



Fixed Price Website Load Testing

Can your website handle the load? Don't be the last one to know. For as low as \$4,500, and in many cases within one week, we can remotely load test your website and report how user experience changes with load and determine if, and under what load, problems occur.

Work is performed by seasoned performance engineers, on shore, using open source load testing software we support and service. We generate load from Amazon cloud computers which minimizes server and network costs. After the work has been completed, optional training and mentoring on tools and methodology is available should you wish to make further use of the test assets we developed.

Terms & Conditions

1. A maximum of 3 user scenarios accessing up to a total of 15 unique web pages will be created to apply load to the system under test.
2. Three testing sessions lasting up to 4 hours each will be conducted. A preliminary report will be provided after the first two tests showing user response time vs. Users, throughput vs. Users, Network bandwidth vs. Users, and CPU¹ vs. Users. After the first round of testing, you have an opportunity to reconfigure the infrastructure or tune the application prior to the final test. A final report will be issued one business day after testing is completed (see sample attached).
3. Load is generated from our servers in the Amazon cloud or from server(s)² you provide on your LAN.

Pricing (new lower pricing reflects our savings from moving load generators to the cloud)

Virtual Users	Cost (US\$)
1,500	\$4,500
2,500	\$5,050
5,000	\$6,425
7,500	\$7,800
10,000	\$9,175
15,000	11,925
Over 15,000	Call for quote

Additional consulting is available at \$1,200/day

- Create more complex workloads
- Identify and correct bottlenecks
- Evaluate stability under sustained load
- Perform failover testing
- Train your staff on our methods and test assets

¹ Requires remote access to your servers to run performance monitoring tools.

² Windows Server 2003 or XP Pro only, Pentium 2.8 GHz, 1Gb memory, 100Mb LAN minimum. These servers must be on the same LAN as the system under test or you risk measuring the capacity of a network link instead of your server capacity. One server is required for each 750 virtual users to be emulated.

SAMPLE REPORT

Contents

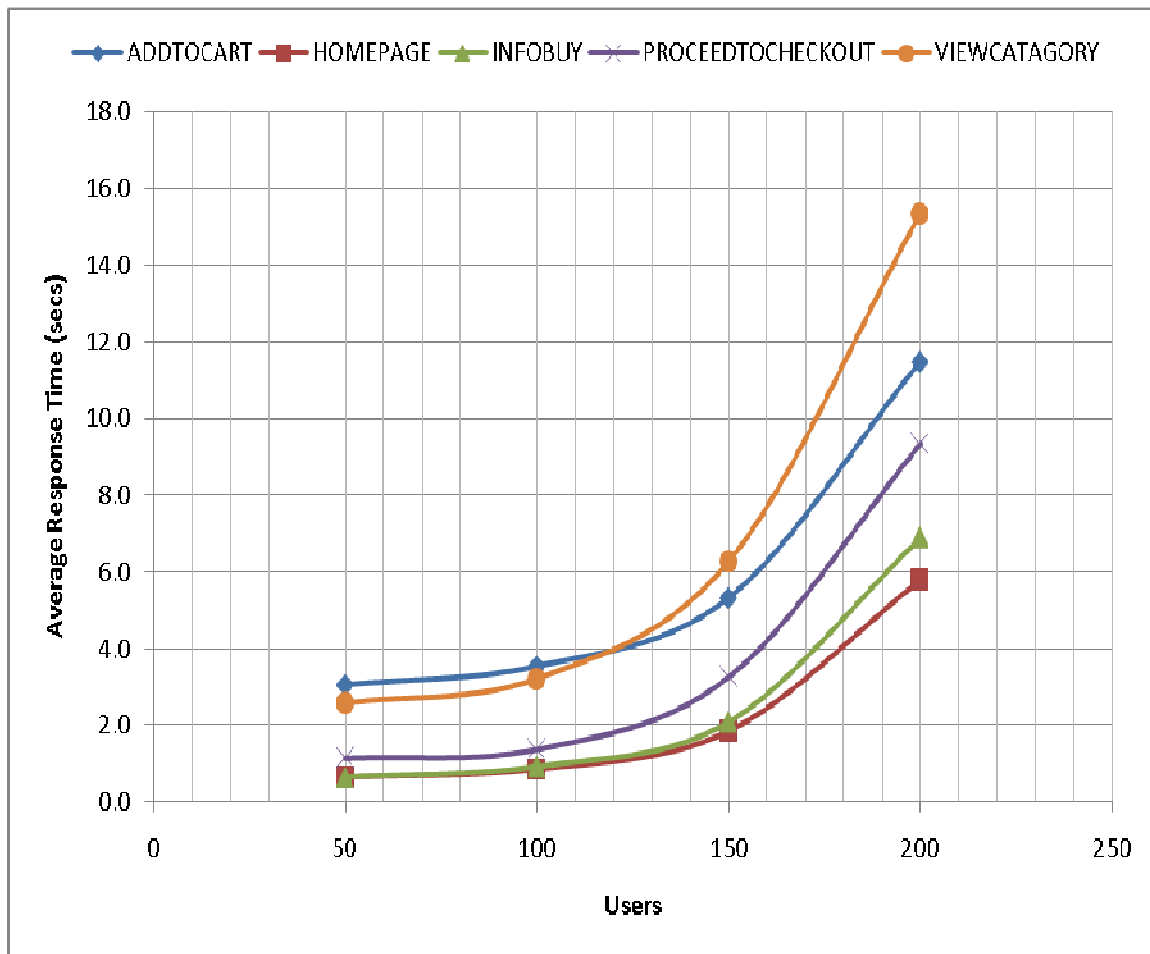
Summary	2
Response time vs. Users graph	3
Throughput vs. Users graph.....	4
CPU Busy vs. Users by server graph.....	5
Network Bandwidth vs. Users	6
Response time tabular data	7
Script Narritive.....	8
Testing Methodology Summary	9

