FLASHBLADE AND GDPR COMPLIANCE

This brief explains how Pure Storage® FlashBlade™ systems help enterprises comply with the data processing and storage provisions of the European Union's General Data Protection Regulation.

EXECUTIVE SUMMARY

The European Union's General Data Protection Regulation (GDPR) takes effect on May 25, 2018. The regulation defines handling, use, and transfer requirements for entities that deal with the personal data of EU residents, as well as those individuals' rights with respect to their data. GDPR applies to all entities that handle the personal data of EU residents, regardless of whether the entities are located in an EU member country. In particular, the regulation applies to the processing and storage of data in digital form, regardless of where the processing and storage occur.

THE SCOPE OF GDPR

Fundamentally, GDPR declares that EU residents (called *natural persons* and *data subjects* in the regulation) own their personal data, and specifies both their rights with regard to it, and obligations of entities that acquire and process it, particularly the obligation to keep it secure and available.

Individuals' rights to their personal data include the right to access it, the right to rectify errors, the right to know how it is being processed, the right to restrict the types of processing it undergoes (within certain legal limits), and the often-cited *right to be forgotten* (i.e., to request that data destroyed when it is no longer required for legitimate processing).

Article 1

Subject-matter and objectives

1. This Regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data.
2. This Regulation protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data.
3. The free movement of personal data within the Union shall be neither restricted nor prohibited for reasons connected with the protection of natural persons with regard to the processing of personal data.

Excerpt from GDPR Article 1 defining the regulation's purpose and scope

GDPR classifies entities that deal with individuals' personal data as *controllers*—entities whose missions entail personal data handling—or *processors*—entities that carry out processing tasks on behalf of controllers. A single entity can fulfill both roles. In the context of the regulation, the term *processing* encompasses both manual operations such as acquisition, filing, alteration, and disclosure, and automated processing, storage, and transmission of data in digital form. The regulation restricts controllers and processors as to what data they may acquire and the purposes for which they may process it, and specifies protections they must provide against both unauthorized access and loss or destruction during processing, storage, and transfer. Additionally, the regulation obliges processors to disclose what data they store and process to its owners, to rectify verifiable errors, and to destroy personal data when it is no longer required for its intended purposes. Finally, it specifies procedural mechanisms for compliance and potential penalties for non-compliance.

Thus, GDPR deals both with *policy*—what data may be collected and for what purposes it may be used, the rights of EU residents with respect to their data, and so forth—and *technology*—how data in digital form should be secured against unauthorized access and inadvertent or malicious destruction as it is processed, transferred, and stored.

COMPLYING WITH GDPR

As of May 25, 2018, entities that acquire and process the personal data of EU residents are required to comply with the GDPR. Compliance requires that controllers and processors have in place verifiable procedures to prevent *personal data breaches*, defined in the regulation as events that lead to "*the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed*."[1]

In terms of policy, compliance encompasses organizational structures, data handling procedures, physical security of data repositories and processing facilities, and hiring, training, and auditing operations. From a technology standpoint, compliance is largely concerned with computing hardware and software, storage, and communication facilities that, when properly managed and maintained, provide high barriers to theft, unauthorized disclosure, alteration, and inadvertent and malicious destruction of EU residents' personal data.

PURE STORAGE FLASHBLADE™ SYSTEMS AND GDPR COMPLIANCE

As a supplier of digital data storage systems, Pure Storage makes every effort to keep the data stored in its systems both available to authorized users and secure against electronic intrusion and physical misappropriation. For example:

- Data stored in FlashBlade systems remains intact and accessible in the face of all single-component failures and many types of concurrent failures of multiple components
- Administrative access to FlashBlade systems is restricted to credentialed individuals. The architecture provides no facilities by which administrators can accessing modify stored data.
- Every blade in a FlashBlade systems uses a hardware implementation of the well-accepted AES-256 algorithm to encrypt *all* data and metadata it stores in flash and stages in NVRAM. Encryption is "always on"—it cannot be disabled. Systems manage both per-blade and system-wide encryption keys autonomously, and store them in such a way that data on a single misappropriated blade cannot be decrypted.
- Immutable snapshots provide unalterable records of data sets as they existed at user-specified key points in time.

