E-PAPER

Nebulon **smart**Infrastructure

a cyber-resilient cloud operating platform for on-premises application infrastructure

nebulon®

THE CYBER-RESILIENCE CHALLENGE: THE FIGHT TO PROTECT OUR DATA

Every 2 seconds. That is how often a ransomware attack on a business, consumer or device is expected to occur in 2021, costing businesses a whopping $256 Billion.[1] While many organizations will turn to a cybersecurity vendor for help, one thing is true: no single vendor can protect a business against the litany of current and future cyber threats. Even if there was a vendor up to the challenge, most cybersecurity companies focus on protection and detection at the perimeter, but not recovery within the infrastructure. When the worst happens, cyber-resiliency is left to the super-human efforts of recovery teams who must rebuild the server infrastructure, restore backup data, and recover applications all while the business is down—the average ransomware recovery taking 22 days.[2]

The reality of the situation has driven a reprioritization of IT spend unimaginable just a few years ago. One large health care company recently froze all IT investments with the exception of but data security spending until they catch up to the current threat level. At a hospital in the Southern U.S., the director of IT and an outspoken advocate for cybersecurity improvements, had just reported that a neighboring hospital had been hit by ransomware when he was promoted on the spot to CISO and directed to resolve the problem. And there have been midsize businesses who have been forced to shutter their on-premises IT operations and move to a cloud service provider because they do not have the skills needed to set up an appropriate cyber defense. These are all actions of desperation, understandable given how high the stakes are, but unsustainable over the long term.

THE NEW THREAT VECTOR

To make matters worse, there is an emerging threat vector—unpatched software and firmware deep in the physical infrastructure. IT Infrastructure is complicated and multi-layered. You can divide it into two parts – 'shallow' infrastructure and 'deep' infrastructure. The shallow infrastructure is the stuff that is easy to do, essentially from the server operating system software up to the application stack, and many solutions and vendors are available to manage that today. Deep infrastructure operations is the difficult part, especially to do it non-disruptively in a way that is friendly to the application owner. There can easily be a dozen different components within the server, each with its own firmware, version, etc.—SSDs, expanders, IOC, NICs, UEFI, LOM, GPUs, and many more. Maintaining all of these with the latest security patch level can be challenging—particularly when you have dozens, even hundreds, of servers across multiple sites and multiple server vendors.

Threat actors are actively looking for these kinds of weaknesses in the infrastructure. According to a global survey from Microsoft, more than 80 percent of the firms surveyed experience a cyber-attack as a result of neglected firmware. Further validating the point is that 70 percent of organizations that lack firmware upgrade plans will be breached due to a firmware vulnerability by the end of 2022 according to Gartner.

THE ANSWER: THE ON-PREMISES CLOUD DATA CENTER (BUT NOT THE WAY YOU THINK)

There are three things every IT organization should consider to improve the cyber-resilience of their infrastructure:

