SIGHTLINE™ SYSTEMS

SightLine™ Best Practices

Capacity Management Focus

SightLine Systems Corporation 11130 Fairfax Blvd, Suite 200 Fairfax, VA 22030

Disclaimer

Please direct and questions and comments to:

SightLine Systems Corp.
11130 Fairfax Blvd.
Suite 200
Fairfax, VA 22030

Phone: 703-563-3000
Toll-free: 877-744-4854
Fax: 703-563-4000

www.sightlinesystems.com

Contents

1.0 Preface

This document serves to demonstrate ‘best practices’ for the deployment of the SightLine software suite across an existing large-scale Unix infrastructure to include personnel involvement, programming involvement, technical resources, etc. In particular, this document focuses on the following areas:

- Capacity Management Functions
- Data Management and Retention
- Data Collection
  - Collection Interval and Data Selection
  - Process Data and Filtering
  - Workload Characterization
- SightLine Deployment in a Distributed Environment
  - SightLine Scalability
  - Remote Configuration
  - Framework Integration
- Reporting Options

At the end of this document, a Use Case is presented as a best practices application of SightLine in a large-scale Unix environment within the scope of a Capacity Planning Enterprise License Agreement. Many of the practices detailed in this document are applicable to a Windows or heterogeneous deployment.

2.0 Capacity Management Functions

Capacity Management involves a number of functions including system tuning, server consolidation and capacity planning. Data requirements for capacity management differ from those necessary for performance management. When a specific performance problem arises, a critical time epoch is examined to determine the cause of the performance degradation. The time epoch is identified using a performance indicator, such as when a measured response time exceeds some threshold. Performance data is typically maintained at fine granularity for a short period of time.

For capacity planning, since there are no critical epochs to observe, samples of data should be collected over times that have meaning to the capacity plan. In addition to these peaks, there may be additional cycles, such as month-end batch processing or seasonal fluctuations. The type of capacity plan determines which workload samples should be selected for evaluation, what metrics should be collected and how long they should be retained.

2.1 System Tuning

System tuning is a platform-dependent activity that uses measured workload and configuration information to identify current bottlenecks. The information is used to change the workload or configuration and to try to improve performance prior to consolidating servers or building baseline models for capacity planning. System tuning requires an intimate knowledge of the operating system, environmental software, applications and the parameters and system settings that are used to control them. It often requires the careful adjustment of several such parameters--things like process priorities, placement of disk files and memory management parameters.

SightLine provides the server and process level data required to perform detailed system tuning. Systems can be monitored, in real time, as changes to kernel parameters, disks and workload priorities are made. The SightLine correlation capabilities can assist in providing insight into unknown dependencies within the data to aid in assessing the impact of adjusting various parameters prior to making any changes.

2.2 Server Consolidation

Server consolidation analysis requires the ability to view server utilization metrics over time to identify trends in resource consumption and identify inter-dependencies in data. Long-term utilization reports can help system managers identify candidates for workload consolidation and identify peak processing requirements.

Because SightLine provides functionality for arithmetic and logistic data manipulations, along with the capability to plot data from different servers in a single graph, users can

perform rudimentary server consolidation exercises to identify opportunities for alternative workload placement. SightLine also provides the capability to display historical data for workload and server trend analysis.

2.3 Capacity Planning

Capacity planning does not require an intimate knowledge of the system and the effects of various parameters. Instead, it treats various components of the computer system as "components" and measures their responses to stimuli such as changes in workload or changes to physical characteristics of the components. Normally, a capacity planning study is only appropriate after the system is well tuned and the users are getting reasonable performance from the system.

SightLine ForSight can be used to create baseline models for existing servers and then simulate changes in workload or configuration. Information used for building the baseline models comes directly from SightLine Power Agents managed by SightLine Expert Advisor/Vision (EA/V). The modeling exercise involves the modeling of forecasted workloads on forecasted configurations and identifying components that must be re-configured or workloads that must change to meet performance requirements. Examples of the type of "*What if…?*" analysis available with ForSight include:

- What if I consolidate the workload from two or more servers onto one server?
- What is the impact on capacity if this workload increases by 20% over the next six months?
- What is the impact on response time if an additional processor is added to this server?

3.0 SightLine Overview

SightLine provides a capacity management and real-time performance management solution that assures the service-level management goals of large-scale business applications from end to end.

For capacity management, SightLine offers:

- Tools enabling proactive tuning and optimization of servers, network components, applications or portions of your IT infrastructure, such as web services.
- Consolidated views of enterprise-wide infrastructure, trends and service levels enabling managers and technical staff to identify opportunities for server consolidation or workload optimization.
- Integrated simulation modeling using real data for on-demand capacity planning.

For real-time performance management, SightLine offers:

- Collection and aggregation of any real-time information, end-to-end, independent of platforms, networks, or applications, including information from external partners;
- Critical path monitoring of a business transaction or request across heterogeneous application components, analyzing the flow as a single entity. As a result, it is now possible to provide insight into the inter-relationships between all of the components involved in your application flows.

The SightLine software is a component-based application designed specifically to manage and optimize the capacity, performance and service levels of heterogeneous, multi-tier business applications. The software collects and aggregates any real-time information from your systems, independent of platforms, networks, or applications, including external market feeds or partner information sources. Real-time data is aggregated and consolidated for long-term trend analysis and capacity planning. SightLine also easily integrates information from 3rd party monitoring tools, and can communicate with popular management frameworks and platforms, leveraging current IT investments.

- **SightLine Expert Advisor/Vision (EA/V)** analyses and correlates unlimited amounts of heterogeneous information in real-time. EA/V's AutoCorrelate and AutoAnalyze features highlight critical relationships between IT components, and can also identify previously unknown indicators as they are coming into play in your business environment. Other automatic features provide true cause and impact analysis, recommend and/or take corrective actions and use alerts to identify critical service level impacts. Progressive discovery can be applied to any set of indicators for more in-depth analysis and trending. EA/V's analysis can be applied to real-time as well as historical information for strategic planning and

forecasting. EA/V provides diverse views of technical, management and business-level data in both traditional and web-based formats.

- **SightLine Integration Adapters** are supported for several framework products including Tivoli Enterprise Console, CA Unicenter and Remedy. EA/V AutoAlerts are translated to the native framework format and forwarded to the specified framework console for display and possible further action.
- **SightLine ForSight** is a completely integrated discrete event simulation tool for modeling the performance behavior of various hardware and workload configurations. ForSight supports an automated data feed from EA/V and builds on its analysis of enterprise systems through a powerful ‘what if?’ modeling capability and comprehensive hardware knowledgebase.
- **SightLine Agents** monitor the critical resource utilization and service level indicators of specific business components. Three types of agents are provided to meet different needs. Power Agents provide in-depth monitoring of a wide range of indicators and are designed for situations where in-depth analysis or historical trending may be required. Summary Agents provide monitoring of truly critical indicators, and are designed for use where snapshot "health" analysis is desired. Service Agents test services from across the corporate Intranet or Internet, monitoring the experience that is being delivered. Interface Agents integrate with third-party monitoring tools, protecting current IT investment by collecting information from these tools and consolidating it with SightLine gathered information. Third-party information is analyzed and correlated transparently through the SightLine Expert Advisor.
- **SightLine SDK** enables organizations to quickly and simply create Interface Agents for proprietary applications or management tools.

