WHITE PAPER

Achieving True Zero Trust with Data Consumption Governance

How extending observability, detection and response to the SQL layer mitigates data security and privacy risks

The Definition and Traditional Tools of Zero Trust

The term "zero trust" as a security concept was first coined in 2010 by John Kindervag, who at the time was a principal analyst at Forrester Research. Traditional approaches to security had focused on defending the perimeter of the network; once a user or application was validated at the perimeter, they had more or less unfettered access to anything within the network. Also, long-running connections and default access for both applications and individual users contributed to a security stance that was much too trusting.

Kindervag introduced an approach that was based on a stance of "never trust, always verify." This approach focused on ensuring that every time a user or application wanted to access something, their identity was verified, along with their permission to access that specific resource, even inside of the network itself.

This philosophy, which became known as Zero Trust, caught on, and most cybersecurity technology providers have raced to associate themselves with it. The tools implementing Zero Trust have traditionally fallen into two categories: Identity Authentication and Access Control.

Identity Authentication is all about ensuring that a person or application is who they say they are. This can involve various authentication approaches, including popular multi-factor solutions. Access Control features a more diverse tool set, and can take the form of network micro-segmentation, access management at the application level, and access management inside of applications or even inside the data layer.

By putting up tougher barriers to verifying Identity and closing up access to any resource which that identity does not rightly have access to, it was thought that the issues of data theft and privacy compromise would be significantly addressed.

Cracks Develop in Traditional Zero Trust

Unfortunately, initial approaches to Zero Trust ran into problems in terms of practicality and ability to scale. The availability of much more sophisticated single-sign-on solutions and identity access management (IAM) services has indeed taken steps toward simplifying identity for enterprises. However, the limitations of legacy systems and applications with identity preferences have kept these identity protections from becoming truly simple for most organizations. Multiple identity systems often still exist, all of which must be maintained and synchronized with each other.

Additionally, identity itself remains somewhat porous due to human error, given that users often share credentials, use weak passwords, or leave their passwords or devices unprotected. Even multi-factor approaches often fail to ensure that someone is who they say they are.

Access management provides an even tougher problem in terms of practicality. To truly put up gates and moats around every resource on your network leads to a proliferation of technologies — some in software, some based in infrastructure — that create a growing web of complexity in both configuration and maintenance. Even setting it all up for a moment in time is challenging, and it often becomes completely unwieldy once the constant change of new users, new applications, new data, and new resources comes into play. And of course all of this access control tied to an identity is for naught if that identity itself becomes compromised.

Generally, organizations end up retreating from true Zero Trust to something they can actually maintain, and this often ends up being only slightly more effective than the perimeter approach they had taken historically. When a bad actor is trying all of the doors, closing all of the doors except one can be the same as leaving them all open.

Learning from How Financial Companies Protect Cash

While the Zero Trust concept represents a positive change in terms of high-level strategy, the problem has been that the security industry hasn't fully implemented the new paradigm in practical terms within the tool set itself. For an industry that has been based on authentication and access control at the network level, initial forays into Zero Trust have only meant doing those same things below the network level. That doesn't go far enough. What is important now is that security leaders consider novel approaches to Zero Trust that extends the concept into new areas.

A useful place to look for novel approaches has always been the financial services industry. Strategies for protecting the flow of cash have a head start of many decades in terms of strategies for protecting data, and they have been very successful. These systems reliably protect the accounts of billions of people and businesses every day.

The reason for their success can be tied directly to Zero Trust–type approaches — even before Kindervag articulated the concept for cybersecurity. Access to cash is heavily based on identity authentication, and access limits are placed around the cash that can be accessed. Good examples are withdrawal limits at an ATM or the credit limit on a credit card. However, with identity theft becoming rampant in the digital age, financial services companies have had to go further into strategies based on consumption patterns.

Controlling consumption is different than controlling access, because in many cases people need legitimate access to a lot of cash or a large line of credit. Consumption controls focus on how cash has been consumed historically, working from that basis to determine what type of consumption should count as anomalous and thus should be subject to stiffer controls.

