VMware vCenter™ AppSpeed Evaluator’s Guide
Managing Application Performance and Service Levels Using AppSpeed
## Table of Contents

1. **Getting Started** ................................................................. 3  
   1.1 About This Guide ............................................................. 3  
   1.2 How to Use this Guide .................................................... 3  
   1.3 Help and Support during the Evaluation ............................. 3  

2. **What is VMware vCenter AppSpeed?** ................................. 4  

3. **System Requirements** ...................................................... 5  
   3.1 Hardware Requirements .................................................. 5  
   3.2 Software Requirements .................................................. 6  
   3.3 Software Installation and Configurations ............................ 6  
   3.4 AppSpeed Installation Requirements ................................. 7  
   3.5 AppSpeed Deployment Considerations ............................... 8  
   3.6 AppSpeed Installation and Configuration ............................ 9  
   3.7 AppSpeed Application Mapping ..................................... 10  

4. **VMware vCenter AppSpeed Use Cases** ............................... 14  
   4.1 Application Discovery and Mapping ................................ 14  
   4.2 Application Performance Monitoring ............................... 17  
   4.3 Application Drilldown ..................................................... 26  
   4.4 Service Level Management and Notifications .................... 30  

5. **Next Steps** ..................................................................... 36  
   5.1 VMware Contact Information ......................................... 36  
   5.2 Providing Feedback ........................................................ 36  

6. **Glossary** ........................................................................ 37  

7. **VMware vCenter AppSpeed Evaluation Worksheet** ................ 38
1. Getting Started

1.1 About This Guide
The purpose of this document is to support a self-guided, hands-on evaluation of VMware vCenter™ AppSpeed. This document is intended to provide IT professionals with the necessary information to manage application performance and service levels using VMware vCenter AppSpeed in a VMware vSphere™ 4 environment.

The content includes a product overview, configuration options, and key-use cases to demonstrate how AppSpeed can help you manage performance and service levels of your applications running within virtual machines.

1.2 How to Use this Guide
This guide will walk you through some key-use cases for AppSpeed to help you conduct a successful product evaluation. A worksheet provided at the end of this guide will help you track your progress.

A glossary is also available to help explain unfamiliar terms.

1.3 Help and Support during the Evaluation
This guide is not intended to be a substitute for product documentation. For detailed information regarding installation, configuration, administration, and usage of VMware products, please refer to the online documentation. You may also consult the online Knowledge Base if you have any additional questions. Should you require further assistance, please contact a VMware sales representative or channel partner.

Below are some links to online resources, documentation, and self-help tools:

VMware vSphere and VMware vCenter Server Resources:
- Product Overview
- Product Documentation
- VMware vSphere Documentation (including hardware compatibility guides)
- White Papers and Technical Papers
- VMware vSphere Evaluator’s Guide

AppSpeed Resources:
- Product Overview
- Product Documentation
- Installation Guide
- User’s Guide
- AppSpeed Community
- Support Knowledge Base
2. What is VMware vCenter AppSpeed?

VMware vCenter AppSpeed enables VMware vSphere administrators to gain deep insight into the performance of their virtualized, multi-tier applications and ensure application service levels. VMware vSphere administrators traditionally lack the ability to determine the performance of applications running in VMware environments. They are often the first to be blamed for performance issues unless they are able to show that the issues are not caused by the virtualization platform. AppSpeed is designed to measure the latency experienced by application end users and correlate it to the different tiers of the underlying physical and virtual infrastructure. This ensures SLAs for quality of service are met, and reduces perceived risks of migrating enterprise applications to a virtual environment.

AppSpeed analyzes the traffic that flows between end users, Web applications, and back-end servers. This analysis provides visibility into multitier applications and enables rapid identification of performance issues originating from inadequate resource allocation and application problems.

AppSpeed will enable IT administrators to:

• Manage application performance and uptime
• Dramatically improve problem isolation times in vSphere environments
• Improve consolidation ratios by ensuring that virtual machines are not given more resources than needed to meet SLAs

AppSpeed is installed on a VMware ESX™ server that resides within the datacenter network and integrates into the VMware vSphere Client. AppSpeed Server is installed on a centralized server, with one or more AppSpeed probes installed on each host that is monitored.

Figure 2.0 a: VMware vCenter AppSpeed Architecture Overview
3. System Requirements

3.1 Hardware Requirements

Note: AppSpeed can also be configured in an existing environment and does not require an isolated dedicated environment for evaluation purposes. The recommended hardware and software stated below is intended for sites that want to evaluate AppSpeed without having access to an existing virtual infrastructure setup.

The minimum hardware requirements for a successful evaluation include the following components:

- At least one physical server – This is used to install VMware ESX. Please see the VMware ESX 4 hardware compatibility list for supported server models. vCenter Server and AppSpeed software can run as virtual machines on this ESX server.
- One Windows workstation or laptop – This is to connect to the vCenter Server using the locally installed VMware vSphere Client.

