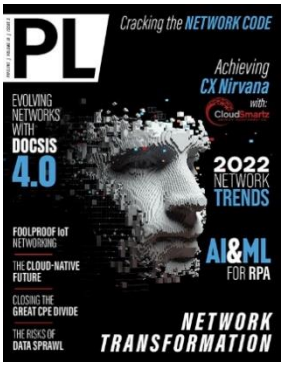

www.pipelinepub.com
Volume 18, Issue 3

Tackling the Risks of Data Sprawl

By: Kevin Johnson

We know data is king for companies worldwide, so in the age of remote work, hybrid IT environments, and the Internet of Things, how can we keep it protected and easily accessible? Most companies have a vast network of apps and devices that are constantly collecting and storing data for employees to access, analyze, and leverage for their everyday operations. This network is growing exponentially and in a more distributed fashion. According to Gartner, three-fourths of all enterprise-generated data will be generated beyond the centralized data center or cloud by 2025—up from only about 10 percent in 2018.

Between CRMs, billing platforms, and BSS/OSS systems, plus cellphones, laptops, and tablets across the workforce, networks are feeling the pressure from all angles. Enterprises are becoming increasingly good at optimizing data management costs and features via hybrid IT and multi-cloud environments, but those solutions equate to a broader network that's more difficult to manage.

The serious implications associated with mismanagement of data are growing right along with the volume and complexity of that data. From operational headaches such as proper access to the latest data, to the most critical of concerns like security, data sprawl has challenged many organizations to better manage the data that helps their businesses stay afloat.

The challenges: efficiency and security

Collecting vast amounts of data is only truly helpful to enterprise teams if they have access to easily accessible, accurate data. The truth of the matter is the faster data is collected, the harder it is to maintain the integrity of that data and keep it organized. Where this plays a major role in

production and sales is with the loss in overall efficiency. Employees are often forced to manually sift through mountains of data to find current information. This inefficiency in accessing data, to an extent, defeats the purpose of the automated systems pulling that data in.

The other and potentially more concerning consideration is the security issues that are inherent to poorly managed data sprawl. Never have enterprises and service providers had to spend so many of their resources on maintaining data security, and even with the boost in data protection-related spend, the numbers show the struggle.

A recent survey of 400 IT executives found that the marriage of rampant data growth and a lack of visibility is worsening security risks. Every one of the surveyed executives reported data storage in "informal repositories like email, collaboration portals, and local devices," which they admit are among the hardest sources to protect.

Ransomware has emerged as the top security concern for companies with more than 1,000 employees, and rightfully so. We've witnessed a meteoric rise in breaches throughout the COVID-19 pandemic, with a few of history's largest, most costly, and most brutal attacks making headlines—from Kaseya to Scripps to Colonial Pipeline to Ferrara Candy (right before Halloween). And unfortunately, almost all ransomware attacks involve a data theft component these days. The fact that we are seeing successful attacks on organizations that are highly regulated, security-focused, and considered critical infrastructure illustrates that every organization is vulnerable. Certainly, data sprawl is contributing to an already significant risk of breach.

The solutions: data tiering and microsegmentation

So, what can companies do to mitigate these challenges, knowing that data sprawl will continue? Many businesses are treating all data the same when that's far from the truth. Take, for example, a medical facility. HIPAA-protection patient records shouldn't be stored the same as years-old call notes or employee schedules. The easiest answer is to apply the gold standard across all data, but it's also the most expensive response.

Data tiering is the cost-sensitive solution that works for both data accessibility and security. You can view automated data tiering the way you might look at your own everyday items like clothing. The items that you wear all the time are likely sitting ontop of your drawer where you can quickly access them. On the other side of the spectrum, your more nostalgic favorites might not qualify to be in the drawer anymore since you're not reaching for them on a regular basis. That ugly Christmas sweater you only wear once a year is well-placed in a box in the attic.

From a networking perspective, centralized data is the starting point; then from there, companies should layer on multilayered security through endpoint protection boosted by the latest generation of firewalls that can detect anomalies and enforce policies. Security monitoring is also key, making sure to leverage encryption and set policies for who can access files and data. (Most executives are guilty of forgetting to VPN in at the coffee shop, so it's best to remove the risk altogether!) This is most strategically achieved through user access policies and firewall access so teams can containerize their endpoints and map addresses to access this data.

Network monitoring is also a key component to a secure network. Automation of device inventories is the starting point to ensure you know what devices are really on your network. Adding a HIDS (host-based intrusion detection system) will help with file integrity monitoring, rootkit and malware detection. Include a LIDS (log-based intrusion detection system) to automatically sift through log files created by network devices and servers. These systems should also include self-healing capabilities to take action when unwanted behaviors are identified.

Network microsegmentation, arguably the most granular approach to segmenting a network, is also extremely effective. With a different approach than network segmentation and application segmentation, microsegmentation focuses on a very granular division of individual servers and applications to protect them separately. Microsegmentation applies more barriers and safeguards throughout the environment, allowing for easier damage isolation in the event of an attack and a smoother recovery process. Instead of centering on north-south external traffic (attackers moving in and out of a network), microsegmentation focuses on internal east-west traffic (attackers moving within a network). This approach, which is most important for organizations with sensitive workloads that are tasked with high regulatory compliance, adds additional visibility to traffic even within the same subnet for further security. Instead of only deploying firewall rules to a particular IP or network, security policies apply to the virtual machine itself and enable intra-subnet traffic filtering. As a workload migrates, that security follows it throughout the entire application lifecycle.

When it comes to preventing breaches, the companies that are doing it right are consolidating data into the cloud and investing in edge-based security devices so they can protect remote sites. They're also using single sign-on and enforcing password policies. The companies that are doing it even better are moving to controlled devices, such as setting up remote desktops so one bout of foul play won't infect the whole network. And the companies that are on top of their games have set up network monitoring systems and ensured they're budgeting for consistent network upgrades. With technology changing so rapidly, companies need to keep investing to stay current.

We've worked with several companies who have unfortunately faced ransomware attacks during the pandemic, and a few of the common denominator risks include a lack of good endpoint protection, solid firewalls, network monitoring or strong password policies. A critical note is that no one piece of this security puzzle will keep breaches at bay; only the multilayered approach of a thoughtful data protection strategy will do the trick. And even then, cyberattacks are not entirely preventable. Ensuring recoverability of data, applications, and critical systems is the best way to hedge against a crippling breach.

32110 serves as a good guide: 3 different copies of data across 2 different media, 1 of which is offsite and 1 of which is immutable or air-gapped, with 0 errors after backup recoverability verification. By applying the 32110 principle and looping in proper data tiering and microsegmentation best practices, businesses can stay out of the headlines and smooth out their processes.