SightLine Systems offers a full range of expert services to assure the success of your SightLine deployment. Our consultants are recognized within the industry as experts in the enterprise performance management arena, with knowledge spanning all aspects of the business infrastructure, across mainframe, UNIX and Windows systems and applications, as well as network domains.

- **SightLine Assurance Services** include Service Level Assurance planning, capacity analysis and sizing, service level and performance monitoring, SightLine customization, and training.
- **Best Practices Deployment Services** are designed to optimize your SightLine deployment to meet your specific needs and requirements. SightLine consultants partner with you to optimize your business infrastructure, configuring SightLine to maximize your infrastructure’s resources while enhancing your management personnel’s productivity.
- The **SightLine Integration Services Group** can design and build custom agents to meet your specific monitoring needs.

4.0 Data Management

SightLine EA/V acts as the central management console within a SightLine deployment. It aggregates, manages, displays and analyzes heterogeneous information from networks, platforms, operating systems, databases and applications.

The capacity to store collected data directly on the monitored system in a host trace file (HTF) provides a temporary store of data which facilitates flexible strategies for downloading data to the EA/V server.

By default, individual HTFs are uploaded or transferred to an EA/V server and are maintained at the raw data collection interval in circular data files called “trace files” on specified local network drives. Trace file configuration includes the ability to specify the location of the files, the maximum size of the file, and the summary interval, among other things. Data consolidation can also be applied, which allows for the selection of a subset of metrics to be used in the creation of a single trace file representing data from multiple, possibly heterogeneous, systems. As previously discussed, data (raw or consolidated) can be served up to higher EA/V tiers as part of a data management strategy.

4.1 Data Retention

An important consideration is a data retention policy. Some data consumes a tremendous amount of space and loses its value quickly, such as process data. The granularity of data also loses its value over time. For example, for analyzing bottlenecks, detailed performance and process data is quite useful. For Service Level Reporting a small subset of data is required, and no process data is needed. For capacity planning, only highly summarized workload and resource utilization data representing peak periods is necessary.

A sound data retention policy provides for storing different types of data for different periods. It may be advantageous to store detailed performance and process data for two weeks to one month, while data summarized hourly will be stored indefinitely. It may also be desirable to store daily capsules of peak demand periods that contain fine granularity and process information. For example, Month-End processing demand may be worth storing permanently as a baseline of peak demand.

SightLine offers users flexibility in how they store data. The primary data storage location is on the SightLine EA/V server or in a centralized repository. This offloads a significant portion of performance management overhead. A comprehensive regime of data summarization should be employed. The regime uses four different data files:

- HTFs – Host Trace Files are stored on the managed objects. Their size can be specified as a function of time or space. This file usually holds 1 to 3 days of data, and is used as a temporary back up for network or EA/V server outages.

- VTFs – These are live, circular files on the SightLine EA/V server or local network drive(s) established as a centralized repository. VTFs have fine granularity and contain process information in a companion Event file (VEV).
- Daily VSQs – Compressed VTFs covering a specified (possibly peak) period. A separate Post-Operation command can be used to accomplish additional file management operations, such as move, copy, rename, etc. These files are ideally located on the SightLine EA/V server or local network drive(s) established as a centralized repository.
- History VSQs – Highly summarized (1/2 – 1 hour sample rate) compressed data files containing no process information aside from workload metrics. History VSQs are typically configured to hold 2 to 10 years of data. Again, these files are ideally located on the SightLine EA/V server or local network drive(s) established as a centralized repository.

4.2 Archival Support and Retrieval of Historical Data

When a user specifies a connection in the **Network Host Sessions** dialog box, a trace file is created that can hold at least one day's worth of detailed data. If EA/V's AutoHistory feature is selected, a history file is also automatically created. The default summary interval is 30 minutes, but this value is configurable. The file is automatically sized to hold approximately two years' worth of history data. As the live data is collected in the live trace file, this data is automatically appended. A filter option can optionally be to specify a subset of data elements. Historical information is accessed by EA/V in the same manner as real-time data.

The *datamgr* process is part of the SightLine Power Agent. It connects to the data collector, and records blocks of metrics into a data file which is usually, but not necessarily, stored locally. The configuration file for the Power Agent contains information about the various data sources on the system, such as host name, port number for connecting to the collector, database directory name, and the number of days of complete data files to store. The *datamgr* can connect to and store data for as many machines as are listed in configuration file. This allows one *datamgr* to manage all of the databases for all your systems.

4.3 POSTOP Utility

The POSTOP Utility provides basic file manipulation to assist in managing the daily or historical files created with the TraceFile Manager or SightLine AutoHistory feature. POSTOP provides basic file manipulation capabilities such as copying, moving, deleting, and date-based renaming, using military or Julian dates.

5.0 Data Collection

SightLine uniquely enables the accumulation of system, network, and workload data. The table below summarizes the major Processor data classes available through the deployment of the SightLine Power Agent for UNIX Systems.

BIO Statistics	Kernel Resource	Segment Map Metrics
Buffer Cache	Logical Volumes	Semaphores
Caché Statistics	Memory Metrics	Sequent Volume Manager (SVM)
CFS Filesystems	Message Queue	Shared Memory
Clusters	Name Cache	Streams Statistics
CPU Utilization	Network	Swap
CPU Wait State	Network NFS Client	System Calls
Disk Controllers	Network NFS Server	System Usage
Disk Partitions	NFS Mounted Fs	Tape Units
Disk Units	Paging	TCP/IP IP
EFS Filesystems	Processes	TCP/IP TCP
Ethernet Device Statistics	Process Info	TCP/IP UDP
FDDI Device Statistics	Pseudo Terminals	TCP/IP ICMP
File System Info	Resource Regions	Token Ring Devices
File Statistics	RPC Stats	TTY
Interrupt Activity	SDI Adapters	Users Logged On
Kernel Metrics	SDI Domain Mgmt Layer	Virtual Memory Ops
Kernel Memory Structures	SDI Hosts	Workloads

5.1 Transferring Subsets of Collected Data

By default, all data classes are collected by the Power Agent and transferred to EA/V for display and analysis. Alternatively, users can define download services that transfer only a subset of data to EA/V. Although the Agent collects all data, a subset of metrics can be specified for transfer across the network to EA/V. This serves to minimize network traffic, minimize data storage requirements at the user end, and focus the collection of data on the requirements of the analyst. Changes to the metric set transferred to EA/V can be made on the fly or a new connection to the entire metric set stored on the managed object can be initiated if additional data is required. SightLine Agent Administrator (SLAA) can be used to make changes to data collection policies on one or more managed objects from a single client. This is discussed in more detail in Section 6.2, *Remote Configuration*.

5.2 Collection Interval

SightLine Agents can be configured to collect system metrics as often as once per second. However, typical near real-time data collection occurs at 30-second intervals. This data (in tracefile format) is uploaded to one or more EA/V servers at the raw collection interval. Individual trace files at any granulation and from any source can be archived “as is” or automatically rolled up into higher granulations in multiple history

files or into a single consolidated tracefile. Data from multiple granulations can be displayed in a single plot or environment.

