

Inside Citrix chapter twenty four – The one with sizing and storage considerations

With XenApp, sizing is all about how many users we can fit onto a single machine, physical or virtual, without compromising, or affecting the overall user experience. With a XenDesktop VDI it is always one user per desktop and preferably it needs to perform as close to a physical PC as possible, or better (although this applies to RDSH as well, of course). Unfortunately there is no ‘one size fits all’ when it comes to sizing your RDSH and/or VDI machines; since there are multiple factors to consider. It will greatly depend on the type of applications used (video, audio, number crunching etc.), the intensity with which your users will use them, the server and/or desktop Operating System used and so on.

The underlying storage infrastructure also plays an important role in all this. While your server and desktop machines can have all the CPUs, GPUs and memory in the world, if your storage platform does not deliver the requested amount and types of IOPS (which can be very workload-specific) and/or storage throughput needed (never mind the GBs needed to actually build your machines) you are set up for failure. The types of machines you deploy (non-persistent and persistent) and the provisioning mechanism used will need to be taken into consideration as well.

This is why it is very important to understand the different types of workloads you have to deal with and the different forms of provisioning at your disposal. As we have seen in some of the previous chapters, there are some distinct differences between MCS and PVS, for example, and they both impact storage in different ways (write cache considerations, differencing disks).

This is also where things like application and baseline performance statistics come into play to help you better understand what is expected and needed within your newly built XenDesktop / XenApp Site. Perhaps this is data you already have or it’s something your customer can provide and of course talking to certain applications vendors might help as well. However, in most cases it will come down to real-world testing, and then test some more.

Do not overdo it

Although calculating storage needs (free space, IOPS, throughput, latency included etc.) and defining application performance profiles (CPU, GPU, memory etc.) can take up quite some time and effort, be careful when doing so. Users can be very unpredictable and even if you spend weeks on calculating your potential storage and compute needs it won’t guarantee anything. The same applies to load testing, by the way. It is an indication and nothing more. Real-world testing, as highlighted earlier, is the only true way of validating what you have built. See what happens and talk to your users.

Some help

Luckily, when it comes to sizing your XenDesktop / XenApp compute and storage needs, we can use multiple resources to draw inspiration from. For one we can learn from others: what was their approach in the past? Did they run into any ‘specials’, and if so, how did they handle them? I’ll include some helpful links at the end of this chapter. I can also recommend having a look at the Citrix XenDesktop Design Handbook, and, although it is a bit outdated (it is based on version 7.6), it does offer a whole bunch of best practices with regard to building and configuring your Site, including compute and storage sizing considerations for building up your Delivery Controllers, StoreFront servers, database server, VDAs and more.

Also, next to one or two community-powered sizing calculators, Citrix also offers us Project Accelerator.

Citrix Project Accelerator

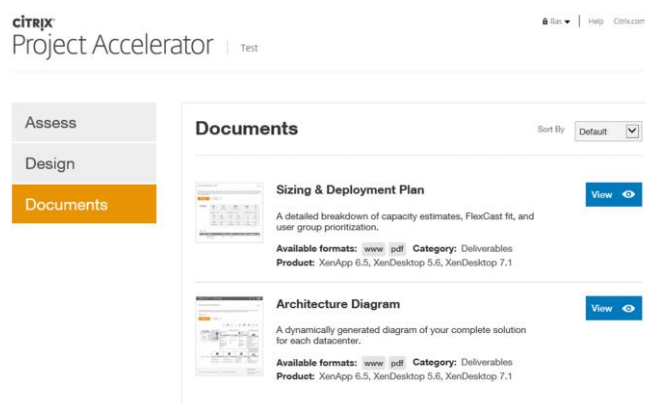
I won’t go over Project Accelerator from start to finish since it is still, and will continue to be, a work in progress. However, the basics and concepts remain the same, as they often do. Project Accelerator is a web-based application (project.citrix.com) that guides you through various questions, requirements, and common design choices.

It is founded on the practices used by the experts within Citrix Consulting and it is regularly updated with new information from the Citrix consultancy and lab guys keeping it fresh and current, although at the time of writing it currently supports up to XenDesktop / XenApp version 7.5. But don’t let that hold you back from giving it a try.

Once you have signed up and started a project you will first enter the Assess Phase. Here you will need to define your organisation: which is a five-step process. First you provide information on your business, the type of project, user groups and applications, and finally you will need to link those applications to your user groups.

