Covax Polymer

Beyond Data Security…Data Control

© 2020 All Rights Reserved

Introduction

The Environment

For as ubiquitous as the term “data security” has become, the definition only resides at the conceptual level. At a more granular level, it is unique to every organization. For larger organizations, it is layer after layer of specialized products. For smaller firms with significantly fewer resources, it may be reliance on native security solutions. Every firm’s data may be valuable for different reasons, but it is valuable, and that is what makes you a target. Differences aside, the common thread for organizations across the spectrums – small to large, industry to industry – is the need to protect the data.

As the world progresses further into this digital age, the infrastructure required to support the current paradigm is growing exponentially more complex and expansive. At best, the perimeter is blurring, in all likelihood, it is dissolving. This is even before the world was impacted by COVID-19. Never before has the enterprise network been as extended as it is now, and by all discernable indications it is never going back. The world is becoming more virtual, less tangible. Organizations need to adapt to these changes and do so by deploying solutions that are flexible enough to automatically or easily adapt to the next change. If we have learned anything in recent months, it is that change doesn’t always come slowly.

These complex infrastructure architectures are often more patchwork than framework – a fragmented network of disparate productivity, compliance, and security solutions cobbled together across on-premise and cloud environments – leaving significant gaps in function and security. Even the newest modernized solutions have significant gaps. Cloud infrastructures are under attack like never before, and errors and misconfigurations are leaving organizations vulnerable. Perhaps the largest such remaining gap is at the data itself. The term “data security” is almost always a misnomer or at least more of a targeted derivative of a focus on something else entirely. In reality, current data security is more network security – end-point monitoring, anomaly detection, malware prevention – than it is data security – creating a false sense of security. In reality, data breaches are more frequent, more expensive, and last longer than in previous years.[1]

Additionally, if the threat of data breaches is not enough to keep you up at night, an increasingly complex web of compliance regulations is being implemented. These regulations are often burdensome draconian measures that cost businesses significant amounts of money to implement and maintain. Compliance with existing standards, as well as the adaptability, to quickly meet new regulations requires a significant level of control over your data. This is important because compliance needs to be woven into the fabric of your data security solution from the beginning, allowing you proper insight and control into your data to meet ongoing and changing compliance requirements with minimal burden.

[1] IBM Security, Cost of a Data Breach Report, 2019

© 2020 All Rights Reserved

Gaps created by the siloed approach to data security

The current approach to data security is to layer disparate, specialized solutions on top of native security to monitor, and hopefully, close gaps created by today's complex systems architecture. The problem is that organizations can only go so far before destroying productivity, rendering the solution counter-productive. As gaps remain, bad actors will always find ways to exploit them.

Network Security: Necessary, but insufficient

End-point monitoring, threat and anomaly detection, anti-malware – all necessary components of a modern systems architecture. However, even when combined, these solutions fail to see the full threat environment, and therefore should not be relied upon to provide adequate protection. End-point monitoring, for example, fails to determine and stop threats from internal or trusted sources, leaving you exposed to significant vulnerability.

Current Data-Centric Solutions: Incomplete

The idea of protecting the data itself is nothing new, and there is no shortage of solutions available in the market. From encryption to audit trails to identity and access control – organizations need to weave together many individual solutions to provide the required insight and control at the data level. And even then, simply based on the current designs of available products, significant gaps continue to be present. Encryption, as a common example, may protect the data in some ways but fails to provide immutable insight into who has accessed the data, what trusted sources are bad actors, how many copies of that data are living, and renders the data less valuable because it cannot be easily searched or mined. Many of these solutions also weigh-down the system infrastructure.

