Your cloud strategy may just have changed, but did you realise it?

A while ago I wrote an article about the problems with how people think about cloud and how it is implemented. Peter Staddon can usually be relied upon to say my predictions in that article have come to fruition and he is of the opinion it is time to make a new prediction; this one will affect the public cloud in fundamental ways and last for years.

Over the New Year there were a couple of news articles about some problem with Intel chips. I am sure most looked at it and moved on with their day, others may have wondered about patches, but I suspect most will assume it will be fixed in a patch from somewhere and anyway it isn't in the news anymore so the problem is fixed.

To recap what happened over the new year; it turns out that in June 2017 Intel were informed of an issue in the design of every chip it produced since 1995. This issue was called "Meltdown". A couple of weeks later Intel and most other chip producers like AMD and ARM where told of an issue which is called "Spectre". The actual mechanics of these problems and the solutions are at the end of this article rather than here, but I would encourage you to read the issues. Issues with chips are frequent and fixes are normally put in something called a chip errata but both issues are both big problems and undermine the security models that all computer security is built on which is why they made it to the news.

Meltdown whilst serious has had a patch issued by Intel and the problem is gone however the sting in the tail of the Meltdown patch is that depending on your workloads you may lose between 5% and 30% of your total processor capacity. On a desktop you are not going to notice this unless you are a gamer or doing something with video editing; With a server farm where by design you run at a higher usage capacity you will probably notice it, especially with database or virtual environments but hopefully you are rarely running at such a capacity that you can't cope with this loss of overhead. Now imagine that you are not running these workloads on your own computers but in the cloud and in effect on a meter. Your cloud bill just went up by 30% for nothing new as you are billed for processor time and same work load takes longer after the patch was applied, but at least you are secure again. I suspect Intel will slowly optimise the patch to have less of an effect over the forth coming years but you will never get back to the way the processor used to work.

Spectre effects pretty much every chip from pretty much every manufacturer and this means that the issues doesn't just appear in your computer but your tablet, mobile phone or internet enabled heating system, anything with a modern chip that can be programmed for. There is no fix for Spectre and there never will be, you may get mitigations but we will be playing whack-a-mole with this issue for years until every single chip in use today has been removed. Unfortunately the chips we need to use to allow the replacement of the existing chips haven't been designed yet and won't be available for at least a couple of years, if we are lucky; and there is no way of turning off the branch prediction of a processor. However even if you could that expensive fast chip in your pc or server will start running slower than something produced in early 1990, before cleaver things like virtualization; don't forget in the early 1990's we didn't have DVD's and Amazon was a river it was a long time ago and things were slow.

So as mitigation we will be applying the whack-a-mole patches and set about isolating our processes which may be fine for your mobile and desktop but what about the servers? If you are running on your own dedicated hardware, be that on premises or co-located in the cloud then you can continue this, but if you are in the cloud which by its very definition is a shared resource how do you do this?

You can point to the fact this is a theoretical issue at present, but if you believe that people aren't actively trying to exploit this then I have a unicorn for sale. Ransomware has proved very lucrative to criminal gangs and they how have teams of software developers working for them, just last year USD225m was stolen from one crypto currency so some criminals have the resources to try for an exploit. The spies of multiple governments around the world will also be working on exploits both for offense and defensive reasons and they can never seem to hold on to their secret tools.

There are going to be tools dropping that allow a virtual machine running in the cloud to access the processor of the underlying hardware and from there to all the virtual machines running on that processor. A problem if you are on a shared resource in the cloud and I can get a credit card out and potentially run my virtual machine on the same hardware as your secure workloads; it is the prefect breach and you will never know. So how do you live with this issue? I think the only reasonable way forward is to speak your cloud account manager and start to say I want dedicated hardware for my workloads and to be removed from shared hardware and infrastructure, a discussion most cloud providers will not want to have.

Security Researchers are starting to look at processors now and it is likely to start a windfall of these errors and each one many require a processor redesign or a patch over the next couple of years, so more Meltdown patches should be planned for and Spectre whack-a-mole patches will be required but whatever your cloud bill has just gone up and it is going to keep going up for years.

