4.2.8 Performance

Return to Non-Functional Requirements

- ☐ **[char]Please Review**

Performance is the ability of a system to accomplish the required functionality at or under the required specification limits. The limits are generally provided relative to time (i.e., so-many transactions per second, so-many updates per millisecond, so-many recorded entries per second, etc.) The specifications can also include accuracy, precision, precision or even efficiency of other dependent systems as requirements. The following are some examples of performance specifications:

- Time for a data-entry window to appear
- Number of units produced in a given amount of time
- Time to react to a given event
- Amount of energy required to perform an activity
- The amount of computing resources required (i.e., Central Processing Unit (CPU), Random Access Memory (RAM), Read-Only Memory (ROM), Storage Device, Bandwidth, etc)
- Time to process a computational task such as compression-decompression, encryption-decryption, generate a Unique ID (UID), serialize-deserialize, calculate an area, calculate a new trajectory, etc.
- Time to process a storage task such as store records, index the records, retrieve records, etc.
- Time to access memory or cache such as direct memory access(DMA), the hit ratio
- Time to transfer data between components of a computer such as from memory to Central Processing Unit (CPU), from CPU to graphics card, etc

Platform Performance

Return to Top

Platform Performance is concerned the abilities of a computer system matched with an Operating System (OS) to meet or exceed the requirements of a specific system (or project). The platform could be a real system or a virtual system running locally or in the cloud. Often this comes down to the cost of the hardware and OS, but this can also include social responsibility requirements such as energy consumption, geopolitical concerns, or even ethics which prevent the use of certain products.

Figure 1 show the salary of a Senior Software Engineer in the USA in July 2020, around $105,000. This is the take home pay, not the fully burdened cost.



Figure 1: Average Sr. Software Engineer / Developer / Programmer Salary - June 2020[1)]

Figure 2 shows the cost of either owning a Server or using a Infrastructure-as-a-Service (IaaS) solution. The Total Cost of Ownership (TCO) cost is about 1o times that of an IaaS.

	On-premises	Cloud	Savings $	Saving %
Year 1	39,347.18 $	3,766.80 $	35,580.38 $	90%
Year 2	9,063.19 $	3,766.80 $	5,296.39 $	58%
Year 3	9,063.19 $	3,766.80 $	5,296.39 $	58%
Year 4	9,063.19 $	3,766.80 $	5,296.39 $	58%
Year 5	39,347.18 $	3,766.80 $	35,580.38 $	90%
Year 6	9,063.19 $	3,766.80 $	5,296.39 $	58%
Year 7	9,063.19 $	3,766.80 $	5,296.39 $	58%
Total:	124,010.31 $	26,367.60 $	**97,642.71 $**	**79%**

Please note that inflation is not factored into these results

Figure 2: Average Total Cost of Ownership (TCO) versus Cloud [2)]

If your application is not performing well and it is estimated that it might take one year to upgrade the software, you could get a senior Software Engineer for the $105,000 or you could upgrade your server. If it is a purchase, you could get almost three servers for the cost of and engineer. If you decided to use IaaS, for the cost of the engineer, you could get almost 20 servers.

It is recommended to follow the guidelines for sizing a server provided on-line.[3)4)5)]

Application Performance

Return to Top

Application Performance are measures of real-world performance and availability of applications. It is often used to describe remote and cloud computing applications, however, there is a lot of time spent on tuning applications running locally to enhance or improve their performance. For example file access

time, database, infrastructure, graphics, human interfaces, network access, etc.

Most applications today are not Brownfield versus Greenfield systems. In addition, many programs rely on incorporating reusable components as part of the cost reduction strategy. Reusing components such as operating systems, databases, Access Control etc. makes a lot of sense since the functionality of these components is not an area of expertise for the developers of the system. For example, if the system is for medical devices, then designing new virtual page swap software has nothing to do with medical technology or medicine. Another example might be developing middleware software for accessing other computers on the network. This kind of development might seem “fun” but it is not within the scope of the project that is developing the system. These re-used components also have the benefit that they are used by a wide range of other projects and systems accessed by many people. Collectively the communities can find and fix software faster that it can be done locally.

