**Document: Stratusphere SpotCheck #'s for VMware VDI**
**Last Updated:** 08/29/2019
**Document Purpose:** Define metrics and thresholds for a SpotCheck as it relates to User Experience in a VMware VDI environment utilizing Liquidware Stratusphere UX.

This document is designed to bring together the recommendations from many experts in the industry about the metrics that need to be monitored and the thresholds that are deemed acceptable as it relates to user experience. This document does not make recommendations on changes needed due to the many industry, usage, costing, and application variables that are in play.

**What is a "SpotCheck"**

A SpotChecks is basically a point in time health check that focus on key user experience metrics with known levels of acceptable performance. Review of data from multiple dates and times is critical before making recommendations or changes in the environment. The thresholds represented below are at a 1 hour level of granularity unless otherwise specified in the description and are key areas that affect user experience. Normal and High Usage dates and times should be examined based on the industry and user requirements.

**CRITICAL NOTES:**

**A) Know your company!**

Knowledge of the industry/company/department work habits, loads and applications are critical for data interpretation and threshold evaluation.

Example 1: Moderate/High Storage Latency may be acceptable during shift changes with lots of people logging in and out. BUT, Not Acceptable during work hours as this impedes productivity.

Example 2: Law firms and Hospitals generally demand sub 10 second login times whereas most other industries are satisfied with under 30 second.

**B) Know your data!**

There are many monitoring and diagnostic systems on the market. Each of the solutions available collect data in different ways and with different levels of granularity. They all also render/report the data in different ways and granularity with roll ups that can vastly change the data and perspective for the viewer. For this reason, the metric values represented in this document are only for Liquidware Stratusphere and may not apply well to other products.

Examples:

Depending on the view, you could be looking at averages, peaks or peak averages.
Did the data come from the Broker, Hypervisor, "In-Guest", "In Band" or "Out of Band"?
How much impact did the "In-Guest" Agent put on the OS?
How much impact and time lag is on the "Out of Band" Agent, Broker and Hypervisor?

**C) Good Blogs**

SpotCheck Methodology
Grey Matter is Required – Automated Solutions don't work
Monitoring vs. Diagnostics

**D) Liquidware Community Site**

Answers to common and advanced questions.
**Liquidware Community Site Link - community.liquidware.com**

**Documentation needed for analysis, conclusion and recommendations:**

**Multiple Spot Check Dates:**

MM/DD/YYYY (Monday), MM/DD/YYYY (Wednesday), MM/DD/YYYY (Friday)

**Multiple Time Frames Examined on each date:**

(Time frames for review are based on business requirements)

9-10AM, 10-11AM, 2-3PM, 4-5PM

The system(s) should be examined on multiple dates and times for the following information based on the max values shown below. **Don't make a change based on a single data point.**

**Critical Sections for Review:**

**1. ESX Host Criteria: (VMware Best Practices)**

CPU % - (Max 75% Average over 60 minutes)
CPU % Ready (Max 3% Average over 60 minutes)
Memory Utilization (Max 85% at any point in time)
Memory Swapping and Ballooning (Should always be ZERO)
Datastore Latency (Should be under 10 Milliseconds – Max 15 Milliseconds)

**2. Machines/OS Criteria:**

Machine Last Boot – Critical Question – How long has the machine been running?
- See Appendix "R" for more details.

Login Delay (Industry Average is under 30 Seconds – This is a company preference)
- See Appendix "L" for more details.

Application Load Time (Industry Average is under 3 Seconds – Company Preference)
CPU Utilization (Max 80%) – Higher than 50% generally is bad over 60 Minutes
- This generally denotes a stuck or run-away process on the machine.

CPU Queue (Should not be more than 2 per vCPU assigned to VM)
https://technet.microsoft.com/en-us/library/Cc940375.aspx

Memory Usage (Should be less than 75%)
Best Practice – to reduce Windows Paging

Page File Usage (Should be as close to zero as possible)
You cannot stop Windows from paging.
Soft Page Faults are in memory – Hard Page Faults are to disk.
Do Not turn off the paging file in windows. Set to Minimum Size.
Do NOT use "System Managed" – Set the page file start size to ¼ the memory.
Windows Paging causes extra CPU and Disk Overhead and should be reduced whenever possible.

Disk Queue (Should be ZERO for 99% of Users)
Disk Queue shows that the OS is waiting on disk reads and writes.
This can be caused by virus protection holding up the IO or latency of the disk sub system.

