



# Scalability Testing with Login VSI v16.2

**White Paper** | Parallels Remote Application Server | 2018

# Table of Contents

- Scalability ..... 3
- Testing the Scalability of Parallels RAS ..... 3
- Configurations for Scalability Testing ..... 3
- Login VSI Workload ..... 3
- Testing Process ..... 4
- Task Worker Workload Findings ..... 4
- Conclusion ..... 7

## Introduction

Parallels® Remote Application Server (RAS) is a comprehensive virtual application and desktop delivery solution that allows your employees to use applications and data from any device. Seamless and easy to deploy, configure, and maintain, Parallels RAS supports the delivery of applications and desktops via Microsoft RDS and major hypervisors.

This document presents an analysis of the scalability testing of Parallels RAS v16.2 using Login VSI.

## Scalability

### Testing the Scalability of Parallels RAS

To validate Parallels RAS configurations, Parallels engineers conducted a series of performance tests. The goal was to analyze the scalability of Parallels RAS sessions running on VMware vSphere virtual machines. As part of this testing, Login VSI was used to generate user connections to RD Session Host servers simulating typical user workloads.

In a typical Parallels RAS deployment, users connect through a Parallels Client application to access remote applications and desktops. Login VSI clients simulate user connections, while RAS Publishing Agents distribute them and set up service connections between end users and RD Session Host servers.

### Configurations for Scalability Testing

For the Parallels RAS scalability testing, a total of two HP DL360 consisting of the following hardware components were used:

CPU	2x Xeon E5-2670 v1, 2.6GHz, 20 MB L3, 115W TDP
RAM	128 GB, 16x 8 GB Micron DDR-4-2100 at 1600MHz
HDD	Western Digital Blue 1 TB SSD

Parallels RAS was deployed on VMware vSphere 6.5 on Windows 2016 Server as follows:

Parallels RAS Component	Total VMs	vCPU in Each VM	RAM in Each VM
RAS Publishing Agent	2	2	4 GB
RAS Secure Client Gateway	2	2	4 GB
High Availability Load Balancing	1	1	2 GB
RD Session Host	6	6	24 GB

### Login VSI Workload

The following Login VSI configuration was used to simulate the workload on Parallels RAS:

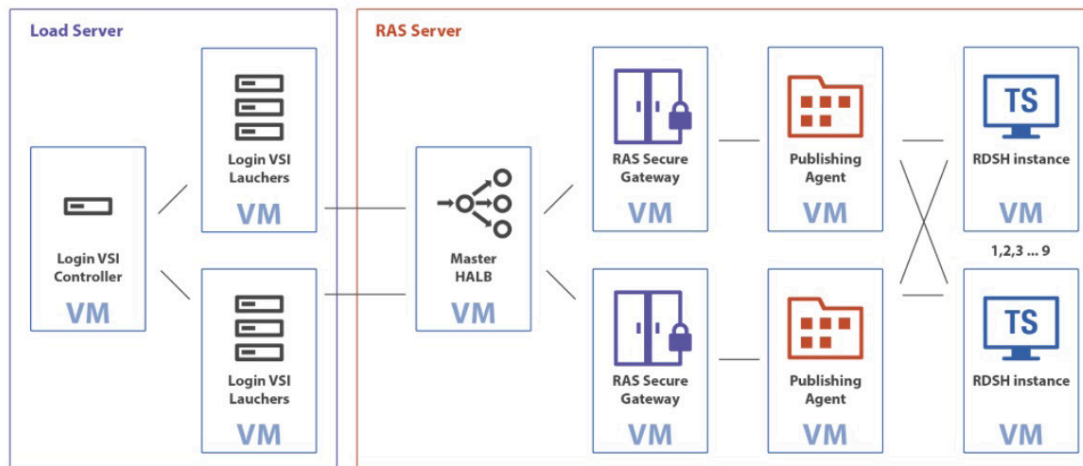
- 250 Task Worker users
- Launch one session every 10 seconds
- Steady state: 2,500 sec

More info about the Login VSI Task Worker workload can be found at the following links:

- [loginvsi.com/documentation/index.php?title=Changes\\_old\\_and\\_new\\_workloads](http://loginvsi.com/documentation/index.php?title=Changes_old_and_new_workloads)
- [loginvsi.com/documents/documentation/Login-VSI-41-Workloads.pdf](http://loginvsi.com/documents/documentation/Login-VSI-41-Workloads.pdf)

All virtual machines comprising the testing environment are siloed on the same virtual network.

Figure 1: Testing Environment Diagram



## Testing Process

In the scalability testing, Login VSI 4.1.32 was used to run a user load on Parallels RAS build 16.2.0–19048. Login VSI helps to gauge the maximum number of users that a desktop environment can support. Login VSI categorizes workloads as Task Worker, Knowledge Worker, Power Worker, and Office Worker.