<table>
<thead>
<tr>
<th>Hardware Requirements for VMware vCenter AppSpeed Evaluation Use Cases</th>
<th>Minimum</th>
<th>Recommended</th>
<th>Hardware used in this guide</th>
</tr>
</thead>
<tbody>
<tr>
<td># ESX Server hosts</td>
<td>1</td>
<td>Minimum of 2 ESX Servers for a small VM population environment</td>
<td>2 (the AppSpeed virtual appliances and the sample applications shared the two servers)</td>
</tr>
<tr>
<td>CPU</td>
<td>2 processors of 1500 MHz</td>
<td>2 dual-core processors of 2500 MHz</td>
<td>2 quad-core processors of 2500 MHz</td>
</tr>
<tr>
<td>Memory</td>
<td>4GB</td>
<td>8GB</td>
<td>10GB</td>
</tr>
<tr>
<td>Disk Space</td>
<td>24GB (SAN/Local)</td>
<td>64GB (SAN)</td>
<td>300GB (SAN)</td>
</tr>
<tr>
<td>Network</td>
<td>1 Gigabit NIC</td>
<td>2 Gigabit NIC</td>
<td>2 Gigabit NIC</td>
</tr>
</tbody>
</table>

For AppSpeed deployment, take into consideration the following system resource requirements:

- AppSpeed Server:
  - 30GB HDD
  - 2 vCPUs
  - 4GB RAM
- AppSpeed Probes:
  - 5GB HDD
  - 1 vCPU
  - 0.5GB RAM
3.2 Software Requirements

The installation of AppSpeed is supported for the following ESX and vCenter versions:

- ESX, version 3.0.x or above
- vCenter Server, version 2.5.x or above

Please see the AppSpeed product page for the latest compatibility matrices.

For the purpose of this evaluation, you can download the latest versions available of the following software:

- VMware vSphere 4 (Standard/Enterprise/Enterprise Plus) Evaluation copy
- VMware vCenter 4 Evaluation copy
- VMware vCenter AppSpeed Evaluation copy

VMware offers a free 60-day evaluation of the below VMware software. Follow the instructions at http://www.vmware.com/go/try-vcenter-appspeed to download the necessary licenses and the following binaries:

<table>
<thead>
<tr>
<th>Binary</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMware vSphere 4.0</td>
<td>You can create a boot CD from this iso file and use this CD to boot your server and install ESX 4.</td>
</tr>
<tr>
<td>(iso file)</td>
<td></td>
</tr>
<tr>
<td>VMware vCenter Server</td>
<td>The zip file includes an installer for VMware vCenter Server 4.</td>
</tr>
<tr>
<td>4.0 (zip file)</td>
<td></td>
</tr>
<tr>
<td>VMware vCenter AppSpeed</td>
<td>The zip file includes an OVF for vCenter AppSpeed 1.0. Extract the contents to temporary folder and launch executable file to start the install.</td>
</tr>
<tr>
<td>1.0 (zip file)</td>
<td></td>
</tr>
</tbody>
</table>

3.3 Software Installation and Configurations

Before you embark on the evaluation exercises, you need to install and configure the software listed in section 3.2. The table uses time estimates based on the hardware used to write this guide.

<table>
<thead>
<tr>
<th>Step</th>
<th>Installation and Configuration details</th>
<th>Time estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Provision two physical servers with the hardware profile described in Step 3.1</td>
<td>Varies</td>
</tr>
<tr>
<td>2</td>
<td>Install ESX Server software on two physical servers, including storage configuration. (See ESX Server 4 Installation Guide for more details.)</td>
<td>45 minutes per ESX host</td>
</tr>
<tr>
<td>3</td>
<td>Install VMware vSphere Client on a supported Windows machine (e.g. a laptop)</td>
<td>20 minutes</td>
</tr>
<tr>
<td>4</td>
<td>Create a virtual machine on one of the ESX hosts using the VMware vSphere Client and install Windows Server 2003 on it. Virtual machine configuration: Operating System: Windows Server 2003 32-bit Enterprise Edition (SP1 or higher)/ Standard Edition (SP1 or higher) CPU: 2 vSMP Memory: 4GB Disk: 40GB thin provisioned Network: 1 NIC card</td>
<td>60 minutes</td>
</tr>
</tbody>
</table>
5. **Install VMware vCenter Server on the newly created Windows Server 2003 virtual machine created in Step 4.**

6. **Start the VMware vSphere Client on your client machine and connect to the VMware vCenter Server just created in Step 5. Add the ESX hosts provisioned in Step 2 to the VMware vCenter Server inventory.**

7. **Create test virtual machines (preferable from already created template virtual machines to save time) that will be used to run sample applications. These virtual machines can be any combination of the supported operating system types.**

8. **Setup and configure multi-tier applications in the test virtual machines you created in Step 7. You need these applications to evaluate AppSpeed.**

   For this guide, two different applications were configured – SugarCRM and Online DVD store. For each application, the Web/application server and database server were set up in separate virtual machine. Another virtual machine was created for each application to generate load against these applications.

---

After you have successfully configured VMware vSphere, vCenter, and virtual machines running sample multitier applications, you can now proceed with AppSpeed installation.

