

PerfCap Corporation's

eCAP Software

For

IBM pSeries

Capacity Planning and Server Consolidation



**76 Northeastern Blvd, Nashua, NH 03062, USA;
Ph +1 603-594-0222; www.PerfCap.com; Info@PerfCap.com**

Introduction

This paper is an introduction to Server Capacity Planning using eCAP Monitor, Analyzer and Planner products to provide a methodology for consultants to forecast the following:

- Performance improvement in migrating from non-IBM hardware to pSeries servers
- Upgrading from pSeries 660 class-servers to P6xx class servers
- Performing saturation analysis and identifying system bottlenecks
- Performing Server Consolidation
- Identify potential problems using workload classification techniques

Need for enterprise-level toolsets

Enterprise-level Capacity Planning has evolved from the mainframe environment where server capacity planning studies have been performed. The toolsets used in these studies have been hardware-vendor specific. Therefore, alternate platform coverage has not been readily available to easily predict the effect of migrating from one platform to another. Customers are reluctant to accept studies from hardware vendors where results indicate dramatic performance improvement over existing hardware. Frequently, such studies do not use actual customer data, and rely solely on 'industry-standard' benchmarks, such as TPCC.

Background

The eCAP capacity planning software is an independently established toolset covering all major Operating Systems including the pSeries market. This planning software uses a highly efficient performance data collector supported on all the major UNIX platforms, Windows, Linux and legacy Operating Systems such as OpenVMS. The eCAP Planner software product provides consultants or vendors with a tool which answers the following frequently asked questions:

- What server do you recommend?
- Why?
- How much headroom will the new solution provide?
- What is the upgrade path for the new solution?

eCAP Methodology

The following section describes a capacity planning methodology and how various components of the eCAP products are used to fulfill various steps of the methodology.

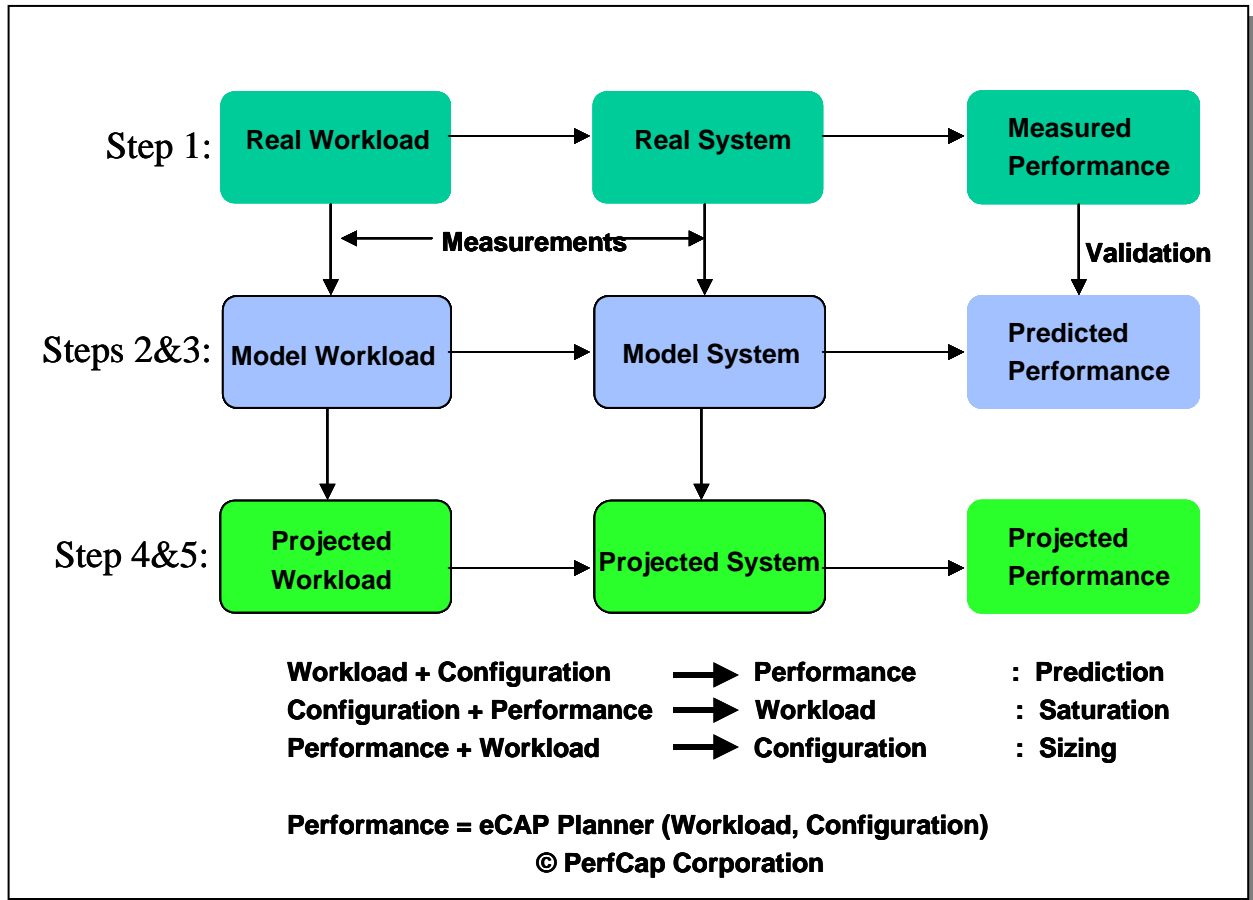


Figure 1: eCAP Capacity Methodology

Step1. Data Collection:

During this step performance data is collected for an extended period of time. Subsequently a subset of the collected data is selected which represents peak activity. This step provides measured performance of the system during peak periods and base data to create current system and workload models.

Step 2: Performance and Workload Analysis:

During this step various workloads/applications running on the system(s) are grouped into various classes to create groups of workloads. This classification can be based on process,

commands, business functions or users. Each workload class specifies the total resources (CPU, memory, disk etc.) required to perform a specific task

Step 3 Creating base model and validation:

Use the eCAP planning software product to automatically create a base model of the current workload and system(s) configuration. Compare the predicted performance from step 3 against the actual measured performance from step 1 to verify that the variation is within acceptable limits.

Step 4 Saturation Analysis:

During this step workload intensity is gradually increased until one or more components reach their maximum capacity. A set of graphs depicting various performance metrics (response time, utilization, throughput rates etc.,) can be generated which displays the respective metrics at various levels of increased workload. This step is instrumental in identifying bottlenecks on the current system(s).

Step 5: Prediction and Sizing:

Once system(s) bottlenecks (CPU, memory, disk, network) are identified, the saturated component(s) can be upgraded. Subsequent models can be run repeatedly using the upgraded components until a configuration is obtained that will support the required workload.

This methodology can be summarized as follows:

- Components
 - Workload (applications, transactions, users)
 - Configuration (system hardware)
 - Performance.

These components can be considered as three variables of an equation. Given any two components, eCAP can predict the third component. For example:

1. If the workload is given (e.g. we know how many users must be supported, or how many transactions are to be processed) and configuration is defined (e.g. hardware to run the workload is fixed) then eCAP can **predict** what is the best possible performance.
2. If the configuration is defined and a certain performance is expected then the maximum workload that will **saturate** the system can be predicted.
3. If the workload is defined and expected performance is specified then the required system **size** can be determined.

eCAP Product Set

The following is a list of various product components required for performance monitoring and capacity planning:

- eCAP Monitor
- eCAP Analyzer or PAWZ Agent/Server
- eCAP Reducer
- eCAP Planner

