

Tools and Technique for Measurement of Computer Performance

Javed Ahmad Shaheen^{1*}, MS (CS) Mian Ali Asghar² and Abid Hussain³

¹Computer Science Department, Virtual University of Pakistan, Lahore–Pakistan

²Computer Science Department, Global Institute Lahore, Lahore–Pakistan

³Computer Science Department, Virtual University of Pakistan, Lahore–Pakistan

¹javedmatyana@gmail.com, ²aliurooj143@yahoo.com, ³ms140400004@vu.edu.pk

Abstract

The monitoring of the system means the performance measurement of computer system and computer performance is the amount of work accomplished by a computer system and its all depends on the context, so high computer performance may involve one or more of the following work, short response time for a given piece of work, high throughput that is rate of processing work, low utilization of computing resources, high availability of the computing system or application, fast data compression and decompression and high bandwidth, and it may be short data transmission time . It is based on workload. So the suitable selection of workload resultant is meaningful measurement. So to attain the meaningful measurement the following points must be kept in mind. First, it is to understand the various type of workload. Second, which kind of workload is mostly used by other analyst? The third one is how we select an appropriate workload. Fourth, how we can summarize the data for measured workload. The 5th step is how we can monitor the system performance. The next is how can the desired workload be placed on the system in a controlled manner and the last one is how can we presented the result of evaluation.

Keywords: *Kernels, different Monitor, Workload, Bench marking*

1. Introduction

A computer system performance analyst must understand computer system as well as various analysis techniques such as statistic, probability theory, experimental design, simulation and queuing theory. This research paper only focuses on performance analysis. it emphasis and integrate all these aspects of performances analysis. As performances analyst with several design term at Digital Equipment Corporation and an instructor of the course on performances analysis and Computer performance metrics include availability, response time, channel capacity, latency, completion time, service time, bandwidth, throughput, relative efficiency, scalability, performance per watt, compression ratio, instruction path length and speed up. CPU benchmarks are available. A CPU designer is often required to implement a particular instruction set so he cannot change number of instructions. Sometimes a designer focuses on improving performance by making significant improvements in frequency while not sacrificing too much average cycles per instruction leading to a speed-demon CPU design. Sometimes a designer focuses on improving performance by making significant improvements in CPI, while not sacrificing too much clock frequency leading to a brainiac CPU design. Fast, inexpensive computers are now essential to numerous human endeavors. But less well understood is the need not just for fast computers but also for ever-faster and higher-performing computers at the same or better costs. Exponential growth of the type and scale that have fueled the entire

* Corresponding Author

information technology industry is ending in addition, a growing performance gap between processor performance and memory bandwidth, thermal-power challenges and increasingly expensive energy use, threats to transistor density, and a broad new class of computing applications pose a new set of challenges to the computer industry in sense of computer performance or say workload.

2. Types of Workload

There are two types of workload one is **real workload** and the second is **synthetic workload**. The real workload may be defined as a workload is one observed on a system being used for normal operation so it is not repeated therefore it may not be used for test workload. The synthetic workload is also characterized as real workload but applied as repeatedly and controlled manner so it is used for test workload. Now we discuss the types of workload

2.1. Addition Instruction

Historically the processors were most expressive and most used component in the computer system therefore the performance of computer was highly depend on the CPU in which the most frequent instruction was Addition instruction so the addition instruction was sole workload in this way the addition time is sole performance metric.

2.2. Mixes Instruction

With the passage of time processors are growing up so the only addition instruction was no longer sufficient therefore mix instruction are introduced for more detailed workload description. Such as Gibson mix was developed by Jack C. in which the following table depict the whole scenario

Table 1. Gibson Mix

1	Load and Store	31.2
2	Fixed-point Add and Subtraction	6.1
3	Compares	3.8
4	Branches	16.6
5	Floating Addition and Subtraction	6.9
6	Floating Multiply	3.8
7	Floating Divide	1.5
8	Fixed-point Multiply	0.6
9	Fixed-point divide	0.2
10	Shifting	4.4
11	Logical, AND, OR	1.6
12	Instruction not using registers	5.3
13	Indexing	18
Total		100

2.3. Kernels

As we know that in the computer system various instruction are high timely variable due to the mechanism of various address translation, instruction caching and pipelining. Therefore a sole of instruction is not to be considered as long but considered a set of instruction which play a role of higher level function. So the researcher try to list out those function which are frequently used such type of function are known as **kernel**. So kernel is induction of instruction mix. The word kernel means nucleus. The disadvantage

of instruction mixes is applied to kernel while there is some limitation due to parameter values.