Understanding the Threat…

To adequately defend against data breaches, one must fully understand the scope of the threat. When most people hear the term "hacker," their mind immediately goes to the Hollywood version of a foreign-looking individual working under the glow of computer monitors. In reality, the threat is much broader than that and much closer. Eighty-five percent of victims and subjects reside in the same country.[2] It is the call-center representative making money on the side selling the personal information of your customers. It is the engineer you have dutifully employed on your most important project taking your intellectual property to your competitor and leveraging it for a job offer. It is your overworked network administrator making a simple configuration mistake. And yes, it is the hacker in some not-so-far-away land breaking into your network, sometimes with unwitting accomplices in the form of careless employees giving away their credentials. While fewer than one-in-twenty breaches exploit vulnerabilities, sixty-seven percent were caused by credential theft, errors, and social attacks.[1]

…and the solution

When the threat comes from both the inside and outside, through the front door, as well as the back, you need a complete solution that gives you security and transparency at each of those levels. You also have to understand that, in reality, true zero trust - a common buzz-phrase touted by many security platforms - is impossible to implement. This means you need both defensive and offensive weapons in your arsenal. Data remains accessible in the capacity required for your organization to function, but if removed from your ecosystem, by anyone, it is rendered inaccessible and destructible.

[2] Verizon, 2020 Data Breach Investigations Report

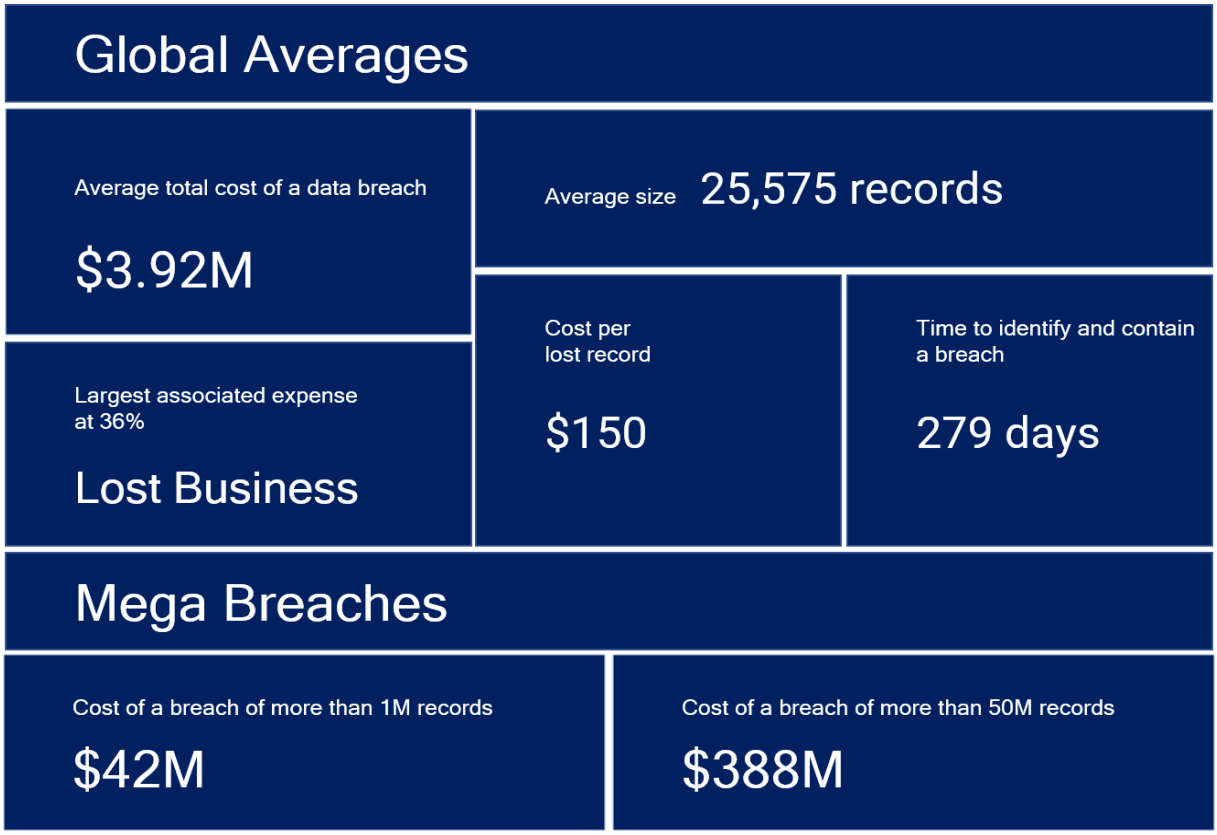
© 2020 All Rights Reserved

The Cost of Being Unprepared[3]