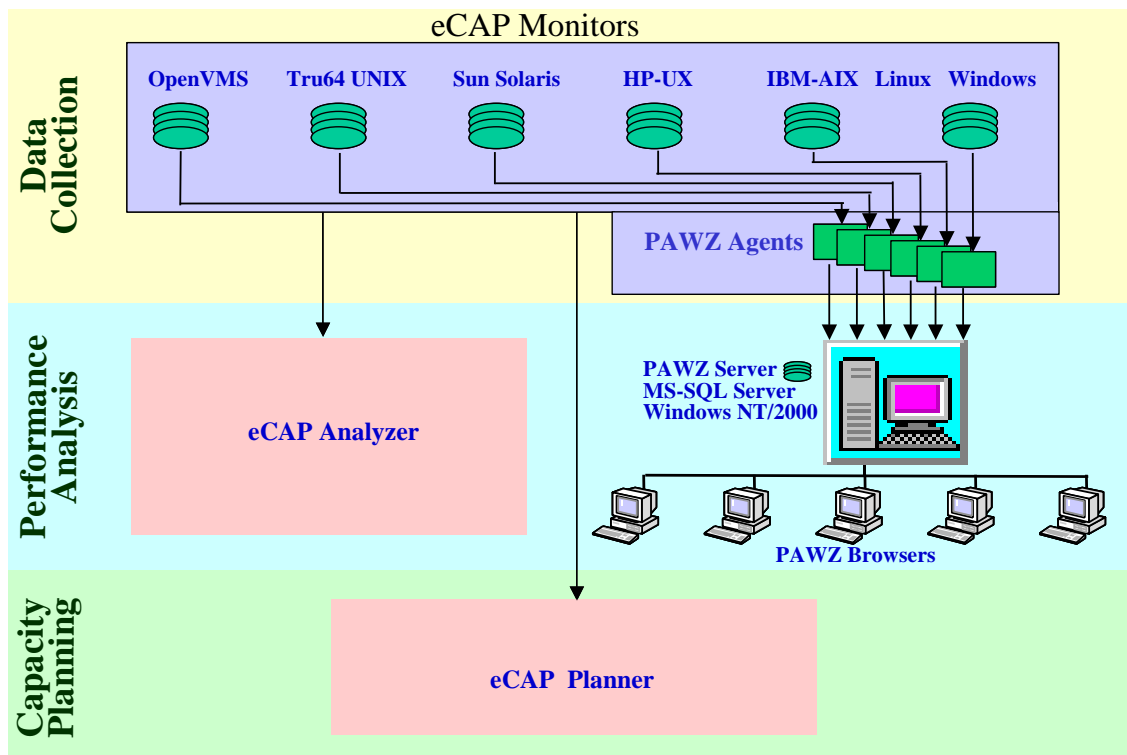


Figure 2: eCAP and PAWZ Product Components

eCAP Monitor

The eCAP Monitor is used in the first stage of the capacity planning process to gather system wide performance data for CPU, Memory, DiskIO, and network components. It is a stand-alone product and its primary function is to collect system information such as cpu utilization, per-processor metrics, memory utilization, command and user resource information, network io, etc. Unlike standard toolsets, such as **sa**, the data collected by the eCAP Monitor can be classified into system workloads. The eCAP Monitor runs continuously on the target system with the *cpcunix* command and is designed to use minimal overhead, to record process events (such as process creation or deletion) and general resource metrics (CPU usage, per-disk IO rate/ queue).

eCAP Monitor output can be viewed by the UNIX-based eCAP Analyzer for interactive system analysis, or by PAWZ (Performance Analysis Web Zone) for automated web-based reporting and trending analysis. eCAP Monitor data can also be used by the eCAP Planner for system sizing and saturation analysis.

The following Table 1 summarizes the current eCAP Monitor supported platforms and their characteristics:

Platform	OS Version	File prefix	Default Sample rate	CPU overhead %	File size/ day
AIX	4.3.3,5.1,5.2	ecpa	to 2 second	1-3	100mb
Solaris	2.6,2.7,2.8,2.9	ecps	to 2 second	1-3	100mb
HPUX	10.20,11.00	ecph	to 2 second	1-3	80mb
Tru64	4.x, 5.x	ecp	to 2 second	1-3	80mb
Intel Linux	2.4 kernel	ecpl	to 2 second	1-2	40mb
Windows	NT,2000,XP	ecp	to 2 second	1-3	30mb
VMS	VAX,Alpha	n/a	n/a	0-1	70mb

Table 1: eCAP Monitor Supported Platforms

All eCAP Monitors collect process cpu, memory, io, network data by default.

Sampling rates can be controlled by the user. A good compromise between file size and granularity can be seen by the following command line to start the eCAP Monitor for a typical capacity planning study

cpcunix -poll 2000 -dump 60

The eCAP Monitor is the core of the PerfCap capacity planning methodology, and must be deployed on each node used for the sizing exercise. Installation on any supported platform does not require a reboot or other system interruption. We recommend that you use a file system large enough to accommodate the raw data files for one week's worth of data for the study.

Licensing

eCAP Monitor is licensed on a per nodename basis, with a text file holding a license key. There are no processor limitations in the product licensing and pricing is independent of processor speed or number of CPUs in the monitored server.

eCAP Analyzer/PAWZ

Performance data analysis is the second stage in the capacity planning methodology.

PerfCap offers two products, eCAP Analyzer and PAWZ to accomplish this task. They present system performance in graphical format, and can be exported into csv format for correlation with other data sources.

eCAP Analyzer is supported on a UNIX-based platform whereas PAWZ is a centralized web-based console with a fully automated data loader, analysis engine and a Microsoft SQL database.

Both products analyze performance data off-line. Specifically, the data is analyzed away from the physical system. The UNIX-based product, eCAP Analyzer, can be installed off-site so that eCAP Monitor files can be analyzed away from the site of the study. eCAP Monitor files can be compressed considerably for portability using *compress* or *gzip*.

Data can be exported from the eCAP Analyzer motif interface into CSV for use with excel etc.

Like the eCAP Monitor, the eCAP Analyze runs native on all UNIX platforms supported by PerfCap. eCAP Analyze does not have to run on the same machine which is collecting data. Moreover, an IBM installation of eCAP Analyze can examine eCAP monitor data from AIX, Solaris, and NT.

In this way you can utilize one eCAP Analyzer installation to size a server to a pSeries from another vendor.

The example uses eCAP Analyze rather than PAWZ. PAWZ is designed for enterprise wide performance analysis. Typically for a capacity plan, eCAP Analyze would be used by the consultant for detailed analysis tasks because eCAP Analyze does not have to be installed on the local client site

eCAP Reduce

The third element in the PerfCap Capacity Planning methodology is the eCAP Reducer. The eCAP Reducer takes the peak workload data provided by the eCAP monitor and identified by the eCAP Analyzer and reduces its volume by aggregating the data for each process/user etc. At the end of this step data collected by eCAP Monitor (.cpc file) is also converted to a platform independent format (.merg file) so that it can be read by the eCAP Planner on any platform.

eCAP Reducer also produces a text file (.proc) that provides resources consumed by each process. This file helps to identify process resource consumption and characterize workloads into different classes.

The eCAP Reducer for UNIX can convert any UNIX or NT file to .MERG format, whereas the eCAP Reducer for OpenVMS converts all eCAP Monitor files.

eCAP Plan

eCAP Plan is used to answer all “what if” questions such as:

- What will be the impact of upgrading an existing pSeries Server to a new Power IV pSeries Server.
- What will be the impact of upgrading a CPU on overall performance
- What will be the impact of expected increase in workload over next quarters
- What will be the expected performance of consolidating existing multiple physical servers into a pSeries 690 server
- What will be the expected performance when a set of applications are migrated from other vendor servers to a pSeries server.

All these scenarios are possible, and eCAP Plan has been used over the last 20 years to answer these and other questions relating to infrastructure performance.

How eCAP Plan works

Put simply, eCAP Plan takes the data captured by eCAP Monitor and uses analytical queuing modelling techniques to build a mathematical representation of the computer system, workloads running on the system and network traffic. eCAP Plan contains an extensive infrastructure database giving performance specifications for a wide range of external devices including CPU, Disk, Memory management, Switch, Network, Router, Fiber channel, SCSI etc. These are used in the what-if analysis stage.