Don't feel that this is my only security issue for the future; whilst the above is an interesting problem that will cause issues for years; there is another that has been causing issues for years and will probably run longer than the above. Firms and people are not learning the IT security lessons apparently taught and learnt 30 years ago, more and more industrial control equipment keeps being put on the public internet which is a major problem as they are not patched or maintained. There will come a time when buildings and other infrastructure will get played with by kids for giggles, just hope no one will get killed and before you think I am exaggerating a fake site was created last year that appeared to show a real train line with real trains on it and where signals could be controlled via the interface. This mimicked a real train line and real train movements; thousands of people logged into the site and a minority actively tried to crash trains, none reported the issue though. Before you think they wouldn't attack me. The reason that attacking so called Internet of Things (IoT) is so popular is that they typically run 24 hours a day all year and are not patched because there are no patches and are typically unmanaged or not maintained by a security team at any rate. If you can get a crypto currency miner on an IoT device it may takes days/months or even years to be noticed and it is making money for the criminal who put it there all that time. This will be the largest type of attack in the next few years.

The UK government is slowly waking up the size of the issue and has just announced they may fine infrastructure firms up to £17m if they have insufficient cyber security but will that make a difference?

Also organisations continue to collect huge amounts of data on people and then not secure those collections, we have not heard the end of people data being lost and I suspect a huge loses will happen before it even starts to get better. GDPR may help in this regard after a couple of huge fines are levied. But in the future your data will be lost and don't expect any meaningful recompense for the issues it gives you, so the only way not to lose it is to not give it out in the first place.

Then finally it seems most people are still using the same password for multiple accounts and not turning on 2 Factor Authentication when it is available.

Expect more data to be lost.

What is Meltdown?

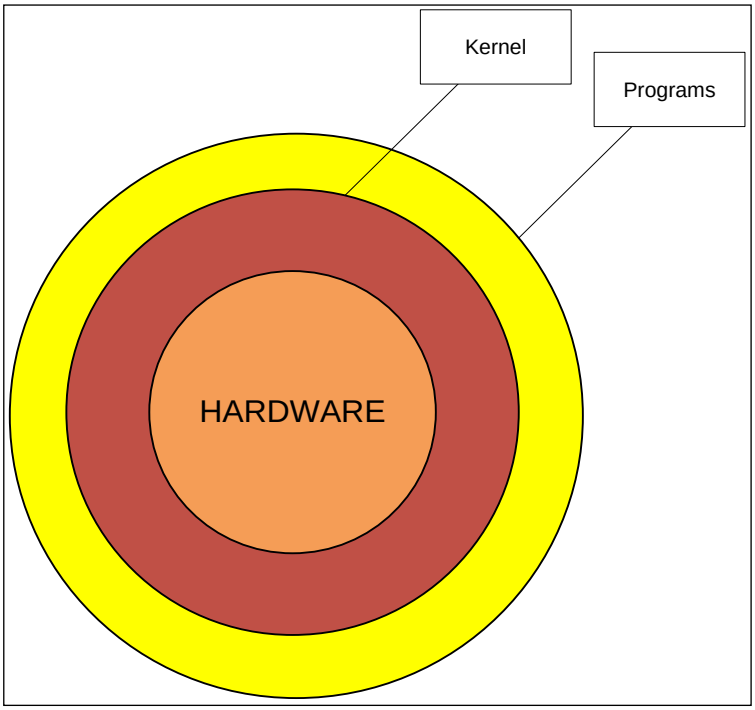
This is an Intel processors issue and it covers most processors from 1995 to now and the foreseeable future. By now patches are available and it is recommended that this patch is applied.

To explain this issue I will have to provide hopefully one simple diagram. This shows how a computer is secured and run. The hardware is run using drivers that are in the Kernel and program that you run talk to the kernel if they need anything, be that memory, a file from the disk or a webpage over the internet. A program is not allowed to access the hardware itself and must go through the kernel.

This system works well, but there is an issue in that how does the program talk to the kernel? In the very old days, what would happen is a program would make a request to access something via a call to the kernel and the processor would then 'freeze' the program and unload it and then load the kernel into the processor and run kernel which would then pick up the request and when the kernel had finished, it would be frozen and unloaded and swapped with the program which would be unfrozen and it could then pick up what the kernel had left. It was not the most efficient way of doing it, so in the 90's in an effort to make processors and computers faster the design was changed.