Graphics Intensity will be noted has high when over 100 for more than 1/3 of users.
- This must be examined to see if graphics off load processor would help.
- See Appendix "G" for more details.

Applications Non-Responsive – (1 per Day/Per Machine/App is OK)
Any more than this and you need to investigate the apps and services needed by the app.

## Appendix A: PCoIP (Remote Display Protocol)

### Image Quality:

This is a good "Base" metric to gage/monitor the user connection quality. PCoIP will lower the image quality if it has packet loss, high latency or a low bandwidth connection to the end user.

- VMware View Version 5 Default Image Quality is 90%
- VMware View Version 6 & 7 Default Image Quality is 80%

Adjusting image quality higher than the default is only needed for clients that may be viewing broken bones and need to see very small fractures.

Image quality directly affects the number of frames per second sent from the VM to the end client. This can dramatically affect the ESX Host CPU and network bandwidth required per user.

### Session Latency:

- General Max Observations:
  - NY to California – 30-50 Milliseconds
  - USA to India – 150-200 Milliseconds
  - Inter Office Same City – 10 Milliseconds
  - Inter Office Same Building – 5 Milliseconds
- HUGE Latency #s in Stratusphere shows users dropping on and off the network. (Huge Denotes 800+ ms)

### Protocol: (Good and Bad… This is Just info for you)

PCoIP is UDP based.

UDP Packets have a lower priority than TCP Packets on most networks.

UDP is dynamic and bursty by nature of the protocol.

UDP is faster than TCP because there is no error-checking for packets

UDP is lightweight. There is no ordering of messages, no tracking connections, etc.

UPD can do error checking if turned on but there is no recovery. Packets must be resent and with no ordering it is up to PCoIP to request large blocks for retransition.

### Packet Loss:

Packet Loss with PCoIP is bad and can cause users poor experiences: Mouse Drag, Artifacts on the screen, slow screen draw.

### General Recommendations:

1. QOS (Quality of Services) should be implemented on all routers.
   - PCoIP should be right under Voice Over IP and Video.
2. Lower the Maximum Image Quality using the PCoIP GPO settings that best fit your use business and applications.
3. USB and Audio Channels:
   Disable and lower the priority of these channels as it suites your business requirements. Disabling USB or lowering Audio quality can dramatically lower VM/Host CPU and Network requirements.
4. There are many options for PCoIP Tuning. Play with them all, consult the best practice guides from VMware and tune/monitor your users for best user experience in your environment.

## Appendix B: Blast (Remote Display Protocol)

**Session Latency:**

- General Max Observations:
  NY to California – 30-50 Milliseconds
  USA to India – 150-200 Milliseconds
  Inter Office Same City – 10 Milliseconds
  Inter Office Same Building – 5 Milliseconds
- HUGE Latency #s in Stratusphere shows users dropping on and off the network.
  (Huge Denotes 800+ ms)

**Protocol: (Good and Bad… This is Just info for you)**

Blast uses TCP and UDP depending on the type data and connection quality.
TCP packets are numbered and have error checking built in.
UDP Packets have a lower priority than TCP Packets on most networks.
UDP is dynamic and bursty by nature of the protocol.
UDP is faster than TCP because there is no error-checking for packets
UDP is lightweight. There is no ordering of messages, no tracking connections, etc.
UPD can do error checking if turned on but there is no recovery. Packets must be resent and with no ordering it is up to Blast to request large blocks for retransition.

**Packet Loss:**

Packet Loss with Blast is bad and can cause users poor experiences: Mouse Drag, Artifacts on the screen, slow screen draw.

**General Recommendations:**

1. QOS (Quality of Services) should be implemented on all routers.
   - Blast should be right under Voice Over IP and Video.
2. USB and Audio Channels:
   Disable and lower the priority of these channels as it suites your business requirements. Disabling USB or lowering Audio quality can dramatically lower VM/Host CPU and Network requirements.
3. There are many options for Blast Tuning. Play with them all, consult the best practice guides from VMware and tune/monitor your users for best user experience in your environment.

**Good Article on Blast Extreme Protocol:**

Blast Extreme Protocol

**Appendix L: Login Delay**

Time consumed with users logging into a machine is a large part of the user experience. Stratusphere can breakdown the machine boot and login processes. Due to the complexity of active directory and the environments we can only offer a few guiding hints in this document. For a complete login breakdown session please engage your Liquidware SE or partner.