Once a mathematical ‘model’ is built and validated against the measured performance data (captured by eCAP Monitor) the analyst can play ‘what-if’ scenarios by varying workload or physical change in all parts of the infrastructure. The model understands all aspects of the computer as a series of queues, and takes into account:

- Page size, memory modelling
- Disk IO change - RAID, Host or Controller-based devices
- SAN - switch and fiber channel infrastructure
- CPU speed
- Multiple processors
- Effect of workload change in
- Number of users
- Transaction throughput
- New applications

For this paper, the easiest way to demonstrate the power of the capacity planning meth-

odology is for a walk-through of a typical study. This study is typical of many performed over the years for all Open system vendors today. Screenshots and explanatory text are provided.

Case Study 1: Capacity Plan

For purposes of this paper, two sample Capacity Planning studies are presented as a walkthrough example of the tools in actual use. Please note that this example is not a substitute for training

There are two types of studies presented:

1. Upgrading a pSeries 680 to newer pSeries hardware
2. Migrating 3 Microsoft Windows physical servers to a single physical VMware IBM server and three virtual servers

Case Study A: pSeries upgrade

The customer scenario is as follows:

Requirement

The customer has an existing application running on a database server and n number of application servers. Figure 3 shows the current configuration as drawn by the eCAP Planner. It is believed that in the next 12 months, over 100% growth will occur within the application. This growth will be driven by increases in application functionality, organic growth in users and report generation caused by wider deployment of the application to end-users.

- The job of the IT department is to size the appropriate database server to meet this expected growth and plan future headroom.
- The current pSeries P680 is saturating at peak time.
- Bottlenecks must be identified in the hardware
- The solution provided must be Scalable
- Allow for 100% growth
- Reduce the total number of CPU's from the current 10-way for licensing reasons

The study will result in a recommendation for replacement hardware. For this study, assume eCAP Monitors and/or PAWZ products have already been deployed. The

following steps ‘walk-through’ all steps indicated above as part of the capacity planning methodology

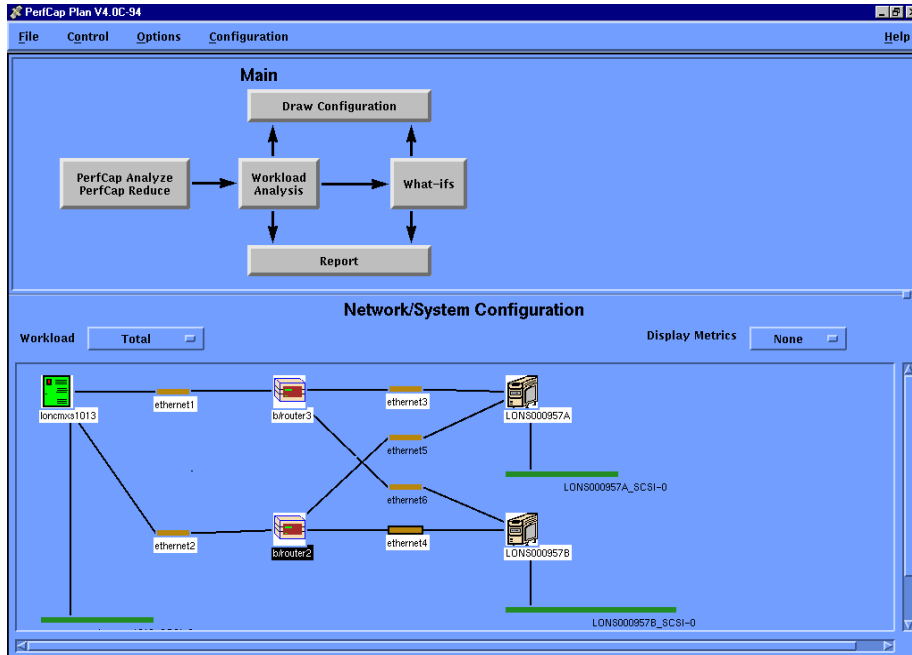


Figure 3: Current Configuration

Step 1: Collect performance data

Products used: eCAP Monitor.

Initially, deploy the eCAP Monitor on the application and database server in question, one for every server. The eCAP Monitor should run for as long as it takes to get an indication of the application’s business cycle. For example, if an application has peak data only on the third Thursday of the month, then the eCAP Monitor must be running during that peak period of time.

In some cases, no one may have any idea when the peak application activity occurs. In this case the eCAP Monitor must be run for a month or more to capture all the business cycles for that application.

During this collection period, it is strongly recommended that the analyst understand how the application works, what are the different components and their dependency on each other, what is happening on a daily basis, etc., to get a general feel for the transaction flow of the application. It is also vital for the analyst to understand the following:

- What are the key system processes of the application?
- What do they do?

- Are they expected to grow?
- What happens to the business when these processes go slowly?
- What is the exact server configuration (CPU, memory, IO, network)?

In our example, we know from the application that the application cycle is monthly. The first two or three days of each month is when the application peak occurs. We also know the batch start-end times, and we also know at what time users require access to the system.

Step 2: Selecting Peak Workload Interval

Products used: eCAP Analyze, eCAP Reduce

Operation: Enter *plan* at the command line to bring up the initial window (see Figure 4 below). This window has two sections. The top section guides one through various stages of capacity planning steps. The bottom window is used by the Planner to draw a picture of the current configuration. Start by clicking on PerCap Analyze/PerfCap Reduce and select the eCAP Monitor files from a local directory. Click the Analyze button as in figure 4 to see the graph data for the entire data collection interval.

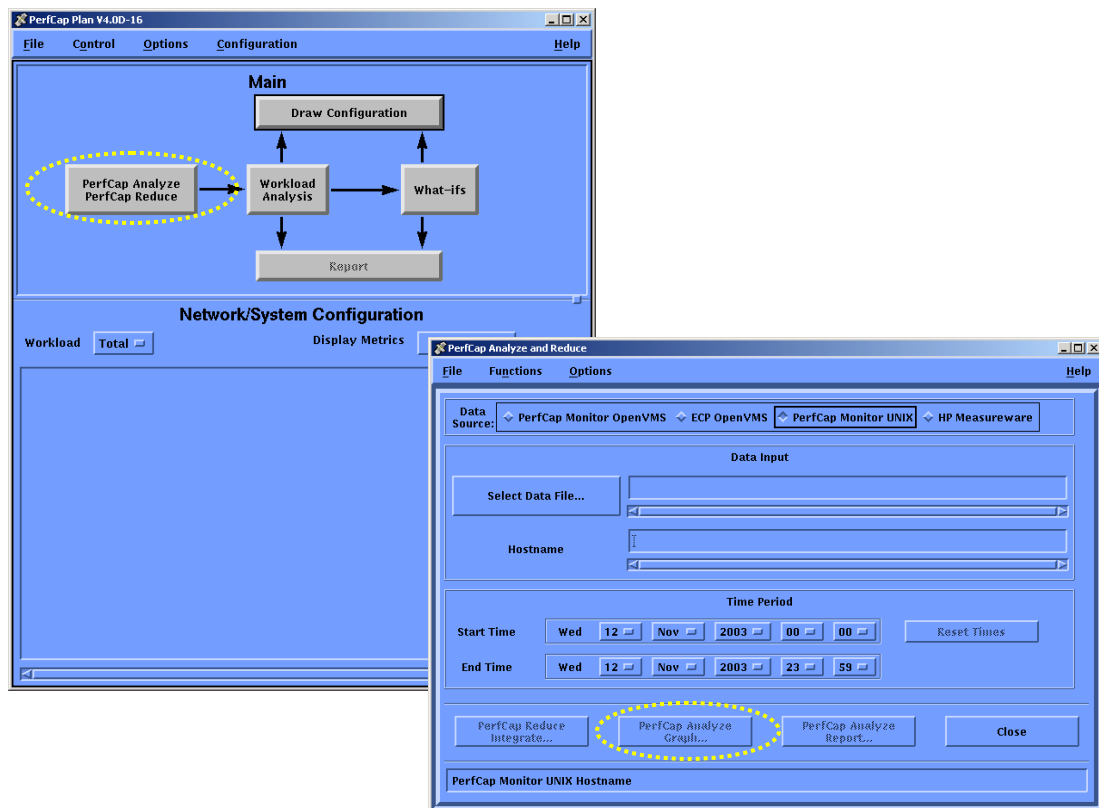


Figure 4: eCAP Planner Main Window and eCAP Analyzer

For this study, the first stage is to identify the data to be used for modeling. In this example, several days' data has been loaded into eCAP Analyze for general examination. PAWZ also offers the same graphical interface through a web-based application.