This phase will include questions on topics like the type of industry that you are in, how many users this particular solution will need to support, primary business priorities like BYOD, being able to work from anywhere, reducing costs, better desktop management and so on.

They will ask you for any existing (Citrix) skill sets (XenApp, XenDesktop, NetScaler, XenServer, XenMobile, App-V, Hyper-V, VMware and so on: it’s a pretty long list) within the team responsible for building and configuring the new Site. Take your time and fill in all the fields as detailed as possible: this is by far the most important step / phase.



Citrix Project Accelerator

FMA fact: When thought through beforehand this phase will probably take you somewhere between 30 to 60 minutes, depending on the number of users and user groups, including the number and the type of applications involved.

Based on the information entered during the Assess Phase, the second phase, Design, will make certain technical recommendations with regard to the products used to actually build the solution, the FlexCast delivery model, hardware and sizing recommendations (compute resources in general, including suggestions around CPU overcommitting and so on), IOPS and more. Do note that you will be able to manually adjust all suggestions made by Citrix, based on personal experience.

This phase also includes recommendations around image provisioning and management like PVS and MCS, including the preferred types of storage.

As soon as you are satisfied you can save your work and skip to the Documents sections where you will find an Architecture diagram and a Sizing and deployment plan based on the information you entered during the first two phases. Use it to your advantage.

IOPS and general storage considerations

It is, or at least can be important to understand what an IOPS actually is, why they (multiple) are important, how they are measured, the different types of IOPS, which factors influence performance, either negatively or positively etc. Being familiar with the IOPS fundamentals will give you a better insight into what is actually going on under the hood and how things interact and rely on each other.

Nowadays infrastructures are being virtualised almost without exception, and as we all know, the success of our projects greatly depends on the end-user acceptance. Ultimately they expect to experience the same snappiness as they do on their physical desktops or even better. Of course this is what we as consultants and system administrators aim for as well, but unfortunately it isn't always that straightforward. The bottom line is, we need speed! If we drill down just a bit further speed closely relates to IOPS, you'll read why in a minute. Bottom line: we need as much (low latency) IOPS as we can get our hands on to drive our VDI and/or RDSH-based infrastructures to their full potential.

Let's start at the beginning, IOPS stands for: Input / Output Operations Per Second, which in general is either a read or write operation. Simply put, if you have a disk that is capable of doing 100 IOPS, it means that it is theoretically capable of issuing 100 read and/or write operations per second. However, being able to issue 100 read and/or write operations isn't the same as actually processing them: reading and writing data takes time. This is where latency comes in. If our disk subsystem can handle, or issue, 100 IOPS but they are processed at around 20 milliseconds per operation (which is slow, by the way), then it will only be able to actually handle 50 operations per second, as opposed to the issued 100.

In the above example, 20 milliseconds is what we would refer to as the latency involved. It tells us how long it will take for a single IO request to take place or be processed.

There is a lot more to it

Remember that a random IOPS number, on its own, doesn't say anything. We can do a million IOPS! Well, ok, that's nice, but how did you test? Were they read or write operations? If mixed, what was the percentage reads vs. writes? Writes are more resource-intensive. Did you read from cache? What was the data block size? How many host and disk controllers were involved? What type of storage did you use? Was there RAID involved? Using RAID will probably negatively impact the IOPS number, but still. The same applies to data tiering. Physical disks? Probably. If so, are we dealing with sequential or random reads and writes? In addition, and this is probably the most important one, how much latency is involved in milliseconds? This will range from around 2 milliseconds, which is comparable to a locally installed physical disk, to 20+ milliseconds at which performance, if any, will be highly impacted. An overview:

- 0 – 12 milliseconds – Looking good: the lower the number, the better off you are.
- 10 – 15 milliseconds – Still acceptable in most cases: users might notice a small delay.
- 15 – 20 milliseconds – Step up and take action: most of your users won't be happy.
- 20 – 25 milliseconds – Get your pen and paper out and shut it all down.

FMA fact: A high number of IOPS is useless unless latency is low! Even with SSDs which are capable of providing a huge number of IOPS compared to traditional HDDs, latency matters. Latency tells us how long it takes to process a single read or write I/O request.

With 'legacy' physical disks, overall speed and latency greatly depend on the rotations, or revolutions, per minute (RPM) a certain disk is capable of: the laws of physics apply. Today we can classify hard disk speeds (HDD) as follows: 5400 rpm, 7200 rpm, 10,000 rpm and 15,000 rpm. A higher rpm equals higher read and write speeds.