### 3.4 AppSpeed Installation Requirements

The successful installation of the AppSpeed Server virtual appliance, plug-in, and probes requires that the following prerequisites be met. Verify the following system requirements before you install AppSpeed:

- Admin-level access to vCenter Server for the automated installation process
- Routable IP addresses for servers and probes need to be defined
- SSL private key is required for HTTPS sites that AppSpeed is to monitor
- Administrator privileges are required on the VI client machine to install Adobe Flash Player 10

In addition, it is recommended that you synchronize the time on the hosts on which the AppSpeed server and probes are to be installed. If the hosts’ time is not synchronized, AppSpeed might lose data and experience delays in the mapping process. The hosts’ time can be synchronized by setting an NTP server in the **Configuring Time** option in the respective hosts’ Configuration tab in vCenter.

Proper AppSpeed installation requires that the following ports be open on your firewall:

- **AppSpeed Server to vCenter Server:**
  - TCP port 443 for connection to vCenter Server (vCenter interface). Port 443 is the vCenter default Web Service port but a different TCP port can be configured in the vCenter.
  - TCP port 80 for connection to vCenter Server (vCenter proxy interface). This port is used only during setup, to verify that the vCenter proxy is configured correctly. Port 80 is the vCenter default Web Service port but a different TCP port can be configured in vCenter.

- **vCenter Server to AppSpeed Server:**
  - TCP port 80 for connection to AppSpeed server (Web interface).
• **AppSpeed Server and probe communication:**
  - UDP port 1194 for connection to AppSpeed server (OpenVPN over UDP connection)
  - TCP port 1194 for connection to AppSpeed server (OpenVPN over TCP connection)
  - UDP port 123 for connection to AppSpeed server (NTP)
  - TCP port 22 for connection to the probes (if you want to be able to access the probes from outside the VPN)

  **Note:** The use of DHCP is not recommended. If DHCP is used, routable IP addresses are not required.

• **VMware vSphere Client to vCenter Server:**
  - TCP port 80 for connection to vCenter Server (vCenter proxy interface). Port 80 is the vCenter default Web Service port but a different TCP port can be configured in vCenter.

### 3.5 AppSpeed Deployment Considerations

When deploying AppSpeed in your environment, you should also consider following items:

• AppSpeed can monitor any application that communicates via one of the following types of application level protocols:
  - HTTP and HTTPS
  - MSSQL
  - Oracle
  - MySQL
  - Exchange

For evaluation purposes, you need multitier applications that generate network traffic using the above protocols. (e.g. Micro-benchmarks such as IOMeter, netperf can saturate virtual machine and ESX resources, but may not generate application traffic that AppSpeed will discover and map.)

• AppSpeed can monitor encrypted traffic for HTTPS or SSL only. AppSpeed will identify SSL endpoints that need certificates. Once you provide AppSpeed the required certificates to decrypt the data, AppSpeed starts monitoring the specific SSL endpoint.

• **AppSpeed Scalability Considerations:** AppSpeed Server supports up to 32 AppSpeed Probes, and each Probe instance can collect data from a maximum of three vSwitches, including both standard and distributed virtual switches.

AppSpeed 1 has certain limits on the amount of data a single instance of AppSpeed Server, with a predefined virtual machine configuration of 2 vCPUs, 4GB RAM and 30GB hard disk, can monitor and process. At this point, it is not easy to determine a priori – that is, before the AppSpeed Server has been deployed in an environment – when this limit will be reached.

Application transactions are a fundamental unit of analysis in AppSpeed. An application transaction roughly maps to unique strings that applications use in communicating with one another. For example:

  - An SQL query skeleton: “SELECT * FROM users WHERE id = “.

Because the number of unique application transactions can vary significantly, even within a single data center (e.g. between production and test/development environment) you will need to deploy AppSpeed Server in your environment for a period of time — typically a couple of days — to discover AppSpeed scalability limits. In an application environment that exhibits high number of unique transactions, you may reach the maximum capacity a single AppSpeed Server can handle much sooner and with fewer virtual machines compared to an environment with fewer distinct application transactions.
When AppSpeed reaches the limit of transactions that can be monitored in an environment, AppSpeed displays a message in the Unmapped Traffic view. You can work around the scalability limits in AppSpeed 1 in three ways:

1. You can increase the CPU resources assigned to the AppSpeed Server to four vCPUs.
2. You can select specific virtual machines for monitoring, while excluding others in your environment.
3. You can deploy multiple AppSpeed servers to work around the virtual machine limit of a single Server instance.

Making these configuration changes requires the involvement of VMware Support or VMware Professional Services at this time.

- VMware DRS and VMware® VMotion™ Considerations: The AppSpeed Server can be migrated automatically between hosts in a cluster using VMware DRS. AppSpeed Probes, however, cannot be migrated across hosts. Therefore, VMotion and DRS is automatically disabled for AppSpeed Probes.

- AppSpeed does not monitor and analyze storage traffic. The Storage I/O latencies are reflected in the Application component of the Latency information displayed by AppSpeed. You should also consult the vCenter Performance tab for ESX and virtual machine level storage I/O latency information.

### 3.6 AppSpeed Installation and Configuration

AppSpeed is provided as a virtual appliance. The installation process consists of installing the AppSpeed Server through vCenter, then using the UI to deploy probe on all monitored hosts.

Before starting to work through the AppSpeed installation and configuration workflows that are outlined in this guide, you should work through the AppSpeed Pre-Install checklist below. Completing this checklist will ensure you are ready to proceed with the AppSpeed setup.