I don't believe there is a company that could pay for all the testing that the user communities provide. Linux is a great example.[6)]

- *About 3 to 3.5 billion, which is pretty much the number of humans on the planet who have regular access to the internet. Let’s break it down….*
- *Directly as a desktop OS? A minority of users, sadly windows is hogging over 85% of that market share - which it certainly doesn’t deserve. The remaining 10 to 15 percent is mostly Apple, who have done a very good job at creating its desktop OS based on a Unix implementation, so it’s kind of a “cousin” of Linux in that sense.*
- *If you use an Android phone or tablet, you’re using the Linux kernel and GNU software indirectly. Many Android devices will show this in the “kernel info” or “kernel version” in the settings app’s About section.*
- *If you have a router, modem, printer, scanner, or any similar device, its firmware is probably based on a small Linux or FreeBSD system.*

It is also true that this many users also marks Linux as a huge target for malicious activities, but pretending that any OS is not vulnerable to attack is a fool's game.

Network Performance

Return to Top

Network Performance captures the statistical metrics and the analytical review of a network. Collectively they reflect the network's Quality of Services.

It is a qualitative and quantitative process that measures and defines the performance level of a given network. It guides a network administrator in the review, measure and improvement of network services.

There are two main ways to connect devices together:

- Ethernet (i.e., wired using network cables.)
- Wireless Fidelity (Wi-Fi) (i.e., wireless using radio signals).

Although it is possible to connect computers directly together, generally the computers connect to a

Network Device such as a Router. There are a number of variables determine actual speed of the connection between the computers. The wired connections are as a rule faster than the wireless connection but the number, length of the quality of the network connections and the kinds of network devices and the number of devices can impact wired connection speed. WiFi connections are more susceptible to interference from electrical devices, physical objects (i.e., metal walls or cages), or environmental conditions (i.e., weather and solar flares).

An Ethernet connection is consequently more reliable especially when Shielding Network Cabling (i.e., Category 6 (Cat-6),Category 7 (Cat-7), Category 8 (Cat-8)) are used. Ethernet is almost always faster than WiFi. The fastest Ethernet speeds today top out at 10Gbps or higher, while the fastest WiFi speeds theoretically max out at 6.9Gbps, though actual speeds are much slower – usually less than 1Gbps.[7)].

Speed

Return to Top

Network speed is for the most part about acquiring the correctly sized physical Network Devices (i.e. Modem, Router, Switches, Network Cabling, etc.) to meet the demands of the system. However, there are restrictions that arise such as the need for wireless connections (i.e., WiFi, Bluetooth, ZigBee, Infrared Wireless Networking etc.), space and heat considerations (i.e., a big problems for planes, ships, labs, hospital rooms, etc.) or when the assets participating in the system are distributed and not under the control of a single source (i.e., blockchains, Distributed Ledger Technologies (DLT), supply chains, etc.).

Most Ethernet connections fall into the following categories: Wired Network and Wireless Network

The wired networks use hardware such a modems, routers, switches, cabling, etc. together.

Wireless cables to connect to each and every one of the computers in the network. The cost of a wired network is lower compared to the wireless network since Ethernet, cables, and switches are not expensive. Wired LAN offers better performance compared to wireless networks.

Wired Connections

Table 1: Summary of Difference between CAT-5 through CAT-8[8)]

Category	Standard	Data rate	Frequency[1)]	# of Conductors
Category 5 (Cat-5)	100BASE-TX	100Mbit	100 Mhz	4 or 8
Cat-5E	100BASE-TX	1Gbit	100 Mhz Duplex	8
Category 6 (Cat-6)	EIA 568B2.1	1-10 Gbit[2)]	250 Mhz	8
Cat-6A	10GBASE-T	10 Gbit	500 Mhz	8
Category 7 (Cat-7)	10GBASE-T	10 Gbit	600 Mhz	8
Cat-7A	10GBASE-T	10 Gbit	1000 Mhz	8
Category 8 (Cat-8)	40GBASE-T	40 Gbit	1600-200Mhz Mhz	8

[1)] **Note:** 1 hertz is roughly equivalent to 1000 milliseconds, 20 killohertz is roughly equivalent to 0.05

milliseconds. See Unit Juggler.
[2] **Note:** Depends on the length and cable type

Wireless Connections

Table 2: Side-by-side comparison of wireless routers[9]

