Understanding and Creating Effective Performance Test Reports

First Presented for:
Florida Institute of Technology
Computer Science Seminar

Scott Barber
Chief Technology Officer
PerfTestPlus, Inc.

Agenda

Introduction

What Stakeholders Want

What Stakeholders (Usually) Get

How to "give 'em what they want" (or at least get close)

Valuable Charts, Tables, Verbiage and Their Uses

Stakeholder Reaction to Charts, Tables and Verbiage

Comprehensive Final Report Template

Want More Information?

Summary/Questions

Introduction

Most people will never read performance test results docs.

Most people don't really understand the underlying components to performance.

It is our job to make it easy for them to understand, and understand quickly.

Being skilled at graphical presentation of technical information is critical for us to help others understand the message we are delivering.

Improper or confusing charts and tables can lead to wrong decisions that cause lost $ and ruined reputations.

What Stakeholders Want

Answers… NOW!   (They might not even know the question)

To understand information intuitively.

Simple explanations of highly technical information.

To be able to make decisions quickly and have the
information to support those decisions.

"Trigger phrases" to use with other stakeholders.

Concise summaries and conclusions.

Recommendations and options.

What Stakeholders (Usually) Get

Complex (verbal) technical descriptions.

Complex technical graphics.

Raw data.

i.e. they get confused!

Timer Name	Baseline		250		500		750	
	95th	Max	95th	Max	95th	Max	95th	
ec_Main_Page	10.46	18.09	6.41	8.22	6.33	8.33	3.8	
ec_logon_help	0.98	0.98	0.56	0.59	0.55	0.55	0	
ec_login	5.35	7.92	6.66	11.84	6.75	17.03	5	
quick_learns	6.66	6.67	5.91	10.98	5.92	11.02	3.92	
view_quick_learn	15.66	17.11	5.53	10.72	3.89	10.61	3.46	10.48
view_faq_window	2.45	2.45	1.47	1.52	1.53	1.66	1.30	6.2
view_faq	0.67	0.67	0.60	0.63	0.58	0.69	0.50	5
view_ec_status	8.08	12.55	1.73	6.66	1.80	1.86	1.60	
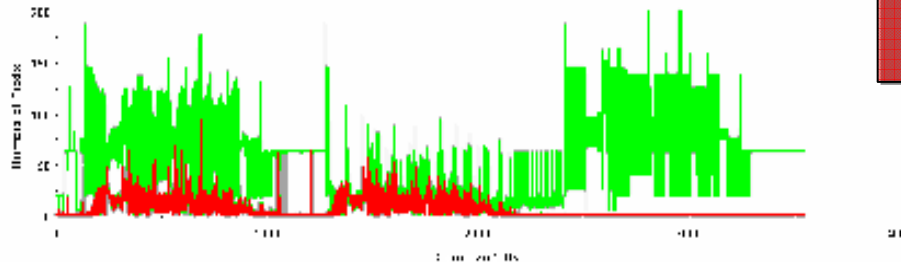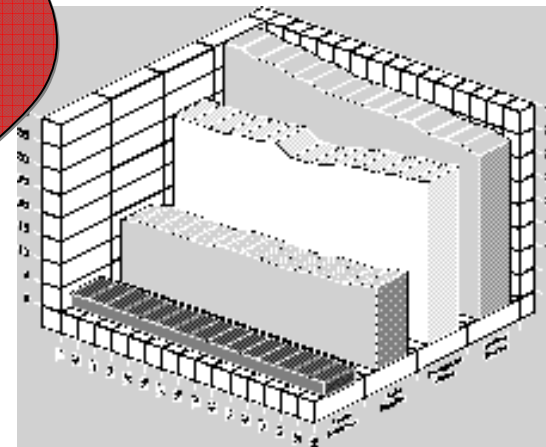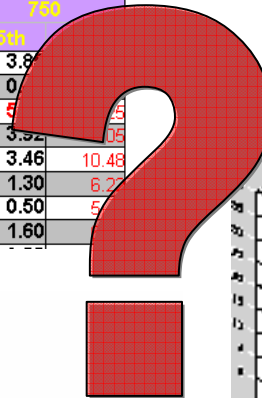
"Give 'em what they want" (or at least get close)

Concise verbal descriptions.

Well formed, informative charts (pretty pictures).

Focus on requirements and business issues.

Don't be afraid to make recommendations or draw conclusions!

Have supporting data available.

Charts, Tables and Verbiage

Performance Report Output Chart

Response vs. Time Scatter Chart

Component (Resource Usage) Performance Chart(s)
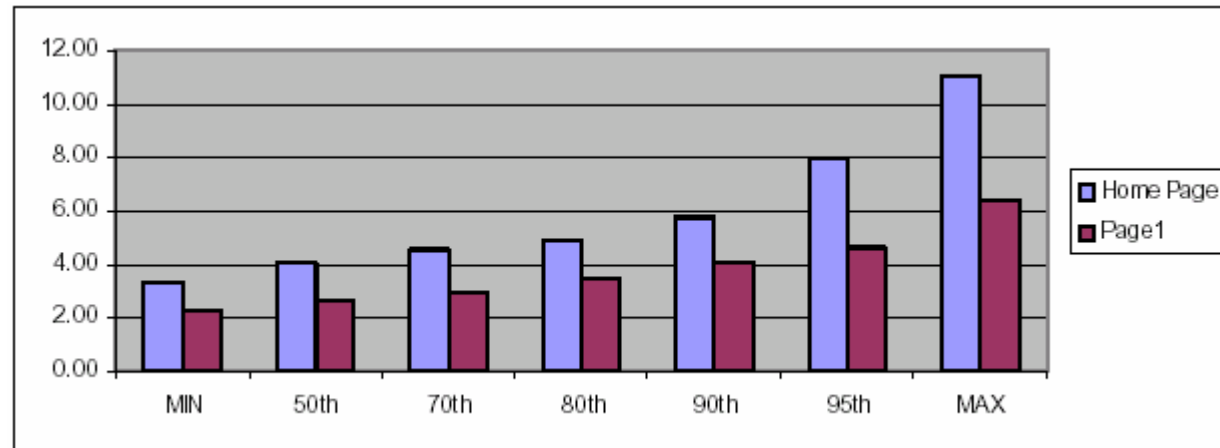
Consolidated Scatter Chart

Response Time Summary Comparison Chart

Response Time by Test Execution Chart

Response Time Degradation Curve

Performance Report Output Chart

CmdID	NUM	MEAN	STD DEV	MIN	50th	70th	80th	90th	95th	MAX
Home Page	99	4.53	1.47	3.33	4.08	4.57	4.87	5.77	7.99	11.05
Page1	100	2.94	0.85	2.26	2.59	2.91	3.48	4.08	4.65	6.44

"This chart shows that the home page meets the 6 second response time requirement 90% of the time under the tested user load. The requirements demand 95% compliance. Page1 was the next slowest page, but it achieved the required response time."