2.4. Synthetic Programs

The processing kernel not makes use of any operating system service or any input output device. I/O Operation on the computer system is essential part of the actual workload. For this purpose the I/O performance can be measured due to simple exercise of simple loop. In this way we are able to measure the average CPU time and elapsed time to every service call. The other advantage of exercise loop is that it is frequently developed and also provides to various vendors. On the other hand its limitation is that it is so small and not makes representative as memory or disk reference.

2.5. Benchmarks

If the computer system is used for a special application such as airline reservation then a representative subset of function for the application may be used such type of function are used for the benchmarks. Because there are so many programs to test a CPU on all aspects of performance, benchmarks were developed. Famous benchmarks are SPEC int and SPEC fp benchmarks; these are developed by Standard Performance Evaluation Corporation. The ConsumerMark benchmark is developed by the Embedded Microprocessor Benchmark Consortium (EEMBC).

2.5.1. Some Popular Benchmarks: The term *benchmark* is always used synonymously with workload. Kernels, synthetic programs and application level workload. But here the process of performances comparison for two or more system by measurement is known as benchmarking and the workload used in measurement is consider as benchmarks.

2.5.2. Sieve: Sieve kernel is used to compare microprocessors, personal computer and high level languages. Sieve is based on Eratosthenes' sieve algorithm and is used to find all the prime numbers which are below from the given numbers.

2.5.3. Ackermann's Function: In this function takes two parameters which are define recursively and it is used to measure the efficiency of the mechanism of function calls. The number of recursive calls in Ackermann ($3n$) shown by Whichmann.

$$(512(512*4^{n-1} - 15 * 2^{n+3} + 9n + 37))/3 \quad (1)$$

The above expression is used to measure the execution time for per call of the calling function.

2.5.4. Whetstone: Whetstone kernel is used to measure dynamic frequency of operation. It consists on eleven modules. Its result is measured in KWIPS which stand for Kilo Whetstone Instruction per Seconds. It is also considered as floating -point benchmarks.

2.5.5. LINPACK: LINPACK is compared based upon the execution rate as measured in MFLOPS. It is used to represent mechanical engineering application on workstation.

2.5.6. Dhrystone: Dhrystone is used to represent the system programming environment. Their results are measured in DIPS which stand for Dhrystone Instruction per Second. It is also used to measure the integer performance.

2.5.7. Lawrence Livermore Loops: This workload consists in the form of set in which twenty four kinds of test takes place huge calculation therefore the result of that test are rather complex for interpretation because the result not in the form in single value.

2.5.8. Debit-Credit Benchmark: It is a de facto slandered benchmark which is used to compare transaction processing system. Therefore it is also known as application level benchmark

2.5.9. SPEC Benchmark Suite: SPEC stand for system Performance Evaluation Cooperative it consist on the following ten benchmarks such as GCC, Espresso, Spice 2g6, Doduc, NASA7, LI, Eqntott, Matrix300, Fpppp, Tomcatv. The given benchmarks primarily focus on CPU, floating point unit and memory subsystem

3. The art of Workload Selection

There are four major types for select the workload which we discussed in detail one by one because workload is the most crucial part to measure the performance of evaluation in the project.

3.1. Services Exercised

As we know that each system gives a number of services. So the first task is to listing these services in systematic way for study. SUT which stand for System under Test is used to denote the complete set of components. One specific component from SUT is known as CUS which stand for components under study as shown in the figure. An important point kept in mind the metrics which are selected should be reflect the performance of services provided at system level rather than component level. Therefore the basis of workload selection is also the system rather than components. As there is time sharing system where a user typically type in a high level request such as in application of the principle outline as shown in the Figure 1.