<table>
<thead>
<tr>
<th>#</th>
<th>Description</th>
<th>Yes / No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>You are using compatible ESX and vCenter versions (ESX 3.0.x and above are supported, as well as vCenter Server 2.5.x and above).</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>You have the IP address of the vCenter Server.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>You can provide admin-level access to vCenter Server for the automated installation process.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>In case you use static IP addresses for AppSpeed Server and probes, you have an IP address for the AppSpeed Server and an IP address for each probe. If DHCP is to be used then nothing is needed in advance. Using static IP addresses is recommended.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>You can provide the SSL private-key and certificate files for HTTPS sites that are to be monitored by AppSpeed.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Note:</strong> This is mandatory only for monitoring HTTPS sites.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>You have adequate disk space and system resources for AppSpeed Server and Probe installations (Server requires 2 vCPUs, 30GB of HDD and 4GB RAM; and each Probe requires 1 vCPU, 5GB of HDD and 0.5GB RAM).</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>You can install Adobe Flash player on the vCenter Client machine (requires administrator privileges on the vCenter Client machine).</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>You have downloaded the AppSpeed virtual appliance (in OVF format).</td>
<td></td>
</tr>
</tbody>
</table>
AppSpeed is provided as a virtual appliance in Open Virtual Machine (OVF) format. The AppSpeed plug-in is installed in the VMware VI Client. Perform the installation and setup process in the following stages:

- Import the OVF and configure the AppSpeed server
- Install the AppSpeed plug-in in the VMware VI Client
- Install the AppSpeed probes using the AppSpeed user interface (within the VMware VI Client)
- Add a valid AppSpeed license key
- Verify the successful deployment of the probes

Please refer to the “VMware vCenter AppSpeed Installation Guide” to complete the deployment of AppSpeed in your environment.

3.7 AppSpeed Application Mapping

Once deployed, AppSpeed analyzes all real user network traffic that traverses vSwitches monitored by the Probes and maps the applications running in your virtual datacenter. At high level, AppSpeed application mapping process consists of:

- Inspecting real user network traffic to decode application layer protocols
- Discovering the relevant end points (virtual machines) and applications running in them
- Learning application transactions
- Correlating these transactions across application tiers to understand any dependencies

AppSpeed maintains the information it discovers in the following structures:

- Applications (e.g. Exchange, MySQL)
- Transactions (e.g. login.php)
- Endpoints (IP + port + protocol)
- Virtual Machines

The application learning and correlation is a continuous process. Once AppSpeed has mapped the servers and applications, it continues to collect performance data until it can infer application correlation and dependencies. The rate of application learning depends on the traffic volume and diversity.

For HTTP-based applications, transactions are the actual user actions performed on the application. For database applications, transactions are identified by the table, method, and the query being used. There is a dependency between the two if the application server talks to the database server.

You can verify that AppSpeed is mapping network traffic in three ways:

- Real time statistics in the Info panel
- Unmapped Traffic view
- Probes Management view

To look at the real time statistics in the Info Panel, click the home icon.
Make sure that the **Throughput** value does not remain at zero. Throughput equates to the amount of mapped HTTP and database traffic flowing to/from virtual machines running on monitored hosts.

You should also note the number of virtual machines and applications in the navigation panel.

Another way to track the mapping process is to look at unmapped traffic using the **Unmapped Traffic** view. The unmapped traffic displays objects that have not yet been mapped to specific servers or applications. Any new transactions that AppSpeed learns about first appear in the **Unmapped Traffic** view and then become mapped transactions. **Unmapped Traffic per virtual machine or protocol** can be viewed by clicking the **Servers** or **Transactions** tab respectively.
You can also verify whether a specific AppSpeed Probe is monitoring network traffic or you have appropriate probes configured from Probes Management view under the Administration tab. The Throughput column indicates the network traffic a particular probe is monitoring.

Figure 3.7 c: AppSpeed Probes Management
This guide uses the SugarCRM application to walk through all use cases. The SugarCRM application has the following architecture: One front-end application server and one back-end database server. AppSpeed will map it as two applications – the front end and the backend. For each application, it will also map the transactions that participate in it.

In the left navigation panel of the AppSpeed main view, you will see the applications and servers that AppSpeed has mapped. In Figure 3.7 d, AppSpeed has mapped four applications and four servers:

- Two HTTP applications – ds2-iis and mysugar
- Two Database applications – db=DS2, an MS SQL Server application, and db=sugarcrm, an MySQL application
- Four servers – DVDStore-IIS, DVDStore-SQL, SugarWS1, and SugarDB1

In the next section, use cases detailing application and service level monitoring as well as application drilldown are addressed.
4. VMware vCenter AppSpeed Use Cases

This section outlines the use cases for VMware vCenter AppSpeed and lists the recommended steps for each case. The most common AppSpeed use cases are covered:

1. Application Discovery and Mapping
   - Auto discovery: Dynamically discover virtual and physical infrastructure elements.
   - Application mapping: Map application and server dependencies, groups workloads into logical applications, and maps application elements to virtual machines.

