



---

## Computer Performance Evaluations in Different Systems and Benchmarks

*Chigbundu Kanu Enyioma<sup>1</sup>, Onwuzo Chioma Julia<sup>2</sup>, Okoronkwo Madubuezi Christian<sup>3</sup>*

*<sup>1,2,3</sup> Computer science Michael Okpara University of Agricultural, Umudike*

---

### Abstract

Business operations need to be effective, efficient, and up-to-date if you want to continue pleasing clients and making an impact within your industry. That is why it's so important to create a performance evaluation benchmarking plan to determine the strengths and weaknesses. Using benchmarks will help to develop better strategies and foster long-term growth for your company and employees. The work briefly discussed the concept of computer evaluation and its uses to compare different systems using benchmark characteristics. It scrutinizes the inept functionalities of each system and compares the results gotten at the end, to know which performs better.

---

**Keywords:** throughput, benchmarking, Global Computer Performance Evaluation, computer architecture, performance modelling

---

### 1.0. INTRODUCTION

A computer performance evaluation is defined because of the process by which a computer system's resources and outputs are assessed to work out whether the system is working at an optimal level. It's almost like a voltmeter that a handyman may use to see the voltage across a circuit. The meter verifies that the right voltage is passing through the circuit. Similarly, an assessment is often done on a Personal Computer (PC) using established benchmarks to ascertain if it's performing correctly. In evaluating a computer's performance, varieties of parameters are used to determine the result. Examples are latency, speed, throughput, bandwidth, and more, which can be discussed next within the lesson. Standards, or points of reference, are used against the parameters, and an assessment is given. This process is understood as benchmarking.

It is not a simple task to make benchmarks for assessing computer performance. The first challenge is that technological characteristics are constantly changing. This suggests that benchmarking has got to be constantly updated, too. This makes computer evaluation a posh process.

---

### 2.0. LITERATURE REVIEW

According to [1], the speed and size of your RAM are extremely important. The quantity of memory your computer has installed can determine its performance. The more memory installed, the faster your computer can access information without having to use its virtual storage which is slower than RAM. The width of the memory's data path also can affect its time interval. The time that the CPU has got to await memory to be accessed limits its performance. The front side bus speed is another factor that will determine the speed of your computer. The front side bus is that the pathway from your CPU to your I/O devices. The clock speed of your processor is vital in determining the general speed of your computer, but not the foremost important factor. The clock, whose speed is measured in megahertz and gigahertz, supplies the CPU with electrical pulses that synchronize. Computer Performance and Evaluation has been a serious topic for many researchers over the years and thus much research work has been administered with regards to the topic matter. Among those that study Computer Performance and Evaluation and further came up with their findings include

[2] in his book titled "The Evolution of Benchmarking as a Computer Performance Evaluation Technique" agreed that The practice of benchmarking as an assessment tool in computer performance evaluation is traced in his article from the first 1960s to this. They traced the evolution of benchmarking practice by examining milestones within the development and important issues raised during this evolution and further suggested a course of action.

[3] Cited that as a fundamental task in computer architecture research, performance comparison has been continuously hampered by the variability of computer performance. In traditional performance comparisons, the impact of performance variability is typically ignored (i.e., the means of performance observations are compared no matter the variability), or within the few cases directly addressed with t-statistics on faith the number and normality of performance observations.

They propose a non-parametric Hierarchical Performance Testing (HPT) framework for performance comparison, which is significantly more practical than standard t-statistics because it doesn't require collecting an outsized number of performance observations to realize a traditional distribution of the sample mean. Especially, the proposed HPT can facilitate quantitative performance comparison, during which the performance speedup of 1 computer over another is statistically evaluated.