Taken together, these FlashBlade properties help data controllers and processors "design GDPR compliance by default" as they implement new processing systems.[2] Pure Storage Technical Brief TB-180101, *FlashBlade Data Security* describes how FlashBlade systems protect stored data from loss and unauthorized access under adverse conditions and attacks. When combined with strong network security to protect data while it is "in flight," and robust data processor and controller system access and data handling policies, FlashBlade systems can be an important component of an overall GDPR compliance strategy that is both comprehensive and cost-effective.

This brief presents a context for securing data starting at the point of digitization, through processing, storage, transmission, and destruction. It discusses the properties and costs of alternative architectures for the information technology component of GDPR compliance, and illustrates how FlashBlade storage can be a key component of an overall compliance strategy.

[1] Official Journal of the European Union, 4.5.2016, Article 4(12).

[2] Excerpt from GDPR Article 25 (Data protection by design and by default): "...the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects."

FLASHBLADE AND GDPR

A MODEL FOR DIGITAL DATA HANDLING

Figure 1 is a simple model of personal information in digital form, from acquisition through processing, storage, dissemination, and destruction. The numbers in the figure represent points at which data must be explicitly secured.

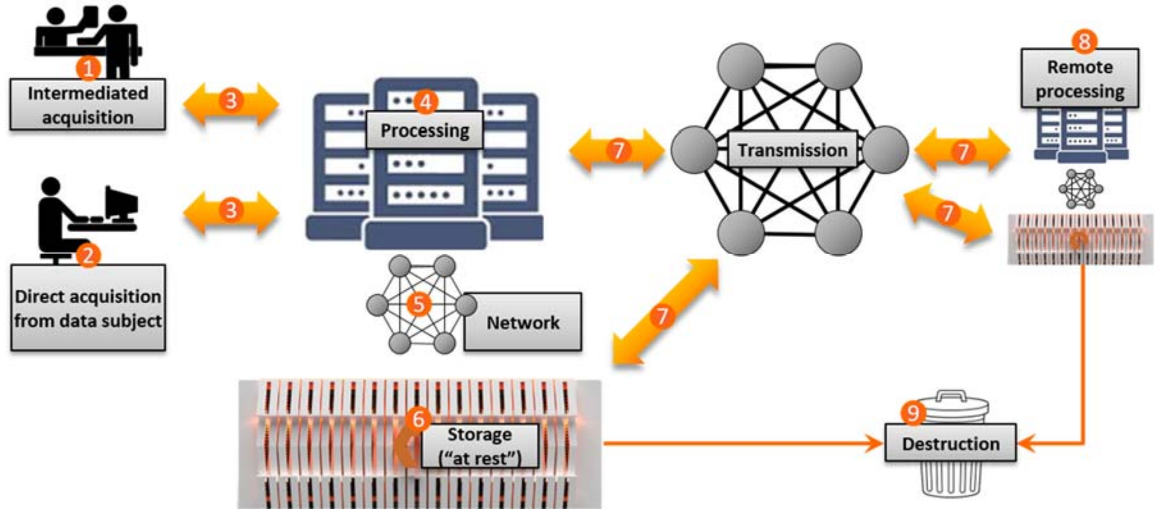

Figure 1: Digital Data Acquisition, Processing, Transmission, and Storage

1. Some personal digital information originates when an individual (*data subject*, in GDPR terminology) interacts with an *intermediary*, such as a bank teller, sales clerk, or government official. The intermediary in turn interacts with a (typically remote) computer system on the subject's behalf. Data privacy and security depend on the trustworthiness of the intermediaries and processing systems, for example in the form of robust vetting and authentication of both data subjects and intermediaries.

2. Increasingly, data subjects digitize their own personal information by interacting with ATMs, online shopping web sites, licensing and permit services, and so forth. Responsibility for securing data lies primarily with the computer systems with which the data subjects interact.

3. Communication links between data subjects and processors' agents and facilities must prevent unauthorized access and passive snooping. Some links are permanent; these tend to have strong security. But increasingly, both data subjects and processors' agents use wireless (e.g., mobile credit card readers) and semi-public internet access points to interact with processing facilities. Data transferred on these links must be secured.

4. It is likely that most processing of personal data subject to the GDPR will take place in physically secure data centers. Thus, security of data while it is being processed depends primarily on (a) trustworthy applications that are well managed and maintained, (b) comprehensive facilities for auditing processing events, and (c) strict control of both local and remote access to data processors' systems.

