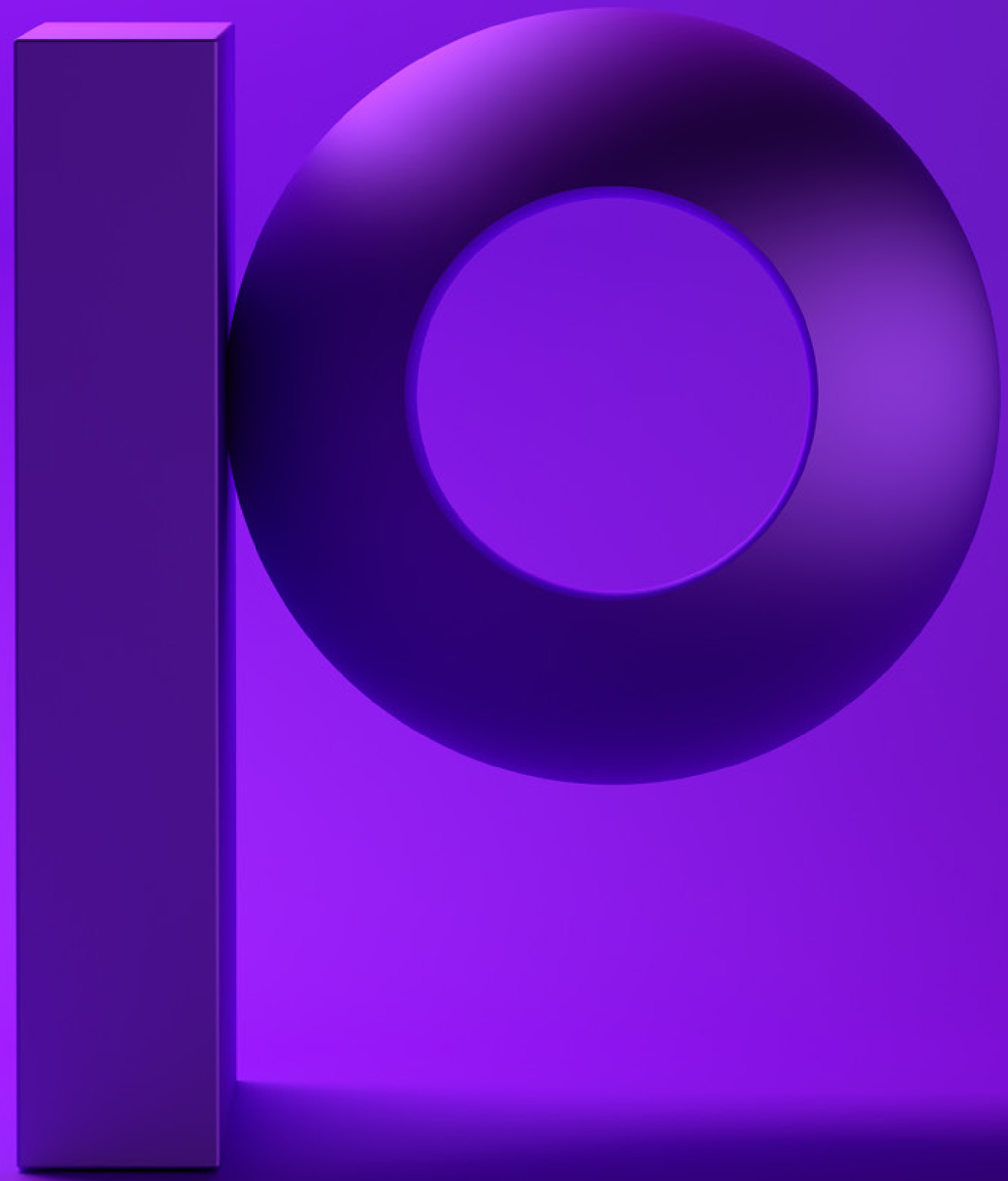
precisely

Five Tips for Secure Data Integration

Key considerations for
scaling up and out securely



Integration and the security imperative

Data and systems integration have become a core responsibility for IT managers in all but the smallest of organizations. It is just not possible for any business to operate without integrating into all sorts of external systems, including hosted IT services, cloud platforms for distributed computing, online retail marketplaces, web advertising services, credit card transaction systems, and many others. No business is untouched by the proliferation of mobile apps and remote working arrangements.