2. Performance Monitoring
   - Monitors latency: Non-intrusively monitor real-time, end-user latency for multi-tier applications.
   - Trend analysis: View historical trend for performance, errors, and usage. Enables users to examine usage trends and their effect on performance levels.

3. Application Drilldown
   - Troubleshooting: Quickly identify the component responsible for a performance issue, down to the transaction level.
   - Correlation: Correlate latency to the different elements of the environment (infrastructure, network, and application).

4. Service Level Management
   - Service level thresholds: Set thresholds based on desired service level targets and historical baselines.
   - Events and email notifications: Generate events and automated email notifications when service levels change.

AppSpeed provides numerous monitoring and analysis views for monitored objects. You can access the applications and servers from the left pane of the AppSpeed main view. AppSpeed also provides several ways to display the different application and transaction specific information:

- Click a different view from the views drop down list to display different data
- Double-click an item (in the main view table) to drill down into more details
- Check and uncheck transactions to add and remove them from the graphs
- Hover over the graph lines to get more details on each data point
- Double-click items in the legend table to highlight specific graph lines

Please see the VMware vCenter AppSpeed User’s Guide for details.

4.1 Application Discovery and Mapping

Due to the ease of deploying virtual machines, VMware Infrastructure (VI) administrators often have no clear knowledge of the various applications running in a virtual environment. With the processor core density and consolidation ratios going up, VI administrators are dealing with more virtual machines than ever and need a quicker way to map all applications and corresponding virtual machines running within a virtual datacenter. AppSpeed helps VI administrators in three different ways:

1. Identifies the applications running in the virtual environment
2. Provides application topology mapping to indicate which virtual machines are running these applications
3. Examine and reveal dependencies between applications and servers in the virtual datacenter

AppSpeed will identify an application in your environment once it has mapped any traffic from the application. VI administrators can look at the applications and servers that AppSpeed has identified in the left navigation panel of the AppSpeed main view. In the screenshot shown here, AppSpeed has identified four applications and four servers:

- Two HTTP applications – ds2-iis and mysugar
- Two Database applications – db=DS2, an MS SQL Server application, and db=sugarcrm, an MySQL application
- Four servers – DVDStore-IIS, DVDStore-SQL, SugarWS1, and SugarDB1
AppSpeed also has two views – **Applications to Servers** and **Servers to Application** – that list application topology mappings. These two views will help VI administrators answer the most common questions about locating applications or servers.

To view servers running an application, click the **home icon** and then select the **Applications to Servers** view from the application view pull down. This view lists all the servers that serve each of the monitored applications. For each server you will also see performance and usage metrics.

---

**Figure 4.1 a: AppSpeed Applications and Servers**

**Figure 4.1 b: AppSpeed Applications to Servers View**
Similarly, you can find out applications that are running on any given server by clicking the **home icon** and selecting the **Servers to Application** view from the application view pull down.

This view lists all applications running on each of the monitored servers. For each application, you will see performance and usage metrics.

**Figure 4.1 c: AppSpeed Servers to Applications View**

AppSpeed also provides graphical topology maps to illustrate these mappings. To see the topology maps, click the **Map** tab at the top. The graphical view provides an application or server focused map that can include any application dependencies AppSpeed has discovered. VI administrators can use this information to navigate through the various applications possibly running in a virtual datacenter, and focus on specific ones as needed.
4.2 Application Performance Monitoring

Traditional performance monitoring tools included with VMware vSphere focus on the system resource utilization such as CPU, memory, network, and storage metrics. VI administrators lacked the visibility into the actual end user experience typically tracked as response time or latency. As a result, VI administrators often cannot verify user performance complaints let alone isolate the issue. AppSpeed enables VI administrators not only to monitor application and server performance, it also provides them with historical correlation between application usage and latency. Using AppSpeed, VI administrators can set up proactive performance monitoring and be prepared even before a customer calls the helpdesk.

To monitor application performance, VI administrators can start with the default view that displays baseline application latency and hit information.

To select this view, first click the home icon. Select Applications: Baseline Comparison from the Analysis Views drop down menu. View the Baseline Comparison tab.
This view sorts the applications in descending order of average latency time. The average latency time indicates the measured time from the start of the client request to the end of the server reply, averaged across all hits in the given time frame. IT organizations often have service level agreements (SLAs) that require this average latency to be within a few seconds. VI administrators can monitor this column to validate if applications meet the SLAs in place. A larger latency value means a slower server response and may present opportunities for performance optimization.

An average latency number often can be meaningless because any transient usage spikes can often degrade the end user performance experience. To alleviate this, one often needs to monitor application usage over time. AppSpeed provides this information using the Applications: Usage Over Time view. To select this view, select Applications: Usage Over Time from the Analysis Views drop down menu.
This view provides application usage information in terms of the number of hits and throughput kilobytes per application. VI administrators can use this information to identify the high throughput applications and plan the virtual environment requirements accordingly.

AppSpeed’s user interface flexibility enables several ways to customize the information that is displayed:

- Check/Uncheck a few boxes in the leftmost column to add/remove data points from the graph.
- Also, move the cursor over the graph lines to see more details on every data point.
- Change the timestamp in the From and To fields at the top of the AppSpeed main view to change the time frame and granularity of the displayed data.