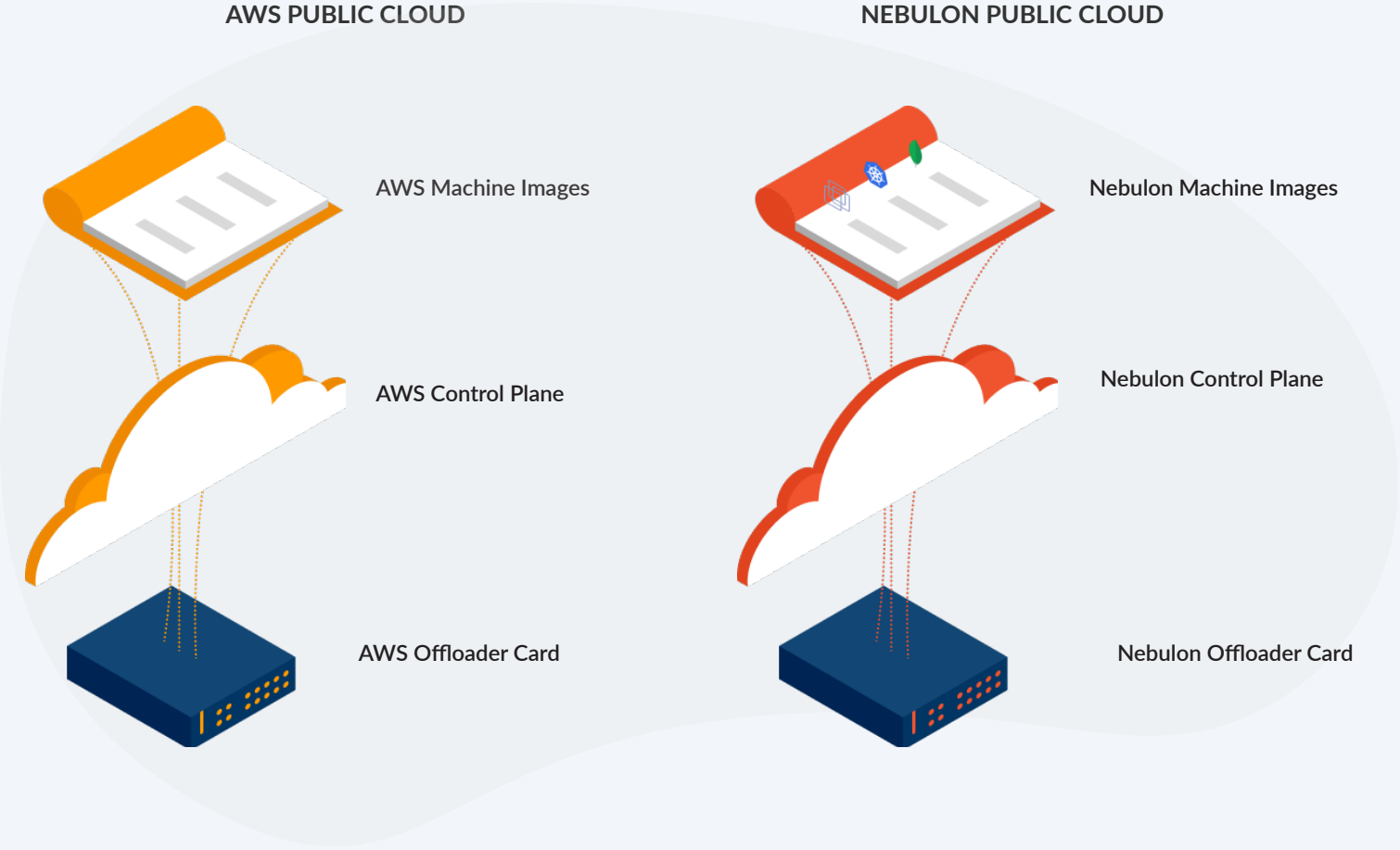
- Move the infrastructure control plane out of the devices themselves and ideally outside of the data center itself. When ransomware strikes, command and control functions remain online and available to initiate recovery.
- Secure isolation — essentially a 'panic room' — within the server infrastructure, and ideally within each server, to protect key data services and pre-attack point-in-time images. When infrastructure becomes infected, the management system (off-premises) can direct local recovery services (in a secure domain or 'panic room') to roll back to pre-attack operating system and application images.
- Reduce emerging threat vectors within deep infrastructure by keeping server software deployments consistent up front and over time (the hard part!) by routinely and automatically reverting to known, good 'immutable' server software images.

There is good news here. This might sound like a heavy lift, but in fact your existing server infrastructure can be equipped to address each of these based on what we know about how the hyperscalers have constructed their cloud operating platforms.

THE BLUEPRINT FOR CYBER-RESILIENT INFRASTRUCTURE ON-PREMISES

Consider AWS's approach for their own hyperscale infrastructure, which gives us the architectural hints we need to improve cyber-resilience on-premises:

- **AWS Management Console**, a centralized console, UI and API allowing infrastructure provisioning and management from anywhere.
- **AWS Nitro System**, a server-embedded PCIe device in each server that offloads storage and other infrastructure services
- **Amazon Machine Images (AMIs)**, which provide consistent and predictable experiences every time users instantiate a new application.



Applying each of these elements to on-premises server infrastructure can result in a more cyber-resilient infrastructure 'out-of-the-box.' Let's take a closer look.

Move the Control Plane Offsite

A ransomware attack can completely incapacitate your data center infrastructure. When it does, one of the things you need to do is get access to your infrastructure console and recovery utilities. However, if these are running in (infected) operating systems, you have a problem. Instead, if these utilities are running offsite — like an AWS Management Console, but for your on-premises infrastructure – you can begin recovery immediately.

Nebulon's equivalent of the AWS management console for on-premises infrastructure is Nebulon ON, a cloud-based software-as-a-service platform, where Nebulon has also addressed the special security requirements of using a cloud management approach for on-premises infrastructure. With this approach, rapid recovery from ransomware is assured, but there are other inherent benefits as well.

Deploy Globally in Minutes: The Nebulon ON console provides rapid self-service composability, so anyone can quickly provision and monitor infrastructure anywhere. In enterprises today, this might be possible for one specific pre-set application environment, like VMware-based virtualization, but not for any application type, including other forms of virtualization, native containerization, or bare metal.

Transparent Deep Infrastructure Management: Much like the AWS console, the Nebulon ON console provides non-disruptive fleet-wide infrastructure management that appears to the application owner to happen behind-the-scenes and transparently. The traditional enterprise experience today requires application users to plan for disruptive scheduled downtime for infrastructure maintenance.