5.3 Process Data and Filtering

SightLine's EventList provides a mechanism for viewing process-level activity on the monitored system both in real-time as well as "replaying" these activities. For UNIX systems, *Summary, CPU, Memory, I/O,* and *Identification* Event Classes are delivered by default. Each Event Class includes information about a select group of processes on the system for each interval. The process or event data can be plotted and included in analysis input and results, such as AutoCorrelate. EventList allows you to define and view as many process classes as desired.

The user has the ability to filter out "uninteresting" or unwanted processes. This helps decrease the overhead of data collection and transfer across the network. There can be zero or more filter statements; all will be included in SightLine's consideration of which processes to keep. For example, the default filters delivered with SightLine will filter out processes that do not meet the following criteria:

- Consume 1% or more of the CPU

  OR

- Have a resident set size of 2000 KB or greater of real memory.

  OR

- Produce 100 KChars or more of I/O

5.4 Workload Characterization

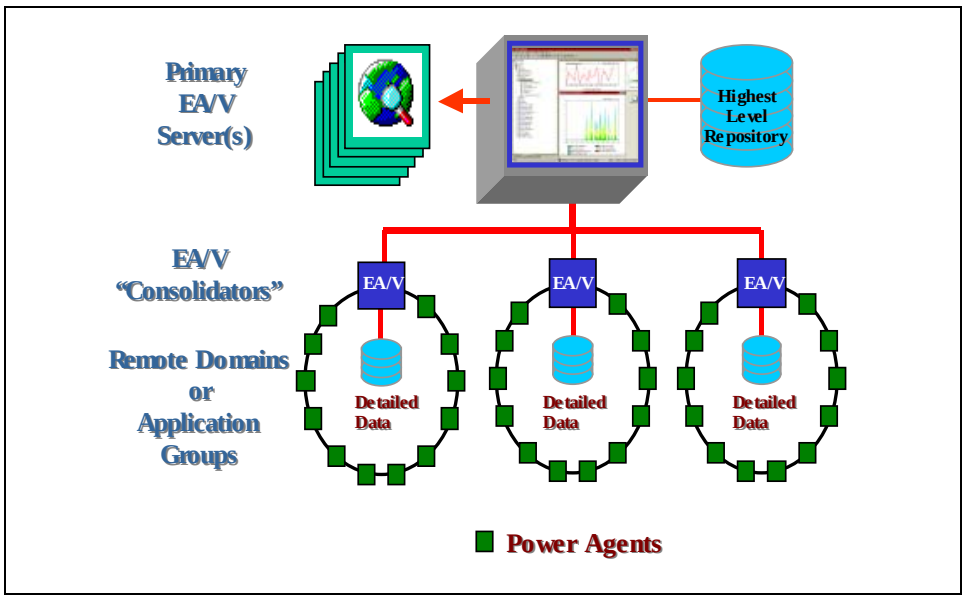
The task of creating logical groupings of resource activity is called Workload Characterization. Business applications are composed of one or more processes working to provide some business-related function. To effectively monitor, analyze, and plan capacity for business applications, it is vital to understand exactly what resources the application consumes. By defining groups of processes as Workloads, SightLine users can obtain application-specific resource demand measurements.

Workloads are defined to capture and subtotal the resources consumed by one or more processes. The Agent can be configured to accumulate these individual process measures into user-specified Workloads. Workload metrics are simply process measurements that have been summed for the entire Workload. For example, The Workload metric "% CPU" reports the percentage of CPU utilization that is dedicated to all the processes that make up a Workload.

6.0 SightLine Deployment

6.1 SightLine Scalability

As shown in the diagram below, each EA/V console can be configured to collect and manage data from a specified set of managed objects based on geographic domain or application group. Detailed data can then be maintained at the local (domain or application) level, while a subset of data is served up to a higher tier EA/V. The detail data can be accessed at any time by directly using the lower-tier (local) EA/V or by pointing to the data with any higher tier EA/V.

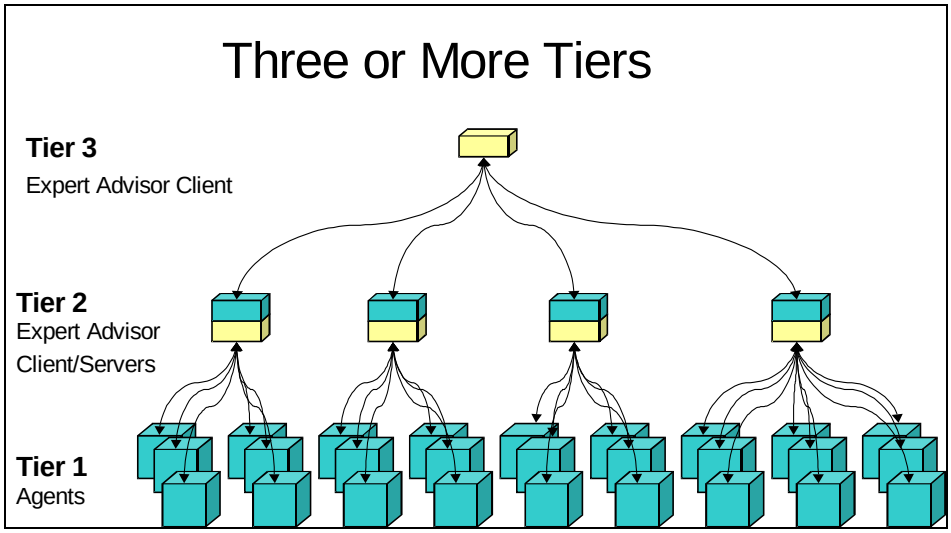


SightLine takes advantage of two technologies, aggregation and consolidation, to satisfy scaling requirements for large-scale servers. SightLine Consolidator provides a method for taking trace files and sending (aggregating) them, in near real-time, to higher-level EA/V servers. This provides a method, using EA/V, to easily visualize and correlate information across multiple servers, with a design goal of 10,000 or more servers. Since the data coming from lower level tiers appear as standard real-time data, there is no additional processor delay incurred as a result of this scaling strategy.

6.1.1 Data Aggregation

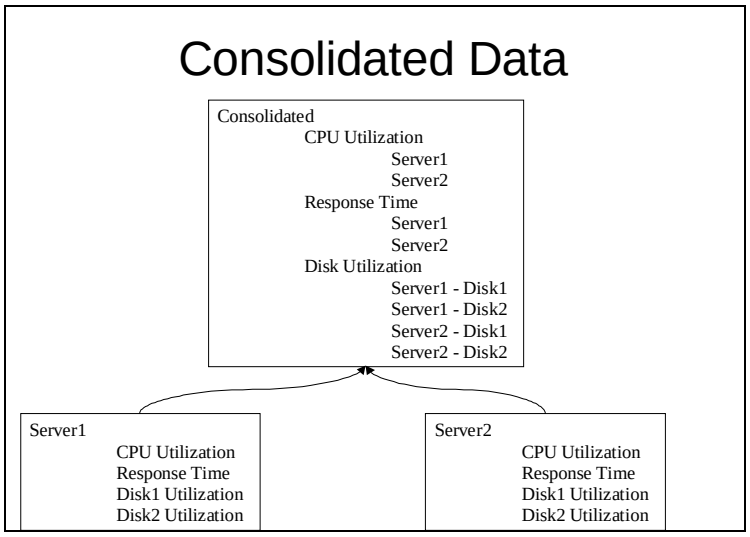
The figure below shows the topology for the multiple tiering or data aggregation. The collector function is embedded in the Tier 2 EA/V server so that the server appears to be a managed object to the Tier 3 EA/V server.