Performance Report Output Chart

Good for:

- Showing response times for high interest pages during a single test.
- Highlighting select instances of poor (or good) performance.

Not good for:

- Results from multiple tests.
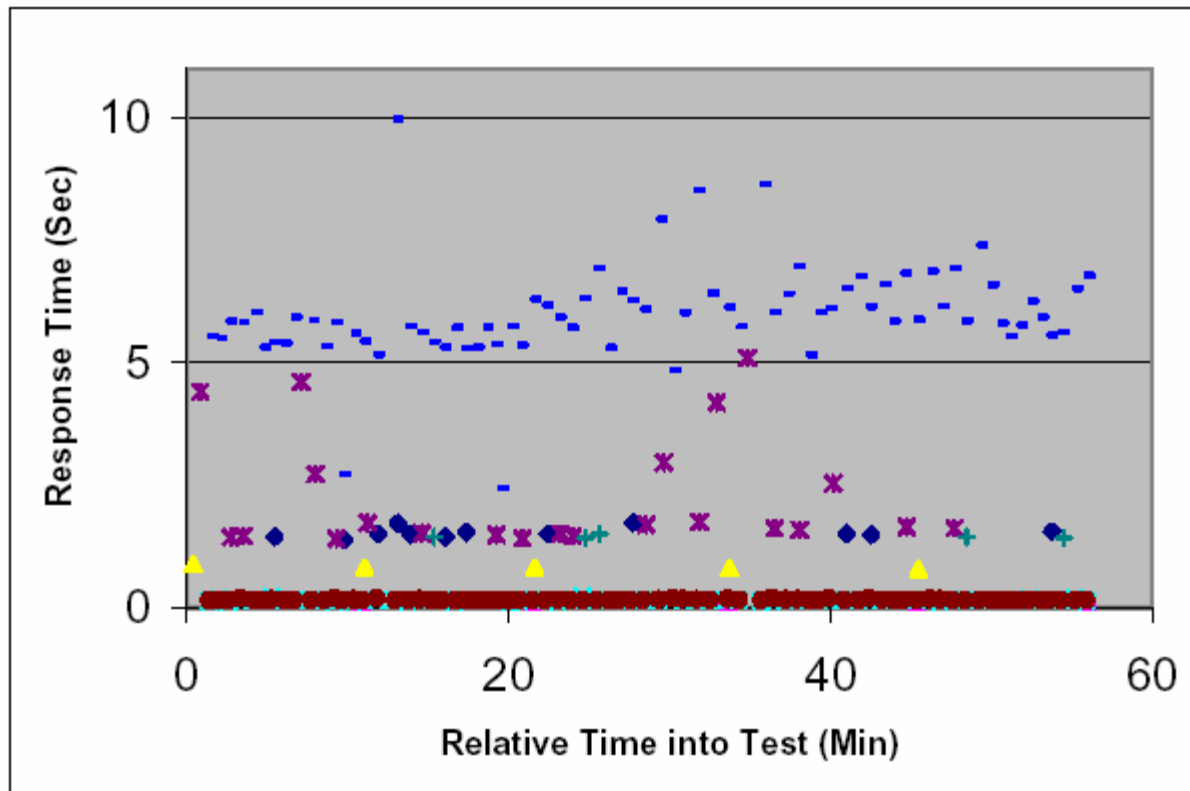- Component/resource based info.
- Results about many pages at once.

Notes:

- This chart and table can be viewed directly in TestManager or recreated in Excel.

Response vs. Time Scatter Chart

"This chart shows that the home page (blue squares) is the slowest page with most response times hovering around the the 6 sec requirement. Page1 (purple X) was the next slowest page, but it achieved the required response time."

Response vs. Time Scatter Chart

Good for:

- Identifying patterns in response times over a whole run.
- Graphically displaying response times vs. goals.
- Highlighting select instances of poor performance.
- Can be overlayed with component/resource data.

Not good for:

- Results from multiple tests.

Notes:

- This chart and table can be viewed directly in TestManager or recreated in Excel.
- Many component/resource metrics can be collected and overlayed directly in TestManager.

Component Performance Chart
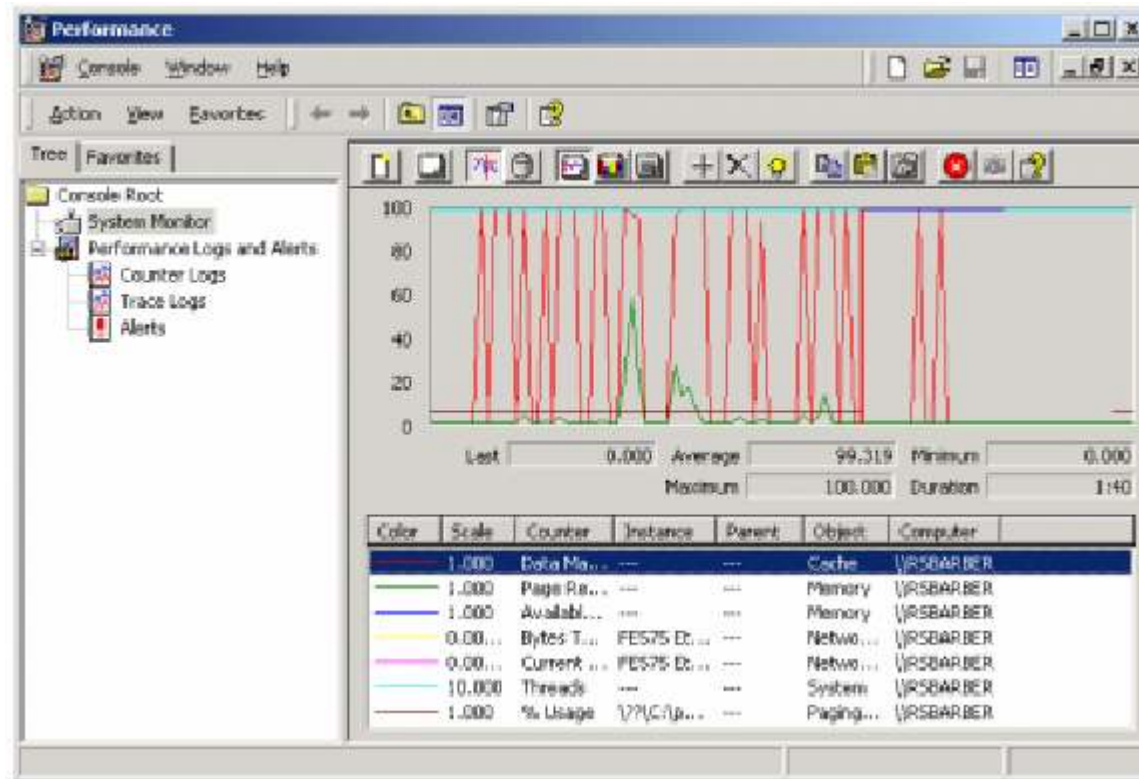
"The red line shows that the CPU of the web server bounced between 0 and 100% utilization during a test run that was not designed to be stressful. This is abnormal and needs to be researched further."