Summary

1. This is the final load test report for <website name> and is based on tests conducted on 1/3/2008.
2. Three changes were made to the system under test cluster since the
 - o The number of web/app servers was increased from 2 to 3
 - o New session management implemented on web/app server
 - o Software mirroring was disabled on the DB server
3. The improvement was greater than expected. Past results showed the server ran well up to 20 users. Now the cluster provides consistently good performance up to 150 users.
4. Submitting orders continues to take a long time. The minimum time observed was 35 seconds. This is a slight improvement over the last cluster configuration where it took 40 seconds to submit an order.
5. The environment tested consisted of the 3 web/app servers, 1 live-DB server, and 1 read-only DB server
6. The website can handle 125 active users comfortably after which it is limited by web/app server CPU capacity.
7. There is ample capacity at the live and read-only database tiers to support further growth. By adding 3 web/app servers the overall website capacity could be increased to approximately 250 active users³.

³ Assuming no I/O bottlenecks emerge on either of the database servers. If expansion is planned, it is recommended that the tests be re-run with additional web servers configured to validate DB server IO capacity.

Response time vs. Users graph

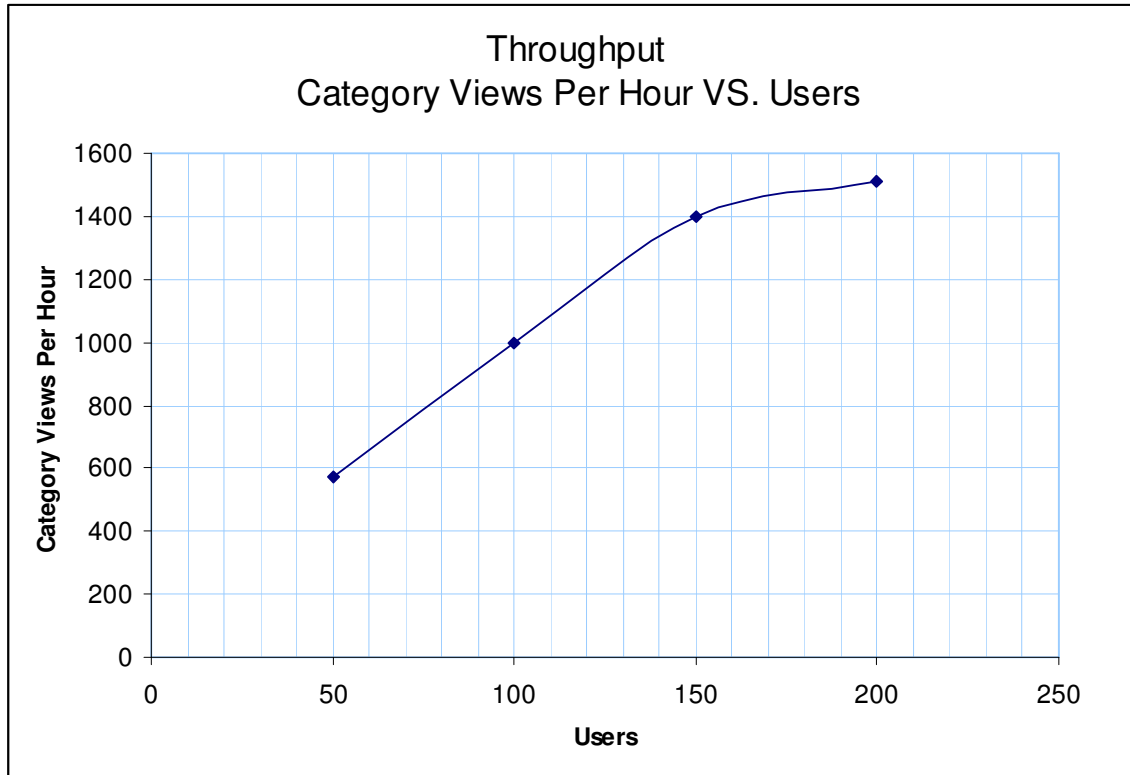


This response time graph plots average response time (in seconds) vs. the number of active users. Users simulate browse and buy activity against your website with an average think time of 30 seconds. A script narrative is included at the end of this report which describes the activity in greater detail.

This graph indicates a bottleneck beginning to develop at 150 users. The time to submit orders is in excess of 30 seconds and if included on this graph would mask the shapes of these curves. For more information on submit orders, see the tabular data which appears later in this report.

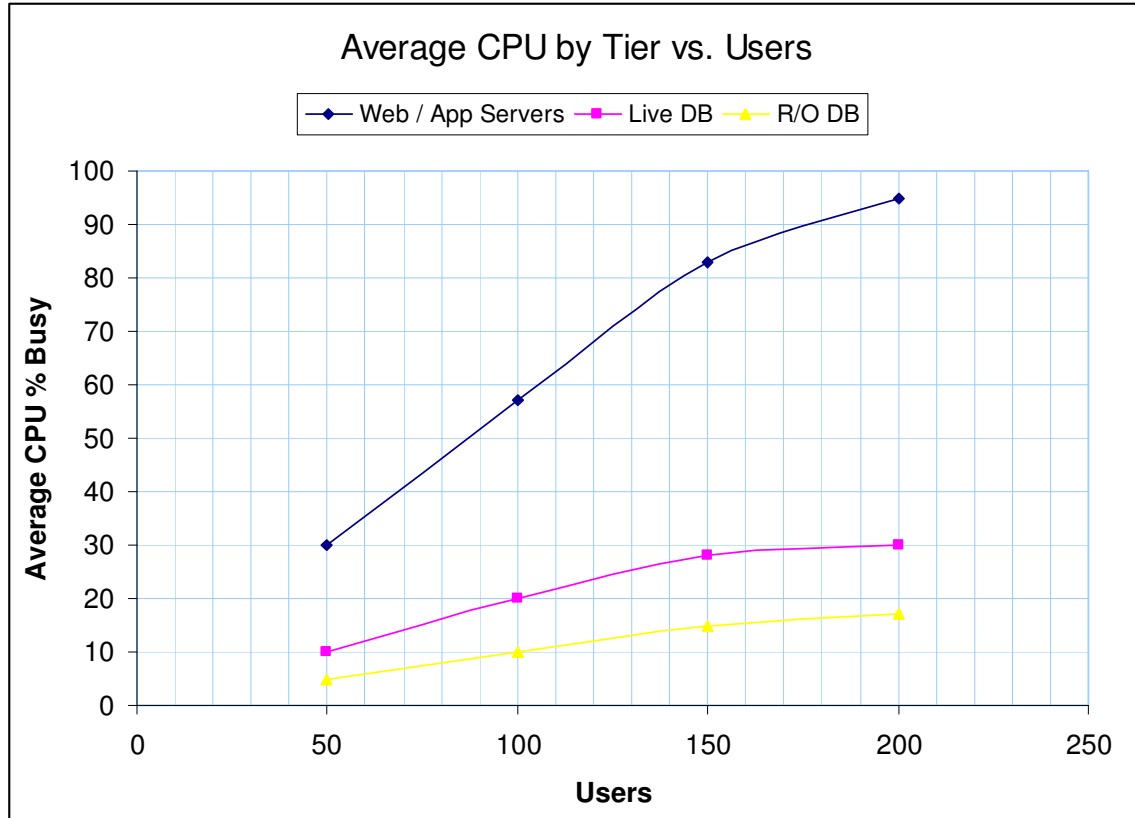
The response time reported includes all network and server processing time but does not include time for the browser to render pages or run JavaScript.