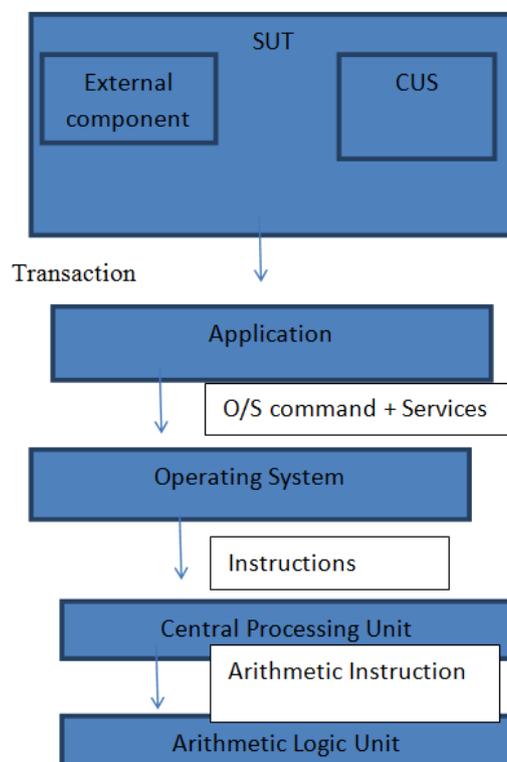


Figure 1. Workload Services

3.3. Level of Detail

After the SUT the next step in the workload selection is to choose the level of detail in recording the request. There is different possibility of types of request such as most frequently request, Frequency of request type, Time-stamped sequence of request, Average resources Demand, Distribution of resource demand. The workload description used for analytical and simulation modeling which is known as **non-executable** workload which is not applicable on real system. While a trace of user commands that are executed directly on the system is known as **executable** workload.

3.4. Representatives

The representativeness is to test the workload in real application which matches with respect to following aspects: Arrival Rate, Resource Demands and Resource Usage Profile which we discuss one by one. Arrival Rate: in which the request should be the same or proportional to the application. Resource Demand: in which the total demands of the key resources should be same or proportional to the application. Resource Usage Profile: in which it is in sequence and the amount in which different resources are used in an application.

3.5. Timeliness

Timeliness is a difficult goal to achieve because user behavior has changed considerably over the years. Therefore interdependence of system design and workload provides opportunity for debate at every performances evaluation presentation so that user's behavior should be monitored on an ongoing basis.

3.6. Other Consideration in Workload

There are more few issues which are Loading Level, Impact of External Components and Repeatability.

4. Workload Characterization Technique

In real user environment is generally not repeatable. Therefore the study of real-user environment is essential. It is all done with the process of **workload characterization** in which requires different statistical technique which are based on probability theory and statistic.

4.1. Terminology

To measure the workload the data consist on services request or the resource demands of a number of users in the given system. Here user is an entity that make services request at the SUT interface. There are different techniques which are used for characterization of workload. Now we try to discuss one by one.

4.2. Average

The simplest way to characterize a workload parameter in a single number that facilitate to summarize the observed values is arithmetic mean which is shown in the formula:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

4.3. Specifying Dispersion

The average alone is unable to if the large variability in the data exist then the variability is specified by the variance.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3)$$

4.4. Single –Paramiter Histogram

A histogram shows the relative frequencies of different values of the given parameters. For continuous parameters values divide a complete histogram into sub ranges which are known as buckets or cells.