There are nice-to-haves and there are need-to-haves, and data security certainly falls in the need-to-have category. But do you need another layer? Decisions around systems architecture and product deployment are like any other business decision, a cost-benefit analysis within the context of which of those two criteria it falls under. Consider this, despite the current model, data breaches are becoming more frequent, more expensive, and taking longer to contain, providing clear evidence the current model is not sufficient. Also, ask yourself, how are governments going to look to replace some of the tax revenue lost due to the COVID-19 pandemic? Regulatory enforcement.

If you are an enterprise-level organization, the answer is obvious. Seventy-two percent of breaches involved large business victims.[4] However, small-to-medium size businesses were not immune, and for them, the costs were significantly higher. While large organizations faced a total cost of $204 per employee, for small businesses the cost was $3,533 per employee, showing a disproportionate burden on smaller organizations. For many of them, data security is a matter of survival.

Table 1. Consolidated Data from IBM Security 2019 Report



[3] IBM Security, Cost of a Data Breach Report, 2019
[4] Verizon, 2020 Data Breach Investigations Report

© 2020 All Rights Reserved

Covax Polymer

Introducing your new insurance policy (*disclaimer: not an actual insurance policy)*. At its core Polymer is a data-centric security and transparency platform, but manifestly it is much more. Borne from the understanding that the largest remaining gap in the network and data security model lies at the data itself, Covax created Polymer to fill that gap. Built from the data up to be a highly functional last line of defense - this is a CISO's platform – designed with you in mind. While feature-rich and results-driven, it is the platform that does the heavy lifting, leaving you with peace of mind and time to focus on other tasks.

When creating Polymer, we recognized several things that continue to steer us in our design and development:

1. The market needed a singular platform for all data types and operating environments – different operating systems, system architectures, consumption models, etc. – and it needed to secure data at both the front-end and the back-end. We needed to create a holistic platform that gives organizations control over their data.
2. The market needed a solution that was not overly burdensome on existing infrastructure, administrators, and users. This means the solution needed to do the heavy lifting without significantly increasing the footprint on both storage and compute resources, and it had to function in a way to adapt to, not change, user behavior. It is critical that a system functions the way an organization needs it to, not the other way around.
3. It needed to be easy to deploy and administer. Organizations could not rip apart their existing architecture, retrain users, and reeducate administrators to deploy Polymer. It needed to be as close to plug-n-play as it could get, and function without requiring the constant attention of an administrator, rather alerting them when and where attention was required, but otherwise, work autonomously.

To accomplish this, Polymer utilizes a distributed architecture that promotes maximum security, flexibility, and scalability, but is executed in a way to not meaningfully impact latency and availability. Covax is designed to provide security and transparency of your data across your entire ecosystem, regardless of what that looks like – on-prem, cloud, multi-cloud, hybrid cloud, remote work, etc.

The Data Molecule

One of the core components of the Polymer platform, the data molecule is an agnostic and malleable virtual encapsulation of your sensitive data. Data molecules are flexible – containing a single file to millions of pieces of data, a single byte to terabytes – and the data can be structured or unstructured; block, file, or object. Policy, along

© 2020 All Rights Reserved

with compliance and privacy attributes are set and managed at the molecule level. Data molecules leverage your existing file or storage systems without requiring data to be relocated.

Data molecules feature innovative and creative properties that allow for certain processes to occur that are not available from standard encryption products. While some of these will be discussed in greater detail later, it is important to point them out now in the context of the process that occurs during molecule creation.

