

## **Article 1: Reinventing the core virtualization platform**

**Client: AWS**

**June 2021**

---

### **Reinventing the core virtualization platform**

*AWS has reinvented the virtualization system with a Nitro System that provides the best price performance in the cloud, the most secure environment, and a faster pace of innovation.*

“With AWS,” Anthony Liguori, principal software engineer of Amazon told us, “we really do take a customer-focused approach. In a lot of places they say ‘We delivered this technology, now how do we get customers to use it.’ Whereas, at AWS, we say ‘We have this customer problem, how do we build something to fix it?’”

That is how AWS realized, long before other cloud providers did, that there was a problem with the tech stack and the only way it could solve that problem was to dismantle the traditional structure and rebuild its base approach from the ground up.

For most other companies, that would have been unthinkable.

“This notion that you would have to throw away your entire existing virtualization stack, something that you built a large business on top of, and completely redesign it ... it requires a mentality where you’re totally focused on delivering value to the customers.” Liguori noted.

### **The virtualization problem**

Around nine years ago, AWS engineers tested their OS and applications in “virtual machines” and used their Xen hypervisor to regulate and isolate each VM. The entire architecture was software-based. When AWS thought of its customers, they realized that this virtualization structure could be improved. After all, it was wasteful (the hypervisor absorbed too much RAM), yielded inconsistent results, and software would inevitably lag behind accelerating security demands.

There was only one way to improve the system: Migrate most of the virtualization functionalities to dedicated hardware.

The new virtualization system would be perfectly minimal.

“When I think of software.” Liguori said. “The more code you have, the more complexity you have, the more bugs you have. So one of our fundamental tenets is ‘How can you simplify the system?’”

It took AWS five years to offload the software onto Nitro hardware. During that time, they worked piecemeal. First, it was the I/O elements of their Virtual Private Cloud (VPC) that were offloaded onto hardware in 2013, known as the C3 instance. The I/O elements of the Elastic Block Store (EBS) that provides block-level storage for data followed in 2014 (this was the C4 instance). Finally, AWS offloaded the last of their components by 2017.

Right from the start, Liguori reminisced “we were able to see orders of magnitude of improvement, not just in bandwidth ... but also in packets per second (PPS). Customers absolutely loved this enhanced networking. And that's what we called it at the time ... ‘enhanced networking’.”

Along the way, Israeli startup Annapurna Labs provided the custom silicon that became an intrinsic part of the AWS Nitro System and accelerated its remote block storage.

By 2017, AWS had also introduced its new Nitro Hypervisor, with the C5 instance type.

Now, all the virtualization was happening in hardware. “So instead of talking to some complex piece of software that’s trying to emulate some piece of hardware, you’re literally talking direct to hardware,” Liguori said. “So if you launch an EC2 instance today ... on a Nitro instance, [it’s as though] you were running your instances directly on hardware.”

Their revolutionary tech stack helped AWS achieve the following:

### **Enhanced performance and price**

The minimal surface of the Nitro System makes for low-latency workloads, reducing jitter to microseconds. Their Nitro Cards divvy up high speed networking, high speed EBS, and I/O acceleration. Savings on management software - the Nitro System makes even the hypervisor optional - results in a drastic reduction in cost.

### **Security**

The Nitro System comes with unprecedented security options, like its Nitro Security Chip that traps I/O to non-volatile storage and its very limited operator accessibility Nitro Enclaves that securely processes highly sensitive data.

In contrast to the traditional virtualization of other cloud providers, the surface area of the Nitro System has only purpose-built hardware and servers designed to run a hypervisor - nothing more - which amplifies security.

## **Innovation**

The Nitro System is a rich collection of building blocks that could be assembled in various ways, helping engineers multiply their compute, storage, memory, and networking options. Results yield a dizzying spate of innovations like instances used to virtualize AWS Outputs. In fact, AWS has introduced nearly four times the number of instances since launching the Nitro System in 2017.

## **The future of AWS virtualization**

For AWS, it all reduces to focusing on the customer.

“Over 90 percent of our roadmap,” Liguori said, “comes from understanding the customer’s needs and that comes from having a lot of conversations with them.”

As long as AWS continues practicing this customer-centric approach, the future of virtualization is limitless. Besides introducing new accelerators, new platforms, and new options, AWS also aims to permeate the industry with its Nitro System, where innovations are ongoing. One recent example is the launch of EC2 instances powered by Arm-based AWS Graviton Processors.

“One of the things that brings great joy to me,” Liguori concluded, “is when I talk to customers who are using instance types that have 100 gigabytes or more ... and they don’t realize that you won’t find an on-premises network that is capable of this. ... And so that’s one of the reasons why it’s exciting to tell the story of Nitro. We’ve been able to achieve something a lot of folks don’t even realize is impossible on existing data centers. What we really aim for is to deliver an experience for customers that’s well beyond what they could achieve on their own. And it’s systems like AWS Nitro that enabled that.”