal appears to have a maximum throughput of roughly 80 transactions per second; the optimal throughput (at 3,000 concurrent users) is roughly 74 logins per second.

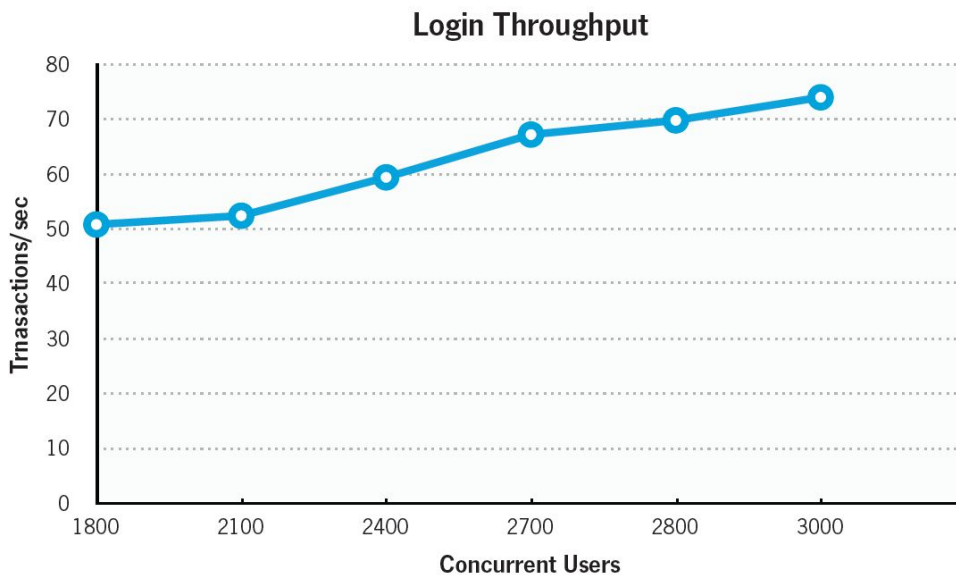


Figure 3 - Isolated Login Throughput

Upon determining the maximum throughput, a second portal application server was deployed. The benchmark results showed that Liferay Portal was able to breach 6,100 concurrent users using two application servers. At 6,100 users, the transaction times remained similar to the times gathered at 3,000 users on a single application server.

Authenticated Content

The authenticated content scenario shows different performance characteristics than the anonymous content scenario. In this scenario, the WCM will leverage the Portal's authentication and authorization system to ensure users have the appropriate permissions prior to viewing content. Thus, the performance characteristics of this scenario are very similar to that of the login test scenarios.

In Table 4, we find the statistics from this test. As we can see from the data, the performance inflection point for this scenario appears to be around 2,800 concurrent users. Although the Portal does have resources to push to 3,300 concurrent resources, the performance characteristics become unstable, driving the 95% transaction time to over 2.8s at 3,300 users. At 2,800 users, 95% of the transaction times are under 850ms.

CONCURRENT USERS	DURATION (MIN)	LOGIN TIME μ (MS)	LOGIN TIME σ (MS)	CONTENT PAGE μ (MS)	CONTENT PAGE σ (MS)	TOTAL μ (MS)	TOTAL σ (MS)	TOTAL 2 σ (MS)	THROUGHPUT (TPS)	CPU (%)
1,800	30	97.9	44.2	22.3	39	120.2	83.2	286.6	76.5	31
2,200	30	119	199	25.4	47	144.4	246	636.4	88.2	35
2,400	30	121	197	33.6	63.4	154.6	260.4	675.4	95.4	43
2,600	30	139	199	55.3	96.6	194.3	295.6	785.5	107.96	55
2,800	30	141	177	79.5	131	220.5	308	836.5	122.42	73
3,000	30	389	391	123	194	512	585	1,682	129.65	81
3,200	30	450	433	154	233	604	666	1,936	131.2	86
3,300	30	655	734	179	288	834	1,022	2,878	133.4	91

Table 4 – Known User Browsing for Content

Figure 6 illustrates that the best performance point for the Portal in this scenario is at 2,800 concurrent users.