For example, the EA/V server at Tier 3 will connect to the EA/V server at Tier 2 through secure TCP/IP and will be able to read one or more Trace files as if the data was coming from one or more managed objects. The transmission of data can be live and continuous, as if the data is coming directly from the server. At the user's discretion, a time interval can be requested from the Tier 2 servers.



6.1.2 Data Consolidation

The figure below shows the basic idea behind consolidation. The user is able to specify one or more variables from a number of heterogeneous trace files. Each variable will then be written to a consolidation file, with the subscript indicating the origin of the data.



In this way, a “Top List” showing the CPU utilization, ranging over a number of managed objects, can be displayed. Correlation of variables within the trace file can be used to show meaningful relationships among the variables and across diverse platforms.

Consolidation can take place at multiple tiers (consolidation of multiple consolidation files) and across heterogeneous systems. In this way, SightLine could theoretically consolidate 10,000 servers using multiple tiers to handle scaling issues. For example, if each level handles 10 collectors (or EA/V servers acting like collectors), a five-tier system could yield a consolidation file that shows high-level health or capacity planning statistics for all 10,000 servers. In practice, consolidation of a large number of trace files, as in the example above, assumes that only those metrics required for performing high-level capacity planning analysis be transferred or consolidated up through the tiers. More detailed data is maintained at the local EA/V consolidation point and can be accessed directly for on-demand analysis or for data selection to conduct ForSight modeling.

6.2 Remote Configuration

The SightLine Power Agent uses an Extensible Markup Language (XML) file format for configuration files. This file format permits remote client GUIs to configure managed objects that are network accessible. The architecture of managed objects and client GUI is called SightLine Agent Administrator (SLAA). With the SLAA architecture, the configuration of Power Agent functionality no longer requires users to physically access each of the managed objects to edit individual configuration files. Using the SLAA GUI, the user can start up an applet within a browser, which enables editing of configuration files from single managed objects or groups of managed objects that are accessed on the site's network.

In addition to improved manageability and serviceability, the design goal for SLAA was to consolidate individual files from various SightLine agents and provide a single, industry-standard format for software configuration. SightLine has implemented a fully compliant XML Schema that enforces validation of element data types, element cardinality, and enumeration of element values.

6.3 Framework Integration

SightLine EA/V supports several Framework Integration Adaptors including Tivoli Enterprise Console, CA Unicenter and Remedy. An alert threshold is set for each type of alert forwarded to the framework console through a simple EA/V AutoAlert dialog. The SightLine Integration Adapter receives alert data from SightLine EA/V in the form of a temporary file. The alert is translated into the appropriate framework event format and forwarded to the framework console for display and possible further action. Additionally, simple SNMP traps can be sent to compatible products and devices.

7.0 Reporting Options

SightLine provides a several reporting options to support both capacity management and real-time performance management functions.

For capacity management:

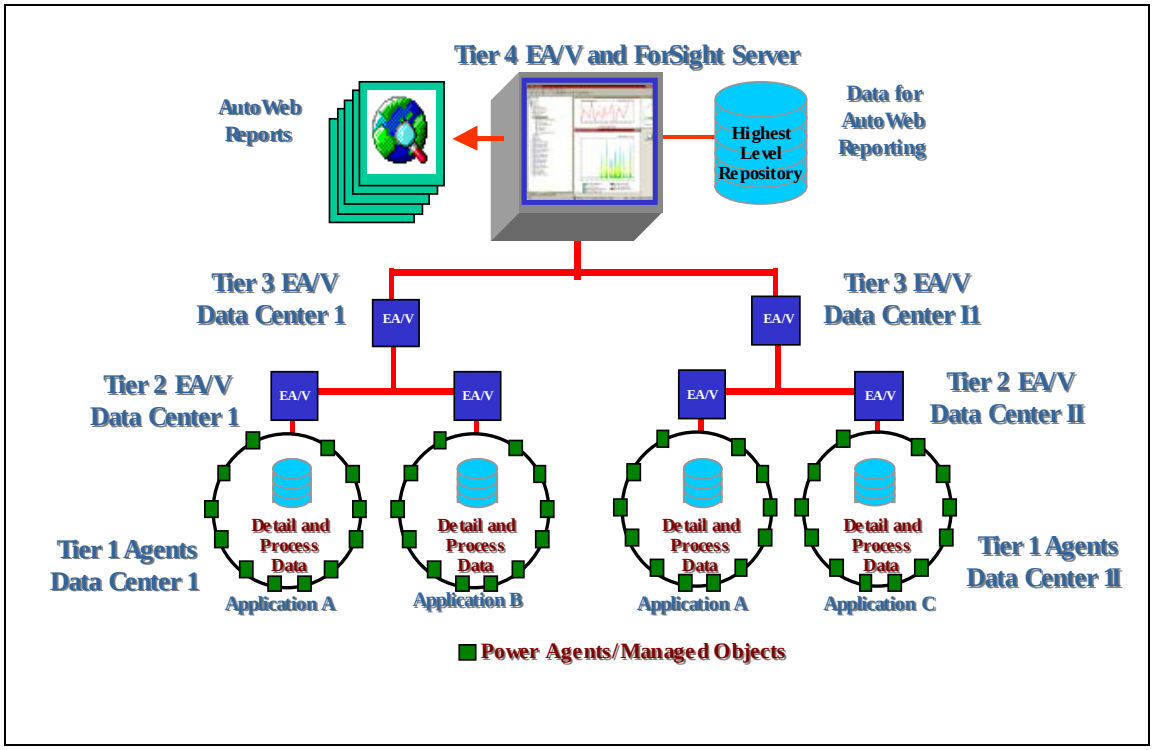
- SightLine supports dynamically updated objects embedded in documents such as Microsoft Word documents and Excel spreadsheets that utilize Windows OLE capabilities. Reports only need to be created once, and can then be scheduled to run daily, weekly, monthly or on-demand.
- Long-term data can be included in reports or exported into a spreadsheet application to aid in trend analysis and server consolidation studies.
- ForSight modeling results can be displayed both graphically (pie and bar charts) and in text based tabular reports to evaluate the relative impact of changes in workload or configuration.
- AutoWeb enables customized reports to be deployed that can be accessed for daily, weekly or monthly capacity reporting. Custom HTML pages can be quickly developed to display SightLine information in a setting that matches the enterprise's look and feel. If requested, AutoWeb will serve up any file stored in SightLine’s AutoWeb directory.

For real-time performance management:

- SightLine is delivered with a fairly substantial set of pre-specified reports in rich text format (RTF) for use in ad-hoc performance analysis. These reports focus on providing summary analysis of key performance indicators and support of built-in analysis features such as AutoAnalyze.
- SightLine plots and environments can be quickly and easily copied into Microsoft applications or email for problem notification or analysis.
- SightLine’s AutoWeb feature automatically creates and serves HTML pages from the pages and plots displayed in EA/V by turning the EA/V server into a Web server. A SightLine environment can be served up for quick load-and-go viewing by any browser with the click of a single button.