Component Performance Chart

Good for:

- Identifying various resource usage.
- Determining issues such as memory leaks.

Not good for:

- Response time reporting.
- Results from multiple tests.

Notes:

- Charts like this one can be created during your script by putting agents on the servers to be monitored.
- Charts like this one can also be created using tools like Perfmon, Perfmeter and Top.

Consolidated Scatter Chart

"This chart shows both response times and resource utilization together. Close examination shows that Application Server CPU Usage and Queue length coincide with degraded Response Time."
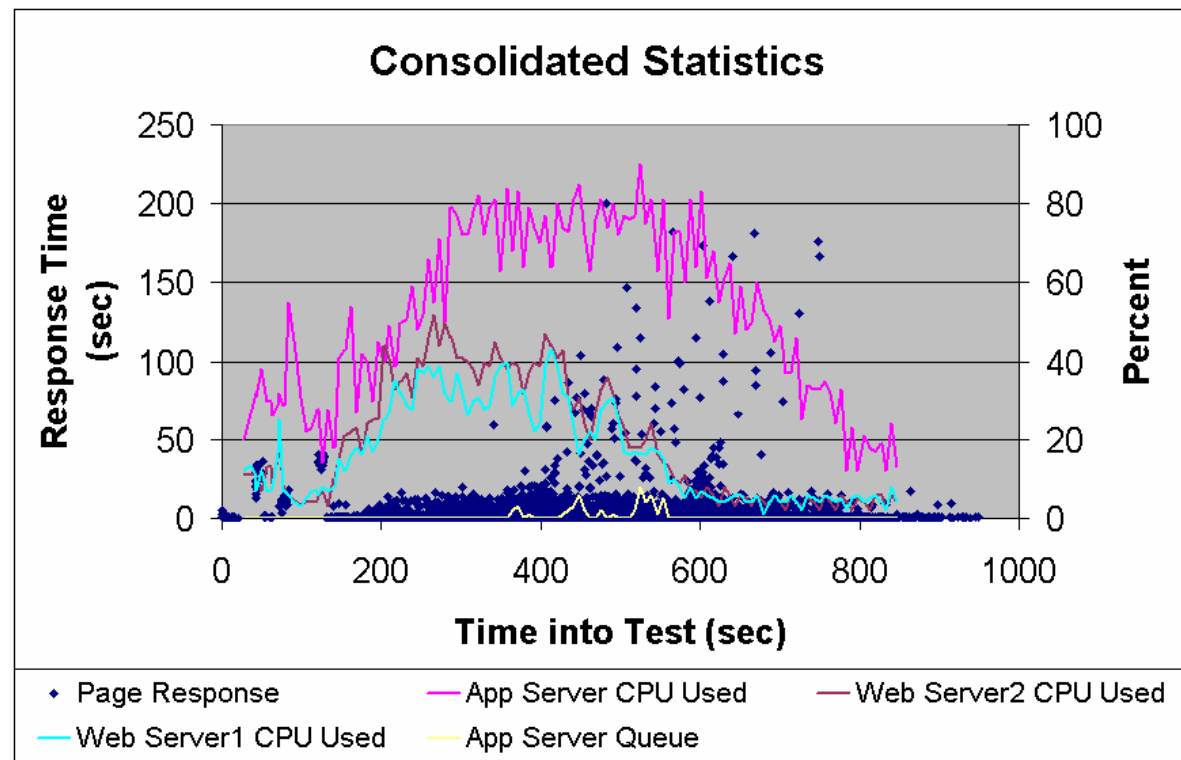
Response vs. Time Scatter Chart

Good for:

- Identifying correlations between response times and resource usage during over time.
- Graphically displaying response times vs. resources.
- Highlighting potential causes of poor performance.
- Technical stakeholders.

Not good for:

- Results from multiple tests.
- "Low tech" stakeholders.

Notes:

- This chart and was created in Excel, but in many cases can be done directly in TestManager.

Response Time Summary Comparison

"This chart shows that the majority of pages met their required response times at up to 150 hourly users accessed via LAN, but that most did not via slower connections."