Another factor impacting performance is disk density. The higher the density, the more data a disk is able to store on its 'platter'; data will be written closer together on the disk, and as a result the disk's read/write head will have to travel shorter distances to access the data, resulting in higher read and write speeds.

This may sound like a small note to some, but imagine having a SAN or Filer holding hundreds of disks: having 15,000 rpm and high-density disks makes a real difference! So when a random vendor tells you that their storage appliance is capable of doing a crazy high number of IOPS, you probably have a few questions to ask them, right?! I think it's also clear that the more IOPS we can actually process, as opposed to issue, per second, the better our overall performance will be!

FMA fact: Latency is king: the less you have, the faster your infrastructure will be! Also, there is no standard when it comes to measuring IOPS! There are too many factors influencing overall performance and thus the number of IOPS.

They are not the same

Not all IOPS are the same: sure, you could boil it down to it being either a read or a write, but that's not the whole truth now, is it? First of all, reads and writes can be random or sequential, reads can be reread and writes can be rewritten, single and multiple threads, reads and writes taking place at the same time, random writes directly followed by sequential reads of the same data, different block sizes of data that get read or written, ranging from bytes to megabytes and all that's in between, or a combination of the above.

As mentioned earlier, it is important to understand your application workloads and their characteristics with regard to the IOPS they need. This can be a very tricky process.

Take block size (just one of many examples): having a huge number of smaller data blocks as opposed to having a relatively small number of larger data blocks can make a huge difference. Having said all that, remember, don't go nuts; you do not have to get to the bottom of it all, all the time.

Ask your storage providers for detailed test procedures, how did they test and what did they use. In addition, at a minimum you will want to know these three 'golden' parameters:

- The latency, in MS, involved
- The read vs. write ratio
- Data block sizes used.

Steady state, boot and logon

We have already highlighted read and write IOPS: both will be part of your workload profile. However, a lot of application vendors will refer to an average amount of IOPS that is needed by their workload to guarantee acceptable performance. This is also referred to as Steady State IOPS, a term also used by Citrix when they refer to their VDI workloads.

After a virtual Windows machine boots up, users log in and applications are launched, and your users will start their daily routines. Seen from an IOPS perspective, this is the Steady State. It is the average amount of read and write IOPS processed during a longer period of time, usually a few hours at least.

FMA fact: Although the average amount of IOPS, or the Steady State, does tell us something, it isn't sufficient. We also need to focus on the peak activity measured between the boot and the Steady State phases and size accordingly.

When we mention the 20:80 read/write ratio we are usually referring to the Steady State. Something you may have heard of during one of the many MCS vs. PVS discussions. As you can see, the Steady State consists mainly of write I/O; however, the (read) peaks that occur as part of the boot and/or logon process will be much higher. Again, these rules primarily apply to VDI-type workloads like Windows 7 & 8.

There are several tools available helping us to measure the IOPS needed by Windows and the applications installed on top. By using these tools we can get an idea of the IOPS needed during the boot, logon and Steady State phases as mentioned earlier, as well as application startup. We can use Performance Monitor: using certain PerfMon counters it will tell us something about the reads and writes taking place, as well as the total amount of IOPS and the Disk queue length, also telling us how many IOPS are getting queued by Windows. Have a look at these counters:

- Disk reads/sec – read IOPS
- Disk writes/sec – write IOPS
- Disk transfers/sec – total amount of IOPS
- Current Disk Queue length – IOPS being queued by Windows.

Here are some interesting tools for you to have a look at: they will either calculate your current IOPS load or help you predict the configuration and IOPS needed based on your needs and wishes.

1. Iometer – measures IOPS for a certain workload: iometer.org/doc/downloads.html
2. ESXTOP – specific to ESX, provides certain disk states, totals, reads and writes: yellow-bricks.com/esxtop/
3. WMAROW – web interface, used to calculate performance, capacity, random IOPS: wmarow.com/strcalc/
4. The Cloud Calculator – web interface, disk RAID and IOPS calculator: thecloudcalculator.com/calculators/disk-raid-and-iops.html
5. Process Monitor – general analyses of IOPS: <http://technet.microsoft.com/en-us/sysinternals/bb896645.aspx>
6. Login VSI – VDI workload generator, simulate user activity on your infrastructure: loginvsi.com/

As we will see shortly, there is a distinct difference between the boot and logon phase. Both (can) create so-called ‘storms’, also referred to as a boot storm and/or a logon storm, potentially impacting overall performance. This is where the read IOPS peaks mentioned earlier come in.