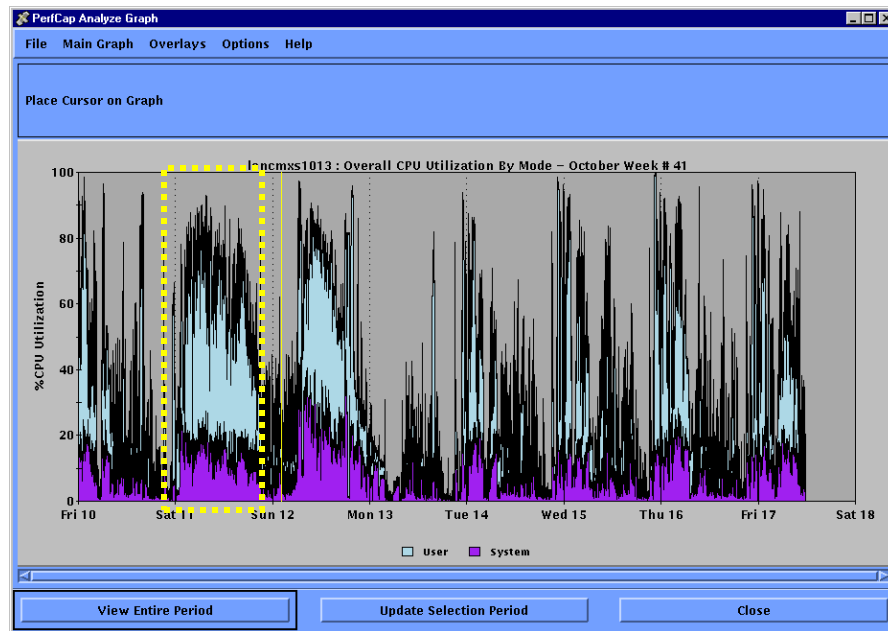


Figure 5: Trend Data and Peak Day

From the trend data in Figure 5, we can select a key day for this application's cycle and examine the data within PAWZ or eCAP Monitor to verify that the peak utilization represents actual workload, and not spurious system events such as an unusually long backup.

In this case, Saturday was interesting for the batch data. Although not the highest peak of CPU over the five days, the weekend is a heavy application workload. Further investigation revealed that between 02:09 and 11:29 Saturday morning (Figure 6) was the peak period for an initial capacity planning study.

Study findings



Figure 6: Selected Day/Interval

From this daily graph of CPU utilization, high resource consumption is seen from early morning until 11-12 noon. The application developers confirmed that this is normal application activity. It provides a view of whether or not the application was running more slowly or more quickly than usual for that period. The answer was 'normal application activity'. However, the disk IO profile was exceptionally high (20k io/sec) - further analysis was performed later in the capacity planning stage to assess if the disk bottleneck was released, how long the CPU could cope with the workload.

Step 3: Define and validate workload characteristics

The output from eCAP Analyze and the .PROC file from eCAP Reduce will show various application processes of interest. The next step is characterizing various workloads running on the system during the selected interval. This is accomplished by creating different classes of workloads based upon processes, commands and users. Using the eCAP Planner click 'Define WC' button and enter key application processes as individual workloads. Use the following screen as in Figure 7, to define workload classes.

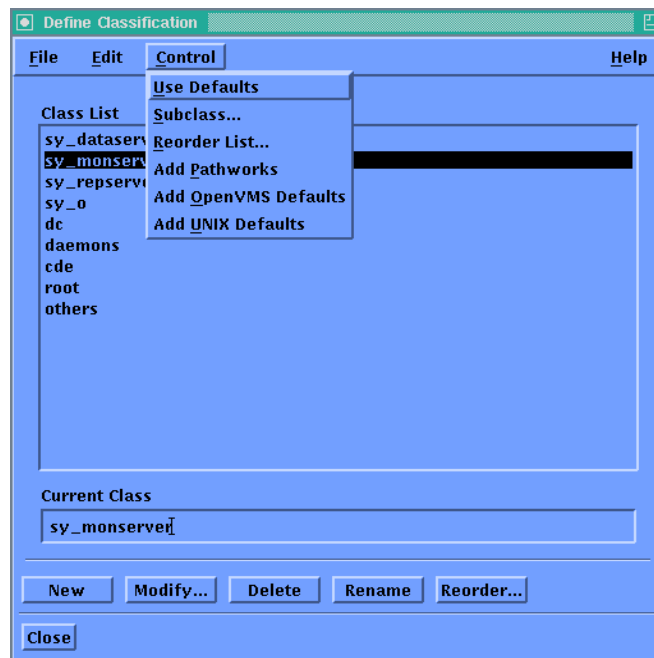


Figure 7: Workload Classification Scheme

The full workload classification of a system will look like the screen below (Figure 8), a top-down configuration. Important application processes must always come first.

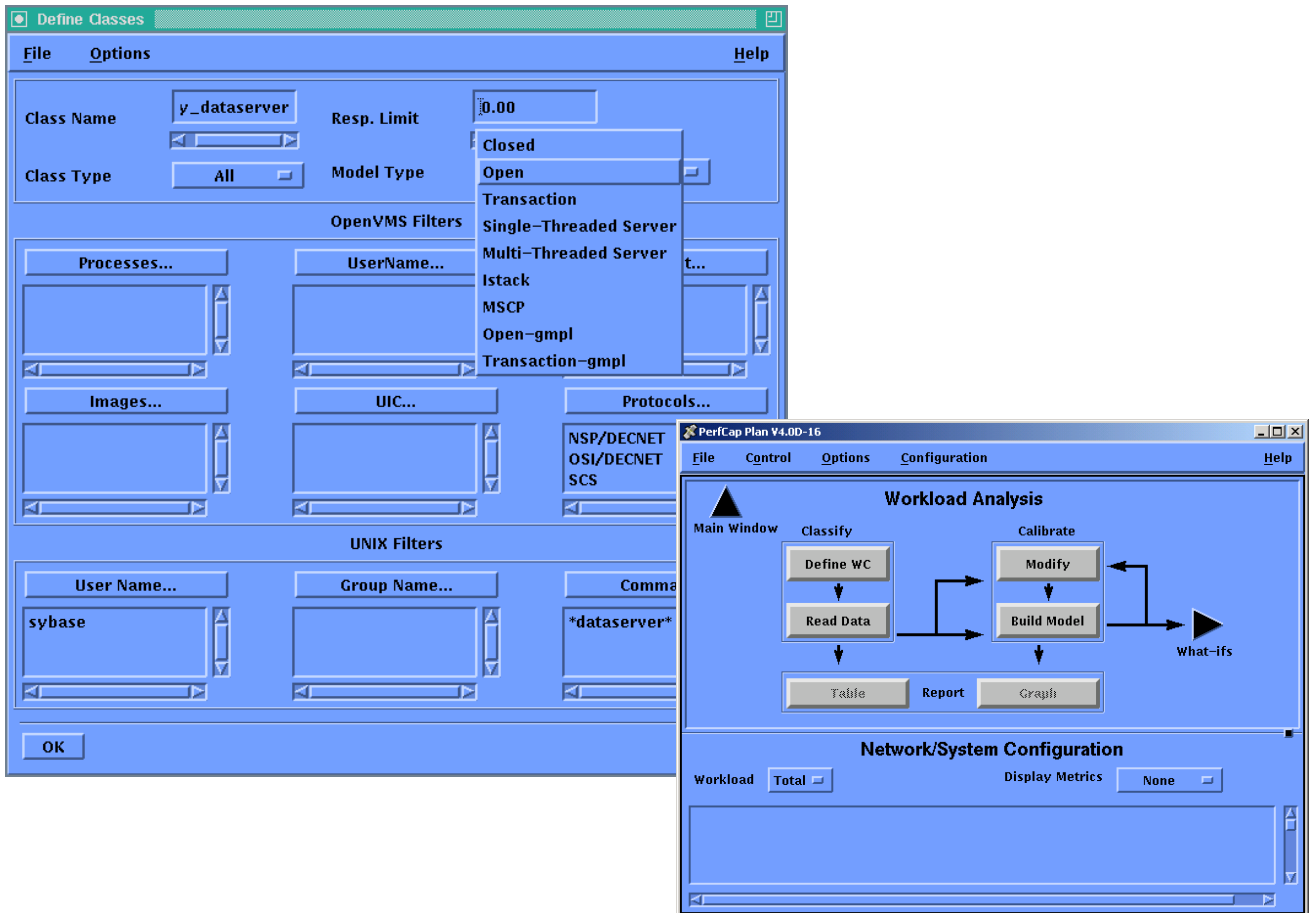


Figure 8: Detail Workload Characterization

Having defined, and saved the workload, use the 'READ DATA' button to select the .MERG file produced by the eCAP Reducer. During this step eCAP Planner will calculate the resource requirement for each workload class. As the tool reduces the data, any unknown devices not recognized by the tool will be highlighted in tabular form. Users can manually define characteristics of a device that are not recognized by eCAP Planner or are not in its database of devices. At the end of this step a workload profile for each class will be generated

