



Terminal Services Scalability Study

Part 1 – The Effect of CPS 4.0

Microsoft Windows Terminal Services

Citrix Presentation Server 4.0

June 2007

Table of Contents

1	Executive summary	- 3 -
2	Introduction	- 4 -
2.1	Overview	- 4 -
2.2	Purpose	- 4 -
2.2.1	Problem definition	- 4 -
2.2.2	Study Objectives	- 4 -
2.3	Intended audience	- 5 -
3	Benchmark configuration	- 6 -
3.1	Summary of hardware and software setup	- 6 -
Table 3.1.1:	Summary of hardware and software setup	- 6 -
3.2	Hardware and Software configuration details	- 7 -
RDP/ICA Client		- 7 -
Microsoft Windows XP desktop client		- 7 -
Windows 2003 Terminal Server		- 7 -
Windows 2003 Citrix Server		- 8 -
Windows 2003 Terminal Server with 4 GB RAM		- 8 -
Windows 2003 Citrix Server with 4 GB RAM		- 9 -
64 bit Windows 2003 Terminal Server		- 9 -
64-bit Windows 2003 Citrix Server		- 10 -
RTO PinPoint (Used to record scalability measurements including round-trip times)		- 10 -
Figure 3.2.1	Software configuration on Windows Terminal Services	- 11 -
Figure 3.2.2	Software configuration on Citrix Presentation Server	- 12 -
3.3	Application and tools	- 13 -
Figure 1.3.1:	Logon Process Topology	- 14 -
Figure 3.3.2:	Terminal server with users connected via RDP	- 15 -
Figure 3.3.3:	Terminal Server users connected via ICA	- 16 -
Figure 3.3.4:	Recommended Max User Sessions	- 17 -
4	Measurement methodologies and findings	- 18 -
4.1	Scalability Methodology	- 18 -
Table 4.1.1:	Citrix Server Test Kit Overhead	- 18 -
4.2	Findings	- 19 -
Table 4.2.1:	Minimum and Maximum sessions for ICA and RDP protocols	- 19 -
4.3	Other Findings	- 20 -
Table 4.3.1:	ICA and RDP Logon Times	- 20 -
5	Conclusion	- 21 -



1 Executive summary

Performance and scalability are essential for your terminal services farms. However, an understanding of the true scalability of the terminal services or Citrix Presentation Server environments represents the Holy Grail for terminal server architects planning roll-outs in such environments. Unfortunately, most testing environments fail to comprehend many of the effects of the complex interactions in each of these environments, and do not accurately reflect the true scalability. Therefore, we set out to conduct a series of experiments in THREE parts.

This paper constitutes PART 1 of the series and is focused primarily on determining the overhead for terminal services clients running desktops. It compares Remote Desktop (RDP) clients and Citrix Presentation Server (ICA) clients while also taking into account the overhead for resources used by Citrix Presentation Server 4.0 (MPS4.0) once it is installed on a server. It does not include applications either published or run on remote desktops and users clicking through them. *PART 2 will focus on how RTO TScale handles memory optimization to substantially increase scalability for terminal server environments and PART 3 will focus on how RTO TScale optimizations scale virtual desktop infrastructure (VDI) environments.*

The test runs proved very consistent in each test scenario. It is important to remember that the number of sessions achieved in this test is not meant to be a “real-world” example of how many sessions you could expect since each of the sessions is effectively “idle”. The goal of this test is to determine the amount of overhead incurred by CPS 4.0 prior to any user activity.

The following findings were recorded:

1. On the Windows 2003 server running terminal services in application server mode all of the test runs ended at either 205 or 206 user sessions.
2. After installing Citrix on that same machine, but still using the RDP client, that machine’s capacity proved to be in the range of 115 to 119 sessions.
3. Using the ICA client the tests showed an increase of 130 to 132 sessions that were allowed to successfully logon.
4. Adding RAM to the Windows 2003 machine allowed 232-243 sessions while running only Windows terminal services;
5. After installing Citrix with the recommended hardware the system allowed only 187-194 RDP sessions or 211-215 ICA sessions which proved that Citrix installation was using more memory leaving less room for user sessions.



In summary, the results across all Windows 2003 platforms showed that a percentage of user capacity is lost by installing CPS4.0. Some of this capacity was regained by using the ICA client, but not enough to compensate for installing Citrix. The ICA client also took significantly longer (approximately 2 minutes more) for each login event than the RDP client. This relatively long average login time with ICA clients compared to RDP clients contributed to a slower overall user experience.

2 Introduction

2.1 Overview

Performance and scalability are essential for your terminal services farms. While business is growing, increasing capacity on terminal services farms is a key requirement to support more users. This paper documents the technical aspects of a scalability study and demonstrates in great detail the impact of Citrix Presentation Server (CPS) and Microsoft Terminal Services software on scalability.

2.2 Purpose

2.2.1 Problem definition

Current scalability reports on Citrix Presentation Server 4.0 Application Server concentrate on performance characterization of terminal services in single-node environments, and performance analysis of scalability is limited to small sandbox environments. Moreover, scalability studies that give a true picture of available capacity of terminal server environments by identifying the terminal server application overhead are significantly lacking.