	Netgear Nighthawk X10 AD7200	Asus RT-AC86U AC2900	Linksys EA6350 AC1200+	TP-Link Archer C7 AC1750	Trendnet AC2600	TP-Link AC2300	Linksys WRT32X
Top Theoretical Speed	4600 Mbps on 60 GHz	2167 Mbps on 5GHz	867 Mbps on 5 GHz	1300 Mbps on 5 GHz	1733 Mbps on 5 GHz	1625 Mbps on 5 GHz	2600 Mbps on 5 GHz

Bandwidth

Return to Top

Bandwidth is defined as the bandwidth data carrying capacity of a network channel or the entire network. Bandwidth measured is bit-rate measure of the network transmission capacity. This is sometimes thought of as the network channel's data transfer speed. Bandwidth can be used to describe wired, wireless or even data buses.

A bit represents a single binary digit either '0' or '1'. The '0' or '1' generally represent yes/no, true/false, on/off, or up/down/ It does not necessarily equate a '0' with false and a '1' with true. When transmitted over a network, the data is sent as a stream of bits (not bytes). A byte is generally used to signify a unit of memory or storage (i.e., RAM or ROM) that usually is eights bits long (wide) and is the smallest number of bits used to represent a character in the original ASCII character set used by most most computers.

Bandwidth is measured as bits per second and is used as a denominator of bits (i.e., kilobits, megabits). When bndwidth is used to describe a network connection (i.e., switch, server or router), it is generally in megabits, however, when it is used to describe the data flowing into the connection then bandwidth referred to as traffic and could be measured in either megabits per second (Mb/s or Mbps) or megabytes per second (MB/s or MBps). Although the nomencalture is subtle, it is important to be aware of the difference. An inadvertent misunderstnding could result in a error of magnitude 8 (i.e., 1 byte = 8 bits).

Since the megabytes figure will be larger than the megabits figure (equation to follow shortly) most industry service providers like to give a total transfer based on this figure – however most bandwidth providers use megabits.

$$1byte = 8\,bits$$

Some examples:

Table 3: Some examples of converting MegaBits to MegaBytes

Mega Bytes per second	bits per Byte	MegaBits per second
8 MBps	*8	64 Mbps [1]
9 MBps	*8	72 Mbps
10 MBps	*8	80 Mbps

Mega Bytes per second	bits per Byte	MegaBits per second
20 MBps	*8	160 Mbps

[1)] **Note:** there are two kinds of units listed Mbps (Mega BIT per second) and MBps (Mega BYTE per second)

> *Today, many cable ISPs are capable of delivering internet speeds over 1 Gigabit per second. That's 1 billion bits per second! Not everyone needs this much speed today (Netflix reports that a connection speed of 25 Megabits per second is all that's required to stream Ultra HD content), but cable ISPs see a future of virtual reality, telehealth, driverless cars, and an internet of things. In that environment, speed requirements are going to increase. Regardless of whether it's necessary today, ISPs are preparing their networks for the needs of the future. So, while we'll likely always measure speed in bits and data volume in bytes, the consistency and speed at which those bits are delivered over the internet will surely rise.*[10)]

In addition to the Internet Service Provider (ISP) bandwidth limitations, it is also important to remember that there are local hardware components to the network. For example, Network Quality of Service (QoS) Policies.

Network Quality of Service (QoS)

Return to Top

Quality of service (QoS) is captures the metrics used to measure the overall performance of a computer network as experienced by participants (i.e., computers, processes, devices, etc) in network.

Internet Protocol (IP) networks, QoS is particularly focused on setting priorities for packet traffic and reserving resources rather than the QoS of the Network Services (i.e, DNS). It helps establish overall priorities for applications, users, or data flows, or to guarantee a certain level of performance to a data flow. For example, setting Voice over IP (VOIP) as a priority over texting.[11)] Some of the common metrics used to describe Network QoS are:

- Throughput
- Latency
- Packet Loss
- Jitter
- Bit Error
- Download Speed
- Upload Speed

DDS Specifics

Return to Top

One of the more common requirements for communications middleware is performance, usually specified as either Latency or Throughput or a combination of both.

Latency is a measure of the total time it takes a data packet to travel from one Network Node to another network node. Commonly, it is the amount of time it takes for a data packet to travel round trip from one node to another node and back again. In essence, Latency refers to time interval or delay in a system to wait for inter-component communications. The duration of waiting time is called Latency.