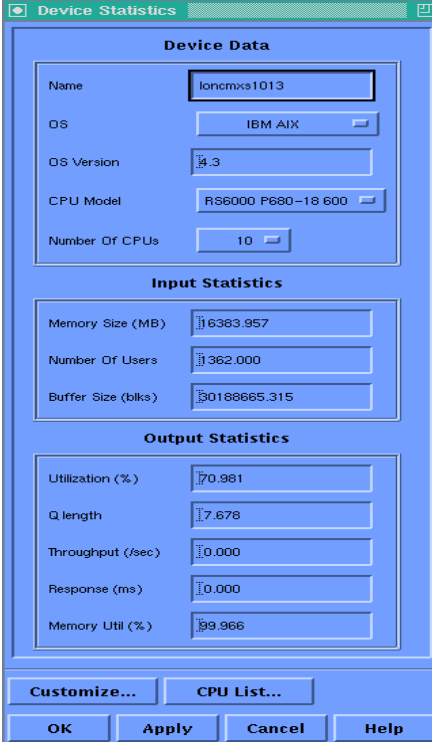
Having read the .MERG file, and identified all devices, check the quality of the workload classification by selecting the 'REPORTS' button, and selecting 'USER REPORT' in the dialog.

This report file will list all process names classified by your classification scheme and the appropriate command line identifying such workloads. 'OTHERS' should be as close to 0.00% as possible. Validating workloads is the most important element to a successful study. Incorrect classification can lead to highly misleading results.

Step 4 Configure System IO, check memory

eCAP Monitor collects data available to the OS level. For IO, AIX may not understand that a logical volume consists of a RAID01/Controller-based volume set. For applications showing high IO throughput, it may be necessary to use the eCAP Planner's drawing toolset to correctly assign logical IO to the actual physical IO devices.

Verify that the CPU and memory have been correctly identified. In this case the disk set defined is a third party disk array. Because of the proprietary nature of the storage, disk sets have been classified as a large RAID01 Controller-based set, and fiber channel cards have been configured with the correct 1Gb throughput.



The screenshot shows a 'Device Statistics' dialog box with the following data:

Device Data	
Name	loncmxs1013
OS	IBM AIX
OS Version	4.3
CPU Model	RS6000 P680-18 600
Number Of CPUs	10

Input Statistics	
Memory Size (MB)	16383.957
Number Of Users	1362.000
Buffer Size (blks)	30188665.315

Output Statistics	
Utilization (%)	70.981
Q length	7.678
Throughput (/sec)	0.000
Response (ms)	0.000
Memory Util (%)	99.966

Buttons: Customize..., CPU List..., OK, Apply, Cancel, Help

Figure 9: Editing Device Characteristics

Step 5. Check workload and build model

Check that the workload CPU looks appropriate for each application, and also verify that the overall CPU returned in eCAP Plan is the same as what was seen using the eCAP Analyzer or PAWZ. Then use the 'PERFORMANCE PREDICTION' button to build a baseline model of the application. This will reveal more workload-based data - not just

CPU per workload, but also IO, memory, and average number of concurrent processes (MPL) for that workload.

Model build

This will build an analytical representation of the system. The CPU utilization shown by the model should match the value measured/reported by eCAP Monitor/Analyzer. New values for metrics will appear for queue length, response time, and throughput.

A stable and validated model has now been built. At this stage, save the model in the DATA EXPORT/IMPORT dialog to a .BASE file, so it can be re-loaded.

Step 6: validate model

Using the REPORT button, select the VALIDATION report dialog. This will compare the actual collected data values with the values generated by the model engine. Model error rates in CPU utilization or other metrics indicate that the workload classification is faulty. See PerfCap eCAP users guide on how to calibrate the model.

Validation Report

Metric	Measured Result	Modeled Result
Node Name	loncmxs1013	Loncmxs1013
CPU Util. %	70.9	70.9

Table 2: Model Validation

Step 7: what-if analysis

Product used: eCAP Plan

Having built a validated model of the application, what-if scenarios such as increased workload or CPU system changes can be run. Use the SATURATION ANALYSIS button in the What-If screen.

Bottleneck Analysis

The first step in saturation analysis is to determine which component on the system will become the bottleneck as the workload is increased. Using the SATURATION ANALYSIS dialog, enter a step size of, say 25%, and give this test a descriptive name. Results of the saturation analysis will disclose when the system will become saturated, and which device will become the bottleneck - CPU, disk, memory, and network?

Study findings - Saturation

IO Analysis

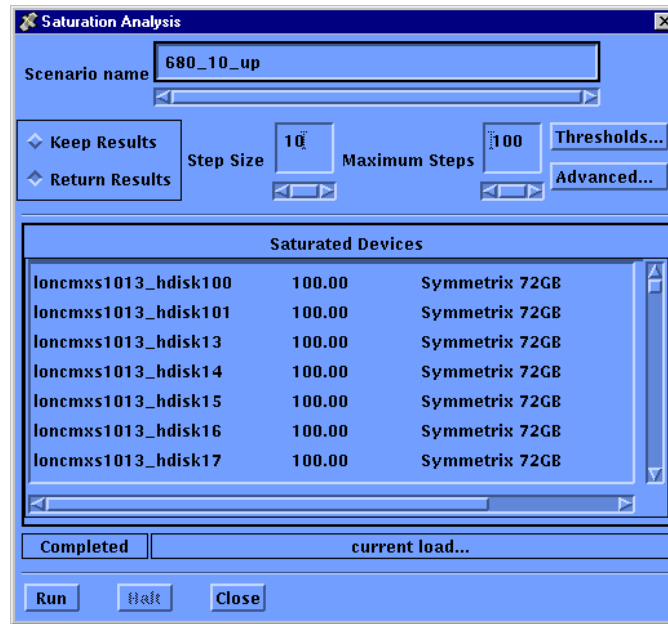


Figure 10: Saturation Analysis – Initial Bottleneck

In our case it has detected that some of the disks are saturating. Figure 10 above shows a list of saturated disks and their utilization.

Response time analysis

Using the eCAP Planner tool, the response time breakdown graphs as seen in Figure 11, further confirm that transactions are spending most of their time waiting for the IO to complete.

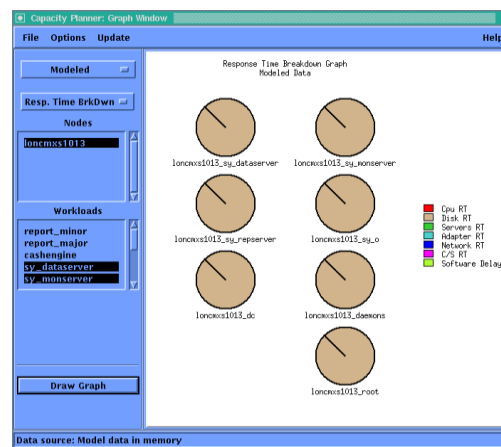


Figure 11: Response Time Breakdown

CPU was queuing an average of 7 outstanding requests/second - fair for a 10-processor configuration. With the apparent IO bottleneck, it was deemed that changing the CPU model without addressing the IO issues would worsen the overall performance - the quicker the CPU, the more IO load.