1. Data is automatically indexed prior to encryption to allow for plain text search throughout the data's protected lifecycle. Searchability is determined by the policy set for the molecule and further subjected to privacy and compliance standards assigned to the molecule.
2. Data is automatically classified, and metadata is extracted.
3. The data classification and metadata are analyzed to look for inconsistencies that could be signs of malware, ransomware, or other suspicious alterations.
4. Policies are set. Policy can be standardized or bespoke; user, administrator, or AI determined. Policies can enforce who, when, where, and from what devices data can be searched, accessed, and/or edited. Policies also set the molecule lifecycle. All policies can be securely modified throughout the molecule's life.
5. Data is encrypted or decrypted on an as-needed basis as data is written to, or read from a molecule.
6. The Chain of Custody begins the moment a molecule is created.
7. The molecule is broken into smaller pieces with the name(s) obfuscated.

The molecule structure also allows for several unique and important features. By breaking the molecule into smaller pieces, brute force attacks can be significantly challenged. While the Polymer system will defeat brute force attempts through traditional and AI-driven methods when "online", these methods do not work when the target data is offline and accessible to the hackers, such as when they've successfully stolen encrypted files. However, if a molecule in the back-end, wherever that back-end may be – local, network or cloud - is comprised of tens to hundreds of randomly named pieces, and those are comingled with the same from hundreds to thousands to millions of other molecules, hackers would need to use brute force access all of the pieces. Once that was complete, they would still need to properly reassemble the pieces to access the data.

Additionally, this structure allows for incremental backups of encrypted data sets for those organizations using file-level backup solutions. While this is typically not an issue for block-level backup solutions, or for single file encryptions, when multiple files, folders, objects, etc., are encrypted within the same "container," file-level system backup solutions cannot understand where, specifically, changes have occurred. This results in potentially massive repeated backups for otherwise small changes.

When data reaches its end-of-life, either pre-determined and set into the molecule policy, or ad-hoc for security or compliance purposes, Polymer's digital shredder guarantees complete sanitation of the digital record. The digital

© 2020 All Rights Reserved

shredder uses a deletion algorithm, exponentially more thorough than the Department of Defense standards, to ensure the data is unrecoverable. Further, our Chain of Custody allows total transparency into additional copies of the data, ensuring full compliance with all deletion requirements. In addition, if copies of data reside in offline molecules, they can be automatically and forcibly expunged at end-of-life.

The Chain of Custody

Built on our patented blockchain light technology, the Covax Chain of Custody is an immutable, searchable, auditable record of virtually everything that happens to your data – create, access, modify, move, copy, delete. Our innovative approach creates a digital record – the Chain of Custody – far more robust than traditional logs or audit trails, but without the compute requirements associated with standard blockchain-style technologies. We have effectively and securely been able to centralize a process whose key attribute traditionally is decentralization.

By quite literally turning the process on its head, we are able to create a highly secure, highly scalable way of tracking every action imparted on your data. Time Stamp Authorities ("TSAs") existing as microservices across a distributed cloud-based framework[5] create a distributed architecture for securing ledger transactions. For an action to occur on a data molecule, the Covax driver at the particular data access point must successfully handshake with the TSA. This required handshake cannot occur without updating the Chain of Custody, while, conversely, an action cannot be carried out on a data molecule without the correct handshake response.

The Chain of Custody can be monitored in real-time by the Covax AI engine, administrators via their dashboard, and users through the Molecule Manager, although users can only see molecules for which they are authorized. Alerts can be pushed based on rules or anomalies as determined by artificial intelligence, with access being revocable at any time either through AI or manual intervention. Conversely, the Chain of Custody feeds valuable data into machine learning models to allow artificial intelligence to detect and respond to anomalous data activities, including such activities from trusted sources.

Auth

Auth is the core of the data security functionality of the Polymer platform. It is a decision engine that controls actions on the data, leveraging both traditional and AI/ML-powered protocols. This system is responsible for enforcing policy and rules, authenticating users, and validating their requests. While actions on Covax-protected data can be rendered from the Covax Molecule Manager, the applications consuming the data, or a CLI, must all receive real-time approval from Auth. When such a request is made, Auth brings together inputs from multiple

[5] For off-prem deployments. On-prem deployments can be set up in several variants.

