Datex

DATA RESIDENCY

A guide for data in the Cloud

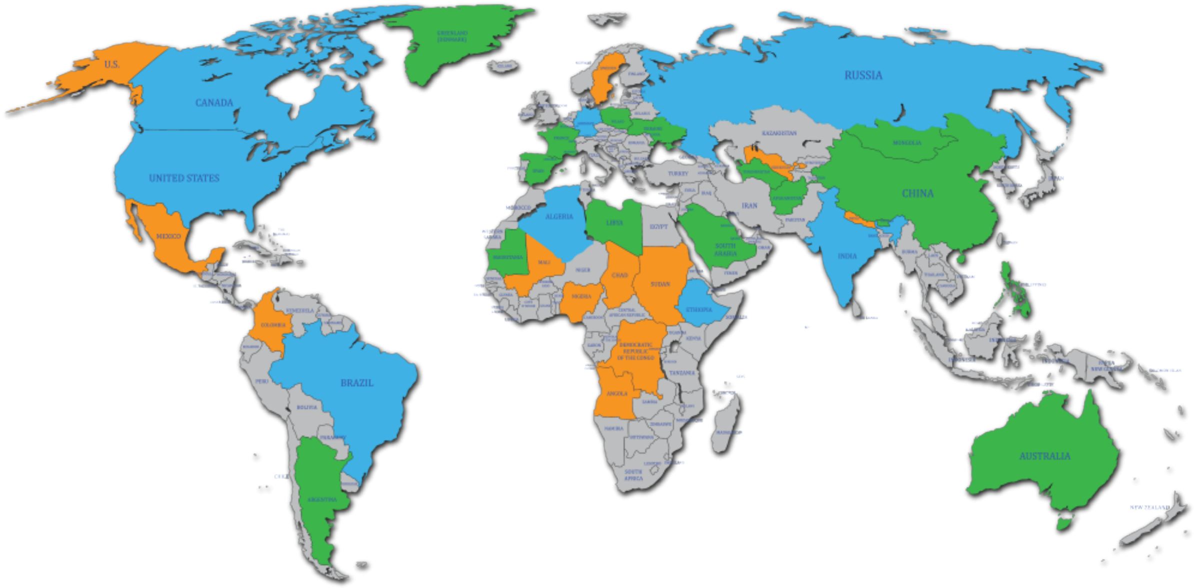
The Situation

Organizations today want to use Cloud storage and Cloud applications, so IT teams are therefore seeking safe and secure environments in which to put their most valuable assets – their data and documents. These environments must be scalable, flexible, robust and resilient, and organizations are starting to realize that Cloud solutions can bring all of these characteristics to the table. Privacy, risk and compliance teams, however, are resistant to moving into the Cloud, arguing that the issues associated with Cloud solutions typically outweigh any benefits.

The Problem

Many companies collect and process vast quantities of customer data, and much of it contains highly sensitive personal information such as dates of birth, social insurance numbers, payment card information, bank account details, online banking credentials, or credit scores. Collected data increases in sensitivity over time when details are combined and related to generate an accurate and complete picture or profile for individual clients. As a result, adhering to a variety of data security and compliance requirements, along with multi-jurisdictional privacy legislation, are of critical importance.

Most companies take privacy and data security responsibilities very seriously and work to remain in agreement and compliance with all applicable and ever changing rules of the game, but when expanding a product or service offering which was developed for one jurisdiction into another, organizations can face privacy challenges from integrating different and overlapping privacy or data protection rules and regulations.



Canada

Canadian companies need to keep in mind that even publicly available personal information is subject to a combination of Canadian privacy and residency laws such as The Privacy Act, The Personal Information Protection and Electronic Documents Act (PIPEDA), The Freedom Of Information And Protection Of Privacy Act (FOIPPA), and the recently approved Digital Privacy Act (DPA), Senate Bill S-4.

The Privacy Act is designed to respect an individual's privacy rights by limiting collection, use and disclosure of their personal information, and PIPEDA is designed to define how organizations may collect, use or disclose personal information as part of commercial activities. The DPA amends PIPEDA to include a new data breach notification requirement, and FOIPPA states that applicable entities must ensure that personal information in its custody or under its control is stored only in Canada and can only be accessed in Canada. Each statute represents a different interest and mandatory breach notification laws introduced by the DPA are enhanced by the FOIPPA which now includes an obligation for organizations to disclose if they receive or become aware of foreign demands for disclosure of personal information.

USA

Laws such as the Foreign Intelligence Surveillance Act (FISA), or the Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act (USA PATRIOT Act), are applicable to companies within the United States. Many countries have comparable laws allowing them access to Cloud-stored data outside their respective jurisdictions, and the current response to these legislative Acts is to follow a lawful course of action, storing data within each jurisdiction (or simply not moving into the Cloud at all).