AppSpeed provides similar information at the server level using Servers: Baseline Comparison and Servers: Least Used views. These views will enable VI administrators to track virtual machine performance and help in any sizing and configuration decisions.

To display server latency and hit information, select Servers: Baseline Comparison from the Analysis Views drop down menu.
You can also view a server’s usage over time. To select this view, select **Servers: Least Used** from the **Analysis Views** drop down menu.
VI Administrators can also monitor the performance of a specific application. This provides the ability to profile an application, identify problems, and track application behavior trends.

To do so, you can either double-click **HTTP mysugar** application in the application grid or click the **HTTP mysugar** application in the left navigation panel. The default view is the **Usage vs. Latency** view. This view correlates the application usage with latency over time. This view can help VI administrators to verify if the application performance changed recently. If so, they can check with other stakeholders for any recent environment changes (e.g. network or storage re-configuration, operating system or application updates.)

Figure 4.2: AppSpeed Application Usage vs. Latency View

This view also lists all transactions that were mapped for the application. Application owners can use this information to pinpoint specific transactions that are performance sensitive and focus their efforts to optimize these critical transactions.

AppSpeed also can provide latency breakdown information – this view is available for both application and virtual machines. You can use this view to further substantiate any recent environment changes. Select the application **HTTP mysugar** from left navigation panel. Select the **Latency Breakdown** tab or choose the **Latency Breakdown** option from **Analysis View** pull down menu.
The latency breakdown information can help VI administrators determine the components that contribute to end user response time. The following table explains the values in the latency breakdown:

<table>
<thead>
<tr>
<th>Name</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Measured from the time the application server receives the beginning of the request until starting to transmit the reply.</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Measured from the time the server starts transmitting the reply until fully replying to the client’s request.</td>
</tr>
<tr>
<td>Network Overhead</td>
<td>The cost of packets retransmissions, errors, and drops.</td>
</tr>
<tr>
<td>Network RTT</td>
<td>The average round-trip time, without failures, from client to server.</td>
</tr>
</tbody>
</table>

The following figure represents the latency breakdown information at a very high level:
Typically, latency will primarily result from network and application components. The Infrastructure component is expected to be only a fraction of the total latency. Also the network overhead component will be minimal as well, because the network reliability is usually not a concern in properly configured datacenters.

High network or infrastructure components indicate issues that usually affect all application transactions. Network or VI Administrators can validate this further by running additional network tests or monitoring ESX and or application virtual machine configurations. Another possibility is that the Application component shows high latency percentage, especially for I/O intensive applications. In this case, you can work with VI administrator or Storage administrators to check disk I/O latencies. Perhaps storage is insufficiently configured for highly consolidated virtual environment. If only a certain subset of transactions are exhibiting high application latency, you will need to drill down into these application transactions. Please refer to the Application Drilldown use case for details.

You can view similar latency breakdown information for the database component of the application. AppSpeed provides detailed information down to the query level.

To view latency breakdown for the Server, click the SugarWS1 in the left navigation pane. Click either Latency Breakdown tab or select the Latency Breakdown option from Analysis Views drop down menu. At the Server level, AppSpeed breaks down latency information by application, infrastructure, network, and network RTT.
IT administrators are also interested in the transactions that occur most frequently for an application. Similarly identifying the slowest transactions of an application helps identify any potential bottlenecks before they occur. AppSpeed provides two pre-defined views to spot the Most Used Transactions and Slowest Transactions in an application. If a transaction is included in both the top ten “slowest” and “busiest” transactions lists, it is definitely worth looking into.

To look at the most used transactions, click HTTP mysugar application in the left navigation panel. From the Analysis Views drop down, select the Most Used Transactions view. The list of transactions is sorted by the number of transaction hits, and the graph displays the top three most hit transactions, that have been checked in the grid, across the selected time frame.
To display the slowest transactions across time, select the HTTP mysugar application in the left navigation panel. From the Analysis Views drop down menu, select the Slowest Transactions view. The list of transactions is sorted by slowest latency, and the graph displays the top three slowest transactions, that have been checked in the grid, across the selected time frame. Once the slowest transactions are mapped, you can drill down into each transaction (double-click the transaction in the table) to see the relevant metrics.
4.3 Application Drilldown

Once a performance problem has been validated, VI administrators need a way to triage and isolate the issue. VI administrators often do not possess the necessary application domain expertise for performance diagnostics and hence find it challenging to resolve application performance issues. AppSpeed can identify the transactions that are part of an application and drill down into specific transactions. Moreover, because AppSpeed can analyze dependencies between transactions across multitier application stack, it can enable VI administrators to identify root cause of performance issue.

After you have isolated the performance issue, you may need to refer to the performance and infrastructure monitoring tools included in VMware vSphere and specific applications to further narrow down problems to tuning or configuration issues in the application, operating system, or hardware infrastructure. In these cases, you will need to use AppSpeed provided information in conjunction with performance monitoring tools and tuning guidance from the vendor of the relevant component to determine the appropriate corrective action.

To drill down into the slowest transactions, click a specific transaction in the list that you want to drill into – in this case drill into `mysugar/sugar/index (module=leads)`. You can now see the transaction latency changes over time.