Throughput is a measure of the quantity of data transferred from one location to another in a given amount of time, and within the Data Distribution Service (DDS) context is the data transferred over the network.

Network performance is one characteristic that will vary between Data Distribution Service (DDS) implementations. However in general, and for most DDS products, Latency is measured in microseconds. For some of the high-performance DDS implementations, the expected latencies are under 50 microseconds for network communication between network nodes and under 30 microseconds for “on-box” communications. For embedded environments, such as rtos, networkstack, or even the network hardware, a Latency is occurs because of various restrictions and limitations which can result in slightly higher Latencies, but still measuring Latencies in the microseconds.

Looking at Throughput, in general DDS will have Throughputs approaching the theoretical maximum limits of the network type. This holds true for networks rated at 1 megabit per second all the way up to 10 Gigabit per second. DDS takes advantage of aggregation and batching of smaller data packets to further reduce overhead.

These Latency and Throughput numbers make DDS a good choice for all kinds of data and all kinds of data streams.

The DDS specifications were written specifically for near real-time systems. Some examples of specification items result in high-performance communications:

- Specifies minimal data copies in memory - Buffer loaning
- Specifies compact binary encoding on the wire - Reducing bytes on the network. There are no large eXtensible Markup Language (XML) or JavaScript Object Notation (JSON) strings being sent over the network.
- Specifies light-weight notification mechanisms - Asynchronous and Synchronous options available
- Data types are specified at compile time - Allowing optimized: marshalling, filtering, searching
- Specifies User Datagram Protocol (UDP) multicast allowing for 1-to-many communications. Additionally, besteffort and reliable protocols can be used. This reduces network overhead and high performance scalability.
- Intermediate brokers are not required - providing direct Peer-to-Peer (P2P) communications

Return to top

Figure 3:

1)

Payscale.com, Accessed 20 July 2020, https://www.payscale.com)

2)

Cost of server ownership: on-premise vs. IaaS, April 2019, Sophie Furnival, Sherweb, Accessed 20 July

2020, https://www.sherweb.com/blog/cloud-server/total-cost-of-ownership-of-servers-iaas-vs-on-premise/
3)

Basic Guidelines for Sizing Servers, Jason Thomas, 24 August 2017, Accessed 20 July 2020, https://www.mirazon.com/basic-guidelines-for-sizing-servers/
4)

3 secrets to right-sizing a server, TidalScale, Accessed 20 July 2020, https://www.tidalscale.com/3-secrets-to-right-sizing-a-server/
5)

Determining Your Ideal Server Size: Which Package is Right for Me?, Media Temple, 7 April 2015, Accessed 20 July 2020, https://mediatemple.net/resources/web-hosting-101/determining-your-ideal-server-size-which-package-is-right-for-me/
6)

Éric Nunya, 9 March 2020, Quora, ccessed 27 July 2020, [https://www.quora.com/How-many-Linux-users-are-there-in-the-world[]]
7)

What is the difference between a WiFi and Ethernet connection?, Spectrum Enterprise, Accessed 27 July 2020, https://enterprise.spectrum.com/support/faq/network/what-is-the-difference-between-wifi-and-ethernet-connection.html
8)

How to Tell CAT Data Cables Apart, Horst Messerer, Machine Design, 8 January 2016, Accessed 24 July 2020, https://www.machinedesign.com/automation-iiot/cables-connectors-enclosures/article/21834690/how-to-tell-cat-data-cables-apart
9)

Connect at Speed with the Fastest Routers, Speedtest, PCMag, Accessed 21 July 2020, https://www.speedtest.net/about/knowledge/fastest-routers
10)

Why Do We Use Bits to Measure Internet Speed but Bytes to Measure Data?, The Internet and Television Association (NCTA), 21 July 1017, Accessed 20 July 2020, https://www.ncta.com/whats-new/why-do-we-use-bits-measure-internet-speed-but-bytes-measure-data
11)

Andrew Froehlich, The Basics Of QoS, 15 August 2016, Accessed 27 July 2020, https://www.networkcomputing.com/networking/basics-qos

From:
https://www.omgwiki.org/dido/ - **DIDO Wiki**

Permanent link:
https://www.omgwiki.org/dido/doku.php?id=dido:public:ra:1.4_req:2_nonfunc:40_performance&rev=1605926006

Last update: **2020/11/20 21:33**