2.2.2 Study Objectives

The objectives of this study are two-fold:

1. Determine the true scalability of Citrix Presentation Server 4.0, and Microsoft Windows Terminal Services.
2. Identify the overhead introduced by the terminal server applications itself that impact available capacity.

In order to achieve the objectives above, RTO designed and set up a test environment to conduct a performance study of the scalability characteristics required of a terminal services infrastructure. To get a complete picture of scalability, we used the following parameters:



- Minimum and maximum number of user sessions
- Overhead of the terminal server software

2.3 *Intended audience*

This paper is intended for IT managers, software engineers, performance analysts, application developers, architects, and system administrators who are responsible for and/or interested in deploying or performance tuning of terminal services solutions in production environments.



3 Benchmark configuration

To effectively demonstrate the near- linear scalability of the terminal services solution, it was necessary to assemble the software environment in a way that accurately reflects a typical scenario deployed in terminal server farms.

This section explains how the benchmarking software was deployed and how the hardware was assembled.

3.1 Summary of hardware and software setup

The following table (Table 3.1.1) provides hardware and software configuration on each of the machines used for this study.

Table 3.1.1: Summary of hardware and software setup

Tier	Hardware	Software
Windows 2003 Server (both pre and post Citrix installation) – win2k3ctx	Memory – 1 GB RAM Processor – 1.13 GHz dual	Windows 2003 running Windows Terminal Services Citrix Presentation Server 4.0
Windows 2003 Server (both pre and post Citrix installation) – win2k3	Memory – 4 GB RAM Processor – 1.13 GHz dual	Windows 2003 running Windows Terminal Services Citrix Presentation Server 4.0
64 bit Windows 2003 (both pre and post Citrix installation) – rtoamd2	Memory – 4 GB RAM Processor – 1.79 GHz dual	Windows 2003 Standard x64 Ed. running Windows Terminal Services Citrix Presentation Server 5.0



3.2 Hardware and Software configuration details

RDP/ICA Client

Client machine

In the real world RDP and ICA clients are installed and used on devices from desktop computers, to thin client work stations, to handheld devices. For the purposes of this test a machine was dedicated to repeatedly running clients, successively starting a new client each time the previous client’s successful connect and user initiate. We took this approach since the client hardware has no effect on the server side of the session and hence has no impact on terminal server scalability.

Microsoft Windows XP desktop client

Client Hardware	Operating System	Software
Microsoft Virtual PC 2004 (Build 582) With 512 MB RAM allocated Running on: Dell Dimension DM051 Pentium(R) D CPU 2.80 GHz 2 GB RAM	Microsoft Windows XP Professional Version 2002 Service Pack 2 Windows Components Remote Desktop Connection Shell Version 5.1 (Build Number 2600) Control Version 5.1.2600.2180	Citrix Presentation Server Client Version 9.00.32649 WinBatch Version 2003.8.0.0

Windows 2003 Terminal Server

Server Hardware	Operating System	Software
Dell Power Edge 2550 Dual Pentium III, 1.13 GHz 1 GB RAM	Microsoft Windows Server 2003 Enterprise Edition Service Pack 1 Windows Components (in addition to default settings) Terminal Services WMI Windows Installer Provider	PinPoint Agent – Version 3.1.0.33 WinBatch Version 2003.8.0.0



Windows 2003 Citrix Server

<p>Server Hardware Dell Power Edge 2550 Dual Pentium III, 1.13 GHz 1 GB RAM</p>	<p>Operating System Microsoft Windows Server 2003 Enterprise Edition Service Pack 1 Windows Components (in addition to default settings) ASP.NET Terminal Services WMI Windows Installer Provider</p>	<p>Software Citrix Presentation Server for Windows – Version 4.0.0 Citrix Server Test Kit 2.1 Citrix Web Interface for Citrix Presentation Server – Version 4.0.45083#28412 Management console for Citrix Presentation Server 4.0 Citrix Presentation Server – Administration Snap-in RTO PinPoint Agent – Version 3.1.0.33 WinBatch Version 2003.8.0.0</p>
---	--	---

Windows 2003 Terminal Server with 4 GB RAM

<p>Server Hardware Dell Power Edge 2550 Dual Pentium III, 1.13 GHz 4 GB RAM</p>	<p>Operating System Microsoft Windows Server 2003 Enterprise Edition Service Pack 1 Windows Components (in addition to default settings) Terminal Services WMI Windows Installer Provider</p>	<p>Software PinPoint Agent – Version 3.1.0.33 WinBatch Version 2003.8.0.0</p>
---	---	---



Windows 2003 Citrix Server with 4 GB RAM

<p>Server Hardware Dell Power Edge 2550 Dual Pentium III, 1.13 GHz 4 GB RAM</p>	<p>Operating System Microsoft Windows Server 2003 Enterprise Edition Service Pack 1 Windows Components (in addition to default settings) ASP.NET Terminal Services WMI Windows Installer Provider</p>	<p>Software Citrix Presentation Server for Windows – Version 4.0.0 Citrix Server Test Kit 2.1 Citrix Web Interface for Citrix Presentation Server – Version 4.0.45083#28412 Management console for Citrix Presentation Server 4.0 Citrix Presentation Server – Administration Snap-in RTO PinPoint Agent – Version 3.1.0.33 WinBatch Version 2003.8.0.0</p>
---	--	---