	1 LAN User	50 LAN Users	100 LAN Users	150 LAN Users	200 LAN Users	100 128kbs Users	100 56.6kbs Users	100 28.8kbs Users
Pages Timed	177	175	175	176	176	176	169	169
Meet Goal	146	160	132	133	6	78	47	30
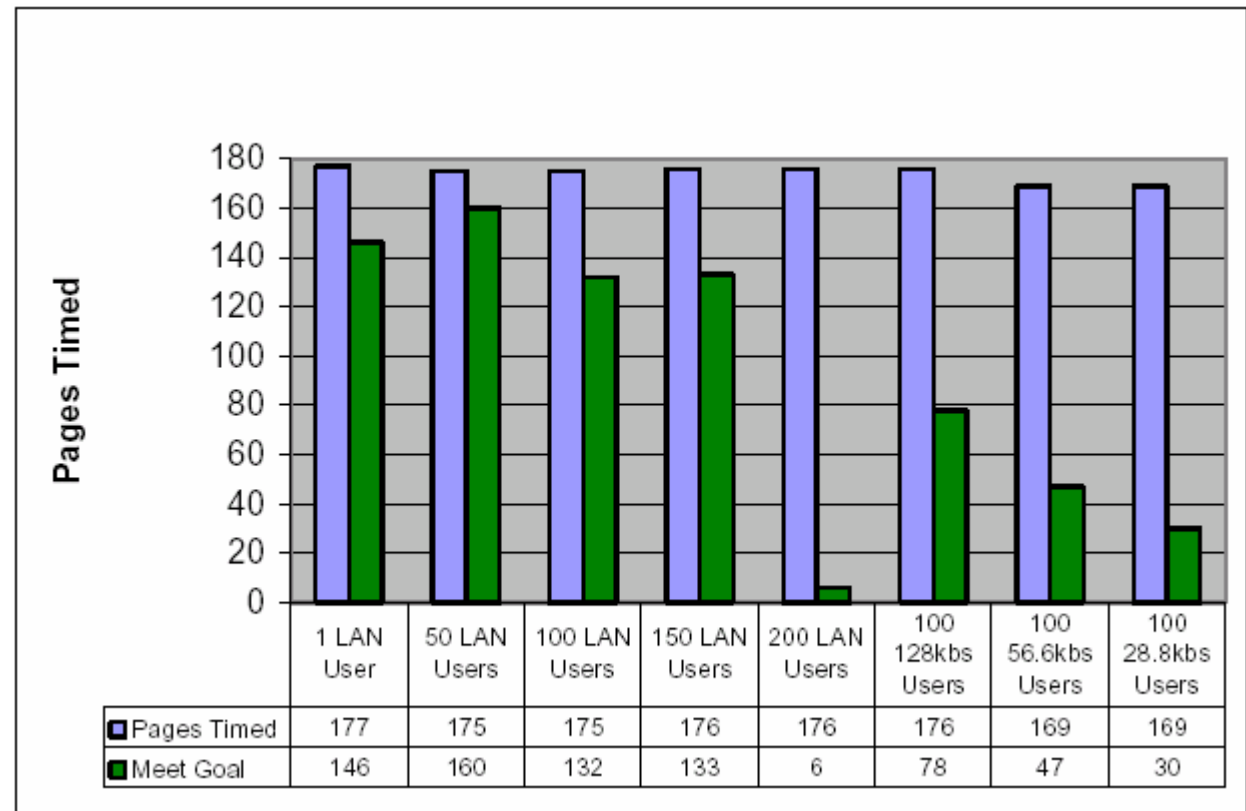
Response Time Summary Comparison

Good for:

- Summarizing requirements met across various and diverse tests.
- Comparing results across multiple tests.
- Highlighting tests with particularly bad (or good) performance.

Not good for:

- Depicting individual response times or component/resource data.

Notes:

- This chart and table must be created by hand in Excel using test execution data and stated requirements.

Summary Comparison								
Statistic	LAN					128 kbs	56.6 kbs	28.8 kbs
Users	1	50	100	150	200	100	100	100
Times Recorded	177	175	175	176	176	176	169	169
Times Under Goal	146	160	132	133	6	78	47	30
% Times Under Goal	82.5%	91.4%	75.4%	75.6%	3.4%	44.3%	27.8%	17.8%

Response Time by Test Execution

	1 LAN User	50 LAN Users	100 LAN Users	150 LAN Users	200 LAN Users	100 128kbs Users	100 56.6kbs Users	100 28.8kbs Users
■ HomePage	5.59	9.08	10.14	10.11	48.21	11.42	24.85	33.89
■ Page1	0.75	0.74	1.49	2.37	41.22	6.80	14.10	27.09

Response Time (Sec) — vertical axis scale: 0, 10, 20, 30, 40, 50, 60
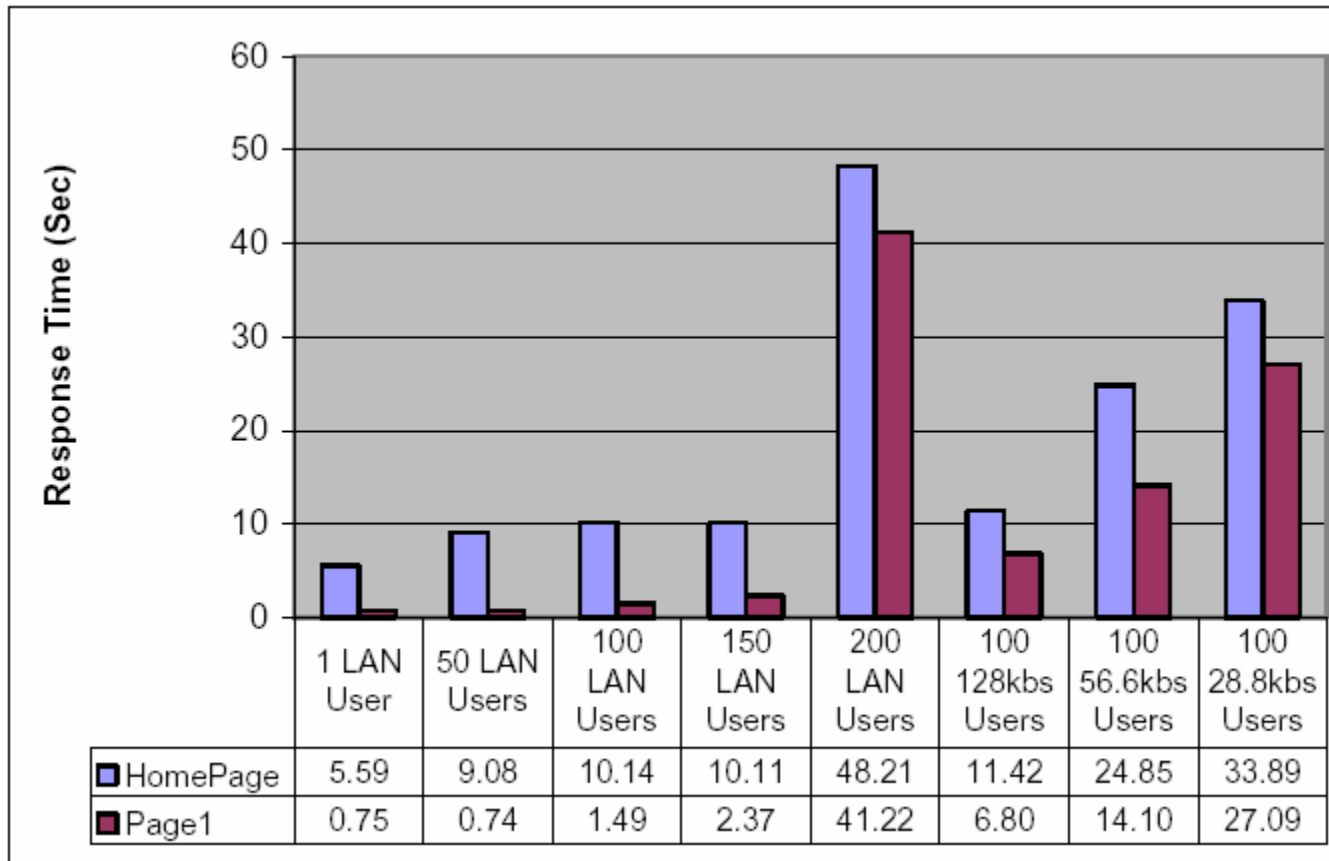
"This chart shows that response times for both Home Page and Page1 degrades dramatically and unacceptably at 200 hourly users and via modem connection speeds."