Maintenance-Free: The Nebulon ON console is packaged in a way that is maintenance-free, with new capabilities appearing seamlessly in the UI like any cloud-based application.

An Isolated 'Panic Room' in the Server

When ransomware strikes, it can infect any network connected part of the IT infrastructure. As operating systems become infected, critical infrastructure is disabled and prevents immediate recovery from data snapshots. The resulting infrastructure restore can take hours—or even days—and lead to unnecessary downtime costs. In the case of hyperconverged infrastructure, data services run in the same shared server--storage security domain, so a ransomware attack on the server infects storage resources and protection software.

The Nebulon equivalent of the Nitro system is the Nebulon Services Processing Unit (SPU) which serves all critical enterprise data services, such as erasure coding, encryption and snapshots. With this approach, there are inherent out-of-the-box cyber-resiliency capabilities as well as density and deployment flexibility benefits.

An Isolated 'Panic Room' Inside the Server: The Nebulon SPU operates in a separate security domain from the server's CPU, memory, network, and operating system, and connects directly to the server's internal solid-state disks (SSDs). Isolating the compute domain from the storage domain prevents ransomware from infecting the data protection software that enables reliable recovery of the operating system and application data.

Workload Density: Like AWS Nitro, the Nebulon SPU also creates efficiencies by leveraging 100% of premium server CPUs for application workloads. The enterprise experience with software defined services consume valuable CPU, memory and networking cycles, unnecessarily increasing the numbers of servers and software licenses that have to be purchased to meet an application's requirements – raising costs substantially.

Infrastructure Re-Use & Future Proofing: Because the SPU operates 'below' the operating system or hypervisor and requires no Nebulon-specific software or drivers installed on the server, an SPU-enabled server has no operating system or hypervisor dependency. Nebulon-enabled servers have the flexibility to run any type of application — virtualized, containerized or bare metal — providing a future-proof platform capable of going wherever the business needs. Unfortunately, the common enterprise reality is that Hyperconverged infrastructure solutions (HCI) often only support one application-type natively, like virtualization, but not bare metal or native containers; or one type of operating system, Linux, but not Windows.

Reducing Threat Vectors in (Deep) Infrastructure

Threat actors are increasingly exploiting old versions of software and firmware to inject ransomware and malware. Although new servers may all be deployed with the same software build, missing a security patch or critical firmware update exposes the entire infrastructure, not just that server. Therefore, there is a huge need to not only deploy servers with consistent, predictable software configurations, but to maintain that consistency over time.

In the case of AWS, they achieve this through the use of Amazon Machine Images (AMIs). With AMIs, users don't make configuration changes or apply software patches to individual instances. Instead, they simply redeploy applications with a new "frozen" or "immutable" instances that have all necessary patches and changes in place. This avoids the configuration drift that can expose the infrastructure to attack or create unplanned outages.

With Nebulon Machine Images (NMIs), Nebulon provides an immutable boot utility which can freeze a known, good version of the server OS and application software stack, then revert the server OS and application configuration back to that version each time the server boots. With ImmutableBoot, recovery teams simply reboot infected servers to restore a known, good instance of their operating system, configuration settings and application binaries. The enterprise experience is substantially different, where the infrastructure itself, with admin access to every server and boundary-based security, can create multiple threat vectors for IT managers.

4 minute Ransomware Recovery: Nebulon rapid ransomware recovery solution, TimeJump, combines the cloud-based (offsite) console, the secure domain in each server, and the immutable aspects of NMIs into the first and only 4-minute ransomware recovery for combined server-storage environments. TimeJump addresses both data and the server operating system, eliminating the manual rebuild of servers and operating systems which can take hours or days. Combined with ImmutableBoot, TimeJump allows you to revert to a known, good version of your operating system in just one four-minute operation.

Consistent, Predictable Instances: Beyond the cyber-resilience benefits, NMIs provide users a consistent and predictable experience every time they provision a new or updated application instance. This is not the case in enterprise data centers where compute platforms are often treated like pets that are patched and updated regularly, drift from validated configurations and as a result, become less reliable and less secure.

ANOTHER SHORTCOMING OF HYPERCONVERGED INFRASTRUCTURE – CYBER-RESILIENCE

Why can't the existing server-based approaches via Hyperconverged infrastructure (HCI) address the needs of the enterprise in this regard? Almost a decade ago, HCI solutions were brought to market that promised to match the capabilities of the public cloud. The claim was that HCI would mimic the public cloud's web-scale characteristics, improve efficiencies and in the process, by going to a server-based infrastructure model, would re-platform the enterprise by eliminating 3-tier architectures.