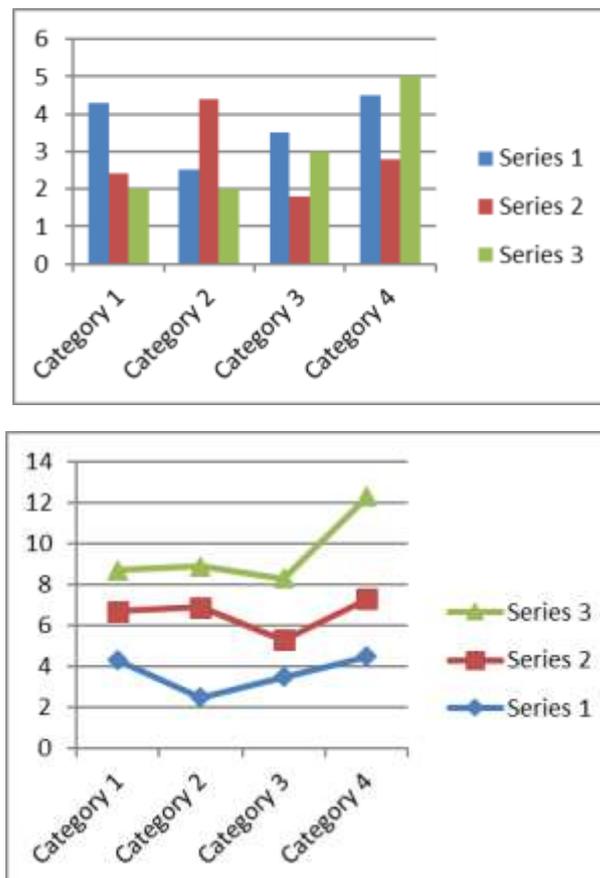


Figure 2. Workload Characterization

4.5. Multiparamiter Histogram

If there is significant correlation among various workload parameters then workload must be characterized to multipart meters histogram.

4.6. Principal- Component Analysis

Usually a technique which is used to listing workload components by weighted sum of their parameters values. So here is a technique to determining the weight in such circumstances is principal- components analysis. If y's are linear combination of x's then

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad (4)$$

here a_{ij} is known as the loading variable x_j on factor y_i

If the y 's is from the orthogonal set then there inter product is zero

$$\langle y_i y_j \rangle y_i = \sum_k a_{ik} a_j \quad (5)$$

y 's is uncorrelated to each other

4.7. Marvoc Models

If it is assumed that the next request is depend only on the last request the request follow Markov Models as shown in the Figure 3.

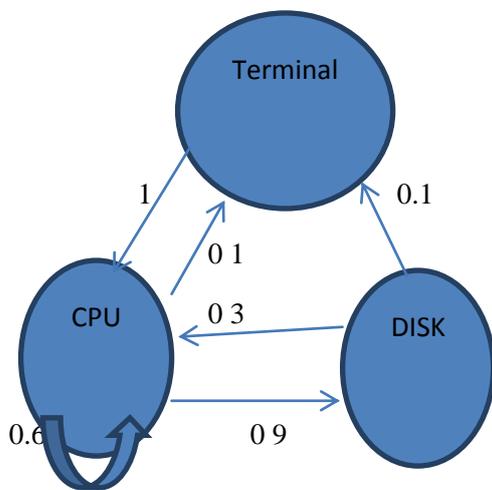


Figure 3. Markov Model

4.8. Clustering

For analysis purpose clustering is useful to classify the components into a small number of classes. To accomplished this task the following steps take place. Take a sample, select workload parameters, transform parameters, remove outliers, scale all observation, select a distance measure, performs clustering, interpret the result, change parameters, select representative.

5. Monitors

A monitor is a tool which is used to observe the activities on the given system. Monitor is useful for performance analyst as well as programmer and system managers. There are different causes to monitor the system such as to seek the frequency used segments of the software to enhance the performance or to measure resource utilization and to find the bottleneck performance or to monitor the tune of the system or to monitor the characteristic of the workload or to monitor the models of the parameters.

5.1. Monitor TEerminology

There are some monitor terminology which is listed as Event, Trace, Overhead, Domain, Input Rate, Resolution and the Input Width.

5.2. Monitor Classification

Monitor is classified on the basis of some characteristic which are implementation level, result displaying ability and trigger mechanism. Depending upon these monitor are **event driven, sampling monitor, on-line monitors** and **batch monitors**. Some particular monitor is classified as **hybrid-sampling-batch monitors**.

5.3. Software Monitors

These types of monitors are used to monitor the operating system and high level software as database and networks. Software monitor are generally lower resolution, lower input rate and higher overhead than hardware monitors.