 © 2020 All Rights Reserved

components to effectuate the most accurate decision possible based on the information at the time. However, should new information become available as a user is engaging the data, the session may be revoked.

Considering that over twenty percent of all data breaches involve the use of stolen credentials[6], it is critical to employ additional layers of security to address this basic process. Multifactor Authentication ("MFA") is the most common additional layer of identity verification available, and in theory, it could block 99.9% of enterprise account hacks, according to Microsoft. In reality, MFA is under attack as well, with its effectiveness constantly being challenged by social engineering and other workarounds. The FBI has issued a warning that even this form of security is no longer enough. It is important to incorporate additional factors in authenticating a user, and their request, according to security experts. Geographic location, time of day, and system analytics are all factors that can and should be utilized in rendering a decision.

Covax incorporates User and Entity Behavior Analytics ("UEBA") powered by artificial intelligence and machine learning to supplement the authorized user decision. These components will be discussed in greater detail shortly, but in terms of rendering a real-time decision on a request, they play an important role. In addition to UEBA, Auth analyzes other data points available at the time of the request - including information provided by third-party independent systems - to determine the most accurate response possible. Even the request itself is scrutinized. Eight percent of breaches were considered misuse by authorized users[7]. While small, it is a growing threat and, to the victims at least, it can be one of the more challenging types of breaches to recognize and contain. By viewing the current request within the context of the user's historical requests and current activity, a more accurate decision can be rendered.

Artificial Intelligence & Machine Learning

Covax employs artificial intelligence across the Polymer platform, allowing the system to operate in a semi-autonomous state, reducing the burden on users and administrators, while increasing the burden on bad actors. The deployment of this technology is architected in a manner to allow the system to learn from multiple sources of truth, enabling a system that learns what is right, not just what is normal - a distinction with a critical difference. As AI models learn behavioral patterns, it is important to establish those patterns virtuously, weeding-out improper activity occurring prior to the implementation of the learning platform that may otherwise be learned as normal. This model of learning allows for a system that can detect subtle signs of anomalous behavior and cause Polymer active components, such as Auth, to react accordingly.

As previously discussed, one of the primary focuses of Polymer's artificial intelligence and machine learning is UEBA anomaly detection. UEBA is used across industries – from verifying identities to creating targeted

[6] Verizon, 2020 Data Breach Investigations Report
[7] Verizon, 2020 Data Breach Investigations Report

© 2020 All Rights Reserved

advertising on websites – with the application goal of gaining as thorough an understanding of the user as possible. The technology looks to create a digital fingerprint for each user to assist in authenticating each user's requests on the system. These users can then be consolidated to an entity level to provide broad insight into the activity occurring within the ecosystem. For example, a failed login attempt may be nothing more than a typing error, but hundreds of failed login attempts in a very short time period is likely a brute force attack.

We have broken UEBA into seven initial focal points. However, it is important to note that as the system collects more data, and as that data is analyzed, modeled, and tuned, it is likely that at least some of these will change, while others will be added.

- Session token tied to user
- Day/time analysis
- Location analysis/user geography
- Data rights/Permissions
- External systems tie-ins
- Resource Risk
- Dwell Time

The second primary point of focus for AI/ML is the behavior of systems and applications consuming Covax protected data. This monitoring and analysis not only improve decision making on the action requests on data molecules, but can provide ancillary benefits of supplementing or even replacing existing disparate systems.

- Application monitoring and analysis
- Metadata monitoring analysis
- I/O operations monitoring and analysis

With every transaction, the user, access point, and the data can each be uniquely identified, creating a granular level of detail necessary to create and maintain data security, auditability, and accountability. These are key tenants in data security and compliance that are demanded by modern organizations.

The Covax Front-end

As previously discussed, there are multiple points of action for Covax protected data. Users can utilize the Covax Molecule Manager, their applications, or even the command line. The system has been designed to minimize the burden on users while protecting the data they are creating and consuming. Covax supports sign-on through the Covax front-end, as well as through single sign-on ("SSO"). Through expanding integrations, Covax will be able to work with credentialed applications directly to handle at least part of the process.