8.0 Use Case

The following Use Case presents a proposed ‘best practices’ deployment plan for SightLine within the scope of an Enterprise License Agreement. The test case will use a hypothetical environment consisting of two remote data centers. The first data center supports 50 Unix servers hosting two major applications. The second data center supports 35 Unix servers hosting two major applications. Data collection and consolidation practices will focus on the support of a Capacity Planning function with SightLine ForSight. The primary reporting mechanism for the display of data is assumed to be AutoWeb.



8.1 Capacity Management Functions

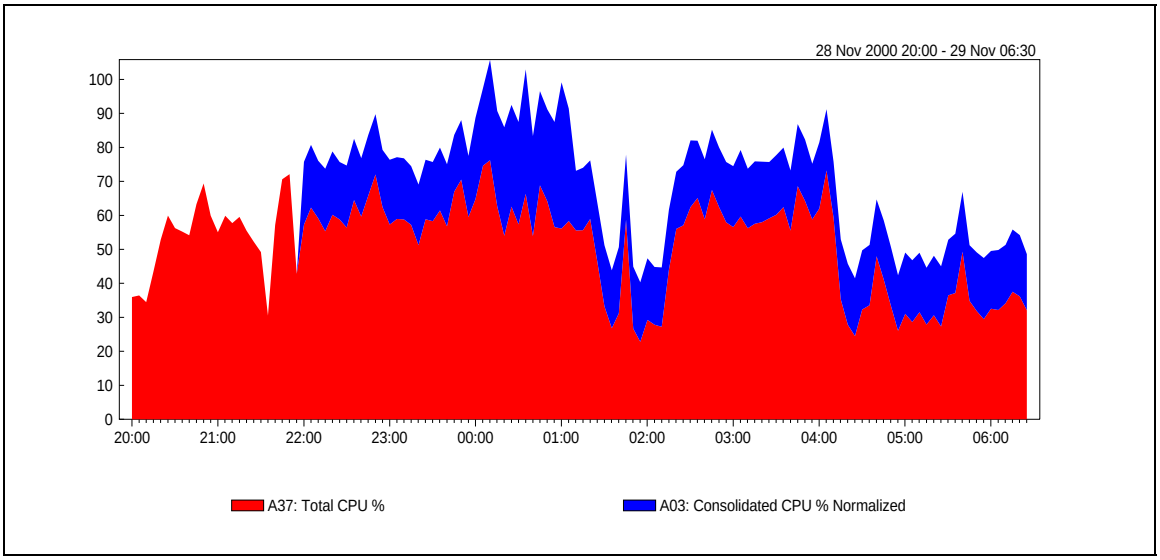
The Use Case is designed to show how SightLine can be used to support the three functional areas of Capacity Management mentioned in Section 2.0, namely system tuning, server consolidation and capacity planning.

8.1.1 System Tuning

SightLine data collection and retention at the Tier 1 and Tier 2 levels will provide the server and process level data required to perform detailed system tuning. Systems can be monitored in real time from any EA/V server as changes to kernel parameters, disks and workload priorities are made. Because data will be maintained in a centralized repository, analysis of system behavior before and after any system change can be conducted. SightLine's correlation capabilities can assist in providing insight into unknown dependencies within the data to aid in assessing the impact of adjusting various parameters prior to making any changes.

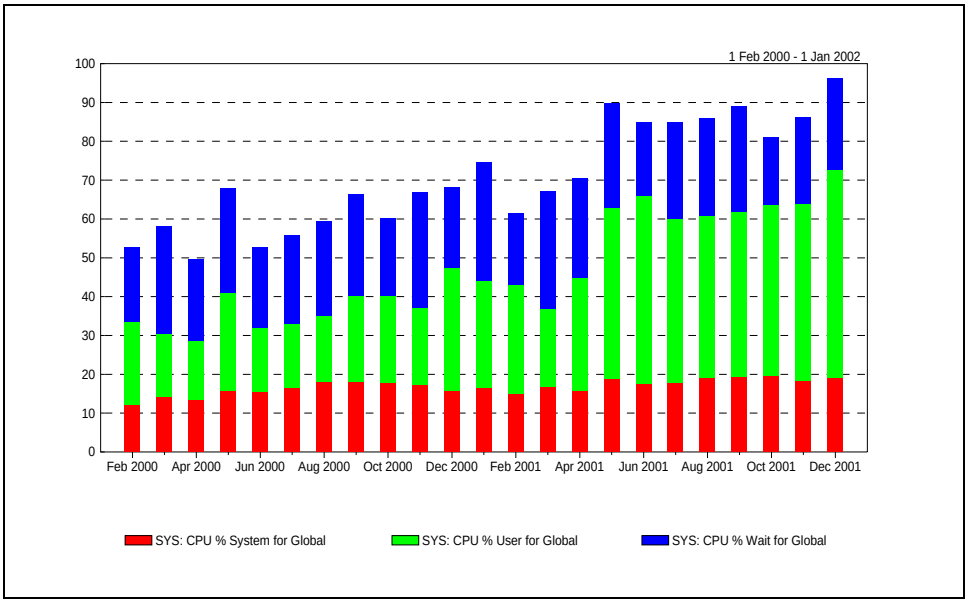
8.1.2 Server Consolidation

Workload and resource utilization data for each managed object will be available within all Tiers at intervals ranging from 30 seconds to one hour. By using SightLine's arithmetic and logistic data manipulation functionality, users can plot data from different servers in a single graph to perform rudimentary server consolidation exercises. The ability to plot data from multiple managed objects over time in a single plot can assist in identifying opportunities for alternative workload placement. The figure below provides a "back of the envelope" approximation of the total CPU utilization that can be expected by consolidating batch workload from two systems. EA/V expressions are used to normalize the consolidated workload to the target server.



8.1.3 Capacity Planning

The collection of resource utilization data over time will enable users to identify trends in workload data and resource utilization requirements during peak periods. The graph below shows prime shift processor (CPU) utilization for a Unix system for the past 24 months, summarized by month. Growth in processor utilization, approaching 100% in the last month, is evident.



SightLine ForSight will be used to create baseline models for existing servers and then simulate changes in workload or configuration. Information used for building the baseline models in ForSight will come directly from SightLine Power Agents managed by the Tier 2 EA/V servers.

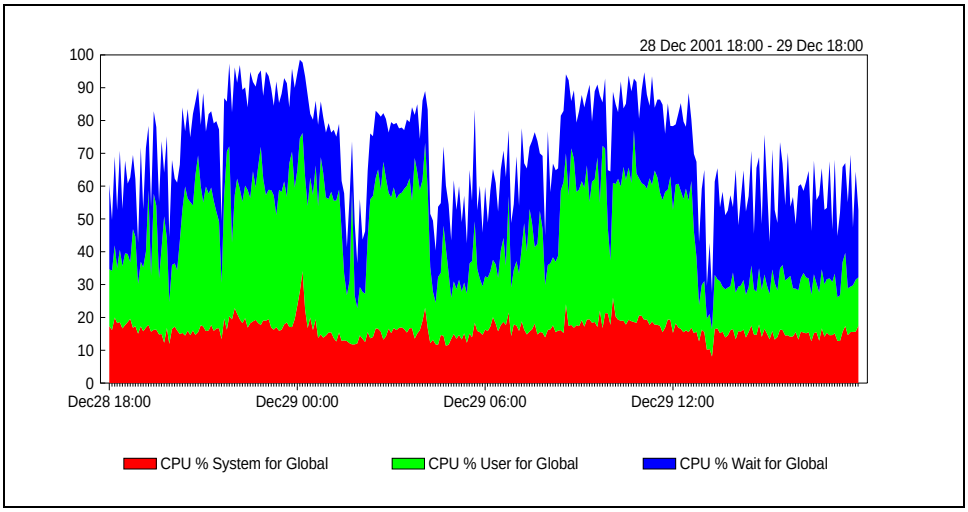
8.2 Data Management

Each EA/V server will be set up to perform centralized management functions within the Use Case SightLine deployment based on data and localization requirements. The EA/V servers will aggregate, manage, and display appropriate data. Data will be stored either on the EA/V server or on a specified network storage area as defined by the site.