5. The overwhelming majority of data processors use storage networks to connect their servers and storage systems. For storage networks completely contained within a data center, security of transmitted data is provided by (a) controlling access to the facility and (b) organizational policies that govern access to equipment. Where this is not the case, for example where servers and storage are located in separate facilities or where data is replicated between remote systems, security during transit should be provided by strong network encryption.

6 The storage systems that hold personal data while it is “at rest” should protect it not only against unauthorized access and alteration, but also against physical threats such as system component failure and hardware theft. In essence, this means that (a) ideally, storage systems should encrypt the data they store, with rigorous key management procedures, and (b) that both administrative and user access to storage should be closely controlled, with all administrative actions indelibly logged and subject to regular audit.

7 Many processors transmit data between systems, either in the form of individual items (e.g., to authorize and confirm transactions), or in bulk (e.g., to replicate, analyze, test with, or archive entire data sets). Some use synchronous replication to maintain identical copies of high-value data sets at two locations. Ideally, network facilities used for these purposes would be private, but increasingly, processors use common carriers to transmit data over distance. Personal data transmitted over public networks should be secured by virtual private networks (VPNs), encryption, or a combination of the two.

8 Backing up personal data and sending it to remote systems creates *copies*. The GDPR declares that data subjects have a right to know where copies of their personal data exist and for what purposes they are used. More importantly, they have a “right to be forgotten”—for personal data items to be eradicated once there is no longer a valid reason to retain them. To eradicate data, it is necessary to know where it is. Thus, data controllers and processors are obliged to track data as it moves among their own and their processing partners’ facilities.

9 Finally, individuals “right to be forgotten” implies that processors must be capable of eradicating personal data upon owner request when it no longer serves a legitimate purpose. At the individual item level, responsibility for complying with this provision lies primarily with (a) the integrity of applications used by the processor and (b) the processor’s policies and procedures for dealing with data subjects. For bulk destruction of large data sets, such as survey results, some storage systems can assist by providing fast, reliable eradication of entire sets of data.

As a supplier of data storage systems, one of Pure Storage’s key contributions to GDPR compliance lies in protecting data that is “at rest” in the company’s systems. Another important contribution is keeping data available and secure as it moves between application and database servers and the company’s storage systems (points 5 and 8). The latter is of particular concern when networks have connections outside the data center, and is typically accomplished in conjunction with storage network or server facilities.

FLASHBLADE SYSTEMS AND GDPR COMPLIANCE: KEEPING DIGITAL DATA AVAILABLE

Arguably, FlashBlade systems offer the most robust facilities for keeping file and object data intact and accessible of any available on today’s market. The systems are designed to continue operating in the presence of hardware component failures, up to and including concurrent failure of two blades. A powerful RAID-HA protection scheme makes data recoverable from all single and concurrent double read failures, as well as in many scenarios that affect more than two blades. In addition, intra-device checksums detect *latent* read errors, both those that deliver corrupt data as good and those that deliver correct data from incorrect locations.

FlashBlade systems also protect against inadvertent destruction of data sets by administrators via an automatic 24 hour *eradication delay* during which destroyed file systems are recoverable.

The systems deliver these protections in conjunction with the most effective file and object data compression available today and with virtually no impact on system performance.

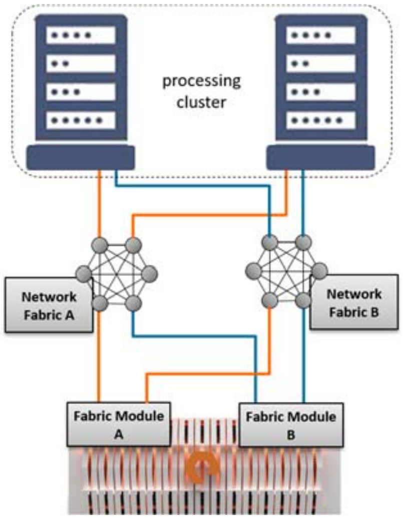

Figure 2: Redundant Sever to FlashBlade Connections

Pure Storage Technical Report TR-160701 describes the FlashBlade architecture, including the techniques employed to protect stored data against loss and unintended destruction.