AppSpeed enables you to figure out the transaction dependencies across application stack. To do this, double-click the specific transaction you want to drill into. In the Dependencies tab, you can also see the back-end database queries that are part of this HTTP transaction.
AppSpeed also provides latency breakdown for the specific transaction. To view this information, click the **Latency Breakdown** tab. The latency breakdown components are the same as reported at the application level.
Furthermore, AppSpeed enables you to seamlessly navigate between the dependent transactions across the application stack. You can switch from a transaction in one application to any dependent transactions in other applications that AppSpeed has identified. To do so, simply double-click a dependent transaction in the Dependencies tab.

Figure 4.3.d: AppSpeed Application Transaction Usage vs. Latency

Similarly, you can view latency breakdown for the dependent transaction. To view this information, click the Latency Breakdown tab.

Figure 4.3.e: AppSpeed Application Transaction Latency Breakdown
Together the usage and latency correlation information can point to possible issues in your virtual environment, and save you time and effort in resolving these issues.

<table>
<thead>
<tr>
<th>Latency</th>
<th>Throughput</th>
<th>Possible Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>Capacity constraints – check if the application and virtual machine configurations can handle the requirements</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
<td>Check the Latency breakdown. If Network and Infrastructure issues have been ruled out, this implies Application issues. The Application needs to be optimized – either configuration limits or tuning or possible code issues. Examples:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Configuration limits may be explicitly set in application parameters in config or properties files</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Tuning may involve creating database indexes or updating statistics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Code issues may involve rewriting SQL queries</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
<td>Make sure you have enough capacity buffer and do not hit capacity constraints</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td>Mostly not an issue, also usually need more information</td>
</tr>
</tbody>
</table>
4.4 Service Level Management and Notifications

As part of proactive application management, IT organizations often establish service levels for key enterprise applications. The IT support agreements are usually tied to these service levels, and hence monitoring them is vital to any successful application environment. Any automation and alerting in service level management is even more welcome by these administrators. AppSpeed helps VI administrators manage application service levels in multiple ways:

- Monitoring and reporting service levels
- Setting service level thresholds
- Event based alerting when service levels change
- Automatic email notification for service level events

This section explains how to monitor application service levels, how to set thresholds for given transactions, how to view events, and how to view alert details.

AppSpeed defines service levels based on latency and error thresholds, and compares the performance of each monitored object to these service levels at all times. By default, AppSpeed defines four different states: OK, Warning, Above SLA, and Unavailable. VI Administrator can monitor the service levels. When an object service level is changed, AppSpeed triggers a state change.

To view the real time service level states of all applications monitored by AppSpeed, first click **home icon**. Then select **States** tab on the top.

---

Figure 4.4 a: AppSpeed Service Level States
You can also view real time service levels for a specific application. Click the HTTP mysugar application from the left navigation panel. Click the States tab on the top.

Figure 4.4 b: AppSpeed Application Specific Service Level States

AppSpeed service levels are defined by setting threshold values for latency and error metrics. You can enable service level monitoring and setting threshold values for specific application transactions as needed. When any one of these transactions breach the threshold settings, the application service level state changes. Furthermore, AppSpeed provides users the choice to set the threshold values manually or suggest automatically based on historical performance data using the following options:

- The Suggest for All button will suggest a thresholds for all transactions.
- Suggest for Undefined will suggest thresholds for transactions which currently do not have thresholds.
- The third option is to set a threshold manually or through a suggestion for an individual transaction.

To set service level threshold values:

- Select the HTTP mysugar application from the left navigation panel. Click the Thresholds link on the top.
- Service level monitoring is disabled by default. You may need to click the Enabled radio button to enable service level monitoring.
• Click any transaction, expand it, and highlight the row with the LATENCY thresholds. Thresholds are already set but thresholds can be manually adjusted. At the bottom of the screen you will see three thresholds: Warning, SLA, and Unavailable. In the Warning input box, change the latency to 500 msec. Click **Update Table**. You should see the new threshold reflected in the table above.

• Click **Apply Changes** to make sure the changes have taken effect.
• You will see the **Latency State** and **Error State** change to **Initializing**. If response times exceed the new threshold settings, events will be issued to alert the user.

Users can see events that are triggered when transactions cross a given threshold. To view the event list, click the **home icon** then click the **Events** link on the top. You may have to select the appropriate timeframe in the **Select Time Frame** pull down menu. All events that have been issued in the selected time frame are listed.

*Figure 4.4 e: AppSpeed Events*

Double-click an event to bring up the event details. From this screen, you can choose the next step from the list of options provided.

*Figure 4.4 f: AppSpeed Application Event Details*
The user can enable alarms for email notifications. By default, this is disabled. To enable email notifications:

- Click the **Administration** tab and then select the **Mail Notification Settings** link in the left navigation panel.
- Check the box for **Send mail notification for events**.
- Type your email address in the **Recipient Address** box and click **Add**.
- Keep the default settings for the rest of the fields.
- Click **Apply** at the bottom of the page.

*Figure 4.4: AppSpeed Mail Notifications Settings*
5. Next Steps
In this guide, the key-use cases of AppSpeed have been presented. Please refer to the VMware vCenter AppSpeed User’s Guide for more details.