© 2020 All Rights Reserved

The Molecule Manager client is an electron app designed to be an easy-to-use point of action for users on the Covax system. Within the application, users have the ability to perform all authorized actions on data molecules – such as creating and accessing. Users also have access to certain analytics and data sets pertaining to their particular usage and permission level. Designed in the same vein as Windows application GUI's, the Molecule Manager requires little training or thought to navigate. Accessing molecules is done in the same manner as accessing files or folders, while Covax provides a wizard-style set of GUI's to facilitate the molecule creation process.

For administrators, Covax was designed to be a semi-autonomous system that does not require constant care and attention. The system runs on its own, pushing alerts, errors, and lockouts as required, and learns from the response. For informational and forensic purposes, administrators have the Covax Dashboard, a web-based application providing substantial amounts of information packaged in an admin-friendly searchable format. Data from the dashboard can also be shared with a Security Information and Event Management ("SIEM") platform, or exported for audit purposes.

Configuration & Deployment

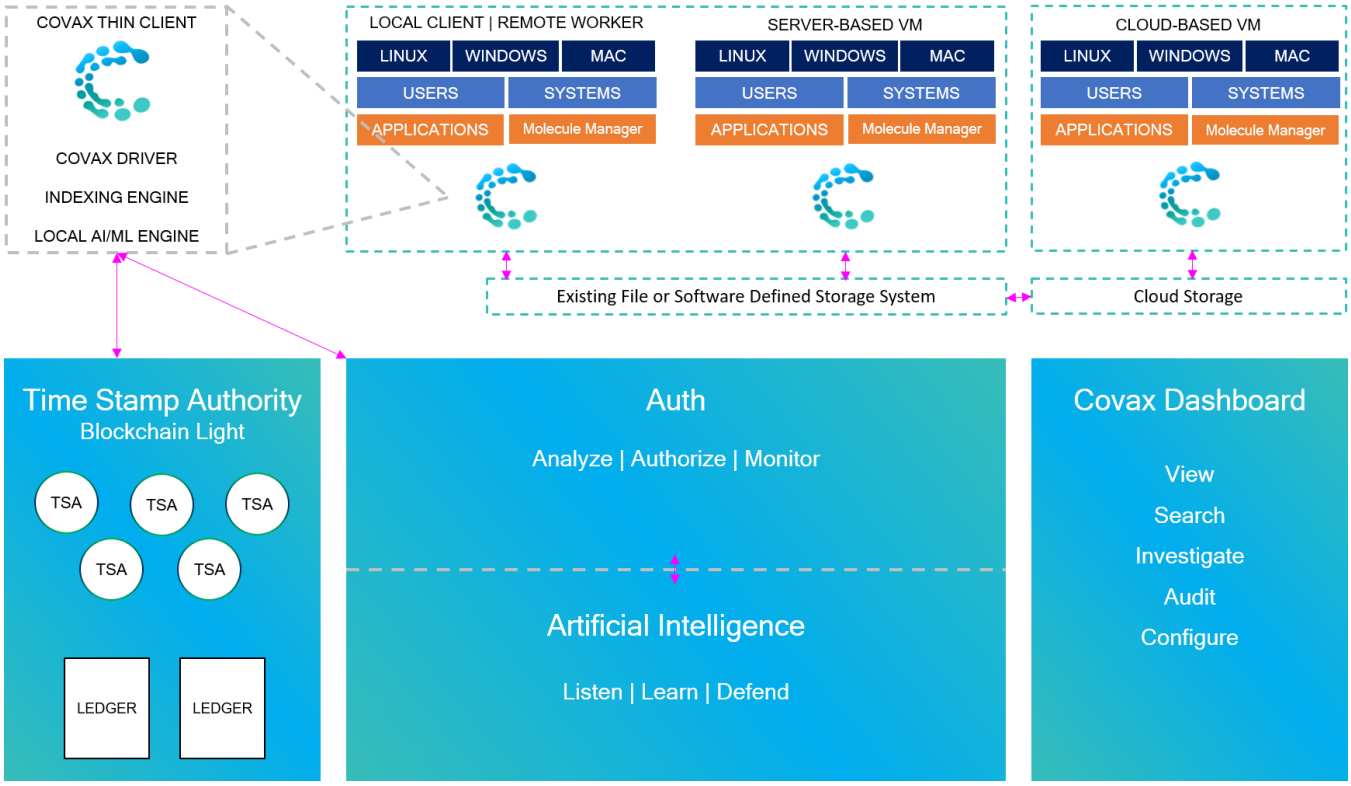
Polymer is a data security and transparency platform, not a storage system. It is designed and deployed as a Software-as-a-Service ("SaaS") platform but is available in on-premise and private cloud configurations as well. The platform is designed to be functionable and scalable for any size organization. We believe strongly that any company – from the sole proprietor to the largest enterprises and agencies – should have the resources to secure their data. This belief is reflected in both our design and our pricing.