European Union

Within the European Union (EU), privacy rules are even more robust. The General Data Protection Regulation (GDPR) is a regulation by which European Parliament, the European Council and the European Commission collectively designed policy to strengthen and unify data protection for individuals within the EU. The primary objective of the GDPR is to return control of personal data to citizens, and it also addresses export of personal data outside the EU, which has become increasingly restricted.

For Service Providers

Cost and overhead for managing services within these new regulatory environments can be significant and restrictive, which in turn often reduces the overall gain of using a Cloud solution. Although multi-jurisdictional organizations are doing their best to decipher and adhere to applicable, region-specific laws and regulatory requirements, maintaining compliance across a multitude of sometimes conflicting requirements is an increasingly complex and frustrating endeavor. When it comes to data residency, though, the single largest obstacle is how organizations plan to maintain control and protection of personal information, regardless of whether it resides on a third-party service that relies on a distributed infrastructure to deliver resiliency, availability and flexibility to its customers.

Solution Overview

In response to such complex data residency regulations, a generally accepted principle has emerged and states that if data has been anonymized to such extent that it is indecipherable in transit and at rest, either by way of tokenization or de-identification, said data can be precluded from most if not all residency restrictions. To this end, Datex has developed a revolutionary new approach that solves the dilemma by keeping organizations compliant using a simple plug and play solution. We call it DataStealth.

At a high level, the DataStealth solution is placed at the egress point of your infrastructure and inspects every element of your data on the way out, identifying and extracting private sensitive and/or confidential data discovered in the stream. On the return trip, at the ingress point to your infrastructure, the data passes through the DataStealth solution again, and the original data is reinserted into the stream prior to presentation to the end user, but only for authorized users, and only for authorized use cases.

What makes DataStealth so revolutionary is that DataStealth is truly plug and play, and is transparent to both end users and to Cloud providers. No app development is required. No code changes are required. No plugins (no API, desktop app, or browser plugin) need to be installed. No software or databases need to be installed and managed. And there is no key management required at all.

Technical Overview

DataStealth consists of three distinct, functional components; Processing, Vaulting and Storage. Processing components are responsible for inspecting data to identify and selectively extract elements based on configurable policies. Vaulting components generate substitute data to insert and replace extracted data, meanwhile securing original data for storage. Storage components, as their name suggests, are where the originally extracted data elements are securely stored.

All three components can be deployed in a single package at a single location, or they can be decoupled and dispersed. By decoupling components, DataStealth can provide Processing and Vaulting within a region of origin, which means that data leaving these regions is effectively anonymized, and that each region can have its own Processing and Vaulting components, along with policies designed for specific regulations or requirements of that region. Data anonymization processes can include various different protection strategies such as tokenization, encryption, de-identification, masking, and a number of other options.

As part of Vaulting processes, metadata for each anonymized data element is collected and stored for later use during re-identification. Metadata may include geo-location, IP address, user information, membership to an active directory group, or one of many other relevant details. Data vaulting strategies are typically made up of multiple complementary techniques working together to achieve a desired result. In Vaulting, extracted data goes through a multi-step crypto-process to secure each individual data element and assign a unique crypto-key before the encrypted pieces of data are then fragmented into a configurable number of pieces (‘y’). After fragmentation, a predefined and configurable minimum number of fragments (‘x’) are required to reconstitute the secured data.