64 bit Windows 2003 Terminal Server

<p>Server Hardware VSP A2210 AMD Opteron™ Processor 244 1.79 GHz 4 GB RAM</p>	<p>Operating System Microsoft Windows Server 2003 Standard x64 Edition Windows Components (in addition to default settings) Terminal Services WMI Windows Installer Provider</p>	<p>Software WinBatch Version 2003.8.0.0</p>
--	---	---



64-bit Windows 2003 Citrix Server

<p>Server Hardware VSP A2210 AMD Opteron™ Processor 244 1.79 GHz 4 GB RAM</p>	<p>Operating System Microsoft Windows Server 2003 Standard x64 Edition Windows Components (in addition to default settings) ASP.NET Terminal Services WMI Windows Installer Provider</p>	<p>Software Citrix Presentation Server for Windows – Version 4.0 Citrix Server Test Kit 2.1 Management console for Citrix Presentation Server 4.0 WinBatch Version 2003.8.0.0</p>
--	--	--

RTO PinPoint (Used to record scalability measurements including round-trip times)

<p>Server Hardware Dell PowerEdge 2550 Dual Pentium III, 1.13 GHz 1 GB RAM</p>	<p>Operating System Microsoft Windows 2000 Version 5.00.2195 Service Pack 4</p>	<p>Software Microsoft .NET Framework 1.1 Microsoft .NET 2.0 RTO PinPoint Analysis Engine – Version 3.1.0.33</p>
--	---	---

Figures 3.2.1 and 3.2.2, in the following pages, provide an overview of how the software components interacted to form the benchmark environment.



Figure 3.2.1 Software configuration on Windows Terminal Services

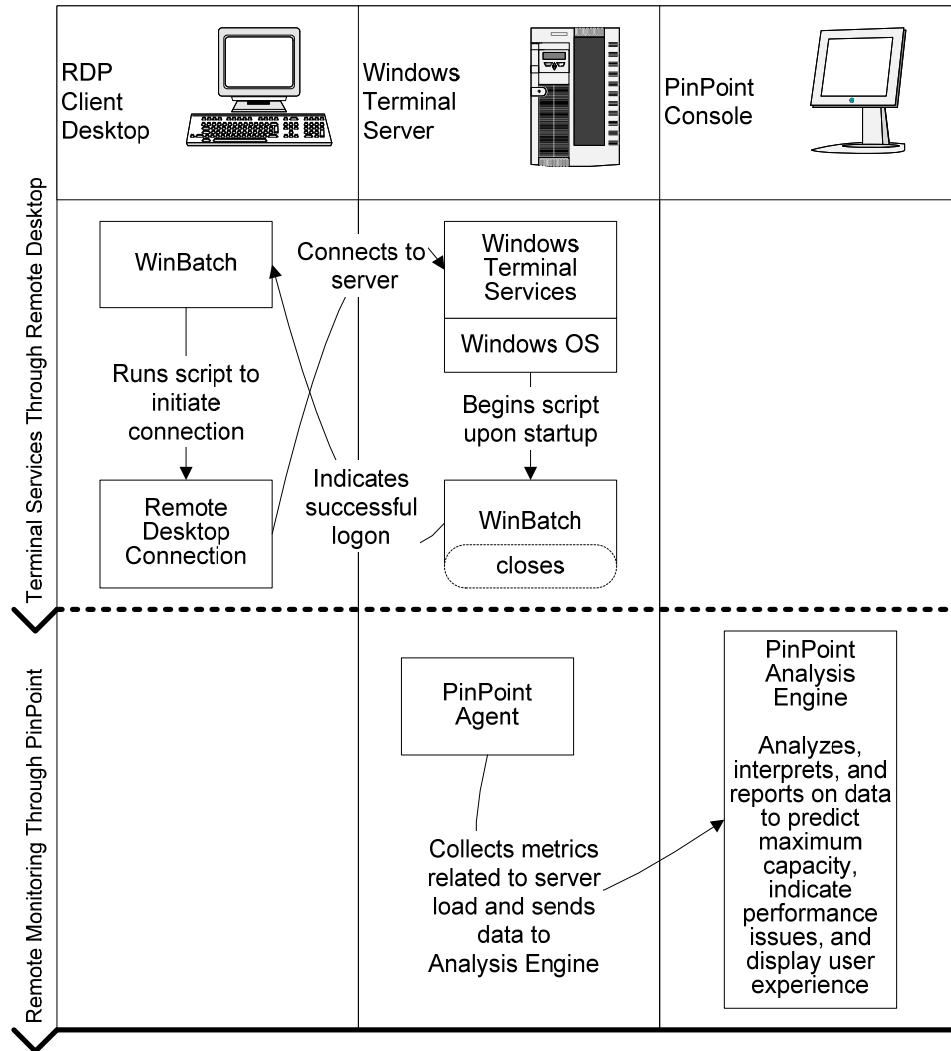
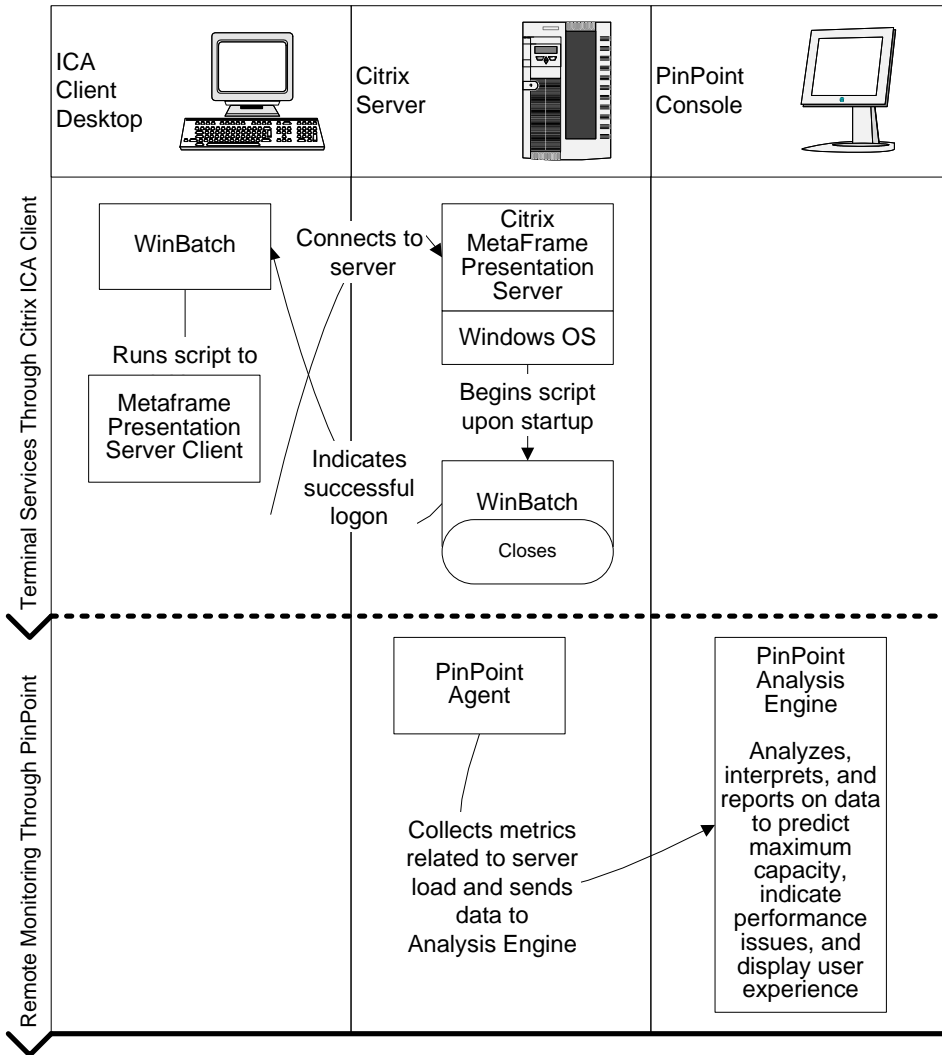