Response Time by Test Execution

Good for:

- Comparing specific page response times across multiple tests.
- Graphically displaying scenarios performance side-by-side.
- Highlighting test configurations with universally poor performance.

Not good for:

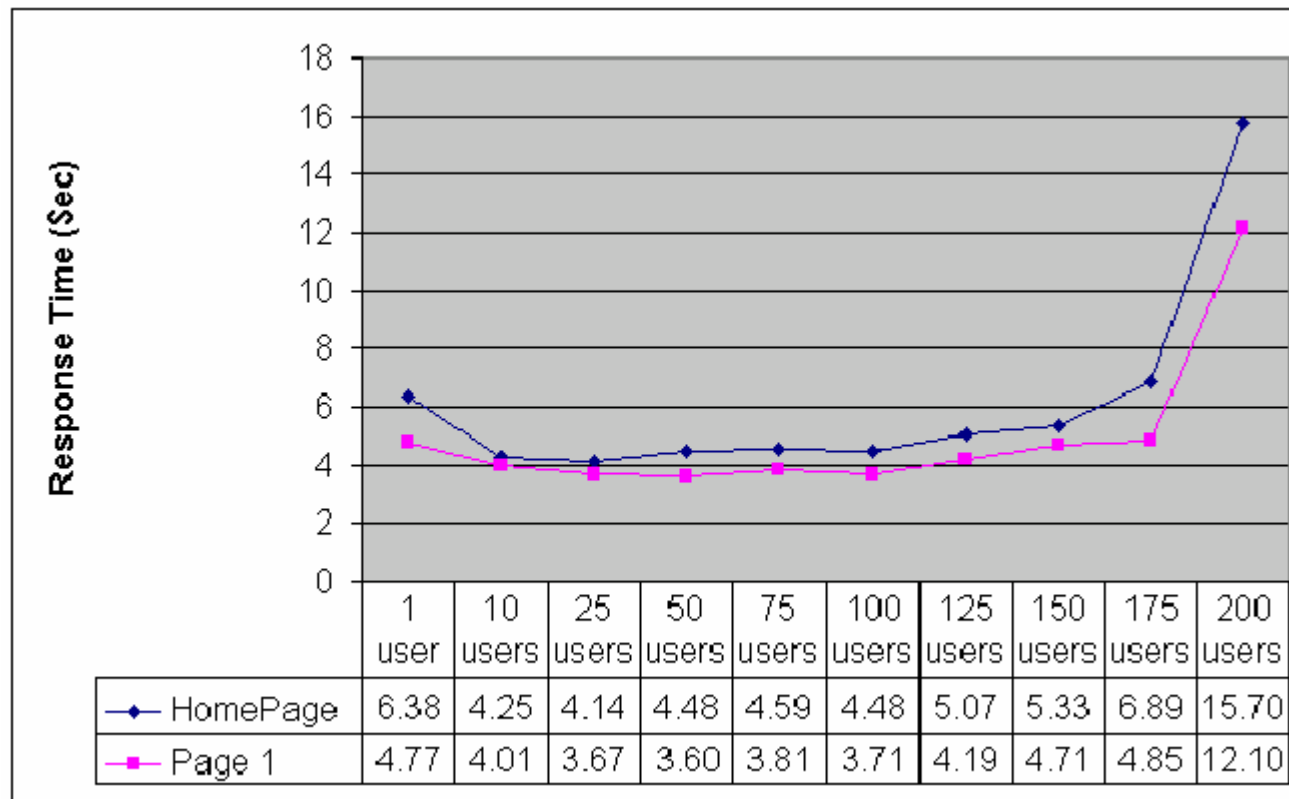- Displaying component/resource data.

Notes:

- This chart and table must be created in Excel from data exported from TestManager.

Page		LAN					128 kbs	56.6 kbs	28.8 kbs
Users	1	50	100	150	200	100	100	100
Home Page	5.72	0.60	1.93	2.94	46.48	7.81	9.02	16.92
Page 1	3.85	0.28	0.36	0.26	17.95	4.51	9.90	19.42
Page 2	0.21	0.20	2.37	0.16	8.53	2.89	6.24	12.12
Page 3	4.99	1.07	3.31	3.70	29.90	28.88	10.72	19.87

95th Percentile Time Comparison

Response Time Degradation Curve

"This chart shows that performance starts to degrade at 150 hourly users and becomes unacceptable between 175 and 200 hourly users."

	1 user	10 users	25 users	50 users	75 users	100 users	125 users	150 users	175 users	200 users
HomePage	6.38	4.25	4.14	4.48	4.59	4.48	5.07	5.33	6.89	15.70
Page 1	4.77	4.01	3.67	3.60	3.81	3.71	4.19	4.71	4.85	12.10

Response Time Degradation Curve

Good for:

- Comparing specific page response times across multiple tests.
- Graphically displaying where performance gets bad.
- Highlighting load where performance becomes unacceptable.