5.3.1. Issue in Software Monitor Design: There are some issues in designing to monitor which should be kept in mind. The first one is activation mechanism in which the following parameters take place such as trap instruction, Trace mode, timer interrupt. The second one is buffer size. The other is number of buffers. The fourth is buffer overflow. The 5th is data compression or analysis. The sixth is on/off switch it is similar as IF THEN ELSE conditions. The 7th is languages such as assembly, C or Bliss. The eight one is priority. The 9th is Abnormal-Event Monitoring.

5.4. Hardware Monitor

These types of monitor consist on separate piece of equipment that is attached to the system being monitored via probs. There are number of hardware monitor are available which contain the following elements such as probes, counters, logic element, comparators, mapping hardware, tiers and tape/disk.

5.5. Software Versus Hardware Monitors

Table 2. SOFTWARE VERSUS HARDWARE MONITORS

Criterion	Hardware Monitor	Software Monitor
Domain	Difficult to monitor OS events	Difficult to monitor hardware events
Input rate	Sampling rate 10 ⁵ per second	Sampling rate is limited by the processor
Time resolution	10 nanosecond	10 to 16 miliscond
Expertise	Require intimate knowledge about HW	Intimate knowledge of software
Recording capacity	Limited by memory	Limited by overhead desire
Input width	Record several events in the same time	It cannot do so
Monitor overhead	None	Overhead depend on the input rate
Portability	It is portable	It depend on the operating system
Availability	Monitoring continuous during failure	It cannot do so
Error	Error exist	After debugs error are rare
Cost	High	Medium

5.6. Firmware and Hybrid Monitors

These types of monitors are implemented by modifying the processor microcode. Such type of monitors is fulfilling the requirement which fall among the software and hardware monitoring level.

5.7. Distributed-System Monitors

Now a day's computer system is distributed and also contains a number of software and hardware components. That work together separately and concurrently. So for this purpose Distributed- System Monitor is required which fulfill these requirements. The following tasks should be done observation, collection, analysis, presentation, interpretation, console and management for such kind of monitors.

6. Program Execution Monitors and Accounting Logs

Program execution monitor, also known as program optimization or program execution analyzer, are software monitors designed to observe application software. Issues related to the given scenario are discussed as under.

6.1. Program Execution Monitors

There are different causes to monitor the execution of the program such as tracing, timing, tuning, assertion checking, and coverage analysis as shown in the Figure 4.

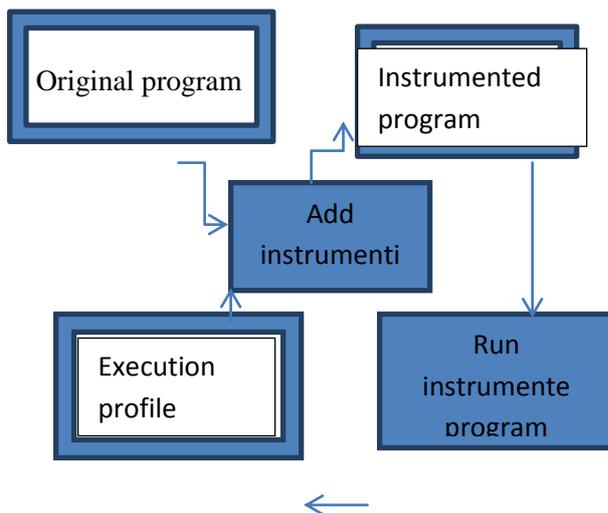


Figure4. PE Monitors

6.1.1. Issues in Designing a Program Execution Monitor: Following are the issues such as measurement unit, measurement technique, Instrumentation mechanisms, and profile Report.

6.2. Technique for Improving Program Performances

There are three common technique of program optimization such as code optimization, I/O optimization and paging optimization

6.3. Accounting Logs

The main advantage of counting logs is that are built in and no extra efforts is needed to develop them and also log analysis program is not supplied.