Figure 3.2.2 Software configuration on Citrix Presentation Server



3.3 Application and tools

To demonstrate the true scalability that is required to support the dynamic conditions prevalent in terminal services, the team chose to simulate a logon process using WinBatch scripts. The script runs in two parts – one on the client and the other on the terminal server.

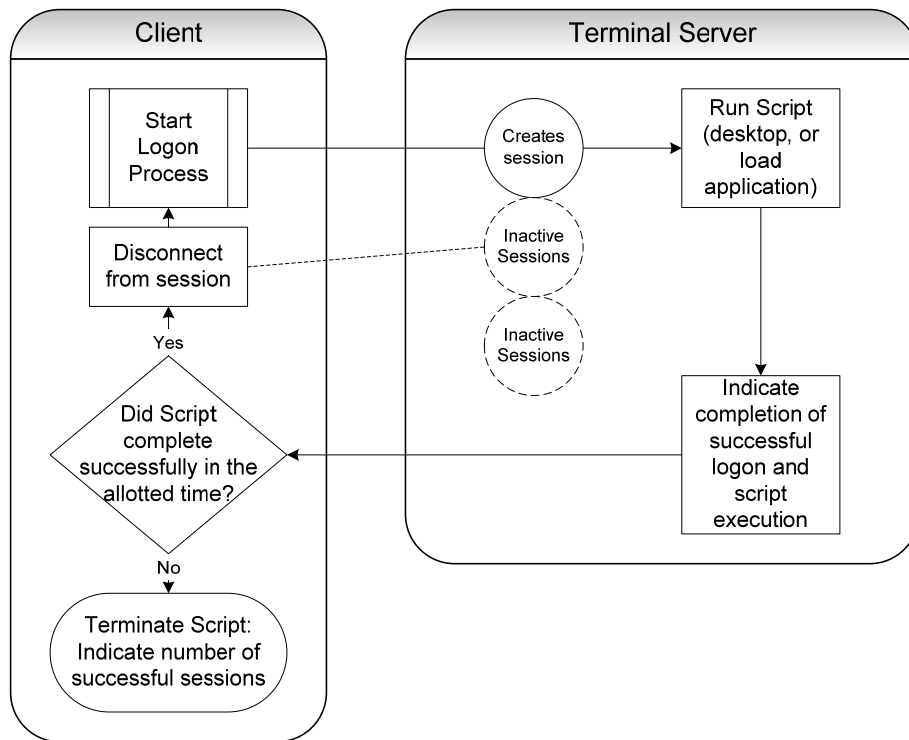
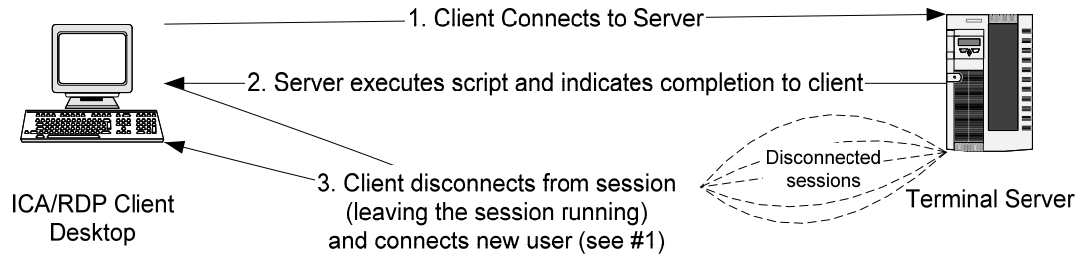
The client half of the script starts a session and then waits for the script on the terminal server to indicate that the user has successfully initiated and run the script. If the script completes before the timeout has been set, the client disconnects from the server, leaving the session running, and initiates a session with the next user. If the script is not successful within the time limit then the script (indicating that the logon was not successful within two minutes) terminates, indicating the number of users that were successfully connected.