[4] In his publication titled "Performance Evaluation: Techniques, Tools and Benchmarks" and stated that Performance evaluations are often classified into performance modeling and performance measurement. Performance measurement is feasible as long as the system of interest is out there for measurement and as long as one has access to the parameters of interest. Performance measurement may further be classified into on-chip hardware monitoring, off-chip hardware monitoring, and software monitoring and microcode instrumentation. Performance modeling is usually used when actual systems aren't available for measurement or if the particular systems don't have test points to live every detail of interest. Performance modeling may further be classified into simulation modeling and analytical modeling. Simulation models may further be classified into numerous categories counting on the mode/level of detail of simulation. Analytical models use probabilistic models, queuing theory, Markov models, or Petri nets. Dragon [5] and Smith [6] discuss the utilization of the acceptable mean for a given set of knowledge. Dragon [4] and Patterson and Hennessy [7] illustrate several mistakes one could make while finding one performance number.

The review in [8] explains that there are two levels of benchmarks. The component-level benchmarks and therefore the system-level benchmarks. supported their composition, benchmarks are often categorized into two types: synthetic benchmarks and application benchmarks.

In [9] different systems were compared about their variance in performance evaluation, one major comparison is that of the MacBook and windows. The authors did several checks on the functionality and drew conclusions of which is best surely jobs. [10] Highlighted Product form networks and therefore the research queuing (RESQ) Package. the invention of product form queuing networks and therefore their properties and the development of efficient computational algorithms for product form networks was a breakthrough in analytic performance modeling. RESQ allowed a user to specify and solve product form networks (initially using the convolution algorithm; the MVA algorithm was added later), but it also allowed a user to specify more general "extended queuing networks" and use discrete event simulation to estimate performance measures.

## **2.1 BENCHMARK ASSIGNMENT IN LISTED SYSTEM**

The list below assesses popular pc benchmarks

**Business Winston:** A system-level, application-based benchmark that measures a PC's overall performance when running today's top-selling Windows-based 32-bit applications. It runs real 32-bit business applications through a series of scripted activities and uses the time a PC takes to finish those activities to supply its performance scores. The suite includes five Microsoft Office 2000 applications (Access, Excel, FrontPage, PowerPoint, and Word), Microsoft Project 98, Lotus Notes R5, NicoMak WinZip, Norton AntiVirus, and Netscape Communicator.

**WinBench99:** A subsystem-level benchmark that measures the performance of a PC's graphics, disk, and video subsystems during a Windows environment.

**3DwinBench:** Tests the bus wont to carry information between the graphics adapter and therefore the processor subsystem. Hardware graphics adapters, drivers, and enhancing technologies like MMX/SSE are tested.

**CD WinBench99:** Measures the performance of a PC's CD-ROM subsystem, which incorporates the CD drive, controller, and driver, and therefore the system processor.

**Audio WinBench 99:** Measures the performance of a PC's audio subsystem, which incorporates the sound card and its driver, the processor, the DirectSound, and DirectSound 3D software, and therefore the speakers.

**Battery Mark:** Measures battery life on notebook computers.

**I-bench:** A comprehensive, cross-platform benchmark that tests the performance and capability of Web clients. The benchmark provides a series of tests that measure both how well the client handles features and therefore the degree to which network access speed affects performance.

**Web Bench:** Measures Web server software performance by running different Web server packages on an equivalent server hardware or by running a given Web server package on different hardware platforms.

**NetBeans:** a transportable benchmark program that measures how well a digital computer handles file I/O requests from clients. NetBeans reports throughput and client reaction time measurements.

**3Dmark MAX 99:** From Future mark Corporation. maybe a nice 3D Benchmark that measures 3D gaming performance? Results are hooked into CPU, memory architecture, and therefore the 3D Accelerator employed.

**SYSMARK:** Measures a system's real-world performance when running typical business applications. This benchmark suite comprises the retail versions of eight application programs and measures the speed with which the system under test executes pre-determined scripts of user tasks typically performed when using these applications. The performance times of the individual applications are weighted and combined into both category-based performance scores also as one overall score. the appliance programs employed by SYSmark 32 are Microsoft Word 7.0 and Lotus WordPro 96 for data processing, Microsoft Excel 7.0 (for spreadsheet), Borland Paradox 7.0 (for the database), CorelDraw 6.0 (for desktop graphics), Lotus Freelance Graphics 96, and Microsoft PowerPoint 7.0 (for desktop presentation) and Adobe PageMaker 6.0 (for desktop publishing).