7. Capacities Planning and Benchmarking

One of the important problem for managers of data processing installation is capacity planning. The term capacity management is used to denote the problem of ensuring that the currently available computing resources are used to provide the highest performances.

7.1. Steps in Capacity Planning and Management

The first step is to instrument the system. The second is Monitor usage system. Third is characterizing the workload. Fourth is predict performance under different alternatives and the last one is select the lowest cost, highest performances alternative as shown in the Diagram 5.

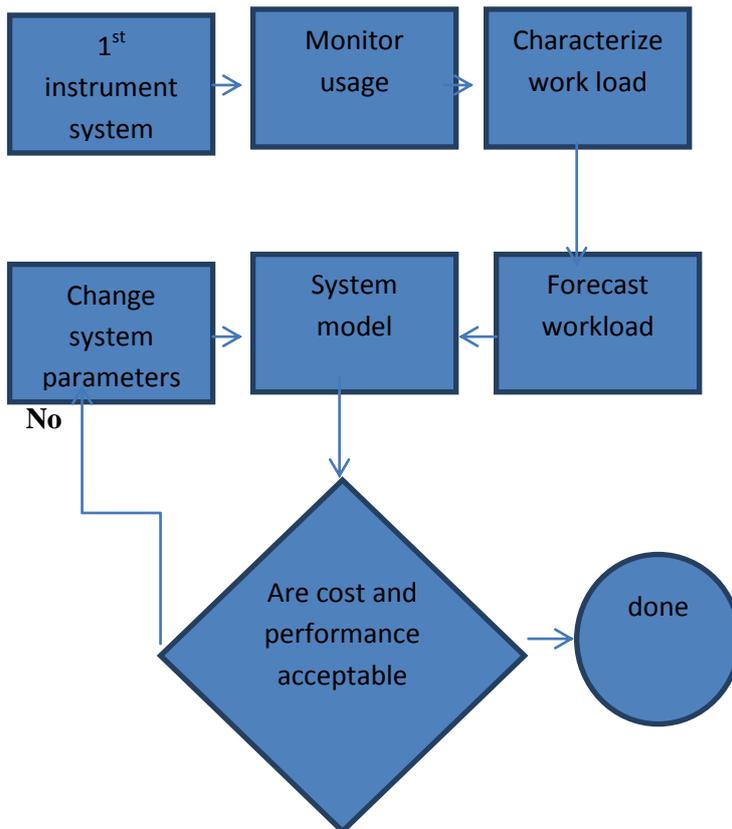


Figure 5. CP & Management

7.2. Problem in Capacity Planning

The following problem exist during capacity planning which are listed below

- There is no standard terminology.
- There is no standard definition of capacity.
- There are number of different capacity for the same system.
- There is no standard workload unit.
- Forecasting future application is difficult.
- There is no uniformity among system from different vendors.
- Model input cannot always be measured.
- Validating model projection is difficult.
- Distributed environment are too complex to model.
- Performance is only a small part of the capacity planning problem.

7.3. Common Mistakes in Benchmarking

There are some common mistakes in benchmarking which are listed as under.

- Only average behavior represented in test workload.
- Skewness of device demands ignored.
- Loading level controlled inappropriately.
- Caching effect ignored.
- Buffering size is not appropriate.
- Inaccuracies due to sampling ignored.
- Ignoring monitoring overhead.
- Not validating measurement.
- Not ensuring some initial conditions.
- Not measuring transient performances.
- Using device utilization for performances comparisons.
- Collecting too much data but doing very little analysis

7.4. Benchmarking Games

The process of benchmarking is to compare two system using standard well known benchmarks. Some of the result which may be misleading during benchmark games

- Differing configuration may be used to run the same workload on two system
- The compiler may be wired to optimize the workload
- A synchronized job sequences may be used
- Test specification may be written so that they are biased toward one machine
- Very small benchmark may be used
- Benchmarks may be manually translated to optimize the performances

8. The Art of Data Presentation

One of the important steps in every performance is to present the data or results. Graphic chart such as line charts, bar charts, pie charts and histograms are commonly used in presenting performances results. There are a number of reasons why a graphic chart may be used to present the results. Such as types of variables e.g. qualitative or quantitative variables. Good graphic charts also contain the following requirement.

