Making Compliance Easy and Cost Effective

Background

Most regulations businesses face today were developed in the past few decades. They span the gamut of industries, such as health, privacy, and finance, and often concern – more recently – data collection. IT organizations are particularly tasked with the challenge of maintaining compliance with these many regulations, given their central role in the infrastructure, operations, and marketing practices of a company.

Compliance requirements have evolved; however, most IT organization processes have not kept up. Rising concerns about personally-identifiable information (PII) privacy has led to new regulations world-wide, including but not limited to: the European Union (EU) and United Kingdom (UK) General Data Protection Regulation (GDPR), Singapore's Personal Data Protection Act (PDPA), the California Consumer Protection Act (CCPA), and many others in-process.

The difference between these regulations and others are that PII regulation's non-compliance fines have real teeth. The case for most non-PII regulations is that non-compliance results in fines, but not debilitating ones. Fines tend to be arbitrary and are frequently so small they can be considered merely a slap on the wrist. Repeated non-compliance fines do increase; however, they are low enough to cause a real decision about the cost of getting into compliance versus the cost of non-compliance.

With PII regulations, it is a different story. For example, in GDPR – the regulation several others are based upon – there are only two levels of fines. The lowest level is **€10 million** or **2% of world-wide revenues,** whichever is greater. The higher-level fine is **€20 million** or **4% of world-wide revenues,** whichever is greater. These are non-trivial fines. When a company evaluates the cost of non-compliance versus the cost of compliance, it is clear that compliance costs significantly less. PDPA and CCPA have different fine structures that also have teeth.

Key Problems

One of the most prominent requirements of PII regulations is "The Right to Be Forgotten," a.k.a. "The Right to Erasure." When any individual requests an organization remove all of their PII information, that organization has approximately 30 days to do so, document it, and inform the requestor that it has been completed.

Completing this task for structured (database-centric) data is relatively straight forward. It's not easy, but doable in a timely manner. Completing it for unstructured data (files) is exceedingly difficult. And completing it in image-based backups, the most common backup in use today, is – for all intents and purposes – impossible with current technologies.

The Challenge with Image Backups

Image backups are so popular today because of their incredibly fast recovery times. When a physical or virtual machine outage occurs, instead of spending hours to recover a single machine – let alone multiple machines – image backups can simply mount the machine image on the backup server or storage. Users are pointed at the new mount and they are back up and running in minutes, not hours. When the hardware is repaired or replaced, the images are failed back. It's easy to see why they are so popular. But, image backups today are not capable of deleting PII data from a master backup and propagating that deletion to all other backups. This is a major problem. It has to do with the very technology that makes them so easy to recover.

Image backups are tied or linked together. Each one after the first is merely an image of the changes that occurred from the previous backup. Point-in-time backups are synthesized into a complete virtual image by tying the backups together. If an administrator removed data from the primary image backup, it would corrupt all of the subsequent backups that have pointers to the original data.

The Challenge with Time

It is a time-consuming, laborious process to effectively remove PII data from primary image backups without corrupting the entire backup series. It requires mounting each and every backup from the most recent to the oldest, finding the files and databases with that particular data – exceedingly difficult for unstructured data by itself – removing the PII, restoring the backup to its dormant status, then documenting what was done, when it was done, and by whom. This process must be repeated for each and every PII request to be forgotten and for when PII data is no longer required for the purpose it was collected.

Compliance becomes a labor-intensive problem when the number of backups requiring PII purging exceeds more than a handful. If the organization is backing up once a day and has multi-year retention requirements for other compliance requirements, there are going to be hundreds to thousands of backups. It will be impossible to complete in the timely manner required, meaning non-compliance and potentially massive fines.

Respecting the Right to Be Forgotten

Some have suggested the "Right-to-be-Forgotten" compliance is only an issue when the data is recovered. There are several problems with that process in today's modern backup. The first is the recovery itself. To remove the PII before it is recovered requires staging the recovery in a sandbox first, then finding all of the PII requiring removal (there are going to be multiple requests every week, adding up over time). This is a time-consuming project for one record, never mind multiple requests a week multiplied by the number of backups taken over a period of time. No organization can afford to spend hours – or days – on searching for and removing PII data from their backups.

And since most IT organizations are now utilizing backups by repurposing virtual copies of the backup data for devops, test-dev, and analytics, there is a very real probability that the PII data will leak into other applications, leaving PII data that has not been masked or removed available in the company's systems. This again risks costly non-compliance.

There is also another crucial PII compliance issue called the "Right-to-Access". This is the right to view all data on a particular individual that an organization possesses so it can be corrected if there are mistakes

or errors. Due to the difficulty of finding PII in unstructured data, the organization is again at risk of non-compliance.

To summarize: compliance with PII "Right-to-be-Forgotten" regulations means being able to locate, delete, and document PII from the backups. Image backups are not conducive to PII access or deletion in the timely manner required, putting most IT organizations at enormous risk of non-compliance fines. "Right-to-Access" PII means being able to locate all of the PII for an individual, to make a copy of the data for them, and to allow them to make corrections. Because finding all the PII in unstructured data is problematic there is additional non-compliance.

ioFABRIC's Intelligent Software Solution

ioFABRIC Software is designed specifically to provide the active intelligence necessary to maintain compliance in imaged backups, working with such products as Veeam and Cloudberry. It also addresses unstructured data.

How does it work? First, ioFABRIC Software indexes backup images to allow for simple search across all document metadata and contents. This makes finding PII data easy. Once the specific PII data has been found and the data subject requests it to be removed, ioFABRIC will erase or replace the data with a null string. The changes are propagated across all of the image backups, eliminating any worry about backup corruption. And ioFABRIC provides a full customizable audit report of who requested the access or erasure, who approved it, where the changes were made, and when each action happened in the process of executing the request.

Best of all, ioFABRIC Software delivers retention locks to add additional regulatory compliance and data protection against malicious acts and data loss (learn more about this in our solution brief on long-term retention and ransomware). And all of this capability is available for a very reasonable low cost subscription license.

For more information, contact us at: iofabric.com/contact or 1-833-IOFABRIC (463-2274).