Not good for:
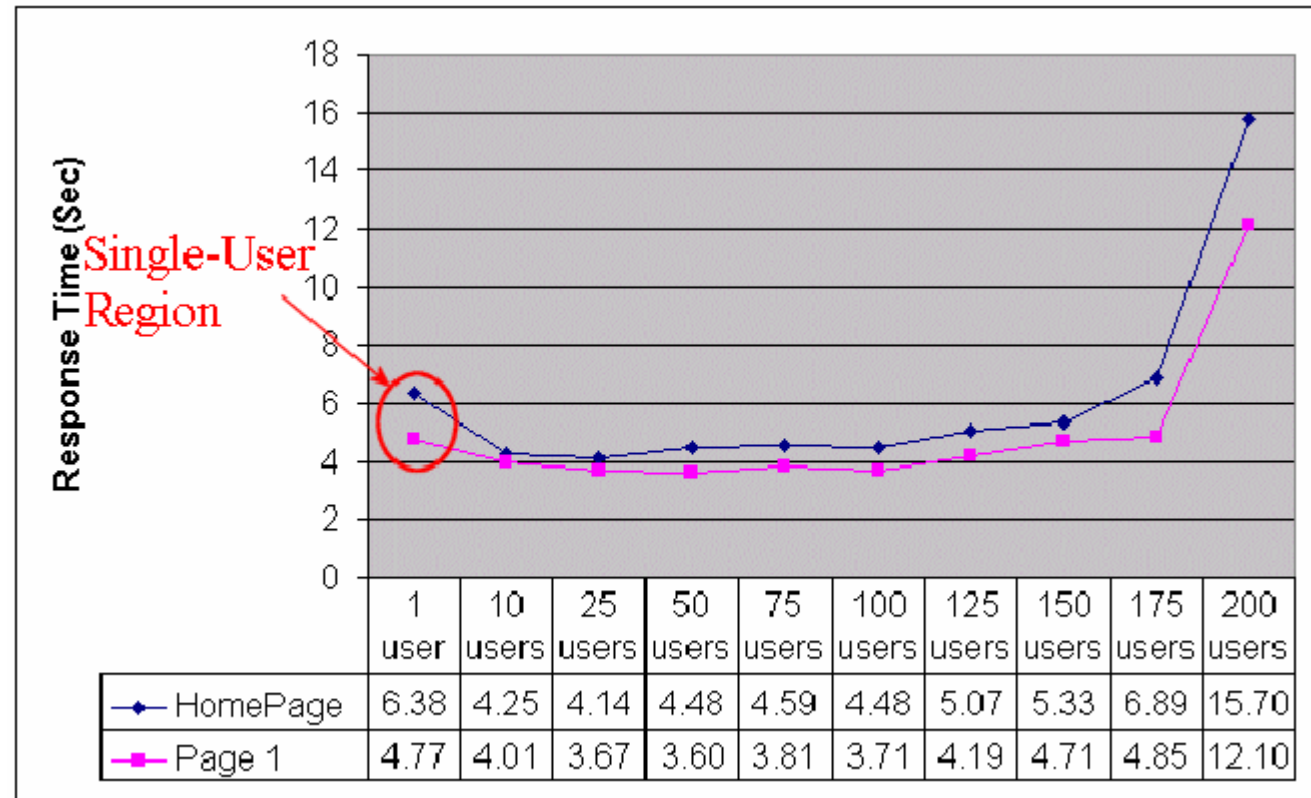
- Displaying component/resource data.

Notes:

- Generally considered the most powerful graph at a performance tester/engineer's disposal.
- This chart and table must be created in Excel from data exported from TestManager.

Response Time Degradation Curve
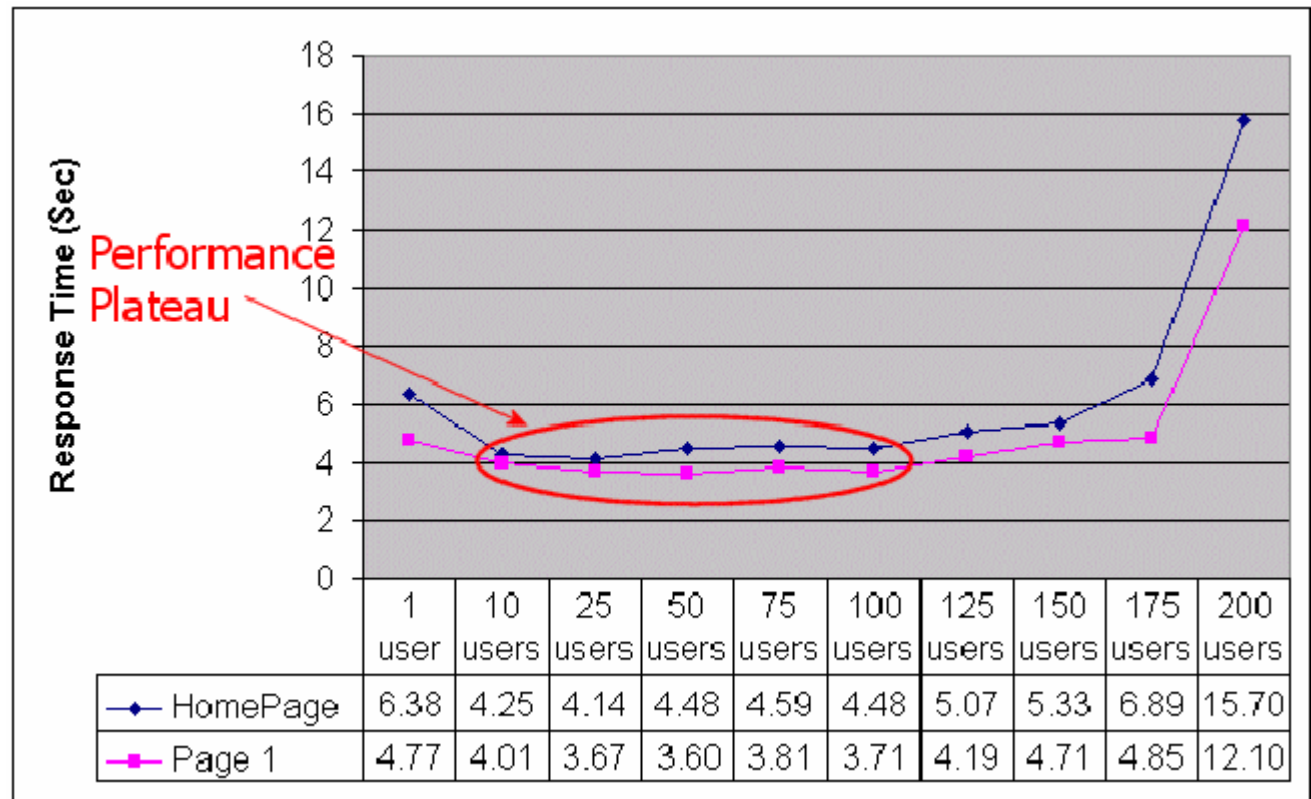
Single User Region:

- Typically slightly slower than best case.
- Think of it like "warming up your car on a cold day."

|	1 user	10 users	25 users	50 users	75 users	100 users	125 users	150 users	175 users	200 users
HomePage	6.38	4.25	4.14	4.48	4.59	4.48	5.07	5.33	6.89	15.70
Page 1	4.77	4.01	3.67	3.60	3.81	3.71	4.19	4.71	4.85	12.10

Response Time Degradation Curve

Performance Plateau:

- The best performance you can expect without further tuning.
- Good candidates for baselines and/or benchmarks.