Not only do FlashBlade systems protect data against faults, they keep it available to application servers ("clients") through multiple storage network ports on separate system fabric modules. When FlashBlade systems are cross-connected to clustered clients through two independent IP network fabrics, as Figure 2 illustrates, data remains accessible for processing if a network path, a client, or an entire blade should fail.

FLASHBLADE SYSTEMS AND GDPR COMPLIANCE: KEEPING DIGITAL DATA SECURE

FlashBlade systems also contribute to GDPR compliance by securing the data they contain against misappropriation, either by electronic means such as unauthorized access by administrators or client computers, or by physical means such as theft of blades from a system. In both cases, the systems combine with other components in the digital information technology "stack" and with data processor policies and procedures to deliver full compliance.

The key FlashBlade features that support GDPR compliance by raising barriers to unauthorized access to data are:

Controlled administrative access

Every administrative login to a system requires credentials. In addition to credentials, every REST call must contain a unique *API Token*, generated by the system for each client-side account for which REST access is enabled. The system itself manages credentials for the single supported account (*pureuser*).

No administrator access to stored data

While authenticated administrators can delete[3] (*destroy*, in Pure Storage terminology) file systems, no administrative interface (command line, graphical, or REST) includes *any* facility that would allow an administrator to retrieve or alter stored data. Moreover, destroying a file system initiates a 24-hour grace period during which an erroneously destroyed file system can be recovered.

Secure administrator communications

CLI interactions take place within encrypted secure shell (SSH) sessions. Communication with browsers uses HTTPS, both during administrator authentication and for encrypting exchanged information. Systems use HTTPS with TLS encryption to transmit log information to the Pure1® Cloud.

Controlled access by client computers

FlashBlade systems only execute I/O commands from client computers that have been explicitly authorized to access specific file systems or object store accounts by an authenticated system administrator.

Logging of events and administrative actions

FlashBlade systems log all administrative commands, regardless of source. Systems store logs locally, and in addition, when they are connected to the Internet (as most are), regularly transmit their logs to the Pure1 Cloud, where they can be viewed by Pure Storage Technical Suppport Engineers (TSEs) or by agents of the system owner through the Pure1 Cloud *Manage* feature. Comprehensive logging of system events and administrator actions enables auditors to analyze all actions performed on a system.

[3] "Delete" is commonly used in information technology contexts to mean either "make inaccessible," "obliterate," or a combination of the two. For example, a deleted file system on a file server becomes invisible to clients immediately, but unless it is *securely* deleted (physically erased), its data may remain in storage until the space it occupies is reclaimed and overwritten.
Pure Storage command line interfaces use the terms "destroy" to mean "make invisible to clients," and "eradicate" to mean "obliterate contents." Destroyed FlashBlade file systems are invisible, but remain intact for 24 hours after destruction unless explicitly eradicated by administrator command.

Encryption of all stored data all the time

FlashBlade systems encrypt all stored data and metadata using custom hardware in each storage unit that implements the AES-256 algorithm, widely accepted as the "gold standard" for digital data encryption.[4] Each of a blade's storage units uses a unique *device key* to encrypt incoming data before storing it in flash or staging it in NVRAM, and to decrypt stored data for delivery to clients. Storage units hold their device keys in encrypted form in special-purpose flash, using a system-wide *key encrypting key* (KEK) to encrypt them. Device keys are never exposed outside storage units.

Systems partition their KEKs mathematically, and store each partition on a separate blade. More than half of the partitions are required to reconstruct a KEK. At system startup, fabric modules request KEK partitions from all blades and use them to reconstruct the system KEK, which they send back to the blades. Blades relay the KEK to their storage units, which use it to decrypt their respective device keys, after which they are ready to encrypt and store incoming data, and to decrypt stored data for delivery to clients.

Immutable snapshots

FlashBlade snapshots cannot be altered. Thus, they protect against inadvertent destruction of data, and provide point in time records that can be used to detect unauthorized alterations or thwart ransomware attacks.

Personal data stored in digital form does not exist in a vacuum. Its inherent purpose is to create a durable record of manipulations of personal information by data processors and (usually indirectly) by data subjects themselves. Storage is therefore inherently integral to the overall digital data lifecycle (Figure 1), particularly whle data is "at rest" (stored persistently) and as it is transferred between processing systems.