Though HCI has been successful in certain pockets of IT, it has, however, fallen short on its promises. There are significant gaps in its operating model compared to the public cloud vendors, particularly in matching the non-disruptive 'behind the scenes' nature of infrastructure operations (think non-disruptive SW and FW updates.) HCI also failed to keep pace with the hyperscaler evolution from using software-defined services to offloading infrastructure services onto cards integrated into the compute platform that leveraged less expensive Arm-based technologies.

And finally, HCI falls short in addressing the growing cyber-security threats taking place at the infrastructure level. As was mentioned above, HCI runs in the same server-storage security domain, so a ransomware attack on the server easily infects storage resources and protection software. This is particularly problematic when infrastructure management tools used for recovery depend on a healthy hypervisor, operating system, and software defined storage (SDS) data services.

THE CLOUD AS AN OPERATING MODEL, NOT A DESTINATION

While the advantages of the public cloud are clear, there are critical workloads that enterprises will continue to operate on-premises because of mismatches in cost, service levels, compliance requirements, and operational transparency. However, those workloads must be protected in an entirely different way going forward.

With this in mind, Nebulon has developed a cyber-resilient cloud operating platform for on-premises infrastructure called **smart**Infrastructure, allowing CISO/CIOs to address the elevated threat of ransomware as well as achieve the same efficiencies and user experience as the public cloud for their on-premises infrastructure deployments. Nebulon cyber-resilient **smart**Infrastructure has extended the hyperscaler 'cloud operating platform' concept in a number of ways specific to the needs of the enterprise and non-hyperscaler cloud service providers, including:

- Handling the edge, not just core data centers
- Supporting the heterogeneity necessary in the enterprise; whether a broad set of server platforms, different operating systems, or different application types
- Addressing the special security needs of the enterprise
- Recognizing that enterprises often prefer solutions that are made available through their existing trusted supplier relationships. Nebulon can already make its solution available, pre-integrated, on a range of server platforms from vendors including Lenovo, Supermicro, HPE and Dell.

Industry analysts like IDC have already identified this trend towards as-a-service-based IT. They go on to say in their July 2021 'Dedicated (Local) Cloud Infrastructure as-a-Service' research that rather than being limited to public cloud services, there are a "new class of offerings designed to bring the cloud experience to enterprise premises." Not only that, IDC predicts the growth of this as-a-service approach to on-premises infrastructure to be 100x—growing to $14 billion by 2025. Nebulon is on the forefront of innovation in this area.

Supporting this trend, 60% of enterprises told another IT industry analyst, Enterprise Strategy Group, that they have already repatriated workloads on-premises from public clouds. In a related report called The Cost of Cloud, a Trillion Dollar Paradox by Andreessen-Horowitz, they have identified an average of 50% cost savings as a result of cloud repatriation, as being one of the leading rationales for workloads moving back on-premises.

WHY CYBER-RESILIENT **SMART**INFRASTRUCTURE?

ZERO-TRUST: Nebulon cyber-resilient **smart**Infrastructure immunizes customer infrastructure with end-to-end hardware-based cryptographic authentication and always-on encryption. It also makes use of an isolated security domain in the server, reducing ransomware attack surfaces and enabling deep data pattern insights for early attack detection. Additionally, Nebulon rapidly recovers both the operating system and application data in the event of a ransomware attack, so customers can bring back online their full server infrastructure (to a previous version of application data and to a known, good version of the operating system) in four-minutes.

ZERO-DRIFT: Nebulon **smart**Infrastructure provides users with templated machine images so they can instantiate their applications consistently every time. The resulting machine instances can be made immutable so that server operating system configurations remain consistent over time and downtime and undetected security breaches are avoided.

ZERO-TOUCH: Nebulon ON templates and machine images provide simple self-service provisioning for application owners, transforming a rack of unconfigured bare metal servers to a full OS and application stack in less than 10 minutes. Nebulon ON also reduces operational overheads by up to 75% with global fleet management and simple infrastructure-as-code integration.

ZERO-DEPENDENCY: Nebulon cyber-resilient smartInfrastructure allows customers to start 33% smaller, with highly available configurations of as little as two nodes — especially valuable at the edge. Customers also stay leaner with up to 25% fewer server resources required with data services offloaded to the SPU. And because no Nebulon software is required on the server it supports any OS, hypervisor or application-type (containerized, virtualized, bare metal).

[1]*CyberSecurity Ventures: Global Ransomware Damage Costs Predicted To Exceed $265 Billion By 2031 (June 2021)*

[2]*Statista: Average duration of downtime after a ransomware attack from 1st quarter 2020 to 3rd quarter 2021. (Nov 2021)*

nebulon.com