Throughput vs. Users graph



Increasing virtual active users increases the rate at which work is requested. Throughput is a measure of how much work was completed. This graph shows the number of times a commonly executed transaction completed a one hour period. The graph should be linear. This graph indicates a bottleneck above 150 users.

CPU Busy vs. Users by server graph

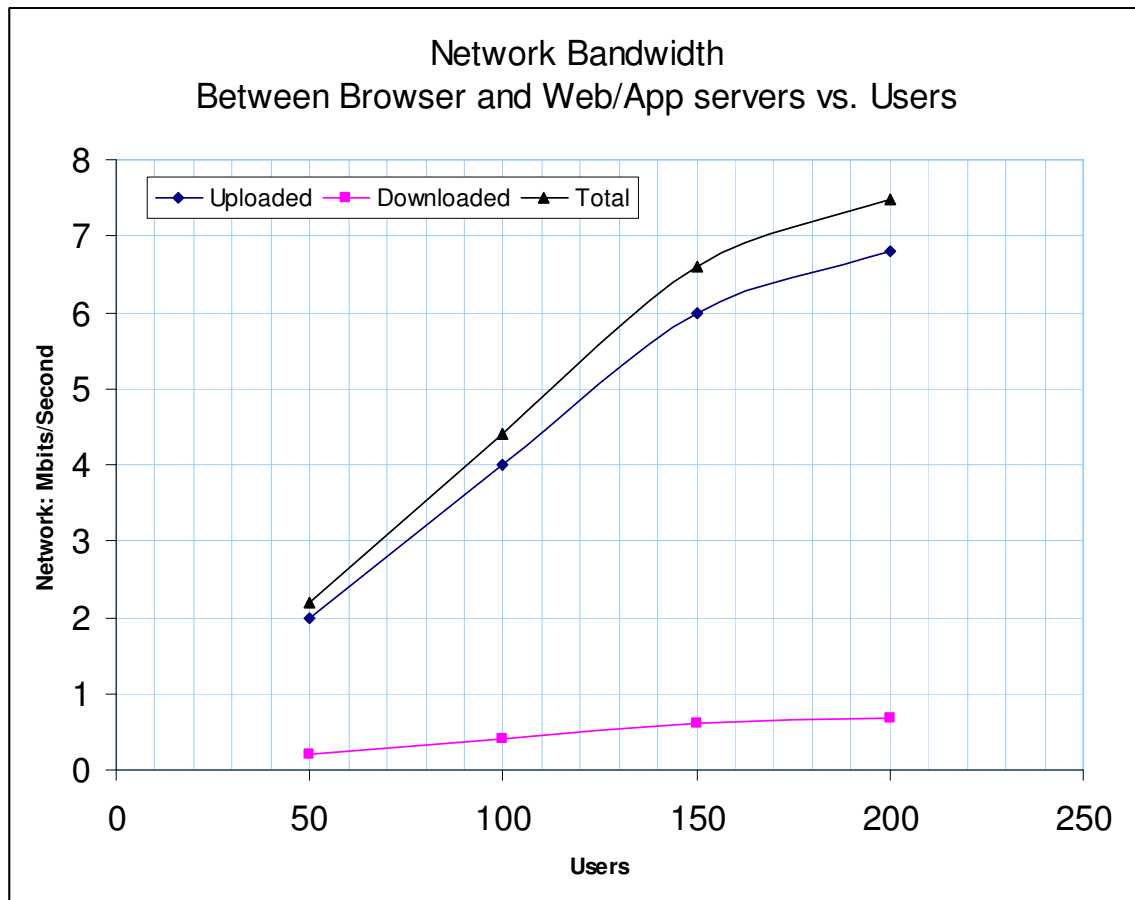


CPU busy was measured using the Windows Performance Monitor. Web/App server CPU utilization (averaged for the 3 servers and is plotted in blue) is past reasonable capacity limits (65% to 80% depending on tolerance for risk) at 150 users and is the primary bottleneck to server capacity.