FMA fact: Storage throughput isn’t the same as IOPS. When we need to be able to process large amounts of data, bandwidth becomes important: the number of GB/sec that can be processed. Although they do have an overlap, there is a clear difference between the two.

Be aware that RAID configurations bring a write penalty; this is because of the parity bit that needs to be written as well. A write can’t be fully completed until both the data and the parity information are written to disk. The time it takes for the parity bit to be written to disk is what we refer to as the write penalty. Of course this does not apply to reads.

When looking at VDI workloads, we can break it down into five separate phases: boot, user logon, application launch, the Steady State, and logoff / shutdown. During the boot process, especially in large environments, dozens of virtual machines might be booted simultaneously, creating the earlier highlighted boot storm. Booting a machine creates a huge spike in read I/O, as such, and depending on the IOPS available, booting multiple machines at once might negatively impact overall performance.

FMA fact: If IOPS are limited, try (pre-)booting your machines at night. Also, make sure your users can’t reboot the machines themselves.

Using this method will only get you so far; there might be several reasons why you may need to reboot multiple, if not all, machines during daytime. Something to think about as your VDI environment might not be available for a certain period of time.

Logon storms are a bit different in that they will always take place during the morning / day. It isn’t something we can schedule during the night, users will always first need to logon before they can start working.

Although this may sound obvious, it’s still something to be aware of. Logons generate high reads (not as high as during the boot process) and less writes, although it’s near to equal. This is primarily due to software that starts during the logon phase and the way that user profiles are loaded. Using application virtualisation, layering, folder redirection, Flex Profiles etc. will greatly enhance overall performance.

As you can see, especially during the first few phases, there will be a lot of read traffic going on. Fortunately reads are a bit less demanding than writes and can be cached quite simple so they can be read from memory when the same data is requested multiple times.

The way this happens differs per vendor / product. Just be aware that, although caching might be in place, it won’t solve all your problems by default. You will still need to calculate your needs and size accordingly. It does make life a little easier, though.

Although we are primarily focusing on IOPS here, note that the underlying disk subsystem isn't the only bottleneck per se. Don't rule out the storage controllers, for example: they can only handle so much; CPU, memory and network might be a potential bottleneck as well. RAID penalties, huge amount of writes for a particular workload, data compression and/or de-duplication taking place, and so on.

FMA fact: Launching applications will generate high read I/O peaks and initial low writes. Chances are that after users log on they will start, either automatically or manually, their main applications. Again, this is something to take into account, as this will probably cause an application launch storm, although it's usually not recognised as such.

Steady State we already discussed earlier: this is where write I/O will take over from read I/O, on average this will be around the earlier mentioned 20:80 read/write ratio. So if we scale for the peaks, read as well as write, we can't go wrong – at least that's the theory.

What else?

Today the market is full of solutions and products helping us to overcome the everlasting battle against IOPS. Some are 'patch' like solutions, which help speed up our 'legacy' SAN and NAS environments using SSDs and Flash-orientated storage in combination with smart and flexible caching technologies, while others focus on converged-like infrastructures and in-memory caching.

These in-memory solutions are somewhat special: they don't necessarily increase the number of IOPS available, but instead they decrease the number of IOPS needed by the workload because writes go to RAM instead of disk, ultimately achieving the same or even better result(s).

Using any of these types of products will greatly increase the number of IOPS available (or decrease the number of IOPS needed) and decrease the latency that comes with it. Just know that no matter which solution you pick, you will still have to determine the number of IOPS needed and scale accordingly.

Even when using the most enhanced IOPS accelerator today won't guarantee that you will be able to boot your entire VDI infrastructure during daytime and won't run into any issues.

FMA fact: By leveraging RAM for writes, a.k.a. RAM Cache with Overflow to Disk in terms of Citrix PVS write cache, we can significantly reduce the number of IOPS needed. In fact, Citrix claims to only need 1 to 2 IOPS per user on a XenApp environment without any complex configurations or hardware replacement.

Give your antivirus solution some extra attention, especially with regard to VDI. You will be glad that you did. I suggest you start here:

brianmadden.com/blogs/rubenspruijt/archive/2013/01/15/project-vrc-antivirus-impact-and-best-practices-on-vdi.aspx

Last but not least, during logoff and shutdown we will see a large write peak and very little read activity. Just as we need to plan for peak reads at the time of boot and logon, the same applies for write peaks during logoff. Again I'd like to emphasise that, although there might be a ton of IOPS available, it's the speed with which they are handled that counts! Less latency equals higher speeds.