5.1 VMware Contact Information
For additional information or to purchase VMware vCenter AppSpeed, VMware’s global network of solutions providers is ready to assist. If you would like to contact VMware directly, you can reach a sales representative at 1-877-4VMWARE (650-475-5000 outside North America) or email sales@vmware.com. When emailing, please include the state, country, and company name from which you are inquiring.

5.2 Providing Feedback
VMware appreciates your feedback on the material included in this guide – in particular, any guidance on the following topics:
• How useful was the information in this guide?
• What other specific topics would you like to see covered?
• Overall, how would you rate this guide?

Please send your feedback to the following address: tmdocfeedback@vmware.com, with “VMware vCenter AppSpeed 1 Evaluator’s Guide” in the subject line. Thank you for your help in making this guide a valuable resource.
6. Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>AppSpeed uses the term “server” to define a virtual machine it discovers.</td>
</tr>
<tr>
<td>Average Latency</td>
<td>The measured time between the start of the client request to the end of the server reply, averaged across all hits in the given time frame. A larger value means a slower server response.</td>
</tr>
<tr>
<td>Average Reply Size</td>
<td>The average size, in bytes, of a single server reply, averaged across all hits in the given time frame.</td>
</tr>
<tr>
<td>Average Throughput</td>
<td>The average number of reply bytes per second in the given time frame.</td>
</tr>
<tr>
<td>Error Percentage</td>
<td>The percentage of error hits as a portion of the total hits in the given time frame.</td>
</tr>
<tr>
<td>Execution Time</td>
<td>The sum of all latencies in the given time frame. This value reflects the total amount of time of all server replies in the given time frame.</td>
</tr>
<tr>
<td>Hit Rate</td>
<td>The average number of hits per second in the given time frame.</td>
</tr>
<tr>
<td>Latency Baseline</td>
<td>The average latency of all the hits in the last year.</td>
</tr>
<tr>
<td>Latency Breakdown: Application</td>
<td>The latency measured from the time the server receives the first packet of the request until it starts to transmit the reply.</td>
</tr>
<tr>
<td>Latency Breakdown: Infrastructure</td>
<td>The latency measured from the time the server starts transmitting the reply until fully replying to the client's request.</td>
</tr>
<tr>
<td>Latency Breakdown: Network (RTT)</td>
<td>The average round-trip time, without failures, from client to server.</td>
</tr>
<tr>
<td>Latency Breakdown: Network Overhead</td>
<td>The cost of packet retransmissions, errors and drops.</td>
</tr>
<tr>
<td>Latency STDEV</td>
<td>The standard deviation of the latency of all the hits in the given time frame.</td>
</tr>
<tr>
<td>Max. Latency</td>
<td>The latency of the slowest hit in the given time frame.</td>
</tr>
<tr>
<td>SLA percentage</td>
<td>The percentage of hits that were not errors and which had latency within SLA levels.</td>
</tr>
<tr>
<td>Total Errors</td>
<td>The total number of hits identified as errors in the given time frame.</td>
</tr>
<tr>
<td>Total Hits</td>
<td>The total number of hits in the given time frame.</td>
</tr>
<tr>
<td>Total Throughput</td>
<td>The total number of reply bytes in the given time frame.</td>
</tr>
</tbody>
</table>
7. VMware vCenter AppSpeed Evaluation Worksheet

You can use the checklist below to organize your evaluation process. The checklist provides a high-level summary of the various AppSpeed workflows and configuration tasks that should be completed during your AppSpeed Evaluation.

Hardware Checklist:
___ Physical server compatible with VMware ESX 4 Qty: ___  
___ Storage (Shared/Local) >= 256GB  
___ One desktop/laptop running MS Windows (e.g. XP/Windows Server 2003)

Software Checklist:
___ VMware vCenter AppSpeed OVF file  
___ License keys for all software components

Installation Checklist:
___ VMware ESX Server 4 installed on physical server(s)  
___ VMware vCenter Server and database installed in a virtual machine or physical server  
___ VMware vSphere Client installed on the desktop/laptop running MS Windows  
___ Sample multi-tier applications installed  
___ VMware vCenter AppSpeed installed

After you have successfully checked off the lists above, you can proceed to perform the evaluation of VMware vCenter AppSpeed. For each scenario, you can use the corresponding checklist below to ensure that you are following the proper sequence.

<table>
<thead>
<tr>
<th>#</th>
<th>Use Case</th>
<th>AppSpeed Views to Use</th>
<th>Yes / No</th>
</tr>
</thead>
</table>
| 1 | Application Discovery and Mapping | • Unmapped Traffic  
• Applications to Servers  
• Servers to Applications  
• Topology Maps |          |
| 2 | Application Performance Monitoring | • Applications: Baseline Comparison  
• Applications: Usage over Time  
• Servers: Baseline Comparison  
• Servers: Least Used to Applications |          |
| 3 | Application Drilldown | • Most Used Transactions  
• Slowest Transactions  
• Application Usage vs. Latency  
• Application Latency Breakdown |          |
| 4 | Service Level Management and Reporting | • Application States Dashboard  
• Setting Thresholds  
• Events and Event Details  
• Email Notifications |          |