Planning for and achieving integration is inherently complex. The list of technical issues involved can seem endless. Just to get on-premise systems and cloud computing platforms connected and playing nicely together involves overcoming data type and format incompatibilities, bridging OS and application differences, managing data sharing, and much more.

As if all those factors were not enough to deal with, there is one more imperative that is overlaid upon the entire endeavor: data privacy and security. As you integrate, you must maintain any and all existing safeguards while also ensuring that data privacy and security measures are implemented in every system and at every point of interconnection.

Essentially, the moment you decide to integrate, you commit to opening up a classic Pandora's Box of detailed security requirements and issues that are, on their own, enough to make any security professional's head spin.

Planning for secure integration

The very first thing you need to do is to engage with your company's security professionals and enlist their help in identifying and solving the security challenges your integration project will inevitably encounter. There is no escaping the reality that integrating systems means creating new potential security vulnerabilities as you add many new connections between systems and applications. Increasingly stringent regulatory requirements and the relentless innovation in the methods and vectors by which bad actors attack make it unwise to proceed without expert guidance.

This becomes especially important if your integration objectives include older legacy environments, such as mainframe or IBM i platforms, or any systems handling highly regulated data. Some legacy platform operating systems, data storage and system security methods were originally architected in the days before cross-platform or cloud integration even existed. Even in their most current, updated state, such systems still carry an integration-resistant core DNA throughout their OS and security structures. So, specialist knowledge and skills are required.

1: Integrating legacy system data

The primary driver of your integration efforts is your organization's need for systems and people to have access to all of its data, whenever and wherever it is needed. And that needs to be the most current data, updated in as close to real-time as possible. Right from the start, this implies the need for real-time change data capture (CDC) and delivery from core business systems to a unified storage platform, such as a data lake or data warehouse.

Providing access to all data requires getting data out of legacy systems and storage as well, which is not usually a straightforward process. While modern applications are designed with flexible data-interchange in mind, many older applications and legacy platforms were not.

One of the hardest challenges is integrating mainframe systems. Traditionally, getting data out of mainframe systems relied upon batch ETL (extract, transform and load) processes, carried out on a periodic basis using the mainframe's own internal query processes, with the data being delivered as output files. This relatively "non-invasive" data transfer method naturally avoided many system and data security concerns.

However, providing data updates to distributed processing platforms via ETL is almost never sufficient. Between the back-level, snapshot nature of ETL output, the cumbersome processes for ingest into data lakes, and the difficulty of ensuring data currency, ETL-based data delivery is too slow, labor-intensive and just hobbles and degrades the value of your big data platform investment.

Change data capture (CDC) for mainframe

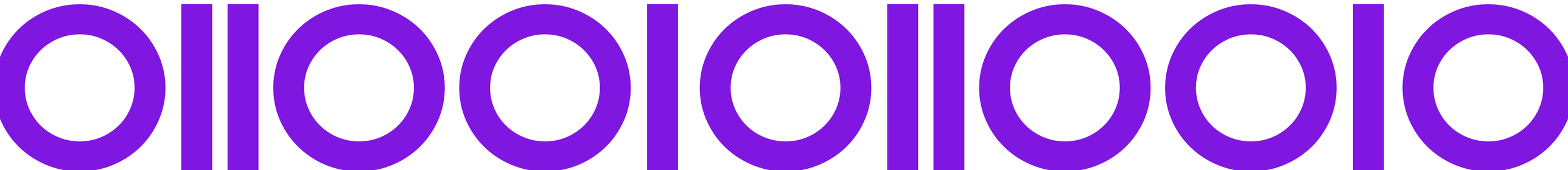
So, in all cases, a CDC approach is required. However, applying CDC to mainframe systems is a daunting task. Mainframe data may be encoded in EBCDIC or packed decimal, and stored in VSAM, mainframe fixed and variable, and other formats. Data is also spread across specialized storage system disks (DASD), and retrieval requires native mainframe OS processes such as channel programs and the use of metadata descriptor files called “Copy Books” which are stored separately from the data itself.

Suffice to say, accessing mainframe data changes in real time is not easy or straightforward. But it can be achieved using CDC software specifically designed for mainframes.