UL works with leading technology companies to make industry-standard benchmark tests that are widely employed by businesses, governments, the press, and consumers.

Each of the benchmark tests is meant for a selected scenario—such as home or office use and a particular class of device—such as a PC, laptop, or Smartphone. you ought to choose a benchmark that best matches the requirements of your end-users. For companies buying PCs, laptops, or notebooks for general office use, PCMark 10 benchmark is suggested.

PCMark 10 measures PC performance with a comprehensive set of tests that cover the big variety of tasks performed within the modern workplace. The tests in PCMark 10 include everyday essentials like web browsing and video conferencing, common office productivity tasks like working with documents and spreadsheets and reach digital content activities like photo and video editing. A PCMark 10 score may be a measure of the general performance of a system for contemporary paperwork. PCMark 10 sub-scores assist you to specialize in performance for specific activities like office productivity or working with digital content.

## **2.2 CHOOSING A REFERENCE BENCHMARK SCORE**

Setting a minimum benchmark score in your RFP helps you judge the relative performance and value of various systems. But how do you have to set about choosing an appropriate score?

You can start by testing a number of your existing systems. Benchmark old PCs that can get replaced and new systems that were purchased recently. Benchmark scores from these systems will offer you good reference points.

If you have already got an honest idea of the specification you are looking to shop for, you'll ask a supplier to supply benchmark scores for the system to supply another point of reference. Otherwise, you can search benchmark results for similar systems on our 3dmark.com website.

Specifying PC performance with PCMark 10

With PCMark 10, you'll use the general benchmark score, which represents the PC's performance for a good range of paperwork activities. otherwise, you can focus on a selected sub-score that's an honest match for your employees' typical work tasks.

- PCMark 10 benchmark score: Overall PC performance for a variety of tasks and activities.
- PCMark 10 Productivity score: System performance when working with spreadsheets and documents.
- PCMark 10 Digital Content Creation score: PC performance when working with digital content and media.
- PCMark 10 Essentials score: System performance for everyday activities like web browsing and video conferencing. Also measures the time to start apps.

Using PCMark 10 scores in your RFP

Setting a minimum benchmark score in your RFP makes it easier to gauge and compare competing offers from your suppliers. Specifying performance with a benchmark score instead of a coordinate system also gives your suppliers more freedom to return up with cost-effective alternative configurations that you simply won't have considered otherwise.

Comparing bids that include benchmark performance scores ensures that you simply won't be distracted by the false economy of a less expensive PC specification that underperforms. once you see PC performance expressed as a benchmark score, you'll even be less likely to overspend on over-specified systems.

---

## **3.0 METHODOLOGY**

Microprocessor on-chip performance monitoring counters: All state-of-the-art high-performance microprocessors including Intel's Pentium III and Pentium IV, IBM's POWER 3 and POWER 4 processors, AMD's Athlon, Compaq's Alpha, and Sun's UltraSPARC processors incorporate on-chip performance monitoring counters which may be wont to understand the performance of those microprocessors while they run complex, real-world workloads. This ability has overcome a significant limitation of simulators, that they often couldn't execute complex workloads. Now, complex run-time systems involving multiple software applications are often evaluated and monitored very closely. All microprocessor vendors nowadays release information on their performance monitoring counters, although they're not a part of the architecture.

Off-chip hardware measurement: Instrumentation using hardware means also can be done by attaching off-chip hardware, two samples of which are: (i) SpeedTracer from AMD and (ii) Logic Analyzers.

Software monitoring: is usually performed by utilizing architectural features like a trap instruction or a breakpoint instruction on an actual system, or on a prototype. The primary advantage of software monitoring is that it's easy to try. However, disadvantages include that the instrumentation can hamper the appliance. The overhead of servicing the exception, switching to a knowledge collection process, and performing the required tracing can hamper a program by quite 1000 times. Another disadvantage is that software monitoring systems typically only handle user activity.