For example, the storage networks that connect application servers to storage systems (④ in Figure 1) routinely carry personal data between the two. For storage networks within the data center, separated from external access by "air gaps," data in transit is primarily protected by data processor policies that limit access to network components and their connections to server and storage systems. Data transferred on networks with connections outside the data center should be protected by combinations of VPNs and encrypting network switches.

Securing personal data within the data center necessarily has a strong policy component. For example, data processor policies should limit administrative access to computing, network, and storage systems to specific trusted individuals. Policies should specify administrative roles narrowly, with non-intersecting responsibilities (for example, different individuals manage storage, servers, networks, and security servers), and automatically create unalterable records of all administrative actions. GDPR compliance within the data center also depends partly on data processors using suppliers whose products provide facilities for implementing and enforcing policies that safeguard personal data.

THE DATA ENCRYPTION DILEMMA

The GDPR does not mandate encryption of stored digital data per se. Within the data processor community, however, it is generally accepted that encryption of personal data throughout its digital lifecycle will ultimately be a practical necessity for compliance. As Figure 1 suggests, "end to end" encryption must occur in multiple stages. For example, data should be encrypted on the path between origin and processor (③ in Figure 1), but it must be decrypted for processing (④), and re-encrypted for transmission to and storage by storage systems (⑤ and ⑥).

As outlined in the preceding section, for storage networks completely contained within the data center, the obligation to protect data from misuse and theft while in transit between servers and storage rests primarily with data processor policies and practices. But practically speaking, some storage networks cannot be isolated, particularly those that use iSCSI over Ethernet. This complicates end-to-end data encryption. One proposed solution is to encrypt data at the application server before sending it to storage, thus limiting the number of points at which it is "in the clear." Some GDPR consultants and IT equipment vendors recommend encrypting all sensitive data before it leaves the application server, and storing and transferring it only in encrypted form. Data encrypted before it leaves the server is protected "downstream"—in primary storage, when it is sent to other servers, and when it is backed up or archived. Protecting

[4] At the time of publication, FlashBlade systems are undergoing certification of compliance with the FIPS 140-2 Level 1 information security standard published by the US Federal Government. Certification is expected to be complete in late 2018. The company is also exploring means by which it can assure compliance with the NIST SP 800-88 R1 standard.

data encrypted at the server necessarily consists of (a) managing encryption key distribution so that copies of data can only be decrypted for legitimate purposes, and (b) implementing policies to ensure that only authorized individuals have access to keys and data, and only for authorized purposes.

Setting aside for a moment the complexities of key management, data encryption at the application server covers many GDPR compliance requirements for digital data, but it comes at a high cost. *Data reduction*—the elimination of redundancy in data prior to storing it persistently—has become a universal expectation of data processors as they evaluate storage alternatives. Increasingly, processors budget for and acquire storage based not on their raw capacity requirements, but on expectations of the degree to which data can be reduced for storage. Reducibility varies widely based on the nature of data, but typically falls in the range of 1.5:1—5:1. Because so much personal data is textual, its reducibility typically falls toward the middle of the range or higher. Put another way, with reduction, processors can store a given amount of data in 20-67% of the physical capacity its unreduced size would suggest. Moreover, when data is reduced, storage acquisition cost, space requirements, and power consumption are all proportionally lower.

FlashBlade systems remove redundancy from data by *compression*—replacing sequences of bytes that occur multiple times within a block with compact descriptors that point to a single occurrence of a sequence.

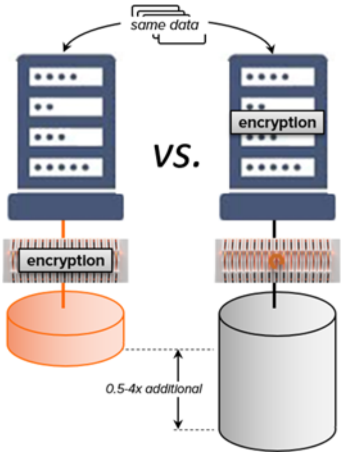

Figure 3: The Cost of Encryption at the Application

Encrypting data, however, inherently produces quasi-random bit patterns (It wouldn't be very effective if it didn't.) that essentially eliminate the possibility of compression. Thus, even ignoring any computational impact of encryption on application servers, encrypting data at the server virtually eliminates the cost advantage of data reduction that data processors have come to depend on.