Saturation Analysis

Figure 12 below shows the average response time, as the workload is increased/decreased. Since the system is already saturated, in order to get a better understanding of the relative response time, eCAP plan has the ability to take a workload into a negative direction. The current system is already past the knee of the saturation curve.

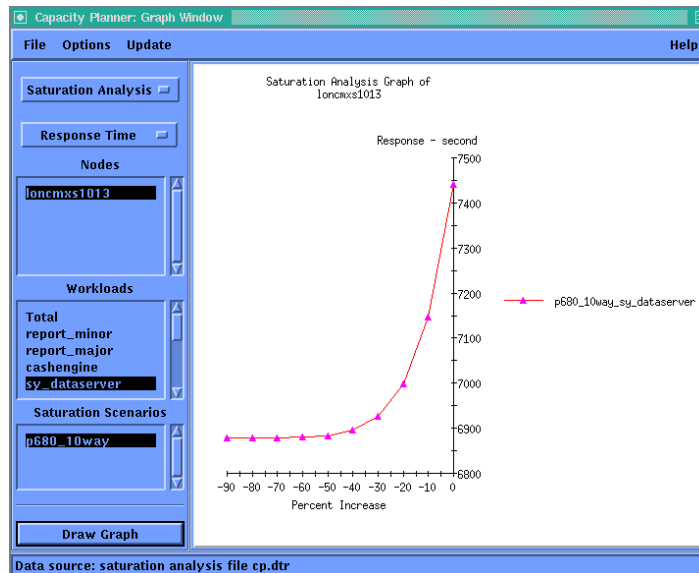


Figure 12: Current Saturation Point

IO Upgrade

There is an improvement in response time once the IO subsystem is upgraded. Figure 13, below shows the improvement in response time. However, simply applying an IO upgrade will provide only about 30% growth compared to the required 100% growth.

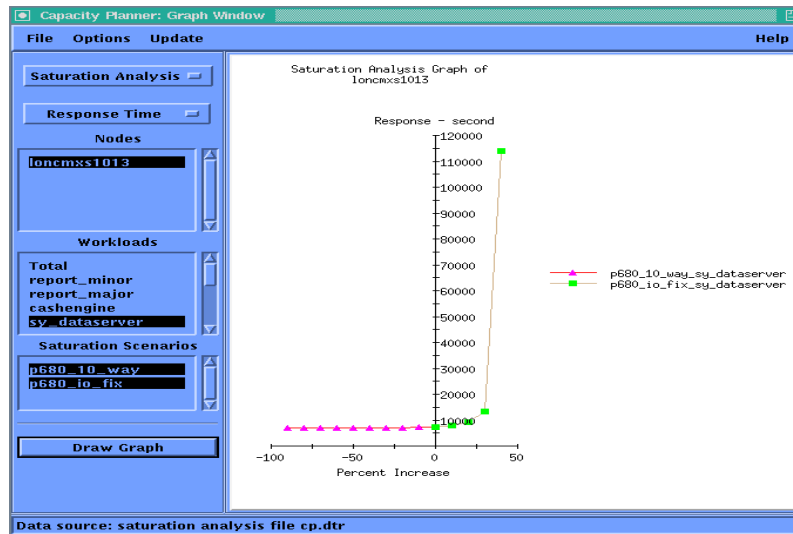


Figure 13: Saturation Point After IO Upgrade

CPU Upgrades

The next step is to upgrade the CPUs. First we will try upgrading to a 4-way P660 and then to an 8-way P660. The Figure 14, below shows the average response times in each of the four scenarios - current configuration, IO upgrade, 4-way P660 upgrade and 8-way P660 upgrade. The final upgrade to an 8-way P660 meets 100% growth and has 50% additional headroom.

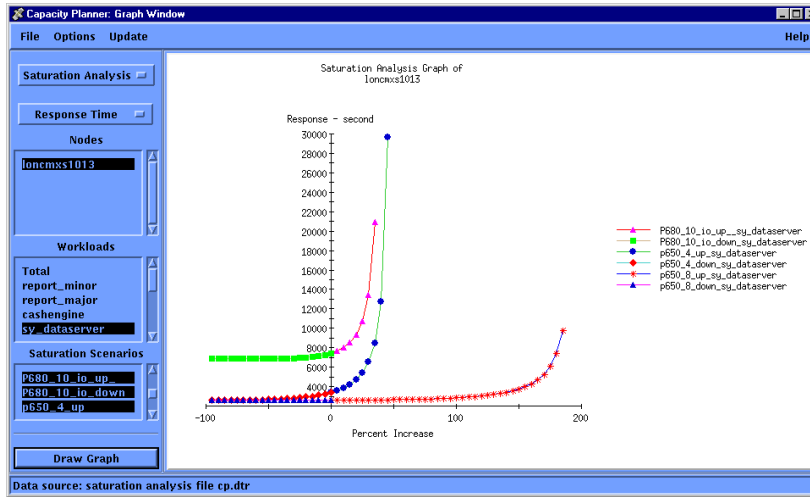


Figure 14: Saturation Points – After IO Upgrade, With 4-Way P660 and With 8-Way P660

CPU Utilization:

The graph below illustrates CPU utilization of the current, 4-way and 8-way P660 server.

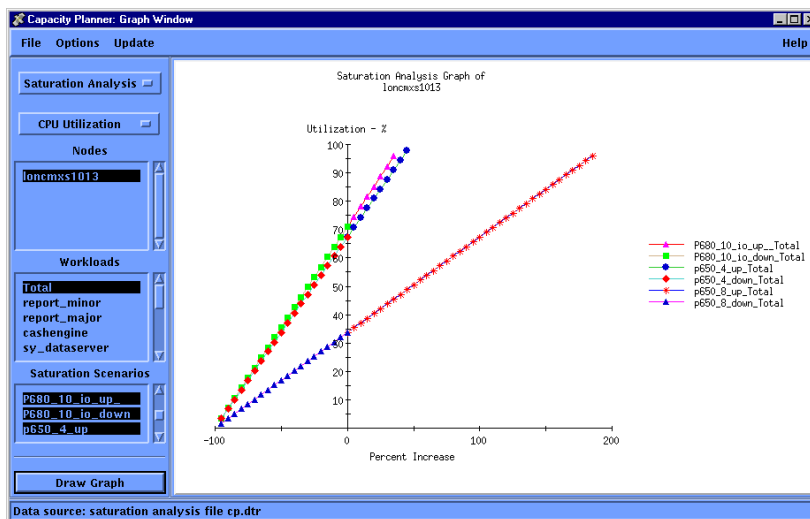


Figure 15: CPU Utilizations – After IO Upgrade, 4-Way P660, 8-Way P660

Conclusion:

The following is a summary of the study:

1. The current configuration is past its saturation point.
2. The first bottleneck is the IO subsystem. Migrating to faster drives will remove this bottleneck however will not provide additional headroom for 100% growth.
3. Upgrading to an 8-way P660 will meet 100% required growth and will have 50% additional headroom.

Case Study 2: Server Consolidation

Background

The following case study is typical of those organizations willing to undergo server consolidation. The following study was purely a technical exercise, to see if a relatively large number of modestly powered servers could co-exist on IBM's server consolidation platform of choice - a P690.

The study did not investigate the business issue of sharing such infrastructure, but did demonstrate how such an infrastructure could be charged back to its users

Consolidation Candidates

For this study the following criteria were applied to the applications deemed suitable for consolidation

- They all had to be capable of running AIX 5.2
- They all had to run the same DBMS
- Production/Development and Staging were considered
- They all had to be database servers.

The following is a list of servers that needed to be consolidated on a P690 Power4 Turbo 1.7Ghz.