Microcode Instrumentation: this is often a way lying between trapping information on each instruction using hardware interrupts (traps) or software traps. The tracing system essentially modified the VAX microcode to record all instruction and data references during a reserved portion of memory. Unlike software monitoring, ATUM could trace all processes including the OS. However, this type of tracing is invasive and may hamper the system by an element of 10 without including the time to write down the trace to the disk.

### 3.1 PERFORMANCE MODELING

Simulation: has become the defacto performance modeling method within the evaluation of microprocessor architectures. There are several reasons for this. The accuracy of analytical models within the past has been insufficient for the sort of design decisions computer architects wish to form (for instance, what quiet caches or branch predictors are needed). Hence cycle-accurate simulation has been used extensively by architects. Simulators model existing or future machines or microprocessors. they're essentially a model of the system being simulated, written during a high-level computer-oriented language like C or Java, and running on some existing machine. The machine on which the simulator runs is named the host machine and therefore the machine being modeled is named the target machine. Such simulators are often constructed in some ways.

Analytical models: while not popular for microprocessors are suitable for the evaluation of huge computer systems. In large systems where details can't be modeled accurately for cycle-accurate simulation, analytical modeling is an appropriate thanks to obtaining approximate performance metrics. Computer systems can generally be considered as a group of hardware and software resources and a group of tasks or jobs competing for using the resources. Multicomputer systems and multi-programmed systems are examples. Analytical models are cost-effective because they're supported efficient solutions to mathematical equations. However, to be ready to have tractable solutions, often, simplifying assumptions are made regarding the structure of the model. As a result, analytical models don't capture all the detail typically built into simulation models. it's generally thought that carefully constructed analytical models can provide estimates of average job throughputs and device utilization to within 10% accuracy and average response times within 30% accuracy. This level of accuracy while insufficient for microarchitectural enhancement studies is sufficient for capacity planning in multicomputer systems, I/O subsystem performance evaluation in large server farms, and early design evaluations of multiprocessor systems

**Table 1 Popular personal computer benchmarks**

Benchmark	Description
Business Winstone	A system-level, application-based benchmark that measures a PC's overall performance when running today's top-selling Windows-based 32-bit applications. It runs real 32-bit business applications through a series of scripted activities and uses the time a PC takes to finish those activities to supply its performance scores. The suite includes five Microsoft Office 2000 applications (Access, Excel, FrontPage, PowerPoint, and Word), Microsoft Project 98, Lotus Notes R5, NicoMak
WinBench99	A subsystem-level benchmark that measures the performance of a PC's graphics, disk, and video subsystems during a Windows environment.
3DwinBench	Tests the bus wont to carry information between the graphics adapter and therefore the processor subsystem. Hardware graphics adapters, drivers, and enhancing technologies like MMX/SSE are tested.
CD WinBench99	Measures the performance of a PC's CD-ROM subsystem, which incorporates the CD drive, controller, and driver, and therefore the system processor
Audio WinBench 99	Measures the performance of a PC's audio subsystem, which incorporates the sound card and its driver, the processor, the DirectSound, and DirectSound 3D software, and therefore the speakers.
Battery Mark	Measures battery life on notebook computers.
I-bench	A comprehensive, cross-platform benchmark that tests the performance and capability of Web clients. The benchmark provides a series of tests that measure both how well the client handles features and therefore the degree to which network access speed affects performance.
Web Bench	Measures Web server software performance by running different Web server packages on an equivalent server hardware or by running a given Web server package on different hardware platforms
NetBeans	A portable benchmark program that measures how well a digital computer handles file I/O requests from clients. NetBeans reports throughput and client reaction time measurements.
3Dmark MAX 99	From Futuremark Corporation. maybe a nice 3D Benchmark that measures 3D gaming performance. Results are hooked into CPU, memory architecture, and therefore the 3D Accelerator employed.