But encrypting data at the application server prior to storing it is only the tip of the storage cost iceberg. Each time data is copied—for analytics, for development testing, for backup, for archiving, or for disaster protection—0.5-4 times as much additional storage and network bandwidth are consumed compared to data reduced and encrypted at the storage system level.

Moreover, the keys used to encrypt data sets at the server are an inherent security weakness. For example, the key that encrypts a production data set at the server must be made available to systems that analyze copies of it, that use copies for development testing, that restore backups, and so forth. Every use of a data set widens the circle of systems and individuals with access to its contents. Worse, when production data sets are re-encrypted with new keys and then copied, the new keys must be distributed to all users of copies so they can correctly decrypt each instance of the data set they process.

Server-level encryption of sets of personal data is thus a "sledge hammer" solution to the digital data security component of GDPR compliance—it does the job, provided that widely distributed keys can be managed, but at significant cost to the processor, and ultimately, to the data subject. It sacrifices data reduction—arguably one of the most important storage technology advances of the past decade. The alternative—encrypting at the storage level as FlashBlade systems do—preserves the data reduction cost advantage and mitigates the key management problem, but it exposes data as it moves unencrypted between application servers and storage systems.

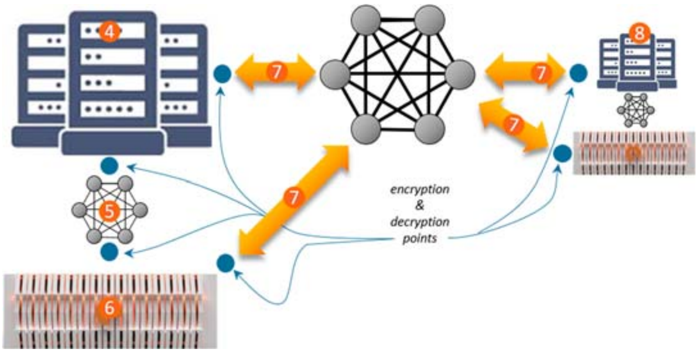

Figure 4: Encrypting Data in Transit

For storage networks contained entirely within the confines of a secure data center, data is protected by (a) controlling access to servers, networks, and storage systems, (b) narrowing administrative roles and assigning them to separate individuals, (c) comprehensive auditing of administrative actions, and (d) scrupulously managing and maintaining application and environmental software and firmware.

However, these measures do not entirely protect data that is transferred over connections to remote servers and storage systems. Data processors may find it more robust and cost-effective to simply encrypt data while it is being transferred, on both internal and global networks. Network equipment vendors offer hardware and software tools for encrypting data in transit, so it is possible to encrypt data as it moves between application servers and storage systems, and between replicating storage systems (⑦ in Figure 4) while relying on storage systems' encryption facilities (⑥ in Figure 4) to protect it while it is "at rest" (stored). Thus, compliance with GDPR digital data security provisions may require IT architecture redesign in some cases, but it is demonstrably possible to provide robust protection throughout data's digital lifetime and still reap the cost benefit of data reduction.

Encrypting personal data during transit provides reasonable protection against "snooping" and "man-in the middle" attacks, but data remains susceptible to *threats from within*—administrators that divulge keys indiscriminately, authorize file system and object store service account access by unauthorized "rogue" servers, and so forth. Thus, whatever the end-to-end data encryption architecture, it must be accompanied by policies that include interlocking safeguards against misappropriation and use by the data processor's own personnel.

A "HEDGE" STRATEGY: SELECTIVE PSEUDONYMIZATION[5]

Application servers can encrypt data at different levels of granularity. Servers may encrypt every block they send to a storage system, all data they write to specific file systems or databases, or even specific files or database columns (for example names or other identifying characteristics). Because GDPR is specific about the types of personal data that are subject to protection, some data processors may adopt compliance strategies that include server-side encryption or other types of pseudonymization for sensitive items only, with non-sensitive items stored "in the clear." Such strategies may enable them to comply with the GDPR's digital data protection provisions while retaining at least *some* of the cost benefit of reduction. Selective encryption is likely to lessen the degree to which storage systems can reduce data, primarily because it tends to reduce the number of compressible strings, but where the percentage of non-sensitive data is significant, some benefit should still accrue. On the positive side, selective pseudonymization is likely to consume less application processing power than bulk encryption of all data that a server processes and writes.