Polymer can be configured for any type of operating environment – Windows, Linux, Mac, and combinations thereof – and any type of architecture – on-prem, private cloud, public cloud, hybrid cloud, and remote work. The platform is quick and easy to deploy, with no need to rip apart existing infrastructure, Polymer resides as a thin client at the data access points, connected to robust and highly available secondary systems. Those data access points can reside on local clients (on-prem and remote), file servers, or storage subsystems, and both physical and virtual machines (local and cloud-based) are supported. The thin client is a specialized driver that behaves as the enforcer of policies and decisions. The driver resides on top of your file or storage systems. Data molecules can only be opened through the Covax driver, which means that should data molecules ever be removed from your ecosystem, they would be rendered useless without a specifically provisioned and commissioned driver.

Polymer's flexible, agnostic design allows data to be continually secured and monitored across platforms. Current models often rely on native security and transparency solutions on each separate platform, causing a bifurcation of protection and record, as well as facilitating the need for yet another specialized solution to protect the data as it is moved between platforms. With Polymer, data can be moved back-and-forth between assets without losing any

© 2020 All Rights Reserved

protection or transparency. Polymer also allows for centralized search of encrypted data across platforms and locations, ending the inconvenience of fragmented search. This is particularly powerful when you consider that data remains encrypted in-flight and at-rest. This is extremely relevant in today's world of remote work.



Conclusion – A better form of data security

There exists no shortage of products marketing themselves under the "data security" buzz-phrase. It is a rapidly growing market, but you cannot help but question if this is what "blockchain" was three years ago, or "the cloud" was ten years ago. The reality is the market is providing a fragmented toolbox that fails to understand the big picture, from the threat to how your system is built and used. Encryption tools that cannot enforce policy, policy tools that utilize encryption as a means of enforcement but fail to see the full threat, audit trails that are simple for hackers to alter – these are not data security tools built for you. Polymer was engineered to be different, designed to truly provide a holistic view of your data to ensure you can secure data and remain compliant with ever-changing threat and regulatory environments. Finally, a system built for the way you do business, Polymer will allow you to throw away some of those disparate tools that are clogging-up your infrastructure, weighing-down your budget, and not fully protecting your data. Whether you face the daily threat of foreign espionage, or you

© 2020 All Rights Reserved

want to ensure your small family practice remains HIPAA compliant, Polymer provides the required strength and flexibility. Contact us today and see for yourself what a better-engineered platform can do for you.

Covax Data, Inc.

Covax is a technology company focused on data security and transparency. Driven by the belief that the current model clearly is not working, Covax takes an innovative approach to this complex problem. Organizations should not have to work around the solution, the solution should work around (and for) the organization. Polymer was built from the ground-up by an experienced and thoughtful team that first analyzed the problem, then went about solving the problem as a whole, not just part of the problem.

www.covaxdata.com | info@covaxdata.com

© 2020 All Rights Reserved