To over-commit, or not to over-commit...

While I could advise you to never over-commit a hosts physical CPU / Core more than 3: 1, or 4: 1, and to always configure at least 4, or perhaps 8 vCPUs when it comes to sizing virtual XenApp servers, for example, I won't. I won't, because it doesn't make much sense, at least not without some more background information.

First you would need to answer questions like; what type or types of workload (s) will I be supporting, are there any specific characteristics that I need to take into consideration (CPU, GPU and/or RAM intensive)? How many users do we want to 'cram' onto a single machine, what is the underlying Operating System used and what about any anti virus and monitoring software? They both have an impact on density / performance as well. What type of hardware are we going to use?

And while some application (and hardware) vendors can give you specific information on what to look out for, unfortunately in most cases you will have to find out for yourself, meaning trial and error.

Having said all that, throughout the years certain (performance) baselines have been established (based on specific hard and software configurations / profiles) and users have been categorized as lightweight, medium or heavy depending on the types of workloads they use and the intensity with which they use them. These 'baselines' advise us on the number of users per CPU / core, the amount of RAM needed, including the number of (steady state) IOPS, again, on a per user basis. Here the medium type user is by far the most popular one.

He or she will use applications like Office, Acrobat Reader, one or two Internet Browsers using multiple tabs at any given time, e-mail in the form of Outlook, Gmail etc. and of a course a YouTube movie will be played every now and again, in short, a typical office worker.

Having this type of information on your users makes it (a lot) easier to start from scratch when it comes to sizing your XenApp machines and/or VDI based VM's. Here my statement made earlier, where I mention that it doesn't make much sense to advice starting out with 4 or perhaps 8 vCPUs for a XenApp server, doesn't really apply.

Although, the number of preferred users, including their intensity levels etc. but also the type of hardware used (are there GPUs in play as well? CPU speed, underlying storage platform etc.) will most certainly influence the total number of compute resources needed, you will have to start somewhere, right? So starting out with lets say 8 vCPUs and 16 / 32 GBs of RAM, for example, on a Server 2012R2 virtual XenApp server might not be such a bad idea. In the end we probably all have our own approach, although most will not differ too much is my guess.

If you go through the XenDesktop Design Handbook it will show you a couple of specific formula's enabling you to get a (ballpark) figure regarding the number of physical CPU cores needed, depending on a certain amount and type of users, including the underlying Operating System used, the same goes for physical RAM, nice to help you on your way. Have a look at the following list of links; it includes the XenDesktop Handbook and Project Accelerator as well.

- The XenDesktop 7.5 sizing calculator by Andrzej Gołębiowski
<http://blog.citrix24.com/xendesktop-7-sizing-calculator/>
- Have a look at this post from Thomas Gamull. It is on sizing (real world) Citrix environments with a specific focus on XenApp 7.x:
<https://www.citrix.com/blogs/2014/04/01/quick-server-sizing-for-xenapp-7/>
- The XenDesktop Design Handbook (direct link to the .PDF document):
<http://support.citrix.com/servlet/KbServlet/download/35949-102-713877/Citrix>
- Sign up and log into mycugc.org and do a search for server / machine sizing. You will come across multiple threads discussing the subject.
- Citrix project Accelerator: <https://project.citrix.com>

Key takeaways

- When it comes to sizing your XenDesktop / XenApp infrastructures: there is no ‘one size fits all’.
- You need to understand the workloads you have to deal with and size accordingly.
- There is more to it than ‘just’ compute resources: don’t forget about your underlying storage platform.
- While sizing is important, try not to overdo it. Real-world testing will always be needed.
- Ask peers for help and/or advice, consult the Citrix XenDesktop Design Handbook, use Project Accelerator and test, test and test some more.
- Load testing can give us an indication of what might be possible, but remember that your users can be (very) unpredictable.
- When conducting load tests, always try to incorporate any exotic applications that you might have. These are the ones you should be most curious about.
- IOPS fundamentals help you in understanding what is going on under the hood. It will also help in understanding what other IT folks might be talking about in other articles / blogs.
- A random IOPS number on its own doesn’t mean anything. What type of IOPS are we talking about: reads, writes, random, sequential, rereads and/or writes, single or multiple threads, block sizes and so on. And even more importantly, what is the latency number in MS?
- Storage providers should be able to provide you with at least the latency in MS, the reads vs. writes ratio, and the data block sizes used during testing.