The client script also keeps track of how long it takes a user to log in. On the RDP script this count starts the moment the connect button is hit (after logon information such as user name, password and domain have been entered). Since the ICA client script is run from pre-existing custom ICA connections which already include user information, this script begins counting the logon time from the moment the Custom ICA Connection is run. Both scripts count the end of the logon from the time the script on the server creates a file indicating it has run successfully. The logon time is calculated the same way for each client type. The start time and finish time are gathered to the nearest whole second and the start time is subtracted from the finish time to give the logon time to the nearest second.

The script that runs in the user session runs on startup and simply indicates that the logon was successful and then terminates, leaving no script overhead on the terminal server after user initiation.



Figure 1.3.1: Logon Process Topology



RTO PinPoint was used throughout the scalability testing process to monitor resource usage and make predictions concerning the maximum user capacity of each machine. It also highlighted resources that became a bottleneck to adding more users.



Figure 3.3.2: Terminal server with users connected via RDP

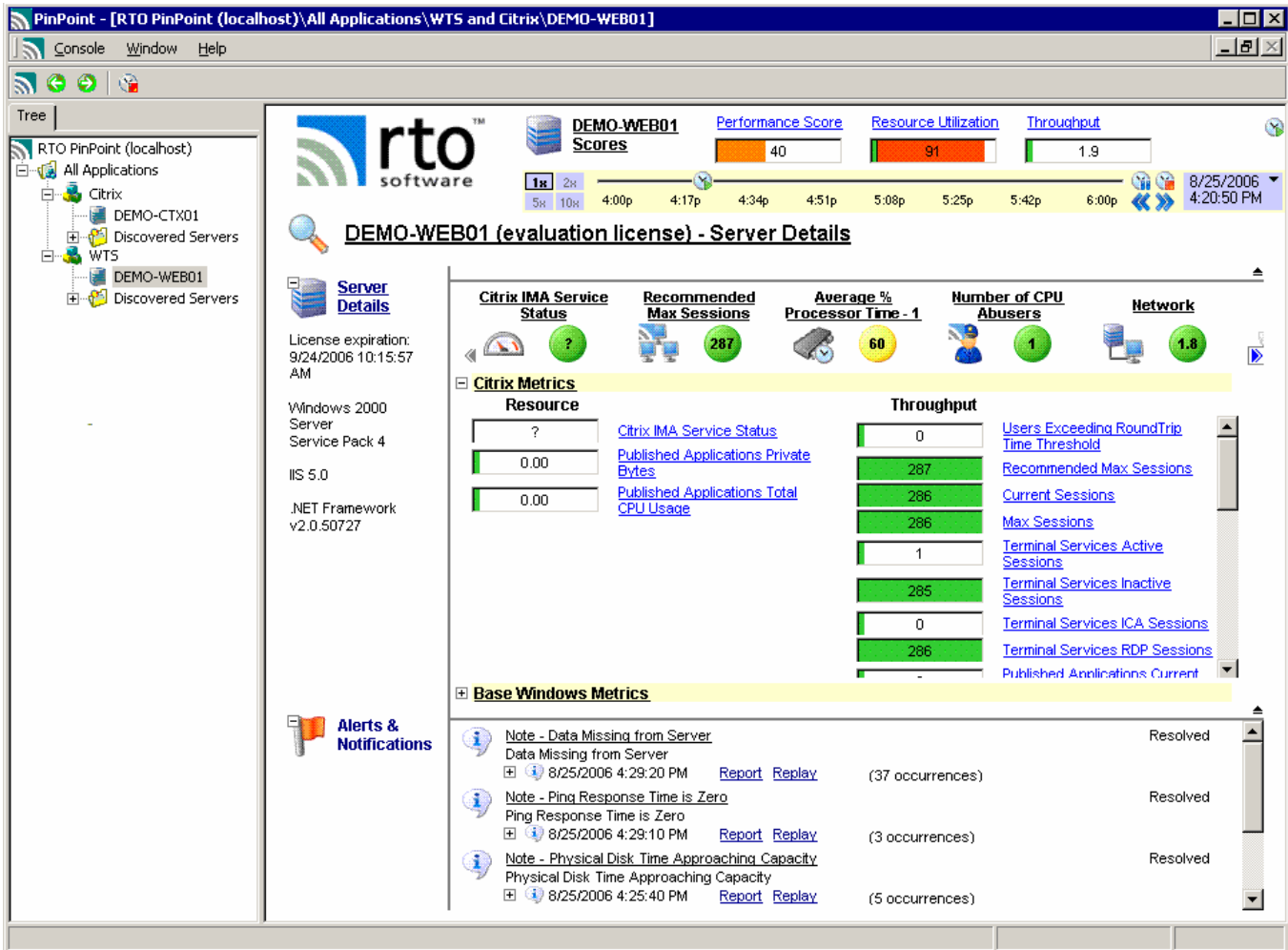
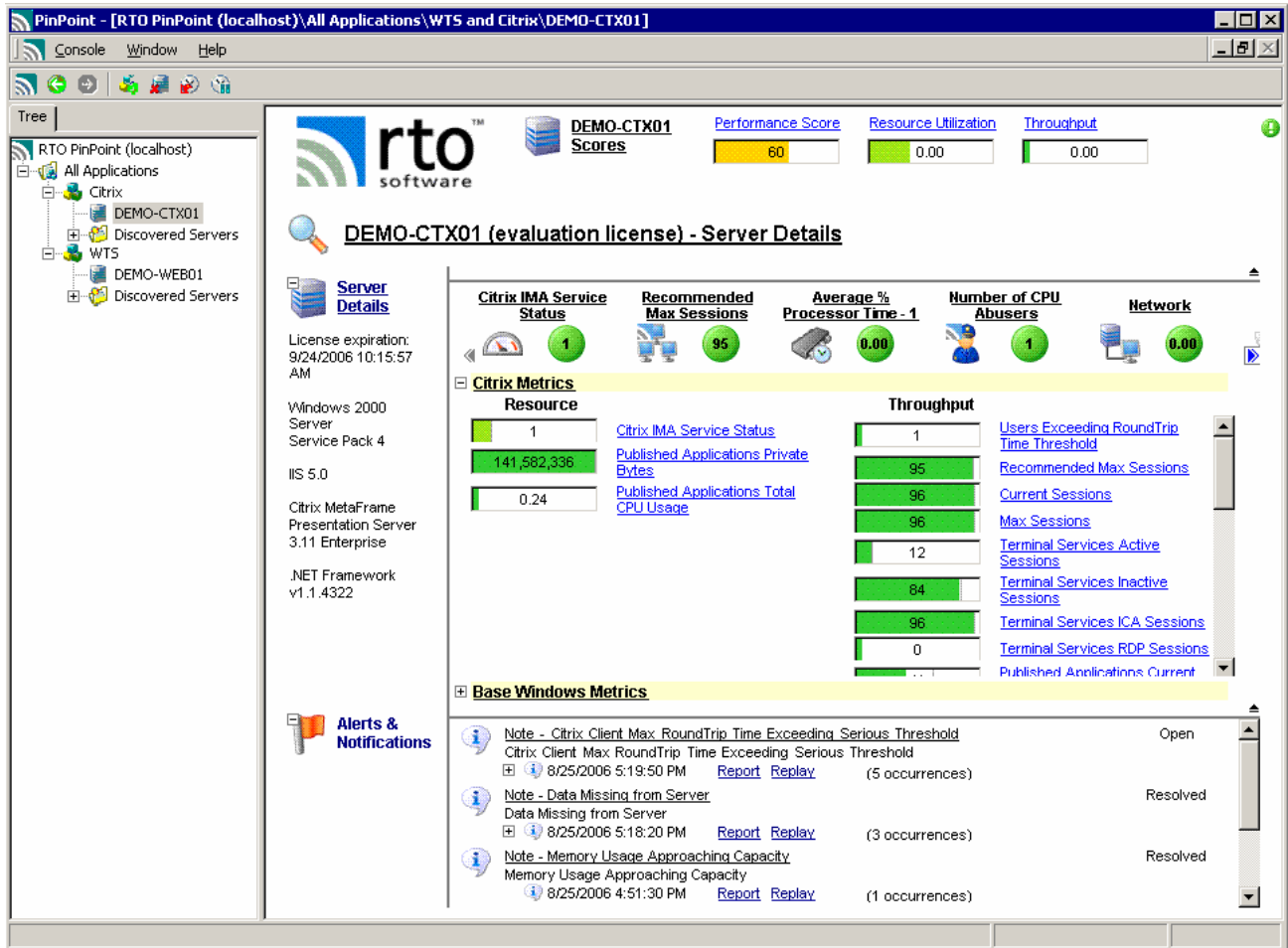