The Task Worker workload was selected for this testing. The workload includes segments with Microsoft Office 2013 Outlook, Excel, Internet Explorer, Adobe Acrobat, and PDF Writer. While being diverse and not focused on one or two applications, the Task Worker workload does not place a very severe demand on the environment and represents users that do not overload the system with heavy tasks.

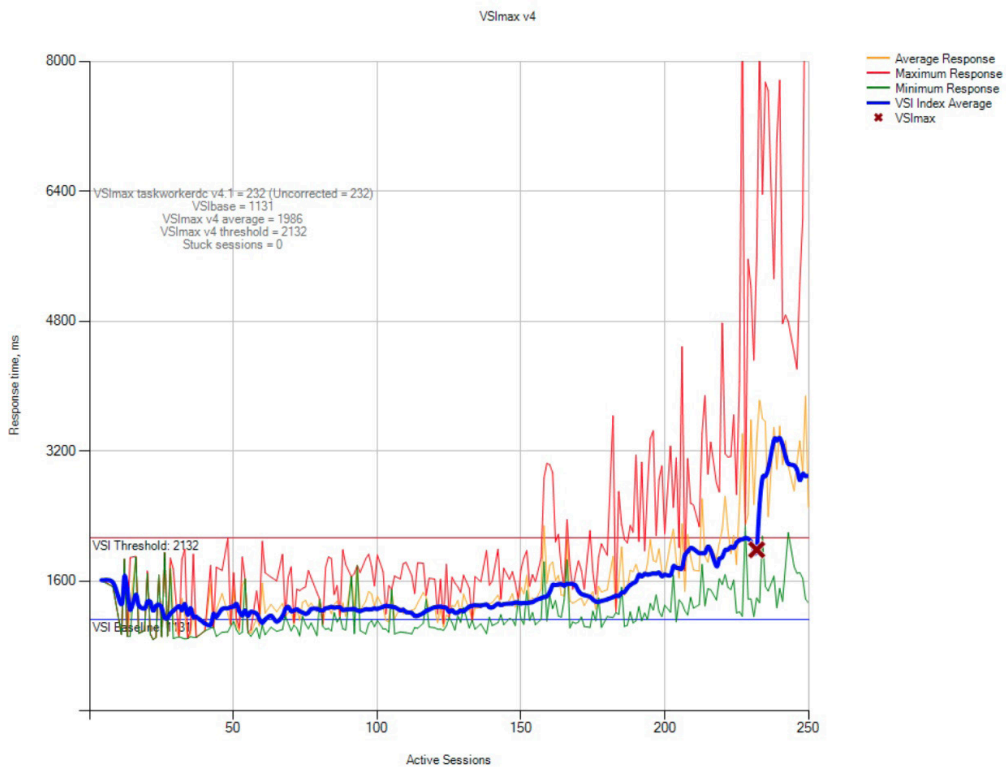
It is important to note that while scalability testing is a key factor in understanding how a platform and an overall solution perform, it should not be inferred as an exact measurement for real-world production workloads. Customers looking to better assess how applications will perform should conduct their own Login VSI scale testing using custom workload scripts. Additionally, such customers could request Parallels RAS POC/Pilot. Since the goal of this testing was to capture a baseline reflecting the densities possible, Login VSI client launchers were configured to go through Secure Gateway in proxy SSL mode.

Performance metrics were captured during user logon and virtual desktop acquisition (ramp-up), user workload execution (steady state), and user logoff. To achieve consistent measurements that would reflect when components were appropriately cached, each workload ran for 48 minutes before Login VSI performance metrics were recorded. VSI tests were repeated three times on each VM instance to get an average number of users who successfully ran the test.

## Task Worker Workload Findings

Following are test results for the Task Worker workload. VSI<sub>max</sub> v4 (which indicates the maximum user density under a specific workload) is determined from the VSI Baseline and VSI Threshold metrics. VSI Baseline represents a pre-test Login VSI baseline response time measurement that is determined before the normal Login VSI sessions are sampled.