One additional serious concern regarding implementing CDC software for mainframe: Can the CDC software itself be trusted? Because they tend to manage the most business-critical applications and the most sensitive data, implementing CDC software directly on a mainframe server could potentially create a hidden but serious security vulnerability.

When evaluating change CDC software for mainframe data integration, it is critical to keep security considerations in mind and to consider the following:

- How the CDC software handles mainframe security protocols and interfaces such as the System Authorization Facility (SAF) and Resource Access Control Facility (RACF).
- Whether the CDC software must be installed directly on the mainframe or if it can run on other servers.
- The ability of the CDC software to transform mainframe data without altering or losing the original data.
- The ability of the CDC software to integrate with other “Guest” operating systems including Linux, UNIX and Windows which can run on modern mainframe servers.
- The level of specialized mainframe coding and systems management skills which may be required for CDC workflow development and maintenance.



2: Protecting sensitive data

Assuring data security and privacy is a basic requirement for any IT environment. Aside from the fact that many regulations demand it, implementing robust security and privacy practices is crucially important to any organization that hopes to gain the highest possible economic value from its data.

Protecting data involves much more than simply "locking it down." Security analysts and consultants advise thinking more broadly about your approach to security, and recommend following principles such as Zero Trust and CARTA. Zero Trust simply states that no entity, whether inside or outside your organization, should be trusted. That means constant, proactive validation of "Who are you and why are you here?" at every point of presentation across the system. CARTA is a more prescriptive framework that centers on "Continuous Adaptive Risk and Trust Assessment." The keyword here is "Continuous." CARTA centers on constant, automated discovery and analysis of potential security vulnerabilities across your entire enterprise.

In the end, the principles and practices of Zero Trust, CARTA and other security concepts can be viewed as subsets of data governance. Data governance focuses on end-to-end management of data quality, usability, availability and security across all systems and processes. It means wrapping proper controls and safeguards around every part of the data lifecycle. Integrating systems and data touches on every one of these areas, so a data governance approach to integration is recommended.

Methods for securing sensitive data

When integrating data and systems, you must identify all sources of sensitive data that may be accessed or moved through each process. You must also identify and evaluate all the possible users or systems that will access the data, so that you can determine the appropriate methods for protection.

The goal and the challenge here are to implement controls and safeguards which protect sensitive data while still allowing systems and users to do their job, that is, protection that still maintains usability. Methods to achieve this start with de-identifying the data using data masking or anonymization techniques, and replacing the data outright or rendering the data unreadable using tokenization and encryption.

Many systems, applications and storage platforms have configurable options for de-identifying data built in. But not all. So, it is important to evaluate the details of what methods and protocols are available, where in the workflow process they can and should be applied, and the compatibility of the methods between the system elements being connected.

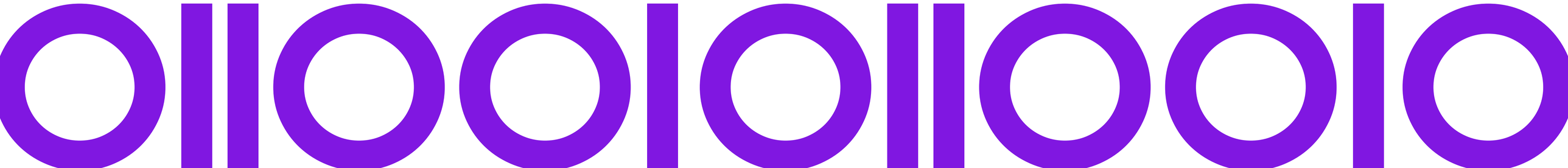
3: Ensuring data lineage protection

As you design and execute your integration plans, it is critical to ensure the integrity of your data lineage, the metadata record of any and all changes to data moving through your systems. This chronological metadata becomes critical for recovery if data is somehow corrupted, lost, duplicated or otherwise rendered incorrect or unusable. It is also critical for regulatory compliance, including annual financial auditing, data privacy controls, and myriad governmental laws and regulations.

From a security perspective, securely maintaining accurate and detailed data lineage is also critical for SIEM systems. SIEM systems rely upon data lineage as part of their automated security monitoring and assessment processes, looking for patterns like unusual frequency or depth of data access by a particular user or system, which could be a sign of malware activity.