At a point near 150 users, a small increase in load will result in a large increase in response time (as is demonstrated by the response time graph a few pages back.)

The CPU graph indicates that live and read-only database servers have ample headroom for growth and overall capacity could be increased by adding additional web/app servers be deployed. Assuming no I/O bottlenecks occur on the database servers, cluster capacity could be doubled by deploying 3 more web/app servers.

Network Bandwidth vs. Users



Network bandwidth was measured using the Windows Performance on our load generation servers. The highest average load occurred at 200 active users and was measured to be just over 7 Mbits/sec. This graph also indicates a bottleneck above 150 active users.

Legend:

“Uploaded” is data sent from your website to our simulated browsers

“Downloaded” is data sent from our simulated browsers to your website.

Response time tabular data

Timer Name	Active Users	Average	Count	StdDev	Min	max
ADDCART	50	3.1	50	0.2	2.7	3.8
	100	3.5	116	0.6	2.8	6.7
	150	5.3	138	1.6	3.0	10.0
	200	11.5	160	3.6	6.2	30.9
HOMEPAGE	50	0.6	172	0.2	0.5	1.6
	100	0.9	336	0.3	0.5	2.7
	150	1.9	462	1.0	0.5	6.5
	200	5.8	498	3.3	0.8	36.6
INFOBUY	50	0.7	516	0.2	0.4	1.7
	100	0.9	1001	0.4	0.4	3.2
	150	2.1	1377	1.1	0.5	9.6
	200	6.9	1520	3.1	0.9	37.3
PROCEEDTOCHECKOUT	50	1.2	16	0.5	0.8	2.5
	100	1.4	21	0.5	0.8	3.0
	150	3.3	37	1.8	1.3	10.0
	200	9.4	39	3.3	0.8	19.0
SUBMITORDER	50	37.9	4	1.3	36.1	39.2
	100	55.9	3	18.3	35.2	69.9
	150	46.3	16	9.9	30.7	65.9
	200	60.6	9	18.2	43.5	91.4
VIEWCATAGORY	50	2.6	518	0.5	1.8	6.2
	100	3.2	997	1.0	1.8	8.1
	150	6.3	1380	2.5	2.0	23.2
	200	15.4	1518	4.7	2.9	48.3

Legend: Timer Name is the name of the event and can be cross referenced with the workload narrative included at the end of this document. Active users are defined as being connected to the system under test and performing work at a rate described in the UI narrative (working with a think time of about 30 seconds in this case). Count is the number of events completed in a 10 minute period (our quasi-steady state observation period). Stddev is standard deviation (<1 means relatively consistent performance, higher values indicate inconsistent performance). Average, Min, and Max are response time statistics in seconds.

Script Narrative

<WebsiteName>.com Catalog browse and buy script

This script emulates a new user visiting the web site and purchasing item at random. The script includes several exit points considering that only a small fraction of sessions end with the purchase of an item.

UI Narrative

Goto <WebsiteName>.com Time: Homepage

Wait 30 seconds

Loop from 1 to 5 times (average is 3)

 View a category at random. Time: ViewCatagory

 Wait 30 seconds

 Press info/buy for a random product (not a package)

 Wait 30 seconds

End of loop

70% of the time, exit the script (30% of original sessions continue)

Add random product to cart. Time: AddToCart

Wait 30 seconds

70% of the time, exit the script (9% of original sessions continue)