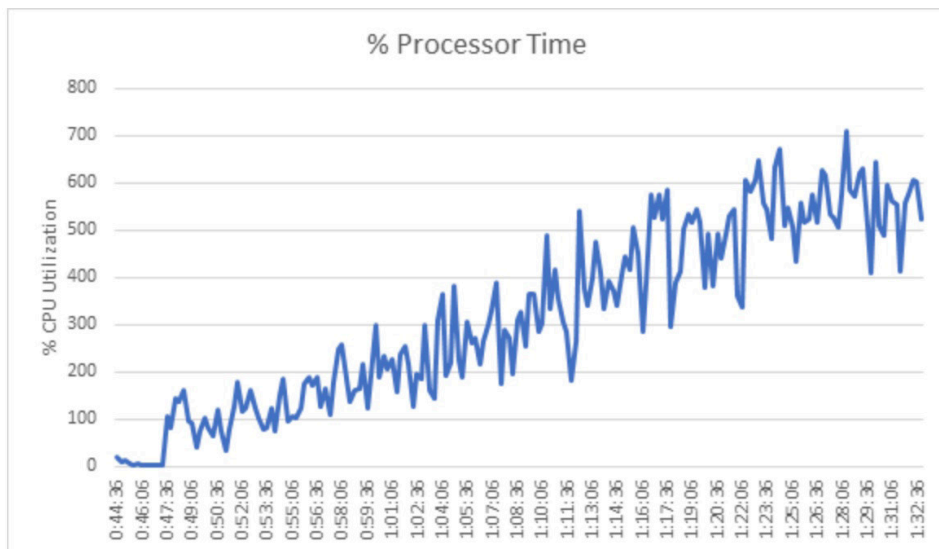
A VSI<sub>max</sub> v4 density of 250 users running the Task Worker workload was demonstrated. In our tests, VSI<sub>max</sub> was reached with 232 sessions. This means that there were already 232 concurrent sessions before any UX degradation was observed based on the current servers specifications.



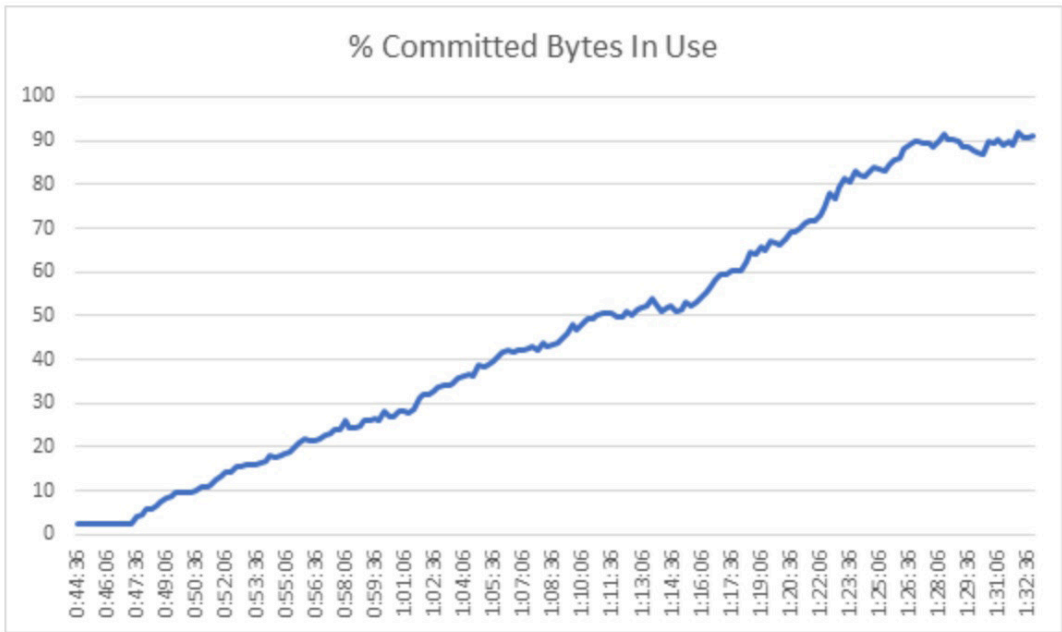
The following test results for CPU and memory consumption and disk I/O response times are helpful in evaluating performance under the test workload. Each chart below shows data collected from a single RD Session Host server. Since there were six RD Session Host servers and 232 simulated users, a single RD Session Host server accommodated about 38 users.

In the following two charts, as user load increases, the CPU and memory usage peak where the number of users approaches VSImax v4.

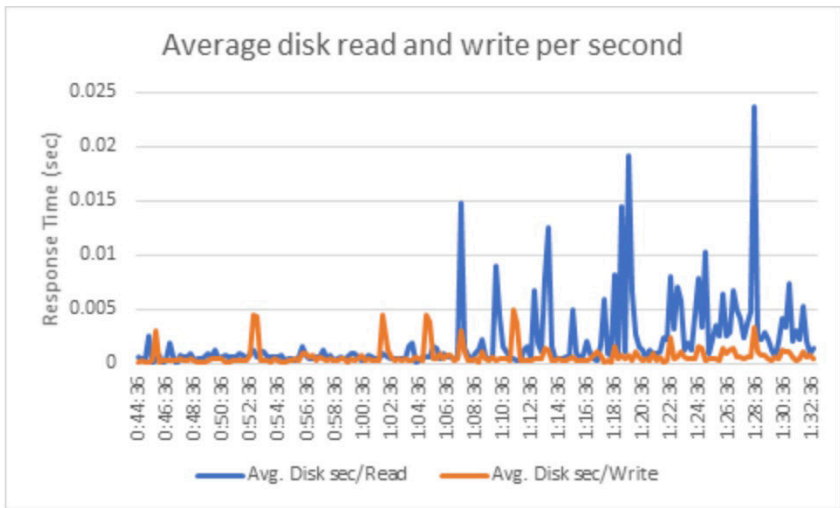
In the following chart, the horizontal axis represents the time (h:mm:ss), while the vertical axis represents the percentage of processor time.