	1 user	10 users	25 users	50 users	75 users	100 users	125 users	150 users	175 users	200 users
HomePage	6.38	4.25	4.14	4.48	4.59	4.48	5.07	5.33	6.89	15.70
Page 1	4.77	4.01	3.67	3.60	3.81	3.71	4.19	4.71	4.85	12.10

Response Time Degradation Curve

Stress Region:

- Area where the application "degrades gracefully".
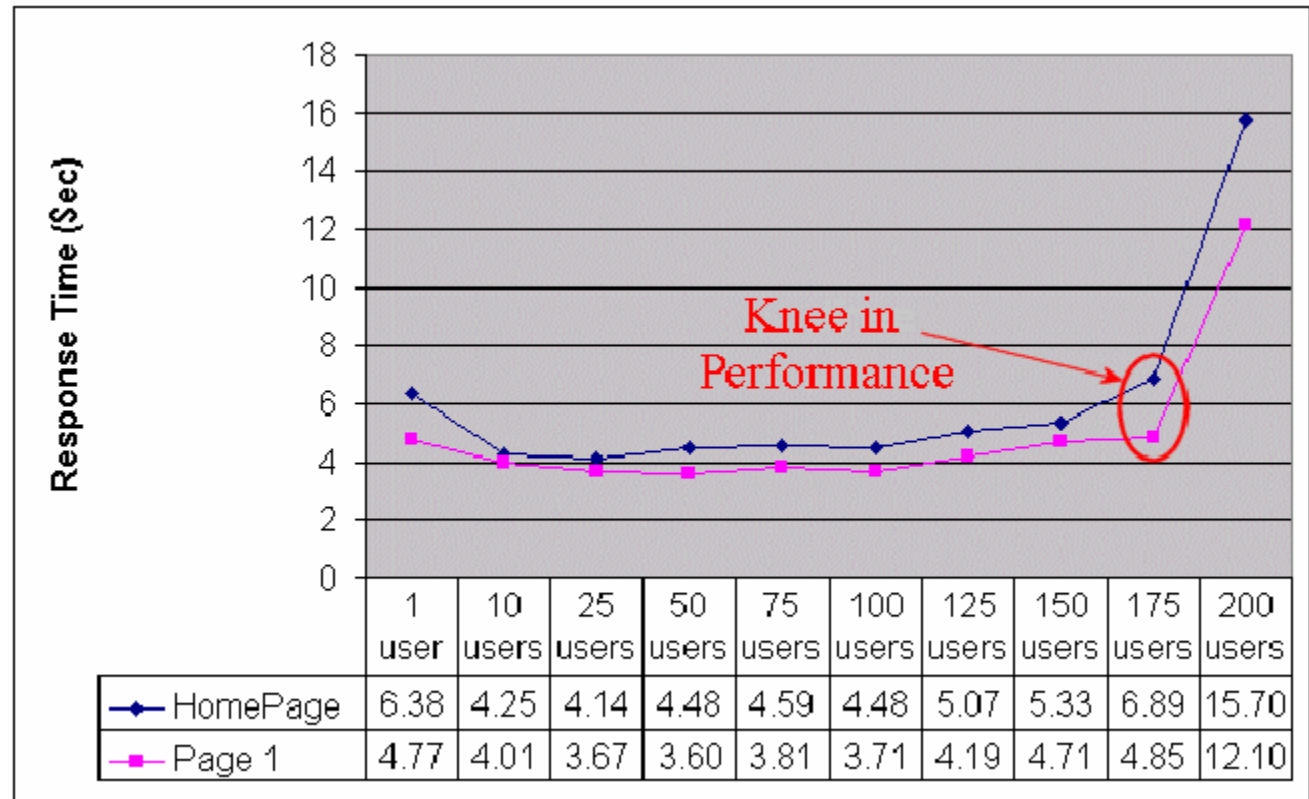- Max recommended user load is the beginning of the stress region.
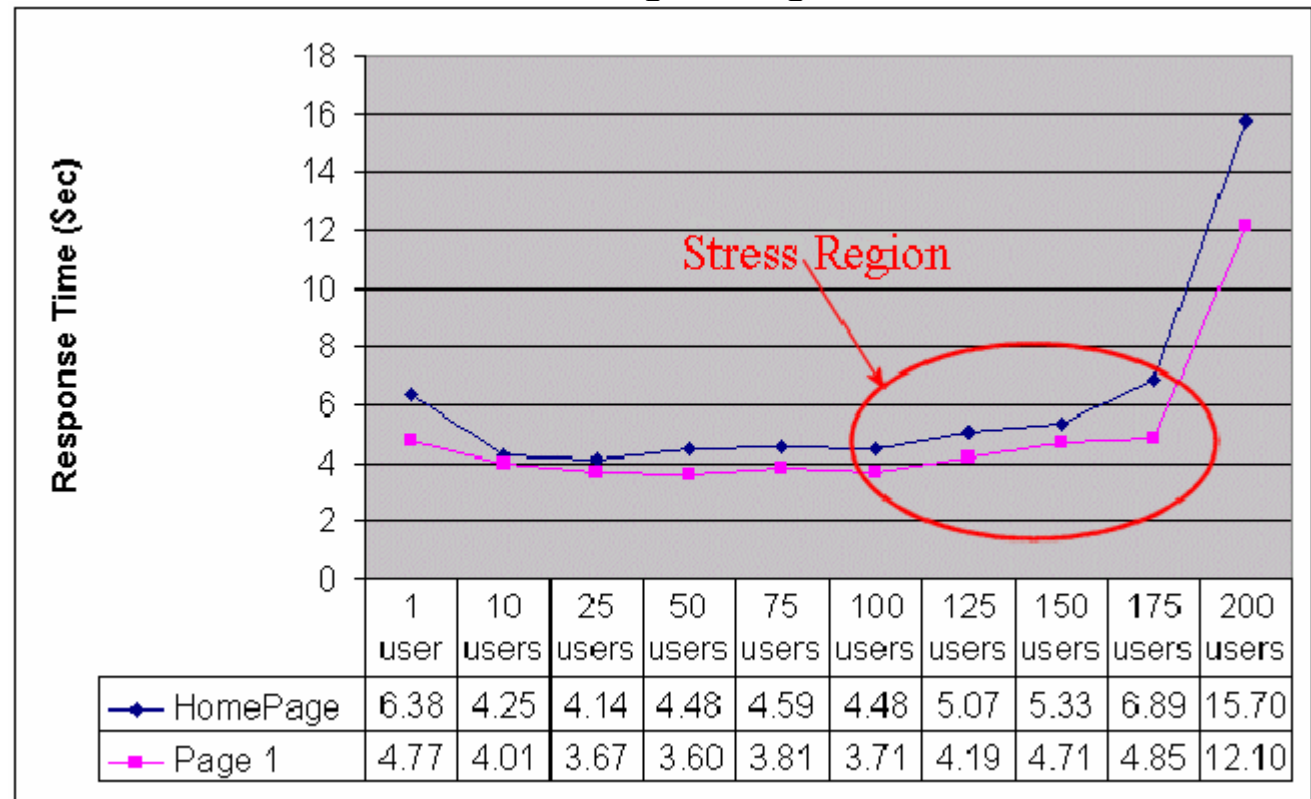
Response Time Degradation Curve