SYSMARK	Measure actual system performance while running common business applications. This set of metrics covers production versions of eight applications and measures how quickly the system under test performs human tasks typically performed while using those applications. Execution time for individual programs is weighted and summed with the total score and performance score for each relevant category. Applications used by SYSMARK 32 include Microsoft Word 7.0, Lotus WordPro 96 for word processing, Microsoft Excel 7.0 (for spreadsheets), Borland Paradox 7.0 (for databases), CorelDraw 6.0 (for graphical desktops), Lotus Freelance (Graphics 96 and Microsoft PowerPoint included). 7.0. Includes. (for desktop viewing) and Adobe PageMaker 6.0 (for desktop publishing).
---------	--

We divide the methods used into three main areas, namely performance measurement, analytic performance modeling, and simulation performance modeling, which we survey within the three main sections of the paper. Although we consider the methods intrinsically, instead of on the results of applying the methods, numerous application examples are cited. The methods to be covered are applied across the whole spectrum of computer systems from personal computers to large mainframes and supercomputers, including both centralized and distributed systems. the appliance of those methods has never decreased over the years and that we anticipate their continued use also as their enhancement when needed to gauge future systems.

#### 4.0 CONCLUSION

One of the challenges for IT managers and logistics professionals is determining IT hardware performance conveniently and affordably to generate competitive bids from computer vendors and bidders. Desktop computers and laptop computers are often provided on the same system to minimize performance requirements. .. But even experts find it difficult to match the performance of different computer systems based on specifications alone. Goals are an essential part of everything. The intentional effort is wasted. Performance appraisal projects are no exception to this rule. In a performance appraisal tool or standard design, you need to specify what you want to measure. A common goal is to measure processor and disk I/O performance. In the current trend of diskless or dateless workstations, the performance of digital computers, and hence networking, are far more important than performance measurement. Benchmark tools usually predict or at least predict, the performance of an unknown system for a particular asset or set of assets. Performance results generally guide purchasing decisions for alternative systems. These devices can also be used as a monitoring and diagnostic tools. The details of the degradation can be accurately determined by running a test program and comparing the results to a known model. Similarly, you can improve or decrease performance by running a test program after the change. The best system performance check is the actual program running on your system. However, this is not always possible, as the program is not always ready to buy the system. Although the hardware existed before the purchase of the system, equivalent computer systems may perform different tasks with equivalent programs because of different input files and other parameters. Another problem is that very few systems involve single files.

#### REFERENCES

- [1] Rajesh Singh et.al, "An approach to reinforce performance of Computer-Literature Review" Journal of Research in Science, Technology, Engineering and Management (JoRSTEM) Volume-2, issue -1, March -2016
- [2] M. C. Merten, A. R. Trick, E. M. Nystrom, R. D. Barnes, and W. W. Hwu, "A hardware-driven profiling scheme for identifying hot spots to support runtime optimization", Proceedings of the 26th International Symposium on Computer Architecture, pp. 136-147, May 1999.
- [3] R. Bhargava, J. Rubio, S. Kannan, L. K. John, D. Christie, and L. Klaes, "Understanding the Impact of x86/NT Computing on Microarchitecture", Book Chapter in Characterization of up to date Workloads, pages 203- 228, Kluwer Academic Publishers, 2001, ISBN 0-7923-7315-4
- [5] H. Cragon, Computer Architecture and Implementation, Cambridge University Press, 2000
- [6] J. E. Smith, Characterizing Computer Performance with one Number, Communications of the ACM, October 1988.
- [7] Patterson and Hennessy, Computer Architecture: The Hardware/Software Approach, by
- [8] Hennessy and Patterson, Morgan Kaufman Publishers, 2nd edition, 1998, ISBN 1558604286
- Ishaan Bassi, "Comparisons of Computer Systems for his or her Performance Evaluation", 2016. Available at <http://wikidot.com>
- [9] Resource Shelf, "Performance Evaluation And Benchmarks Of Systems", International Journal Of Computing Science, 2017.
- [10] C. Barakat, P. Thiran, G. Iannaccone, C. Diot, and P. Owezarski. Modeling Internet backbone traffic at the flow level. IEEE Transactions on Signal Processing, 51(8):2111–2124, 2003.