Host	OS	Model	CPU	Spec95	Memory	Storage	FC	Use
03	5.1	IBM 7026-M80	8	24.1	8192	Remote	2	Production
04	5.1	IBM 7026-M80	4	24.1	4096	Remote	2	Test
05	5.1	IBM 7026-M80	8	24.1	8192	Remote	2	Production
06	5.1	IBM 7026-H80	2	21.7	4096	Remote	2	Development
22	5.1	IBM 7026-6M1	8	39.9	16384	Remote	2	Production
23	5.1	IBM 7026-6M1	8	39.9	16384	Remote	2	Production
24	5.1	IBM 7026-6M1	8	39.9	16384	Remote	2	Test
25	5.1	IBM 7026-6M1	8	39.9	16284	Remote	2	Production
11	5.1	IBM 7026-6H1	4	28.1	8192	Remote	2	Development
13	5.1	IBM 7017-S85	10	28.1	16384	Remote	4	Production
15	5.1	IBM 7017-S85	10	28.1	16384	Remote	4	Production

Table 3: Current Servers to be Consolidated

Host	OS	Model	CPU	Spec95	Memory	Storage	FC
NewA	5.2	P690 Power4 Turbo 1.7Ghz	64	74	64Gb	Remote	32

Table 4: Consolidated Server for Evaluation

Consolidation Requirements

- The target machines (2: one in production, one in DR) must be capable of running each application with at least 100% headroom from its original inception
- The target machines must also have reserve CPUs, memory and fiber channel cards, capable of handling additional applications or strong application growth in the short term. All LPARS must not be in use at initial implementation

Data Collection

All machines in the study had eCAP Monitors enabled continuously. PAWZ was used for trending historical performance, in order to work out the best time for each machine to analyze capacity planning data. Since the study includes production and non-production machines, usage patterns varied considerably.

For each machine, peak workload was selected and fed into eCAP Analyzer. Each workload for most machines was from a different time, and represented peak application workload – on-line users, batch activity or development work. Non-application workload, such as backups, was excluded

Workload classification

All systems under study were running the same database, with different database versions. This meant that a single valid workload classification for one platform would, with minor changes, be valid for all machines in the study.

In this case the workload was classified in the following scheme

- Data server processes
- Replication server processes
- Monitor server processes (database monitoring)
- Other database processes not falling into the first 3 categories
- Message transfer processes
- System data collectors (such as eCAP Monitor)
- Other enterprise management processes
- Other application processes not classified above
- All other processes

Performance Analysis

All systems were loaded into the eCAP Analyzer for identification of specific capacity planning periods of interest.

The hardware configuration for the systems was obtained from PerfCap FindIT inventory collector, and network configuration was validated accordingly.

From PAWZ, trending analysis established application cycles where high application workload could be identified.

Preparing for Capacity Plan

Peak workload data from all machines was reviewed and selected. An extra machine was also read into the tool with low utilization. This additional machine will become our consolidation target, a P690.

A centralized storage model has also been incorporated into the Capacity Plan

All systems were analyzed with eCAP Analyzer, and fed to eCAP Plan.

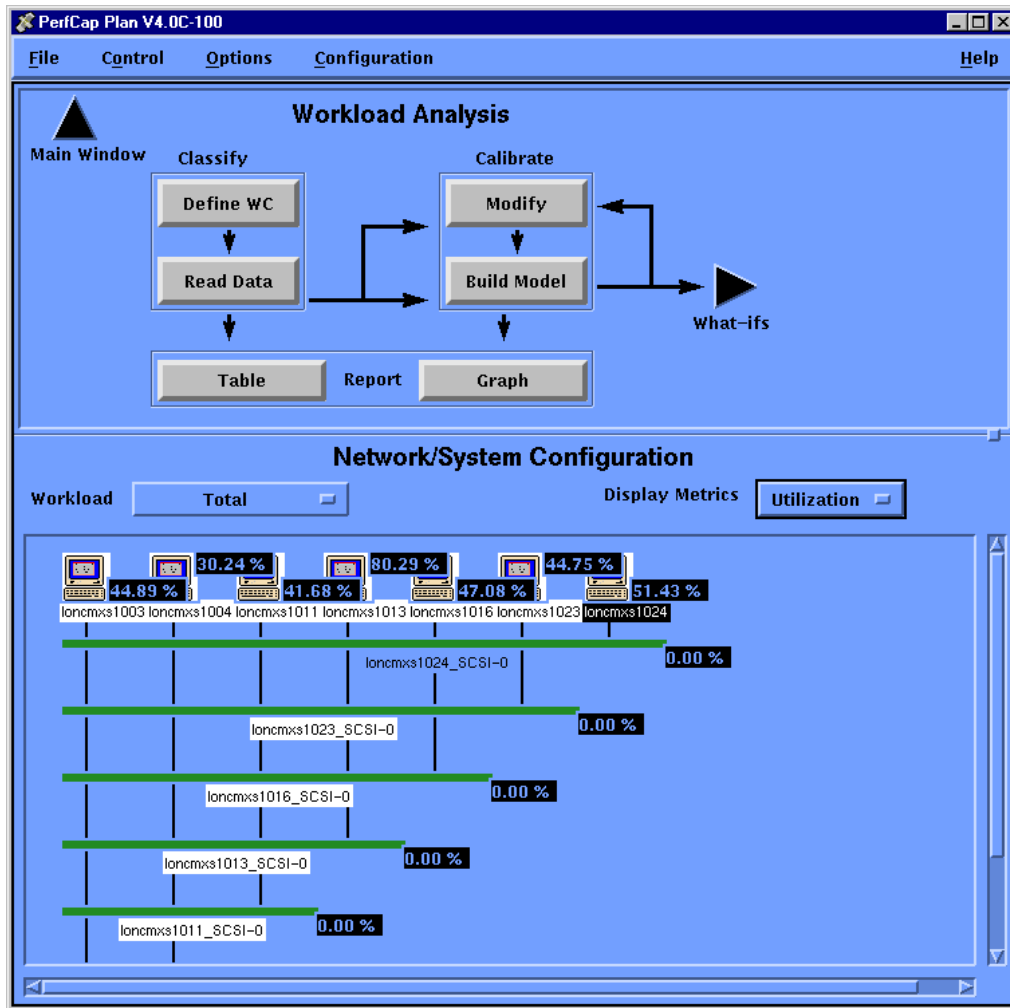


Figure 16: Current Servers to be Consolidated

The above Figure 16 represents a subset of servers that needed to be consolidated.

Methodology

A P690 was simulated within the model, and each node was changed to an LPAR within a 32 way P690 offering current performance and 100% workload increase without changing CPU configuration. Each LPAR would have different numbers of processors, depending on individual node requirements.

P690's do not support LPARs lower than 1 processor and 1 CPU, which is the minimum node allowable.

Results

Storage was already consolidated at the central level. However, the same numbers of fibre channel cards were retained for IO throughput. In the capacity plan, the fiber channel cards were upgraded to 2Gb.

Memory was retained at the current physical levels – the faster CPU's and bus would stretch the new hardware further.

Summary

Specification	Post Consolidation	Pre-consolidation
Physical Nodes	1	8
Headroom	100% +	30% +
Memory	116Gb	116Gb
CPU available	11	16
Rack-space	24" system frame (2 racks)	7 racks
CPU's used	21	78

Table 5: Total Pre and Post Configurations

Host	OS	Model	Original CPUs	Consolidated LPAR	Memory	% Headroom
03	5.1	IBM 7026-M80	8	P690 2 CPU	8192	100%+
04	5.1	IBM 7026-M80	4	P690 1 CPU	4096	100% +
05	5.1	IBM 7026-M80	8	P690 2 CPU	4096	100%+
06	5.1	IBM 7026-H80	2	P690 1 CPU	4096	100% +
22	5.1	IBM 7026-6M1	8	P690 2 CPU	16384	100% +
23	5.1	IBM 7026-6M1	8	P690 2 CPU	16384	100% +
24	5.1	IBM 7026-6M1	8	P690 2 CPU	16384	100% +
25	5.1	IBM 7026-6M1	8	P690 2 CPU	16284	100% +
11	5.1	IBM 7026-6H1	4	P690 1 CPU	8192	100% +
13	5.1	IBM 7017-S85	10	P690 4 CPU	16384	100%+
15	5.1	IBM 7017-S85	10	P690 2 CPU	16384	100% +

Table 6: Individual LPAR Configuration

The following graph shows the result of moving the current physical platforms to a consolidated LPAR P690.