Knee in Performance:

- Point where performance degrades "un-gracefully".
- There is always a knee, it's our job to find it so we can plan around it.

	1 user	10 users	25 users	50 users	75 users	100 users	125 users	150 users	175 users	200 users
HomePage	6.38	4.25	4.14	4.48	4.59	4.48	5.07	5.33	6.89	15.70
Page 1	4.77	4.01	3.67	3.60	3.81	3.71	4.19	4.71	4.85	12.10

Response Time Degradation Curve

Min/Max/Expected Version:

- Comparing performance based on various workload distributions.
- Yields a "confidence interval for predicted performance."
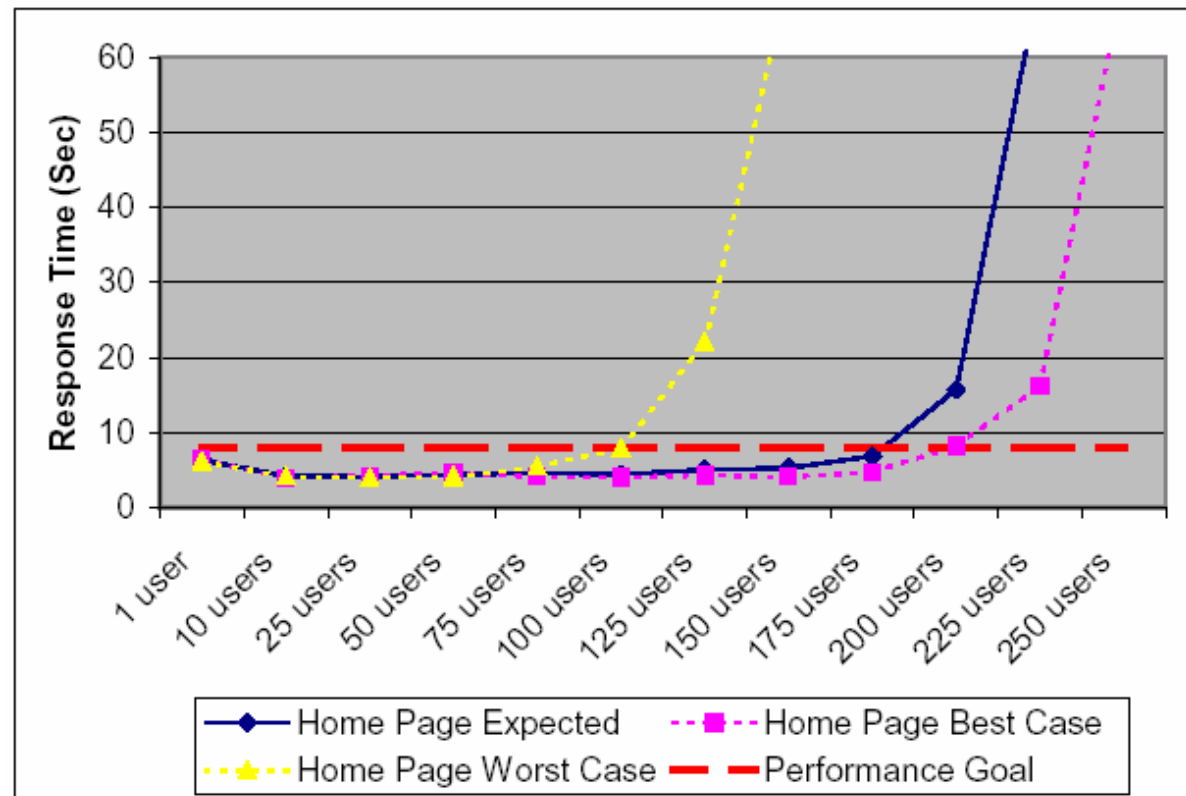
DEMO

Stakeholder Reactions

"These are great, but where's the supporting data?"

- All the data lives in TestManager and/or Excel.
- Include the spreadsheets as appendices with final report.

"Very pretty, but what do they mean?"

- That's why we always have a paragraph that tells what the chart/table means, not what it says.

"Terrific! This is exactly what I wanted! Don't worry about the final report — these will do nicely."

- As tempting as it may be – don't do it. Write the report anyway… trust me.

Comprehensive Final Report Template

Executive Summary!!!

Re-cap Acceptance Criteria (Requirements)

Re-Cap Workloads Actually Used

Summarize Tests Individually

Summarize Tuning Activity (System Modifications)

Summarize Results Collectively

Document Conclusions

Document Recommendations

Want More Information?

Information adapted from User Experience, not Metrics: Parts 6, 8, 9 and10 and Beyond Performance Testing: Part 2 located at http://www.rational.net (RDN) and http://www.PerfTestPlus.com (My site)

Good sources for additional information about Performance Testing:

- http://www.PerfTestPlus.com (Methodology, Templates, Articles, Presentations)
- http://www.loadtester.com (Good articles and links)
- http://www.keynote.com/resources/resource_library.html (Good articles and statistics)

Graphical Presentation of Information – Edward Tufte, PhD. http://www.edwardtufte.com (Books and seminars)

QUESTIONS

Contact Info

Scott Barber

Chief Technology Officer

PerfTestPlus, Inc

E-mail:

sbarber@perftestplus.com

Web Site:

www.PerfTestPlus.com