An example that most people can relate to is the use of a personal credit card. Because these cards — either the physical piece of plastic or the credential information tied to it — get stolen fairly frequently, the issuing banks have become skilled at detecting and responding to unusual consumption. For instance, if most of your credit card charges are within a 30-mile radius of your home and your average transaction size is $50, a $4,000 transaction thousands of miles from your home will be stopped by your issuer, who will typically try various means to contact you and make sure the charge is legitimate

before permitting it. They would do much the same thing if a flurry of unusual purchases closer to your average transaction size were made all in a row.

This is Zero Trust taken down to the individual transactional level. It's powered by sophisticated algorithms that have established a baseline for your consumption of credit to help your bank know in real time when something is happening outside of that baseline. Couldn't we treat data the same way?

Data Consumption Governance

This is the foundational idea behind Data Consumption Governance, which is a powerful new approach to data security that treats every request for data in a Zero Trust model. It's an approach that relies on accurate fingerprinting of an identity to establish a baseline for how a user or application consumes data, but also works independently of identity by not trusting that an authenticated user is who they say they are. By taking this approach, Data Consumption Governance creates a powerful layer of security that succeeds even when identity-based access becomes compromised.

Technologically, implementing Data Consumption Governance is challenging. The tools to govern consumption must have a presence immediately around the data source, with the ability to both observe normal consumption of data and to act immediately to limit data consumption when it becomes abnormal. All of this must be done in a way that doesn't impair the normal flow of data, as timely access to business data can be a make-or-break proposition for data-driven organizations.

Another challenge is that the structure of data stores and the SQL syntax used to access data within them can vary widely, so the technology used must absorb this complexity without creating an added burden on the security team.

Beginning with Observability

It's true that most organizations don't have a very good view into the actual transactional traffic across their data stores. Much of the awareness is in the layers above that, including the identity of employees and which employees have access to which applications. The first step in governing consumption effectively is correcting that by creating granular observability into data flow. Which types of sensitive data flow the most? To which applications and to which roles using those applications?

Observing data consumption allows an organization to outline the risk to its data. Cross-referencing risky data with risky roles in the organization can help identify hot spots where consumption governance should likely be prioritized.

Once a confluence between high-risk data and a high-risk role is identified, the next step is to understand which signals to use to establish reliable patterns of baseline consumption. Is data almost always consumed during working hours? From a specific range of IPs? At a certain rate of flow or in a highly predictable quantity?

Consumption can also be predictable in terms of the way data requests look as they come in. People can often be unpredictable, which is a specific challenge in End User Behavioral Analytics; applications, by contrast, are highly predictable. They present the same SQL statements again and again as those requests are generated by a reasonably static code base. In that context, an abnormally formatted data request from an application can be cause for alarm, and a sign of a potential SQL injection attack.

Using observability to create parameters for predictable consumption becomes the starting point for understanding

where to put up guardrails for consumption, what anomalies look like, and what the corresponding responses should be.

Controlling the Avenues of Consumption

A Data Consumption Governance model, by virtue of it being very closely tied to the data source, is naturally very difficult to flank. However, like any protective measure, it can be vulnerable to privileged network access. For instance, DBA credentials might allow a user to go directly to the database server and access data without being subject to governance.

A way to prevent this is to build at-rest obfuscation of certain sensitive data values into the governance model itself. What this means is that some data is actually disguised, usually replaced with a token inside of the database itself. The same system that provides Data Consumption Governance is the only system that can interpret the token and convert it back into its clear value. Through this mechanism, any user or application that wants to consume particularly sensitive data absolutely must go through a governed pathway to get it; because of that, their consumption is evaluated every time without fail.

Zero Trust means not trusting any credentialed access, and this is a good way to extend that distrusting stance to the most privileged access within an organization.

Data Consumption Governance and True Zero Trust

We work in a world where identities and the access levels associated with them are compromised regularly. We need to augment identity systems alongside other approaches that assume that someone isn't who they say they are. By observing how users and applications regularly consume data and limiting or stopping any abnormal consumption of data in real time, we get very close to a true Zero Trust posture. Going further and creating a governance model so that data requests must flow through it ensures an even tighter mechanism of control.

When we provide fundamental protection and Zero Trust at the SQL layer, everything above it becomes significantly less risky. New employees, new software, new partners, new data platform projects, new approaches to infrastructure — with all of these, when you have a handle on data consumption, your organization's ability to unleash value through data-driven approaches is assured.