Each Monitored Object (Tier 1) will be configured to maintain a Host Trace File (HTF) to provide a temporary store of data to facilitate flexible strategies for downloading data to the EA/V servers, as well as recovery of data missed should a connection be temporarily lost.

Tier 2 EA/V servers will receive data from the managed objects for individual trace files (VTFs) at 30-second intervals and save the files in compressed trace files (.VSQs). The

data transferred to the Tier 2 EA/V servers will include the metrics specified in Section 8.3.1 plus associated Event File (VEV) information. The primary purpose of Tier 2 data is for use in System Tuning and Capacity Planning. The graph below shows processor (CPU) utilization for a 24-hour period at 5-minute intervals. This data is analyzed to identify representative peak processing data for importing the baseline model in ForSight.



Tier 3 EA/V servers will receive summary data at five- to fifteen-minute intervals from the Tier 2 EA/V servers using the ServeClient feature. The data transferred to the Tier 3 EA/V servers will include the metrics specified in Section 8.3.1. This data will be consolidated into two files based on application type. The primary purpose of Tier 3 data is for use in Server Consolidation.

Tier 4 EA/V servers will receive the consolidated trace files (VCF) from the Tier 3 EA/V servers using the ServeClient feature. The data transferred to the Tier 4 EA/V server will include only the metrics required for high-level AutoWeb reporting and will be retained at thirty- to sixty-minute intervals.

8.2.1 Data Retention and Archival Support

SightLine data can be stored in a variety of ways depending on the specific requirements of the user base. The following methodology provides a comprehensive regime of data summarization and locality:

- Tier 1 Managed Objects – Host Trace Files will be stored on the Unix servers and will be sized to hold all available performance metrics for 3 days. This will serve as a temporary back up for network or EA/V server outages or for use in on-demand performance analysis.
- Tier 2 EA/V Servers – Live, circular files on the Tier 2 EA/V servers or local network storage devices will store data as it is downloaded from the server. The Tier 2 VTFs will have a 30-second granularity and contain process information in

a companion Event File (VEV). The live VTF file and associated .VEV file will be sized to hold 24-48 hours of data. This is the data file that will be used to provide information so that ForSight can create baseline models.

- Tier 3 EA/V Servers – Two consolidated history files will be created using the data received from Tier 2 EA/V servers. These files will be summarized at five to fifteen- minute intervals and configured to retain a minimum of two years of data. Data will be consolidated based on major application.
- Tier 4 EA/V Servers – Filtered, compressed, and consolidated data files will be retained by the Tier 4 EA/V servers for use in high-level AutoWeb reports. No process information aside from Workload metrics will be maintained at this level. Data will be retained at thirty- to sixty-minute intervals for a minimum of three to five years, allowing for “Year to Year” reporting.

8.2.2 POSTOP Utility

The POSTOP Utility may be used to provide various types of basic file manipulation to assist in managing the daily or historical files created with the TraceFile Manager or SightLine AutoHistory feature within the different tiers.

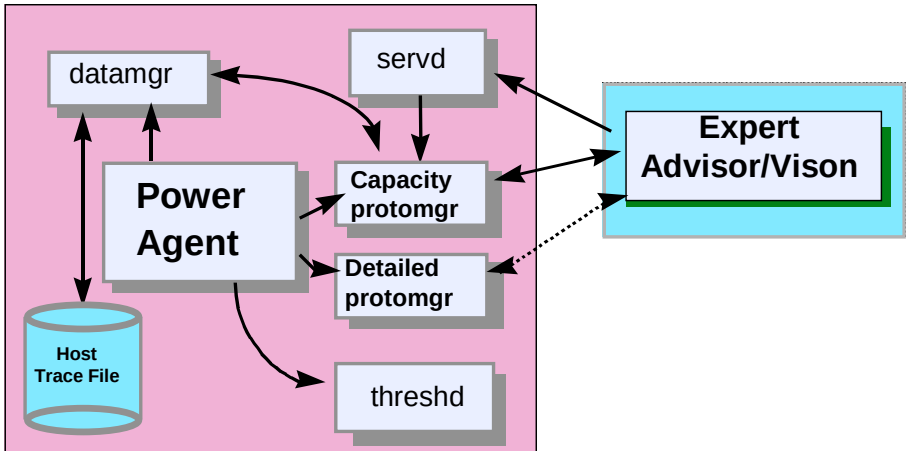
8.3 Data Collection

The SightLine Power Agents for Unix will be installed on each server within the two data centers, with default configuration parameters, to collect data from all processor cata categories listed in the table in Section 5.0.

8.3.1 Transferring Subsets Using Protomgr

While all data classes are collected by the Power Agent, two *protomgr* configuration files will be configured on each managed object. The "capacity" *protomgr* configuration will be implemented to continuously download a key subset of data to the EA/V servers at Tier 2. This will minimize network traffic and data storage requirements at the Tier 2 EA/V server. The "detailed" *protomgr* configuration will be created to download all of the detailed Power Agent data to any EA/V server when requested on an *ad-hoc* basis. Since all data will be collected by the Agent and stored in the Host Trace File, changes to the metric set transferred to the EA/V server can be made on the fly or a new connection to the entire metric set stored on the managed object can be initiated for on-demand analysis and reporting.

Server and SightLine Processes



SightLine recommends that the following types of data be included in the "capacity" *protomgr* configuration for downloading to the Tier 2 EA/V server and retained in long-term history files for trend analysis and capacity planning requirements. This subset is adequate for building the baseline models in SightLine ForSight.