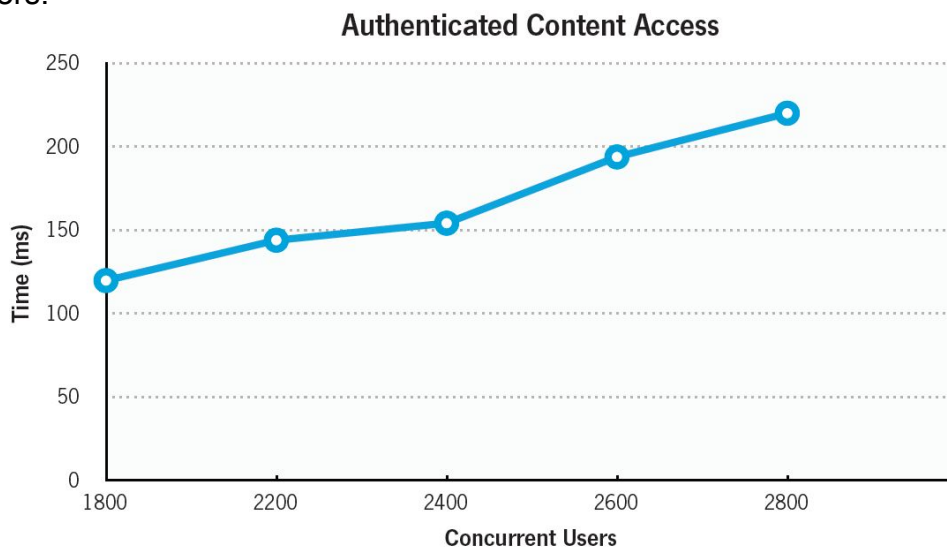


Figure 6 - Authenticated Content Browsing

The throughput of this test case displays significant similarities to the throughput for the login test cases. However, this test case has higher throughput than the login test due to the faster transaction times for content browsing.

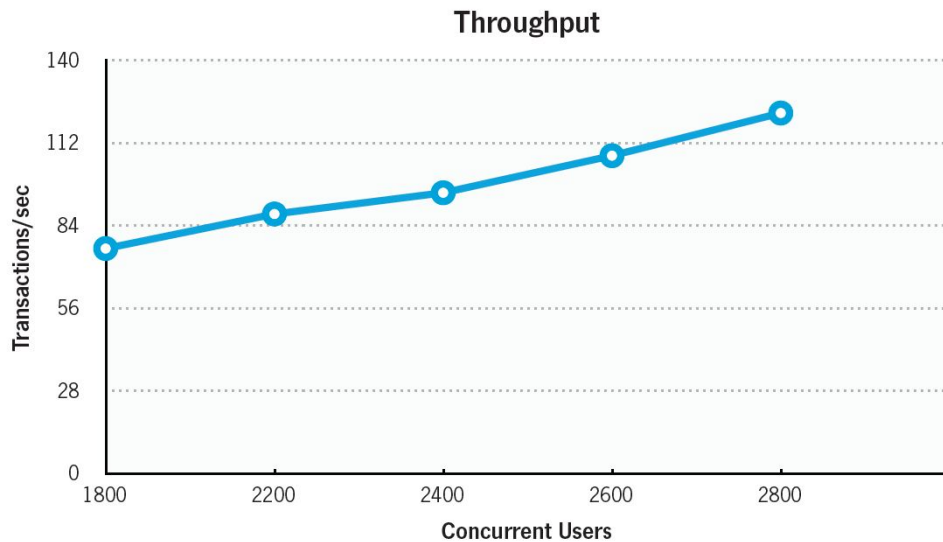


Figure 7 - Authenticated Browsing Throughput

As with the login test case, in the authenticated content access, Liferay Portal scales linearly, doubling the maximum concurrent user threshold when another physical server was added to the cluster.

Other Scenarios

Please see the whitepaper, "Liferay Portal Performance: Benchmark Study of Liferay Enterprise Edition" available here:

<http://www.liferay.com/documentation/additional-resources/whitepapers>

You will find statistics for these and other collaboration-centric scenarios to aid in the sizing of your Liferay system.