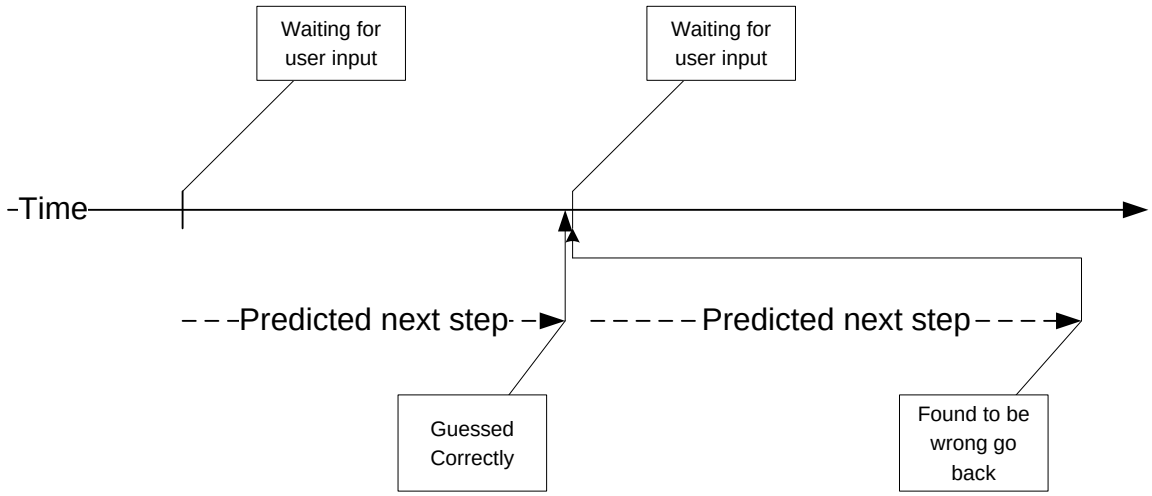
Both the kernel and the program would be run on the processor, but the kernel would be protected from interference from the programs. This left the issue of how does the program talk to the kernel. It was decided that part of the kernel would be allowed to exist in the area of the programs resources as a sort of god. The program can't see the kernel, it can't interact with it etc but if it prayed out loud the kernel could hear it and would do as asked.

The problem Meltdown found was that the program was able to prove the existence of this kernel intrusion and to interact with it. So the fix is to go back to the old way of doing calls to the kernel. This will introduce a penalty of about 5% to 30% or it will use an additional 5% to 30% of processor capacity, not something that you may notice on a desktop but on a server you probably will and if you are a cloud user, then your bill is going to go up accordingly as you are using more processor cycles. The penalty is entirely down to how much you use external things to the memory, so Databases, Virtual Machines etc are going to feel the hit worse.



What is Spectre?

This affects most processors and this issue comes from the 1990's as well. At the time everyone wanted faster computers and therefore faster processors. The problem is that most processors spend a lot of time just waiting for you to do something. So it was thought a good idea was to guess what you wanted and to pre execute this guessed code. If it is wrong then you just unwind the guesses and carry on. It was soon seen that unwinding the wrong guess was slow, so instead the processor just jumped back to where the wrong guess was made and carried on.



This is fine, but with Spectre it was found that if you can persuade the predicted next step to do something it shouldn't that security wasn't applied until it was actually looked at and was told no and go back. The problem is that the predicted instruction could be to copy something it shouldn't to somewhere the program can access it and when the security part catches up and says no, that the copy is not undone and the copy results are left. This allows a program to have read access to memory it shouldn't; for example passwords. So a malicious webpage could access your passwords that are supposed to be protected from exactly this type of attack. This attack will be mitigated in the short term in software running in the OS eg Windows, OSX etc in the medium term by getting programs recompiled and reissued and in the longer term by designing out the fault. The problem is that it takes years to design a modern processor, so even if the fix was fixed in June when processor manufactures were told, those processors won't be available to us until 2020 or 2021 at best. Again the mitigations are going take additional noticeable processor cycles. The problem is that issue effects most processors in use today. Which means as a business you may have to patch every device you own, mobile phones, tablets, pc's servers, storage everything, and be doing it repeatedly as you patch the patch of a patch etc. for years to come until everything is replaced.