- Require minimum efforts from the reader.
- Maximize the information
- Minimize Ink
- Use commonly accepted practices
- Avoid Ambiguity.

8.1. Common Mistake in Presenting Charts

The following mistakes may be occurred during the presenting the charts

- Presenting too many alternative on a single charts
- Presenting many variables on a single chart.
- Using symbol in place of text.
- Placing extraneous information on the charts
- Selecting scale ranges improperly.
- Using a line chart in place of column charts.

Some other characteristic in presenting the data or results in performances perspective such as pictorial games, GANTT charts, KIVIAT graphs, SCHUMACHER charts, decision maker games *etc.*

9. Ratio Games

Ratio games provide a good opportunity for performance perspective. The following parameters should be taken during ratio games Such as choosing an appropriate base system, using an appropriate ratio metric, Using relative performances enhancement, ratio game with percentage, strategies for winning a ratio game and the correct analysis.

Conclusion

Computer system users, administrator and designers are all interested in performances evaluation since their goal is to obtain or provides the highest performances at the lowest cost. This objective has resulted in continuing evaluation of higher performances and lower cost system leading to today's proliferation of workstation and personal computers. As the field of computer design matures the computer industries is becoming more competitive and it is more important than ever to ensure that the alternative selected provides the best cost performances trade-off. Overall, the performance of a computer is dependent on how well it works together as a whole. Continually upgrading one part of the computer while leaving outdated parts installed will not improve performance much, if at all. It should also be noted that the processor, memory and video card are the most important components when determining performance inside a computer. There are many factors which affect how fast your computer can process data and instructions: like the amount of RAM memory, the speed and generation of your CPU (the system clock), the size of the Register on your CPU, the Bus type and speed, the amount of Cache memory and to make computer performance uninterrupted these factors should be in proper working and in measuring the performance these components may affect "the availability, response time, channel capacity, latency, completion time, service time, bandwidth, throughput, relative efficiency, scalability, performance per watt, compression ratio, instruction path length and speed up" aspect of performance.

References

- [1] N. R. Adam and A. Dogramaci, "Current issue in computer Simulation", Academic press, new York, (1979).
- [2] M. F. Barnes, "Measurement and modeling method for computer system performances", input two-nine, Surrey, U.K. (1997).
- [3] E. Gelenbe and I. Mitrani, "Analysis and synthetic of computer system", Academic Press, London. (1980).
- [4] J. Gilbreath, "A high level languages benchmarks", byte, vol. 6, no. 9, (1981).
- [5] C. R. Hicks, "Fundamentals Concepts in design of Experiments", Holt, Rinehart, & Winston, New York. (1973).
- [6] H. C. Lucas, Jr, "Performance Evaluation and Monitoring, computing Survey, (1971).
- [7] Artikel: Benchmarking Parity Games - Blog - Jeroen Keiren
- [8] T. Krazit, "NVidia's Benchmark Tactics Reassessed", IDG News, (2003).
- [9] K. Castor, "Hardware Testing and Benchmarking Methodology". Retrieved 2008-02-24, (2006).
- [10] A. Ehliar and Dake Liu, "Benchmarking network processors" (PDF).
- [11] "VMmark Rules 1.1.1" (PDF). VMWare. 2008.

Authors



JAVED AHMAD SHAHEEN, is presently working as SST (CS) in Punjab School Education Department; he has done his MSCS (Networking) from Virtual University of Pakistan Lahore. He is also teaching in Islamia University of Bahawal Pur "Bahawal Nagar Campus" as "Visiting Lecturer". He has practical experience of Networking.



MIAN ALI ASGHAR, is also presently working as SST (CS) in Punjab School Education Department; he is doing his MSCS from Global Institute Lahore. He has good command over Web Development and Software Engineering.



ABID HUSSAIN, is also presently working as SST (CS) in Punjab School Education Department; he is doing his MSCS (Networking) from Virtual University of Pakistan Lahore. He has good command over Wireless Networking and Programming.