Sizing Your System

Now that we have taken a look at some of the performance metrics of the Portal, we can start to calculate the hardware that is required. Here is the reference hardware that we used, once again:

Hardware platforms:

4. Web Server
 - 1 x Intel Core 2 Duo E6405 2.13GHz CPU, 2MB L2 cache (2 cores total)
 - 4GB memory
 - 1 x 146GB 7.2k RPM IDE
5. Application Server
 - 2 x Intel Core 2 Quad E5430 2.66GHz CPU, 12MB L2 cache (8 cores total)
 - 8GB memory
 - 2 x 146GB 10k RPM SCSI
6. Database Tier
 - 1 x Intel Core 2 Quad E5430 2.66GHz CPU, 12MB L2 cache (8 cores total)
 - 16GB memory

Given the statistics in Table 3, here are the estimated hardware requirements for the three scenarios covered in this document:

ANONYMOUS CONTENT BROWSING HARDWARE REQUIREMENTS

Concurrent Users	Server Type	# of CPUs * per server	Memory per server	Notes / Infrastructure
150,000	Web Server	1 x Intel Core2 Duo	4GB	1 web server, 1 app server, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	
300,000	Web Server	1 x Intel Core2 Duo	4GB	2 web servers, 2 app servers, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	
450,000	Web Server	1 x Intel Core2 Duo	4GB	3 web servers, 3 app servers, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	

* equivalent or better

ISOLATED LOGIN HARDWARE REQUIREMENTS

Concurrent Users	Server Type	# of CPUs * per server	Memory per server	Notes / Infrastructure
3200	Web Server	1 x Intel Core2 Duo	4GB	1 web server, 1 app server, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	
6400	Web Server	1 x Intel Core2 Duo	4GB	2 web servers, 2 app servers, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	
9600	Web Server	1 x Intel Core2 Duo	4GB	2 web servers, 3 app servers, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	

* equivalent or better

AUTHENTICATED CONTENT BROWSING HARDWARE REQUIREMENTS

Concurrent Users	Server Type	# of CPUs * per server	Memory per server	Notes / Infrastructure
2800	Web Server	1 x Intel Core2 Duo	4GB	1 web server, 1 app server, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	
5600	Web Server	1 x Intel Core2 Duo	4GB	2 web servers, 2 app servers, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	
8400	Web Server	1 x Intel Core2 Duo	4GB	2 web servers, 3 app servers, 1 database server
	App Server	2 x Intel Core 2 Quad	8GB	
	Database Server	1 x Intel Core 2 Quad	16GB	

* equivalent or better

REALISTIC SCENARIOS

Of course, in reality, your system utilization will not be 100% anonymous content browsing, 100% isolated login, or 100% authenticated content browsing. It is up to you, as system architects and business analysts to predict the utilization by your end users, and make an educated estimation.

So for example, if you are a 9-to-5 shop, and you know that you are going to get a hit on your system at about 9:00 AM every morning, when everyone comes into the office and logs in, you know that you at least want to cover the concurrent users that you expect to login at that time every day. Then you can mix and match the scenarios from the Liferay Performance

whitepaper, and use the statistics there to make sure you cover the scenario that could potentially become the bottleneck. For example, let's say that you have mostly web content from Liferay WCM, and you know that 75% of your users will be anonymous (unauthenticated), and that 25% will be authenticated. For 10,000 total concurrent users, you can see from the statistics that 7500 anonymous concurrent users will easily be handled, and that 2500 authenticated concurrent users will just make it under the threshold of acceptable performance, given the reference hardware configuration above.

For more scenarios, please see the whitepaper, "Liferay Portal Performance: Benchmark Study of Liferay Enterprise Edition" available here:

<http://www.liferay.com/documentation/additional-resources/whitepapers>

You will find statistics for these and other collaboration-centric scenarios to aid in the sizing of your Liferay system.

In Conclusion

This document is intended to provide only an initial hardware recommendation based on benchmarks that have been performed on the Liferay Portal product. It is up to you to determine the utilization scenarios that your system will need to service and to run the appropriate load tests on your system before production deployment, so that you can identify any significant bottlenecks due to custom applications/portlets, and other unforeseen system and network issues.

We hope that this document and the Liferay Portal Performance whitepaper (<http://www.liferay.com/documentation/additional-resources/whitepapers>) will be able to provide helpful guidance in sizing your system and procuring your hardware.

Moving Forward

LIFERAY ENTERPRISE EDITION SUPPORT

Liferay Enterprise Edition ensures stability and reliable technical support for your Liferay Portal installation and your organization's team. Including a customer portal, product bulletins, security alerts, plus the support of over 60 partners worldwide.

LIFERAY PROFESSIONAL SERVICES

Liferay Professional Services can help you in the design, planning, and implementation of your system. Performance tuning consultation is also available.

Contact sales@liferay.com for more information.