- System CPU Metrics
  - Cpu % User Glob
  - Cpu % System Glob
  - Cpu % User[]
  - Cpu % System []
- System Virtual Memory Metrics
  - Metrics as appropriate for the specific platform.
    - Pgs in/Sec, Pg In/Sec, Pg Ins/Sec, Pg Ins From Pg Space/Sec, Swapins/Sec, Swap Pg Ins/Sec
    - Pgs Out/Sec, Pg Out/Sec, Pg Outs/Sec, Pg Outs To Pg Space/Sec, Swapouts/Sec, Page Space/Sec
    - Pgs Paged In/Sec, Swap Pages Paged In/Sec, Swapin Blks/Sec
    - Pgs Paged Out/Sec, Swap Pages Paged Out/Sec, Swapout Blks/Sec
    - Swap Outs/Sec, Swaps Out/Sec
    - Pgs Swapped Out/Sec
    - Swaps In/Sec, Swap Ins/Sec
    - Pgs Swapped In/Sec
    - Pg Reclaims/Sec
- System Disk Metrics
  - Disk % Busy[]
  - Disk I/Os/Sec[]
  - Disk Blks/Sec[]
- Workload Metrics
  - % Usr[], Ppt1 % Usr[]
  - % Sys[], Ppt1 % Sys[]

8.3.2 Collection Interval

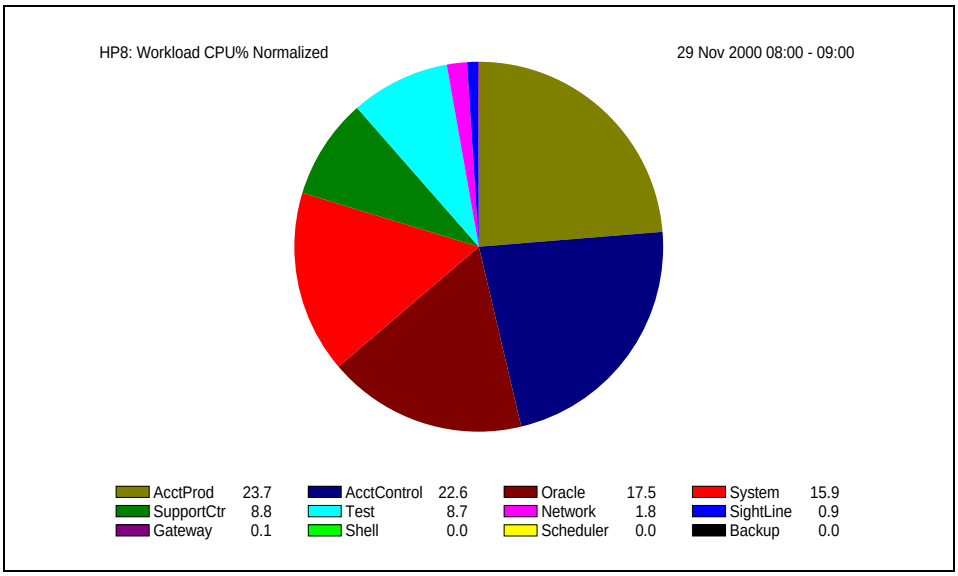
Data will be collected by the Power Agent at 30-second intervals. The primary, “capacity,” *protomgr* configuration, transferring only the subset of data identified in Section 8.3.1, will specify that the data be downloaded to the appropriate Tier 2 EA/V server at an interval of 30 seconds. The "detailed" *protomgr* configuration, transferring the full set of metrics on demand, will also specify a collection interval of 30 seconds.

8.3.3 Process Data and Filtering

The default configuration settings will be used for collecting and displaying *Summary*, *CPU*, *Memory*, *I/O*, and *Identification* Event Classes. Event data for each system will be maintained in the local data repository, in this case, the Tier 2 EA/V servers.

8.3.4 Workload Characterization

Workload definitions will be created for all major applications running on each system. Workload definitions for standard applications will be duplicated across systems as appropriate. Workload names will be specified to create unique application references where required. Additionally, processes will be grouped into workloads that will most accurately represent the business unit impacted during capacity planning studies. The graph below shows processor (CPU) utilization broken out by workload for a peak one-hour period for a single Unix system.



8.4 SightLine Deployment

8.4.1 Data Aggregation and Consolidation

The EA/V server at Tier 4 (located within the Capacity Planning Group) will connect to the Tier 3 EA/V servers within each data center through secure TCP/IP using the ServeClient feature and will be able to read one or more Trace files as if the data was coming from one or more managed object collectors located on the servers within the data centers. It is anticipated that the transmission of data between the data center EA/V servers and the Tier 4 EA/V server will be live and continuous. Data collection intervals and retention policies will be discussed in their respective sections.
To minimize network traffic and data storage requirements Tier 2 will be configured to consolidate the trace file data it receives from all local managed objects into two or more consolidated trace files based on application type. Consolidation up through the tiers will include the application of a standard filter to only transfer the required set of metrics to each level. If additional data is required at any time, the user at Tier 4 can connect directly via secure TCP/IP to the detailed data maintained in separate trace files at any tier or directly to the Host Trace File maintained on the managed object.

8.4.2 Remote Configuration

SightLine Agent Administrator (SLAA) software will be installed at Tier 4 to provide functionality to the user at that level to configure systems that are network accessible. Primarily, this will allow for remote configuration of Agent software. Managed objects will be grouped by one or more factors, which might include application group, server size, or geographic workload. Managed objects within the same group can then be configured using a single configuration file. Managed objects can belong to more than one group or can remain as “stand alone” entities.

8.4.3 Framework Integration

The EA/V servers located at each data center will be configured with the Tivoli Integration Adapter. This will enable EA/V to continuously inspect specific conditions that, when triggered, will trigger an alert in the appropriate Tivoli event format and forward the alert to the TEC for display and possible further action.

8.5 Reporting Options

SightLine’s AutoWeb feature is used as the primary reporting mechanism within the defined Use Case environment. Standard environments will be authored so that AutoWeb can automatically update and serve HTML pages from the pages and plots displayed in EA/V by turning the EA/V server into a simple Web server.

Tier 4 AutoWeb environments will be customized to display high-level reports in compliance with daily, weekly, monthly, and annual reporting requirements. These reports will be the primary vehicle for the transfer of data to management and system users.

Tier 3 AutoWeb environments can be configured to display critical metrics for the past 24-hour period and will be intended only for technical Capacity Management personnel needing to perform on-demand analysis of system resource requirements.

8.6 SightLine Installation

SightLine recommends the installation of the Unix Power Agent on all systems to be included in the Enterprise deployment. The Power Agent provides the greatest level of detail, including both workload and process-level data. Power Agent data also provides the metrics required for a direct data feed into SightLine ForSight for capacity planning and "what-if" analysis. SightLine Customer Support will work with site personnel to explore the use of existing software distribution methodologies as the preferred mechanism for software installation.

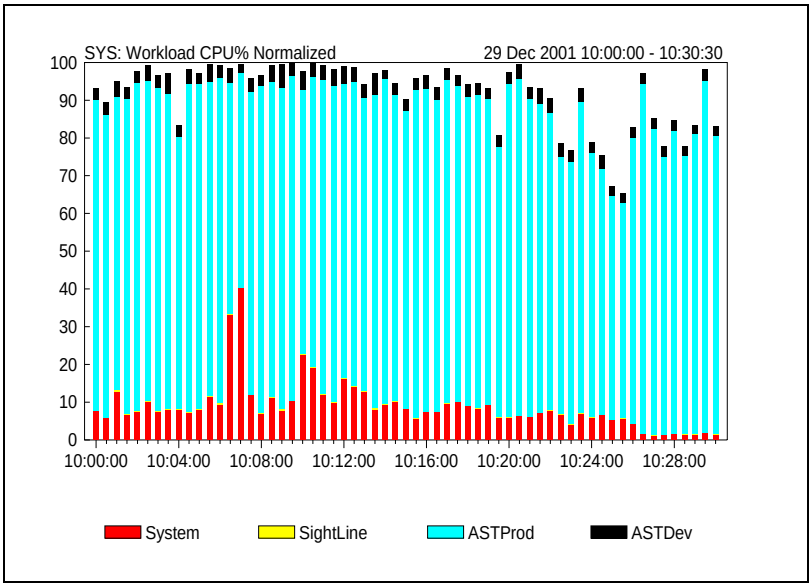
It is recommended that the SightLine EA/V software be installed on at least one workstation per application within each data center and on at least one workstation within the Capacity Planning Group. The primary function of the EA/V server(s) located within the data centers will be a tiered approach to data collection and management at the local level. The EA/V server located within the Capacity Planning Group will be configured to manage top-level data, serve Web-based reports and assist in performing Capacity Planning functions such as trend analysis, server consolidation and modeling using SightLine ForSight. EA/V installation will include the installation of all support files for managed platforms within the enterprise.