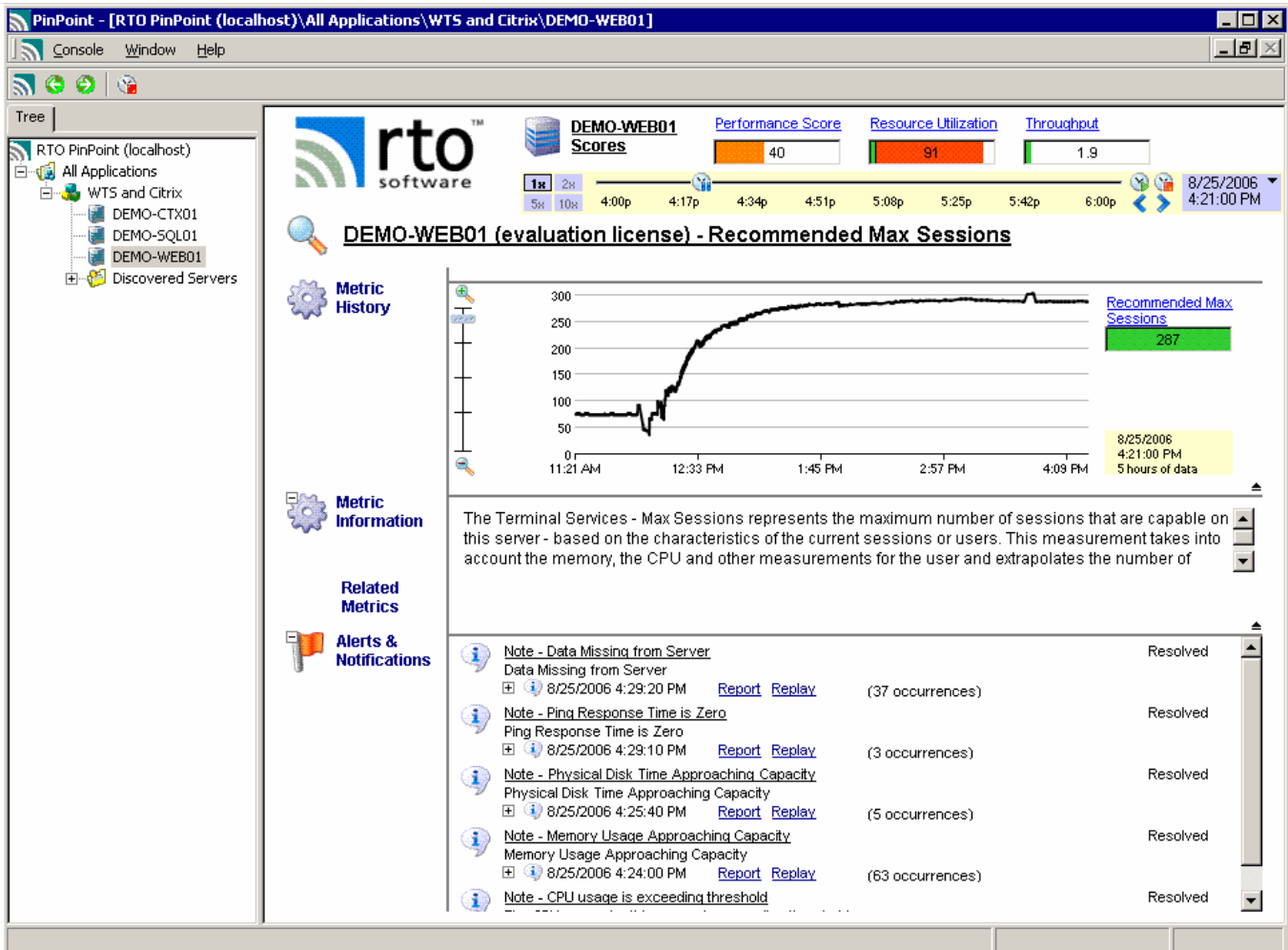
Figure 3.3.3: Terminal Server users connected via ICA



In the test scenario, RTO PinPoint calculated the recommended number of users based on the resources tied up in users already connected and what those users are doing. As more users connect the Recommended Max Sessions becomes more accurate as seen in the graph below (Figure 5) during a scalability test. As more users logged on, the recommended number of users converged to the actual machine capacity.



Figure 3.3.4: Recommended Max User Sessions



4 Measurement methodologies and findings

To accurately measure scalability, the maximum number of connections that could be achieved using the particular connection protocol being tested was measured. Each test was run multiple times to confirm consistency and reliability. A range was established for each test scenario with the minimum being the fewest number of users successfully connected and the maximum being the most users connected under the particular test scenario. All other test runs fell within this range. In very few circumstances, the test was so consistent that all runs ended with the exact same number of users, in these test cases, only the single number is given.

4.1 Scalability Methodology

There is a Citrix Server Test Kit (CSTK) that can be used for scalability and performance testing. However we removed the CSTK from the equation due to overhead on the terminal services server. It required both a single user session to run the CSTK console and a client process to run in each user session.

Table 2 below indicates the approximate Citrix Server Test Kit overhead we identified during the scalability testing.