In this chart, % Committed Bytes In Use means the amount of page file used.

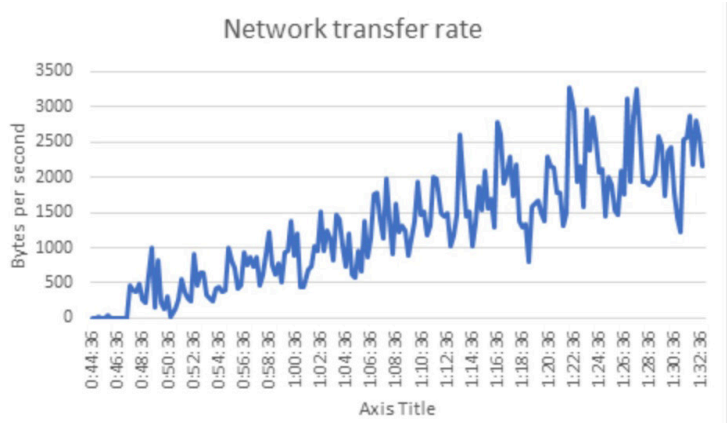


The following chart shows the average disk read and write response time. The write I/O response time averaged about 0.000646 seconds. Read I/O response times averaged to 0.002010 seconds.

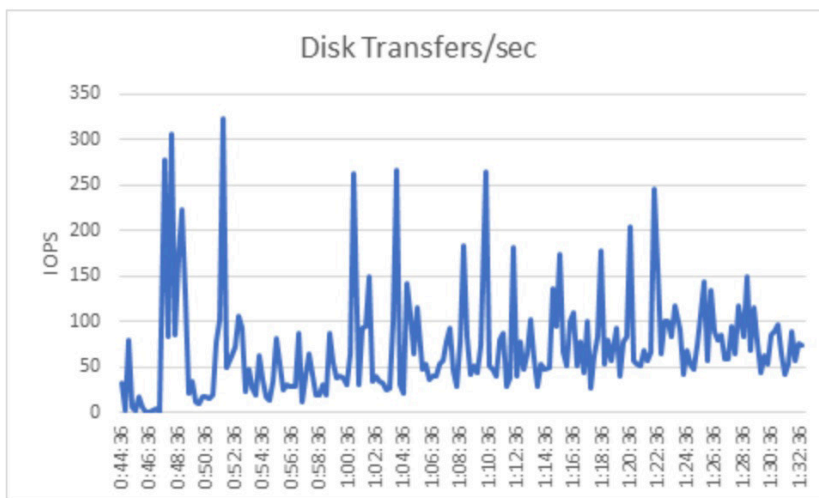


The next two charts show resource consumption for network transfers and disk I/O.

The following chart shows networking transfer rates for data going out. For the Task Worker workload, the average outbound bandwidth at steady state is about 1277 kB/s for our test group of 38 users (232 total users divided by six RD Session Host servers). Therefore, the outgoing transfer rate per user is about 33.62 kB/s.



The chart below shows disk transfer metrics. For the Task Worker workload, disk transfers during steady state averaged about 72 IOPS for the test group of 38 users, or about 1.9 IOPS per user. The peak transfer rate was 323 IOPS for 38 users, 8.5 IOPS per user. User profile data is recorded at logoff, generating disk transfer activity.



### Conclusion

The Parallels RAS scalability results presented in this document confirm that 232 Login VSI sessions using the Task Worker workload can be successfully launched using the given hardware configuration. Specifically, a total of six RD Session Host servers with 6 vCPU and 24 GB of RAM each were sufficient to accommodate these sessions with no user-experience degradation.

Parallels RAS was deployed on VMware vSphere 6.5 on Windows 2016 Server as follows:

Parallels RAS Component	Total VMs	vCPU in Each VM	RAM in Each VM
RAS Publishing Agent	2	2	4 GB
RAS Secure Client Gateway	2	2	4 GB
High Availability Load Balancing	1	1	2 GB
RD Session Host	6	6	24 GB

The results presented here should be used only as guidelines when configuring your Parallels RAS installation. Before making final sizing and deployment decisions, it is suggested that you run proof-of-concept tests using your own workloads.