8.7 Functional Capacity Planning

With workloads defined and data being collected, the capacity planner is now equipped to begin performing the functions of the job. In most cases, the capacity planner’s job is to quantify how changes made to systems or workloadw impact service levels. Examples of future changes would be an increase in transaction volume or upgrades to disks, network devices or CPU power. For example, a simple capacity planning exercise could be:

- Anticipating a 25% increase in transactions for a major application on an already busy system, what is the impact on utilization and service levels with an increase in CPU power?
- Next, determine how the 25% increase in transaction rates affects the upgraded system.
- Finally, can workload from another system be consolidated onto the upgraded system and still provide acceptable levels of service?

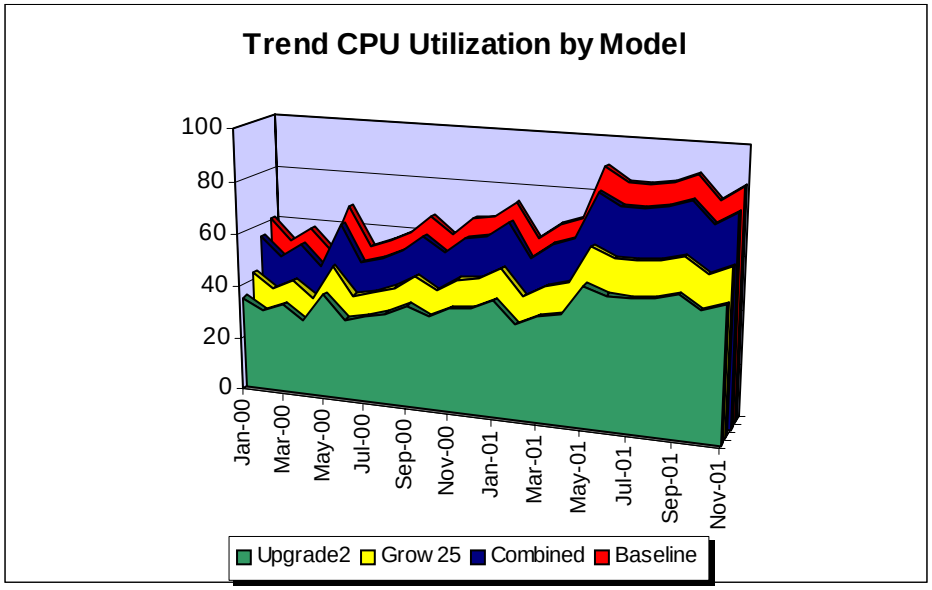
The graph below shows CPU utilization by workload for a simple system under evaluation. The data was collected at 30-second intervals over a one-half hour peak period.

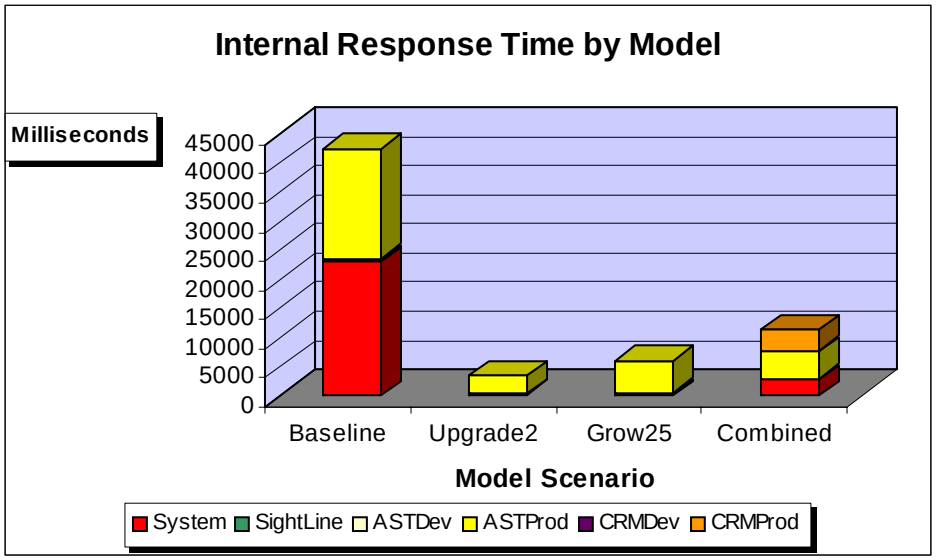


The data was imported into ForSight and a baseline model (*Baseline*) was created to represent the current hardware environment and workload mix. Three additional models were then created to simulate the anticipated changes mentioned above. Namely,

- *Upgrade2* models the addition of two more processors to the existing system
- Grow25 models the anticipated 25% increase in ASTProd transactions, and
- *Combined* models the consolidation of CRMProd and CRMDev onto the system.

The two graphs below show the results from the modeling exercises in terms of anticipated CPU utilization against historical trend data and internal response time (service level) by workload.





9.0 Conclusion

The SightLine software is a component-based application designed specifically to manage and optimize capacity, performance, and service levels of heterogeneous, multi-tier business applications. The software collects and aggregates any real-time information from your systems, independent of platforms, networks, or applications. Data consolidation and aggregation assist in performing trend analysis and SightLine ForSight provides an integrated discrete event simulation tool for modeling the performance behavior of various hardware and workload configurations.

The focus of the above Use Case has been the application of SightLine in a large-scale Unix environment. The methodology outlined in this document proposes best practices for the collection, aggregation, consolidation and retention of data in support of Capacity Management. This methodology supports the use of SightLine for:

- Reporting on and understanding the relationships between detailed operating system, kernel, and application metrics and settings for system tuning.
- Monitoring servers in real time as changes to kernel parameters, disks and workload priorities are made.
- Viewing server utilization metrics over time to identify trends in resource consumption and identify inter-dependencies in data.
- Creating long-term utilization reports to assist in identifying candidates for workload consolidation and identify peak processing requirements.
- Conducting capacity planning modeling studies to measure the relative changes to system resources in response to stimuli such as changes in workload or server configuration.
- Automating the task of daily, weekly, monthly, and annual reporting via traditional or web-based reporting mechanisms.

Although the best practices presented include guidelines for personnel involvement, programming involvement, and technical resources based on the hypothetical infrastructure, the principles can be applied to most any environment.

SightLine Systems Corporation
11130 Fairfax Blvd., Suite 200, Fairfax, VA 22030
703-563-3000
877-744-4854
www.sightlinesystems.com