1. Domain Controller(DC) Discovery Time
    a. DC Discovery happens at boot and login time.
        i. Healthy response times are 300-500 milliseconds.
        ii. Changing of the DC during boot and login shows a potential issue.
    b. DC Discovery Times over 500ms:
        i. DC Overloaded – Cannot process request fast enough.
        ii. Network latency from the machines to the DC.
        iii. Sites and Services – Machine/User is talking to a DC in another location.
2. Long running processes
    a. AD GPOs, Item Level Targeting and Scripts
        i. Need to review these in Stratusphere Login Breakdown.
            1. AD Lookups and Local machine WMI Queries are very slow.
            2. Mapping a drive/printer to a machine that does not exist or the user does not have access to can cost a lot of time.
    b. Antivirus Scanning
        i. Don't forget that batch files, PowerShell, VB Scripts are all interpreted languages. Meaning that each line in the batch file or script is executed one line at a time. AV systems scan each line then all the previous lines of the script to ensure it is not a virus.
        ii. Physical Desktops and Persistent virtual machines must be treated differently than non-persistent virtual desktops.
3. Broken and/or Corrupt GPOs
    a. Yearly (At Minimum) reviews of the GPOs should be done by all companies.
    b. Example: IE7 GPOs should not be applied to Windows 10!!
    c. Review of the GPOs can help dramatically with user login times and security.
4. Sites and Services
    a. This is one of the top issues found at client sites with Stratusphere login break down.
    b. If the machine is in New York it should not be authenticating from a domain controller in Canada.

**Good Videos:**

Boot and Login Breakdown

**Appendix G: Graphics Intensity**

Graphics rendering is a large part of the user experience. Depending on the application it can use MS GDI, DirectX, OpenGL, or many other video interface drivers/protocols.

You may be thinking that you don't run any graphic intensive apps. But this is not true. The Windows OS and normal MS Office applications have a lot of graphics requirements. Think about any desktop/laptop built in the last 10 years. They all have a (GPU) graphic processing unit. These processors are used by the OS and applications to offload drawing of boxes, circles and other complex shapes from the main CPU and rendering them on the monitor.

**Definition:** GPU – Graphics Processing Unit

**Non vGPU Virtual Machines:**

Turn off hardware acceleration for all applications! Even though you don't have a vGPU in the host VMware tools still has a driver that looks like a GPU to the OS and the Apps.

Applications that have the option to disable "Hardware Graphics Acceleration" should be done unless you have a GPU installed in the host. Most modern application have a GPO that can turn this off. Note that this is generally a per user-based GPO. MS Office, Google Chrome and Firefox all have GPO settings to turn off hardware acceleration.

Note: These simple application changes can result in a 10% CPU reduction on your host operating system. Your results will vary based on the OS, Application and Host. – You can monitor this with Stratusphere.

**vGPU Virtual Machines:** – (Machines that have access to a vGPU in the HOST)

vGPUs are expensive and sometimes difficult to determine if you are getting the most out of them. The resources are allocated per machine and most settings are around the RAM allocation. Using Stratusphere data you can determine if the machine and even the app is using the GPU memory that is assigned to it. Example: Allocated 2,048MB or vGPU RAM but only using 512MB with a burst to 768MB. Lowering the Allocation to 1,024MB will allow you to let more machines access the vGPU.

**Good Videos:**

Machine and Application GPU Usage

**Appendix R: Machine Last Reboot Time**

How long a machine has been running is a critical question. Applications can have memory, graphics and CPU process "Leaks" over time that will degrade the OS performance. Machines running longer than one month are missing the installation of critical security/feature patches that put them at risk.

Below is a recommendation only of reboot policies based on experience of Liquidware engineers. This is not a Liquidware Labs Inc. recommendation and to my knowledge there are no official recommendations from Microsoft on this topic.

Note: The below recommendations also must conform to company business practices and change control policies.

***Domain Controllers:***

Monthly Reboot – Primarily for OS Security Patches

***Critical infrastructure machines running Windows Server OS:***

Monthly Reboot – Primarily for OS Security Patches

***User Virtual Machines:***

Weekly Reboot – Your mileage will vary based on the applications being used by the users.
Minimum of a Monthly Reboot – Due to OS Security Patching.