So, from a data and systems integration planning standpoint, you must ensure that data lineage details are created at each step, but more to the point, that they are correctly and securely passed between systems, are stored/retained securely, and are always available on demand. Most of all, data lineage must be an unbroken sequence of records, fully traceable back across all systems and events to the moment the data was first created.

As you build, first ensure that each and every application and system in your integrated environment either provides data lineage functionality directly or can integrate fully with systems that can cover it, such as a third-party application or SIEM system. Also evaluate the data lineage methods and formats employed in each process and validate their ability to forward or transfer data lineage records.



4: Integrating cloud data

There is little doubt that the moment you decide to add public cloud infrastructure to your IT, the number and complexity of security issues escalate rapidly. Many cloud computing security challenges stem from a single, unavoidable truth: Cloud Service Providers (CSPs) are not in the business of ensuring that the data and systems which you put in their cloud are completely secure. In fact, they cannot do so. Here's why:

The CSP has access to your data

Regardless of the security schema your CSP applies to your "instance" in their cloud, they must always have access to your data, in one way or another. Further, if you choose to place them in charge of data protection, they must also possess the passwords, encryption keys, token repositories and whatever else is needed to manage it. That means someone in the CSP organization has the ability to view your data in the clear.

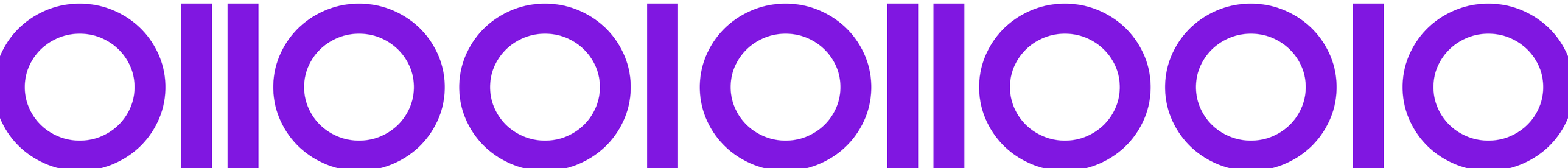
The CSP has control of your data

In order for the CSP to be able to provide compute elasticity and fault tolerance, you have to permit them to copy or move data without notice, to and from any of their systems. Among other concerns, this can lead to regulatory compliance and security issues regarding data residency.

You can't audit or verify CSP security

Because your CSP cannot possibly know where your data is actually stored or applications are physically running, they cannot and will not provide access to their physical infrastructure for security audits. This immutable fact immediately preempts any Zero Trust policies.

Given these kinds of cloud computing realities, how can you even begin to develop a secure data and system integration architecture that includes cloud?



Cloud integration requires modern security methods

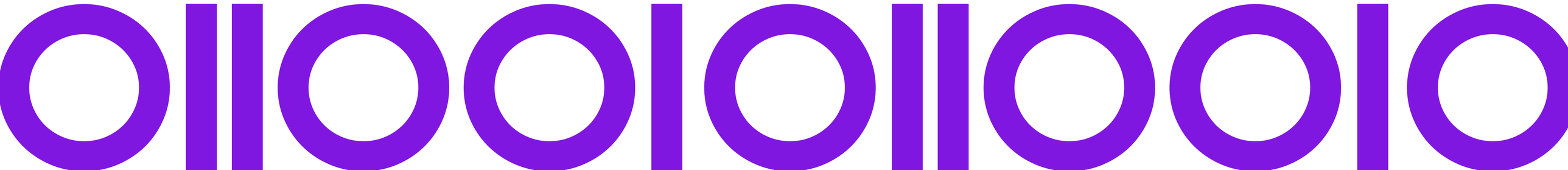
First and foremost, while you cannot avoid allowing your CSP to have basic data access rights (copy, transmit, delete, etc.), you should do everything in your power to limit their freedom. If you don't want them to have access your sensitive data in the clear, do not give it to them. If you do not want them to possess your encryption keys, do not provide them.

What those bold statements really imply is that you take a Bring Your Own Security (BYOS) stance. What this means is that you need to consider investing in technologies that enable you to directly and fully control data security. Such solutions focus on applying security directly to data before it ever moves to the Cloud, using data masking, encryption and/or tokenization, depending on the security level required.

Such systems manage data security control processes via a "Gateway" architecture, which routes requested cloud data through a data security server or appliance within your on-premise infrastructure, before sending it back to the cloud for delivery to the requesting user.