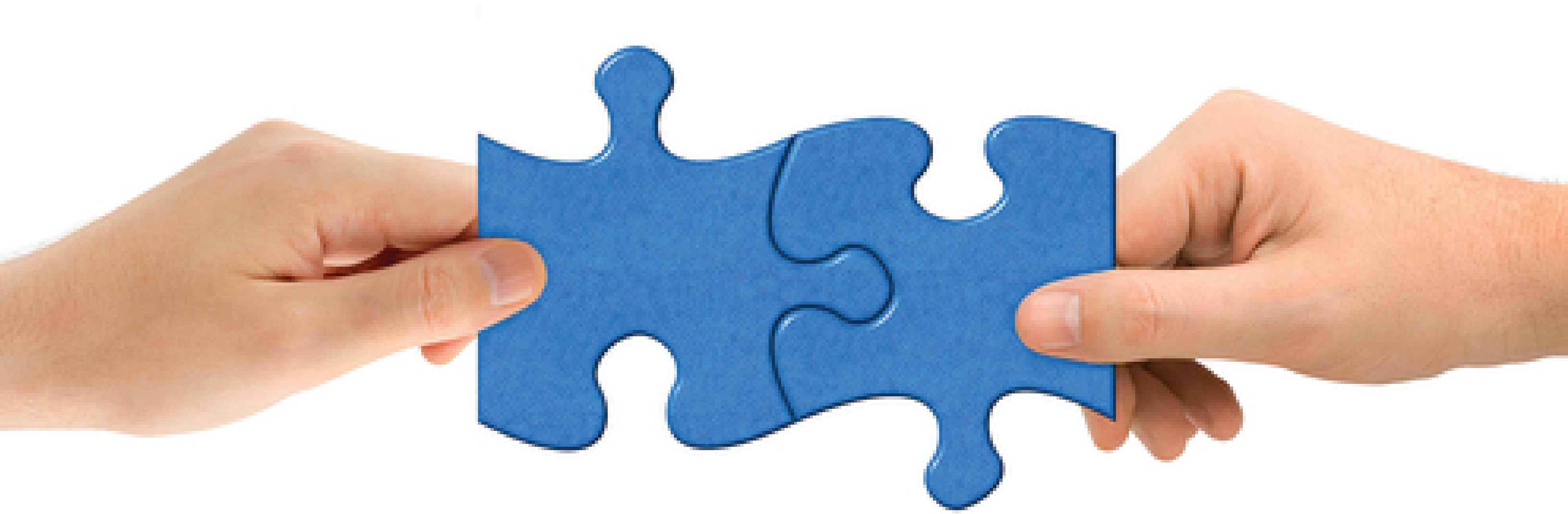
Each of the ‘y’ fragments are then dispersed to a diverse set of storage nodes, and it is important to remember that fragmented data is not considered data until it has been reconstituted. Fragments do not leak any data, and until a minimum number of fragments are obtained, there is no way to even begin reconstruction of the original, encrypted data element. Using this strategy, no single node has enough fragments to reconstitute the original data, and therefore the storage node does not technically contain any regulated data.

This type of storage strategy is inherently fault tolerant, as Storage nodes are typically part of a global storage pool. The main difference between ‘x’ and ‘y’ is redundancy, as only a minimum threshold of ‘x’ fragments are required to reconstitute data. If a client was running a 5 of 9 strategy, they would require only 5 of 9 available fragments to reconstitute the original data, and could lose up to 4 storage nodes without affecting their ability to reconstitute. In a case like this, the ‘x of y’ strategy would allow almost 50% of allocated storage to fail without any loss of data.

From a data residency perspective, distributing storage nodes geographically is an interesting, though perhaps counter-intuitive strategy that also provides data residency compliance. In a typical 5 of 9 configuration where three storage nodes are in location 1, three are in location 2 and three are in location 3, no single location would have enough data fragments to reconstitute data. Since none of the individual geographic locations have enough fragments to reconstitute, even if one location were exposed under a data breach scenario, any storage that was compromised would technically contain no data.

When stored data is recalled for use, re-identification rules can include a number of policies and any combination of source IP, geo-location of Vaulting, user location, or other metadata to help identify and collect fragments. An example of a common policy is that data being extracted in a particular geographic region, may only be recalled by a user in the same geographic region, but policy overrides may also be added for managers or compliance users that might require access regardless of their geographic location.

Numerous options also exist for how reconstituted data will be presented to end users. For example, while a credit card may be fully encrypted and fragmented on the way to an application, a replacement or tokenized value might be displayed in full for some end users, and masked for others. Less privileged users like those in a call center can see a format-preserving, masked version of the tokenized value while a developer or QA team member would see a more complete-looking representative replacement value (i.e. 16 digits, starts with a 2/3/4/5, has a valid BIN, passes a LUHN check) which is not actually a credit card.



The Result

DataStealth allows enterprises to define their data protection and data residency policies to ensure their sensitive data is appropriately secured and protected while they remain in control of data residency. Private, sensitive and / or confidential information never leaves their environment without explicit consent and full control over where it goes, how it is protected, and when it can be reconstituted. NEVER.

Authorized data security administrators can select whether to allow a data going to the cloud should remain in clear text, be encrypted, or to be replaced with a token on a field-by-field basis. When using surrogate token values, sensitive data never leaves the organization's control in any format – making it particularly useful for organizations that need to adhere with Canadian Privacy Laws. Data in the Cloud is either tokenized or encrypted, or both, so it is meaningless when viewed within the cloud and organizations can be confident that their sensitive data is within their full control at all times.

DataStealth is the enabling technology you have been looking for.

Datex Inc.
2333 North Sheridan Way
Suite 200
Mississauga ON L5K 1A7
Canada

+1-855-55-DATEX
www.datex.ca