Proceed to checkout. Time: ProceedToCheckout

Wait 30 seconds

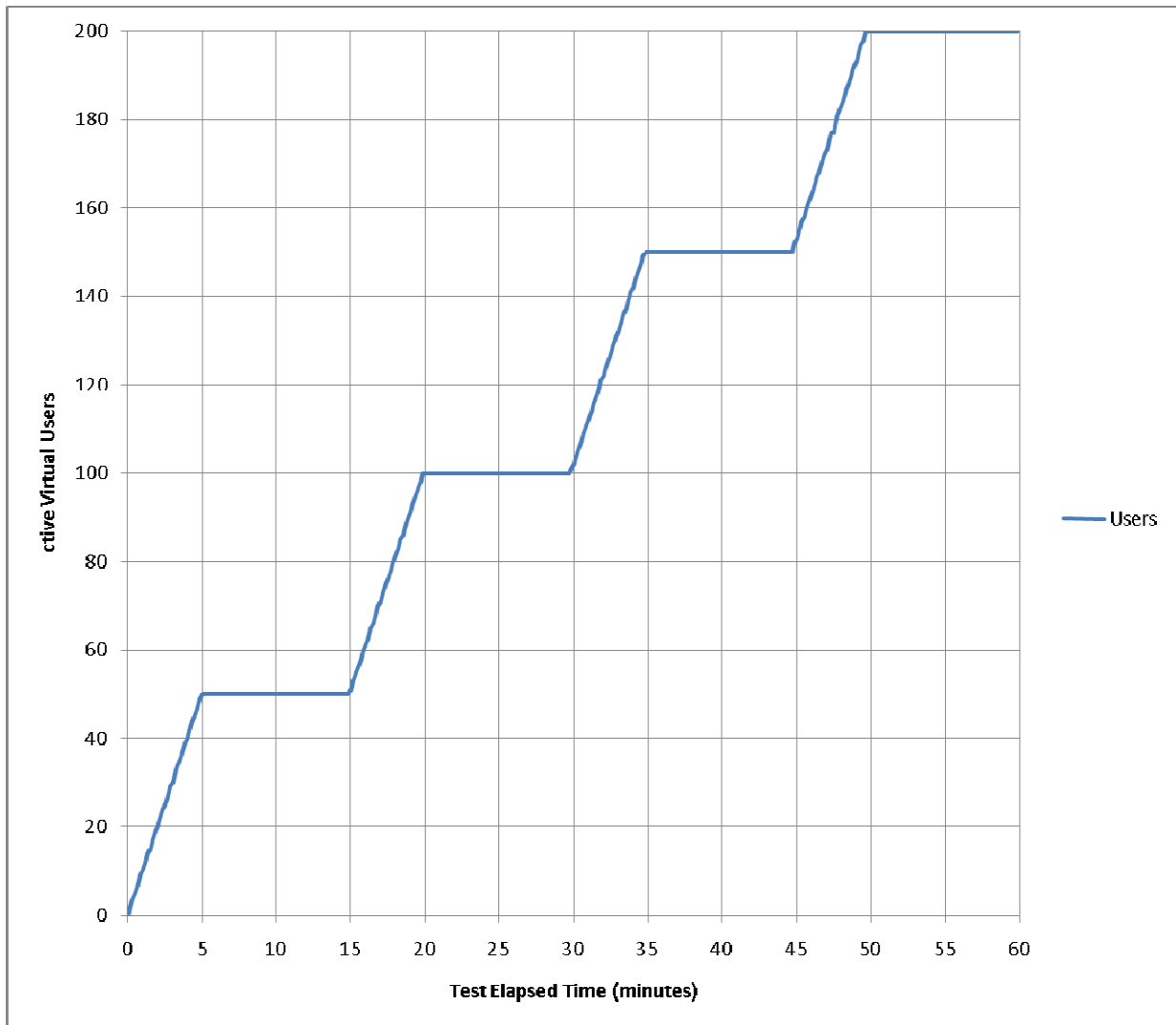
70% of the time, exit the script (2.7% of original sessions continue)

Submit order. Time: SubmitOrder

Testing Methodology Summary

Once the workload has been defined and codified in scripts, a load is applied in a step wise fashion to various hardware configurations so that trends in response time, throughput, and hardware resource utilization can be measured and analyzed.

To allow the system to settle into a steady state and achieve consistent and repeatable results, tests run for at least one hour. Each test was logically partitioned into four 15 minute periods. Each period consisted of a ramp-up lasting 5 minutes where 25% of the maximum workload was added followed by a 10 minute data collection window. An average value was produced for each test metric (i.e. response times, throughput, and operating performance statistics by server) over each of the ten minute intervals. Data collected during the ramp up period was discarded. For example, a 200 user test would look like this;



Contact Bernie.Velivis@iPerformax.com or visit iPerformax.com for more information

© 2009 Performax Inc