As you begin identifying and evaluating options for cloud data security solutions, you will of course find a wide array of vendors presenting many different solutions. Some vendors will have solution options which are designed for use with specific CSPs. They will also probably be part of an ecosystem of technology partners, which may include vendors you are already working with.

It will take time and a lot of work to settle on specific solutions. But given the broad impact your cloud security technology will have on so many other integration security choices, you should address this part of your project early on, preferably as the first phase.



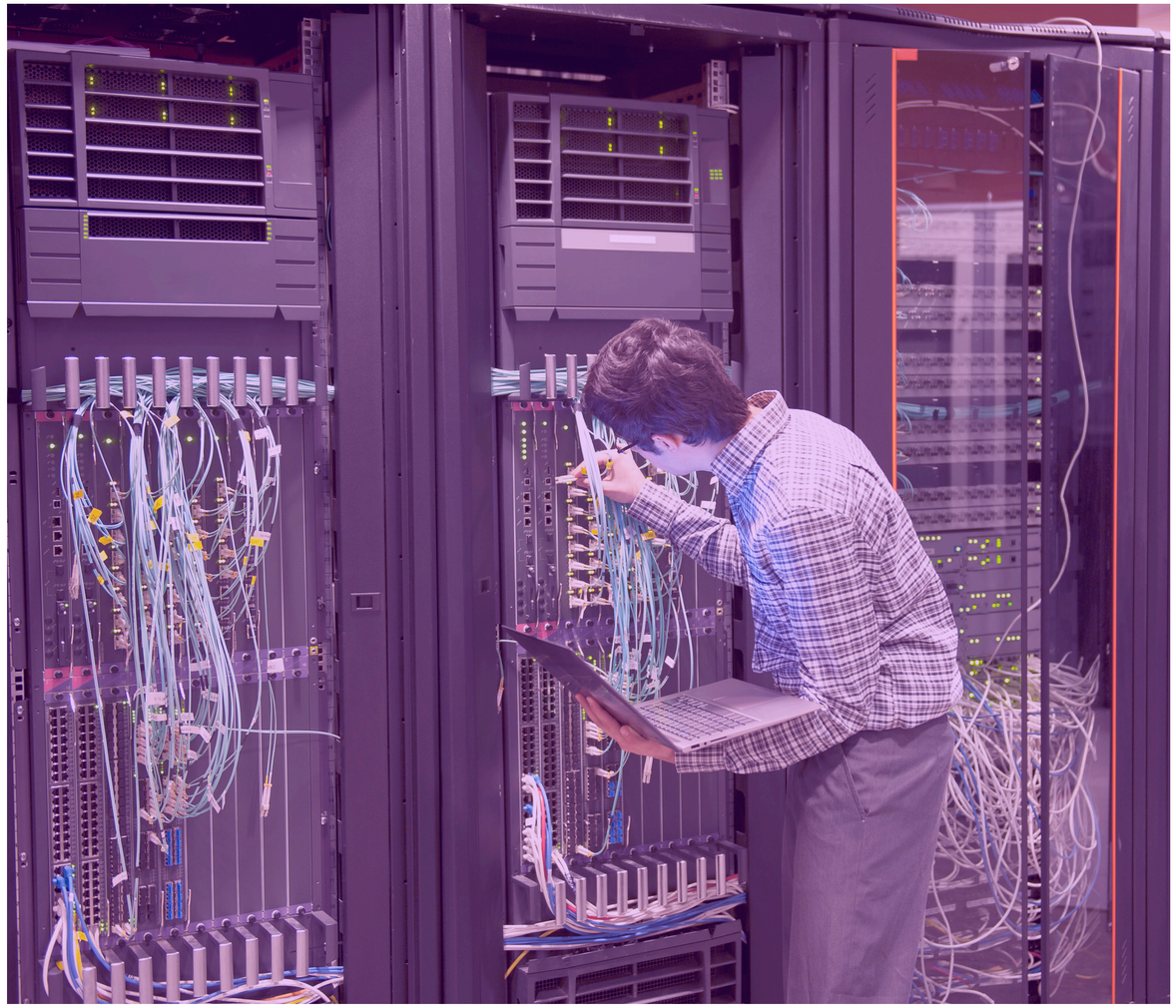
5. Integrating security systems

Data integration can't be achieved without integrating security. Perhaps an obvious observation, but certainly one that leads to a long and diverse list of follow-on questions. And at the top of that list sits "What does that even mean these days?"