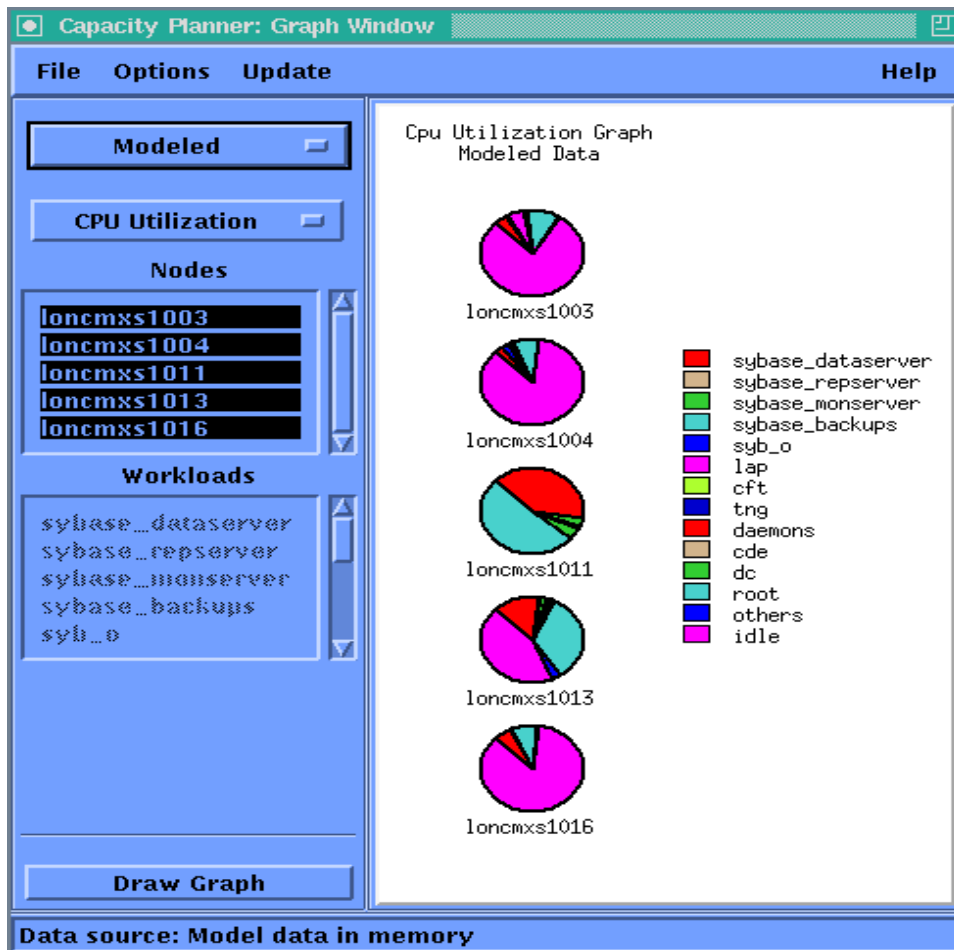


Figure 17: Headroom on each LPAR and Application Utilization

The following graph in Figure 18 shows saturation response time and headroom for the new P690 platform. For all nodes, the total headroom reached was around 160% workload increase.

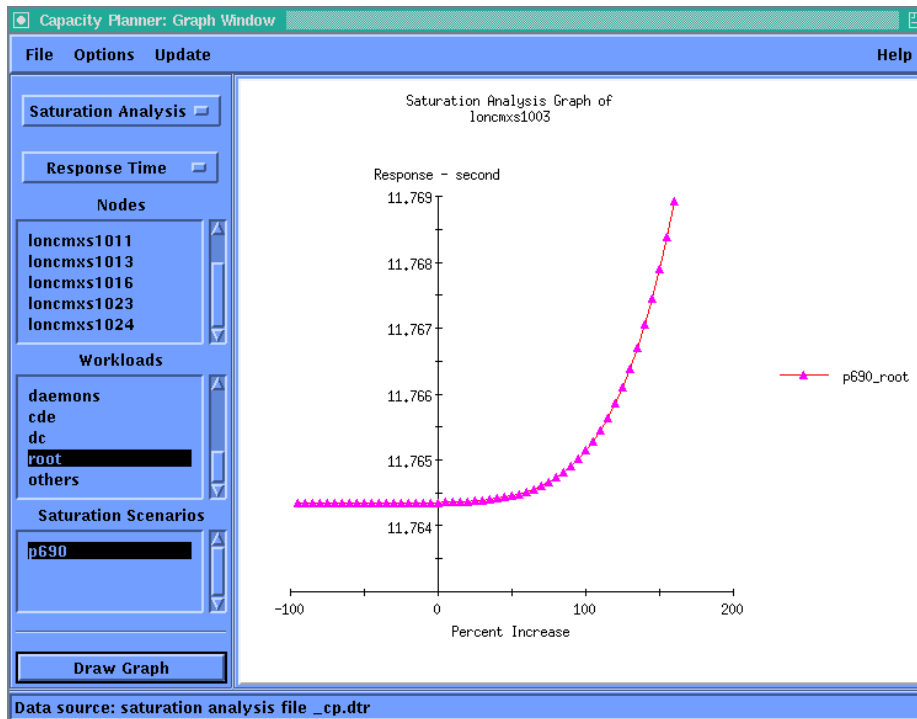


Figure 18: Saturation Graph for Consolidated P690

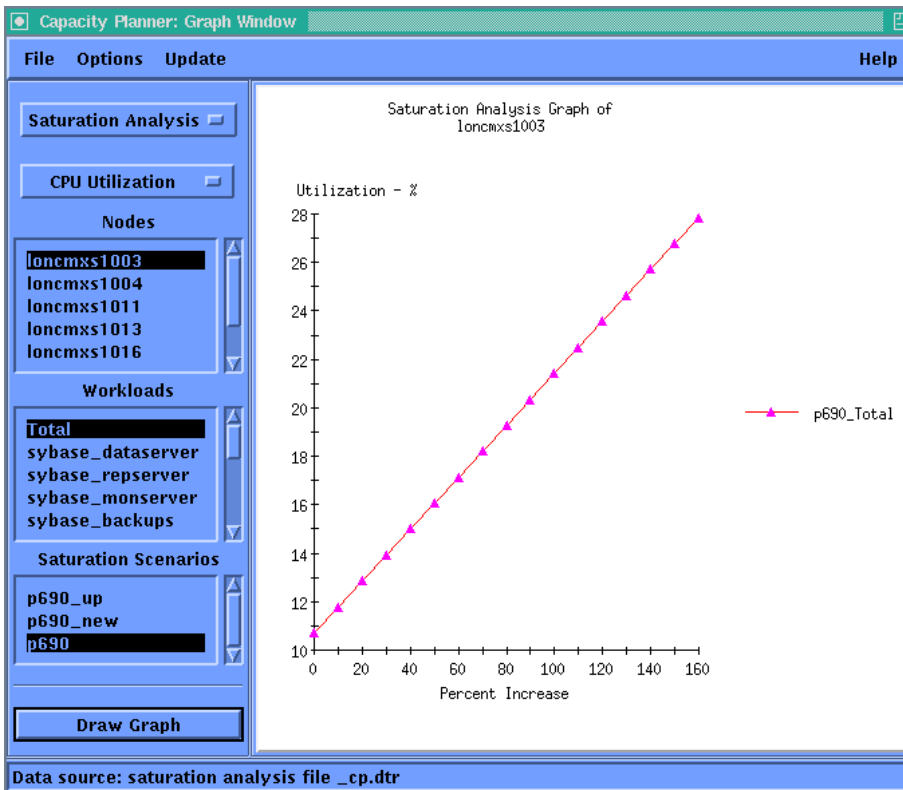


Figure 19: CPU Utilization of Consolidated Server P690

Figure 19 shows that with 160% increase in workload P690 CPU is only about 28% utilized. Further analysis reveals that the saturating resource is a controller channel – which can then be analyzed and upgraded for additional headroom.