Table 4.1.1: Citrix Server Test Kit Overhead

	Memory	Virtual Memory
Console (one time only)	5332K	1676K
Client (once per user)	3412K	1096-1988K

The WinBatch script runs only once at the logon of each user and then terminates leaving no overhead in the user session. Therefore the only resources used for each user session are those necessary to maintain the remote desktop or published desktop session.

The WinBatch based scalability tests were first run against *qa2k3ctx*, a Windows 2003 server running Terminal Services without Citrix installed, using only the Remote Desktop client. Then the same test was run on *qa2k3ctx* after installing Citrix 4.0. Finally we ran the same test this time using the ICA client. This process was repeated on *qa2k3* after adding RAM to the 4GB point, the amount recommended by Citrix and then on *rtoamd2*, which has a 64 bit processor.



4.2 Findings

The test cases proved very consistent in each test scenario. The following results were observed after all the test runs.

1. On the Windows 2003 server running terminal services in application server mode all of the test runs ended at either 205 or 206 user sessions.
2. After installing Citrix on that same machine, but still using the RDP client, that machine’s capacity proved to be in the range of 115 to 119 sessions.
3. Using the ICA client the tests showed an increase of 130 to 132 sessions that were allowed to successfully logon.
4. Adding RAM to the Windows 2003 machine allowed 232-243 sessions while running only Windows terminal services.
5. After installing Citrix with the recommended hardware the system allowed only 187-194 RDP sessions or 211-215 ICA sessions which proved that Citrix installation was using more memory leaving less room for user sessions.

Table 4.2.1: Minimum and Maximum sessions for ICA and RDP protocols

Machine	RDP without Citrix	RDP with Citrix	Using the ICA Client
	Min-Max	Min-Max	Min-Max
Windows 2003 Desktop	205 - 206	115 - 119	130 - 132
Windows 2003 Desktop w/ 4GB RAM	232 - 243	187 - 194	211 - 215
64 bit Windows 2003 Desktop	152 - 170	96 - 102	105 - 109



4.3 Other Findings

Another issue encountered while running the tests was overall performance. The timeout set for both ICA and RDP tests was 500 seconds, which means that if a client takes more than 500 seconds from the start of the connection through successfully logging on and running the script, then the terminal server is assumed to have reached its maximum capacity. However if the script on the terminal server indicates that the logon was successful, then the client script will not wait for the 500 second timeout, instead it will go ahead and initiate a connection for the next user without waiting.

The actual login times were gathered for the RDP and ICA clients on the Citrix machine with four gigabytes of RAM and calculated to the nearest second. Login times were only gathered after Citrix was installed, thus only the client changed while the server remained unchanged.

Table 4.3.1: ICA and RDP Logon Times

	RDP login times	ICA login times
Average	23 seconds	2 minutes 22 seconds
Minimum	4 seconds	47 seconds
Maximum	1 minute 4 seconds	8 minutes 19 seconds
Median	23 seconds	1 minute 7 seconds
Mode	8 seconds	56 seconds
Standard Deviation	15 seconds	2 minutes 16 seconds

The ICA client had a longer average logon time than the RDP client by almost two minutes. The ICA client also took longer in all categories, minimum login time, maximum login time, median and mode login time. The Standard deviation of two minutes sixteen seconds indicates that ICA login times were much less consistent than with the RDP client which had a standard deviation of only fifteen seconds.

The RDP based test took less than half the time of the ICA test to finish because earlier RDP clients logged on much more quickly than ICA clients. Even toward the end of the RDP test, the failure was not with logon times but with insufficient memory to run the script. The ICA clients started with a longer logon time, and at the end of the test the failure was an inability to complete the login. The RDP test completed in little more than five hours, consistently. The ICA tests were much less consistent in their completion times. While some tests completed relatively quickly several took more than 14 hours.



5 Conclusion

In summary, the results across all Windows 2003 platforms showed that a significant percentage of user capacity is lost by installing CPS4.0. Some of that may be regained by using the ICA client, but not enough to compensate for installing Citrix. The ICA client also takes significantly longer for each login time than the RDP client. The average login time through ICA averaged almost two minutes longer than the RDP client making for a slower overall user experience.

The results from **Table 4.2.1** on Windows 2003 indicate that though there is a loss of 43% of server capacity by installing Citrix, almost 7% of that is gained back by using the ICA client for a total loss of only 36%.

The results from **Table 4.2.1** on Windows 2003 with 4 GB RAM, the Citrix recommended RAM, signify that though you lose 20% of your server capacity by installing Citrix, you gain almost 10% of that back by using the ICA client for a total loss of almost 11% on recommended hardware.

The results from **Table 4.2.1** on the 64 bit machine running Windows 2003 Standard x64 Edition are similar to those from the 32 bit edition of Windows 2003 with 39% loss due to Citrix overhead, and only 34% when using the ICA client.

All results were rounded to the nearest percent.

This paper concludes PART 1 of the series of experiments. Please [subscribe to the RTO newsletter](#) to stay informed on new products, partner news, tips and our forthcoming technical whitepapers:

- *PART 2 will focus on how RTO TScale handles memory optimization to substantially increase scalability for terminal server environments.*
- *PART 3 will focus on how RTO TScale optimizations scale Virtual Desktop Infrastructure environments.*

For more detailed information on RTO Software please visit <http://www.rtosoft.com>