Traditionally, the primary focus has been on intrusion protection. But over time, the many "threats from within" have also been prioritized, with emphasis on both integrating and automating internal security.

Password protection and multi-factor authentication are only a small part of the equation now. Beginning with leveraging Active Directory or centralized security appliances, things have moved on to SIEM solutions, CSP-native services and protocols, third-party Security-as-a-Service solutions and a host of other technologies.

From the perspective of data and systems integration, the most complex challenges involve securing cloud and hybrid environments. Even more so if you are tying tightly into key client or partner systems, whether that means reporting portals, just-in-time supplier coordination, or any other deep business system integration. Such integrated business operations require coordinated, integrated security.



To help you chart a pathway toward truly secure cloud integration, here are three ways to get started:

Refresh and extend your security awareness

Make the effort to educate yourself regarding security methods and tools currently being used in cloud and hybrid environments. Your goal is not to become an expert, but to become aware of the kinds of issues that are top-of-mind these days, and to brush up on industry terms and acronyms so that you can be an active part of the conversation.

Understand your current state

Review with your security professionals what systems and tools are already in place. What problems are they currently trying to address, as far as the present state of your systems? Where do they see your organization's security priorities moving next?

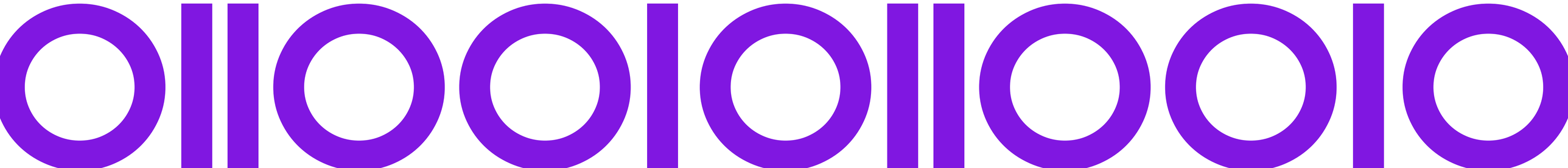
Even if you think you know, do it anyway. Probe and discuss a bit deeper than you ever have before. The benefits of this effort are two-fold: You are equipped with better skills for evaluating your business solution options and you have built a stronger bridge of trust with your security team upon whom your success will definitely depend.

Practice integrated planning

No matter what, don't treat security integration capabilities as an afterthought. It is impossible to integrate systems without generating a technology and vendor ecosystem. And you will be living with your choices for a long time. So, you need to ask the same questions regarding Security integration as you do for operational functionality integration. That means more than just capturing the keywords and protocol acronyms within their offerings.

For every vendor or solution, be sure to ask:

- Who are their ecosystem partners?
- Are they certified with or through their partners?
- How actively do they cooperate on product development?
- What is next on their development horizon?



Integrate securely with Precisely

Precisely's data integration offering, Connect, takes a one solution approach to integrate, prepare, load, cleanse, transform and stream data across your organization's environments. Providing end-to-end support and protection for data lineage, Connect requires no installation of software on the data source systems, so you can safely and securely integrate data from a variety of sources and targets, including mainframe and IBM i, for use across your enterprise.

And Precisely's ongoing partnerships with some of the world's leading security technology developers means Precisely products and services support and enable strong, secure integration across even the most complex IT environments. Highlights of these security technology vendors and Connect include:

- Integration with Protegrity™ security solutions helps to protect data at rest and in-motion across even the most complex environments, including those relying upon highly secure mainframe and IBM i servers.
- Close collaboration with CyberArk™ ensures seamless integration of Privileged Access Management solutions for complete protection against cyber-attacks that leverage insider privileges.

To learn more about how Precisely's Connect can help you achieve complete and secure data and systems integration visit precisely.com

precisely

Precisely is a global leader in data integrity, providing accuracy and consistency in data for 12,000 customers in more than 100 countries, including 90 percent of the Fortune 100. Precisely enables companies to integrate, verify, locate, and enrich their data to power better business decisions. To learn more, visit www.precisely.com.

www.precisely.com