Different processors from Intel are used in the cloud and selecting the right one can make a profound difference when selecting your instance.

In this episode, Darren Pulsipher, Chief Solution Architect, Intel, and Stephen Holt, Cloud Solution Architect, Intel discuss cloud optimization and studies that show that cores perform differently for various workloads.

Intel has been hiring cloud solution architects to help customers move intelligently to the right cloud instances for their workloads. Stephen has recently come on board at Intel and has a long history in databases in general, with the last seven or eight years focusing on the cloud, leveraging its resources for customers.

Some of the most difficult aspects that Stephen has run into in helping customers move to the cloud are the pricing and the expectations for the move. Cloud service providers often claim customers will save money by moving to the cloud, but this is often not the case.

One reason is that there is constant change in the competitive cloud service provider market in features, functions, and abilities. Another reason is ignorance, for example, believing that a core VCP is the same across the board as others. Migrating with a lift and shift mindset can also be very expensive. Optimization is key for intelligent cloud migration.

To decrease costs and potentially save a large amount of money, you must choose the right instances for the right performances. Not all cores, or virtual CPUs, are equal. Within AWS, they have the M4, Intel’s Broadwell, and the M5, which is a mix of Cascade Lake and Skylake, and various workloads can take advantage of those platforms very differently.

**CoreMark Performance**

CoreMark is a single threaded application that tests register moves and simple additions for any number of CPUs. Across three generations of Intel processors, Broadwell, Skylake, and Cascade Lake, this test shows very little difference among them. This information has been erroneously used in guiding customer decisions, as the test was is not an effective way of measuring the differences in the processes themselves. There can be huge differences in performances with different workloads.

**Database WL Performance**

Intel (or Stephen? – not sure who “we” is) performed digital tests on real workloads, focused on open source databases. The Cascade Lake environment was three times better than the Broadwell environment for open source, which makes sense given the advances in speed and on-chip advances that take advantage of low latency access. So going up two generations gives two and three times better performance on these workloads.

The workload and the core must be considered together. For example, a customer might be tempted to switch to a VCPU that Amazon says is 10% cheaper. AWS might, however, pack cores onto these processors and you are now moving from an environment where you’re one
of 48 virtual machines on a box to one where there's 128 virtual machines; your access to memory is limited, and you might get 60% less performance for a 10% savings.

Organizations often believe that if they outsource to the cloud, they no longer need a systems architect. It's true that architects are no longer racking and stacking machines, but they do it now virtually and need to understand what different instances provide for different workloads.

Intel is working on a tool to show which workloads run best on which instances. Meanwhile, Intel has some general guidelines and cloud solution architects to help guide customers.

In addition, a lot of testing is important to understand where to place workloads, but it's also important to test and benchmark to validate that you are getting what you expect from a provider.

Database WL Performance – Non NVMe

Intel did some testing on workloads on fresh installs on 64 core fixed 64 VCPU instance. After getting the initial number, another instance was instantiated and the same test was run again. The results varied dramatically from the same test each time. This could be because there may be a mix of processes that are offered at a certain tier level, or there could be some distance and additional latencies to block storage attachments, for example, or even noisy neighbors. So it is worthwhile to test to make sure the system meets expectations.

In short, to get the best value out of the cloud, you need to educate yourself, test the systems, and take advantage of available help. A study from over 20 years ago on the return of security investment showed the earlier security is built into the development cycle, the cheaper it is than waiting until the end of the cycle. Although we have known this for two decades, it's still a work in progress.