FLASHBLADE SYSTEMS AND THE "RIGHT TO BE FORGOTTEN"

Several articles of the GDPR relate to a data subject's "right to be forgotten"—the right to request that data controllers and processors destroy personal data that no longer serves a valid purpose (within limits imposed by member states' laws). Individuals' rights to be forgotten must be considered in two contexts:

Destruction of a single subject's data items

Deleting the records of a single individual inherently requires knowledge of the structure of the files or databases that contain them. Complying with an individual's right to be forgotten therefore lies with applications that process data, with users of the applications, and with the processor's data set management policies. In particular, the latter must make it possible to identify all instances of a subject's records (e.g., in snapshots, copies, backups, and other transformations of the data set in which they originate) so they can be removed from those derived data sets as well. Thus, the data processor's policies for tracking and managing digital data sets play the dominant role in compliance.

"Deletion" at the application-level typically makes records inaccessible, but does not physically obliterate them immediately. In principle, therefore, it is incumbent upon data processors to physically destroy all copies of data subjects' obsolete personal records within an acceptable time after an application "deletes" them.

[5] The term pseudonymization is used in the text of the GDPR regulation to denote digital data items that cannot easily be identified with their owners without some external means of relating the two.

Destruction of data about entire groups of subjects

Application-level record-by-record destruction of large sets of personal data (e.g., obsolete polling results or no-longer required database tables) is usually too onerous to be practical. FlashBlade systems can *destroy* entire file systems at a single administrator command, so they can expedite destruction of entire data sets as long as they are appropriately organized. Destroyed file systems can be recovered by administrative action for 24 hours (to protect against erroneous destruction), after which systems permanently *eradicate* them and reclaim and the storage they occupy.[6] A FlashBlade system can expedite eradication of an entire file system, but data processor management policies must identify all copies of obsolete data sets and ensure that they are destroyed.

FlashBlade systems present data to clients in the form of files in file systems or objects in an object store service account. The systems have no awareness of the files' or objects' internal structures. Thus, while a FlashBlade administrator can destroy an entire file system, it is not possible to erase or alter data *within* a file system or object store service account without a command from a client computer. Erasing large quantities of data by destroying a file system is therefore feasible only when the items to be erased are the sole occupants of the file system.

THE BOTTOM LINE

- From May 2018, compliance with the European Union's General Data Protection Regulation is expected, for all practical purposes, to be a condition of doing business in the EU.
- GDPR contains both organizational, procedural, and digital data handling provisions. Digital data handling includes both keeping personal data available and protecting it from misappropriation and misuse.
- GDPR does not specifically require encryption of data in digital form, but end-to-end encryption simplifies compliance to the point where data controllers are almost certain to require it, and data processors are likely to adopt it as a standard practice.
- Encryption at the application or database server protects data throughout its digital life, but at a substantial incremental storage cost due to the near impossibility of reducing encrypted data for storage. Storage cost is multiplied each time data encrypted by applications is copied for analytics, backup, testing, and so forth.
- Server-level encryption also increases key management complexity and introduces inherent security risk due to the need to manage wide distribution of keys to every user of any copy of a data set.
- For systems such as FlashBlade, that both reduce and encrypt data before storing it, end-to-end encryption that preserves the cost advantage of reduction and operational simplicity can be achieved in combination with encryption of the network links between servers and storage by network hardware and/or software.
- FlashBlade systems control administrator access, log every administrator interaction, and in addition, encrypt all stored and staged data and metadata *all the time*—FlashBlade encryption is not selectable, and therefore cannot be disabled, inadvertently or otherwise.
- The FlashBlade architecture also helps satisfy GDPR requirements for keeping subjects' data available for processing. In conjunction with redundant storage networks and application servers, it can provide a secure, highly available environment for personal data in digital form as part of an overall GDPR compliance strategy.

Compliance with the digital provisions of GDPR is necessarily an integration of application, server, network, and storage data protection facilities, together with data processor policies and procedures for handling and protecting data while it is in digital form.

[6] An authorized administrator can end the 24 hour period and force immediate eradication of a file system.

APPENDIX
FLASHBLADE AND GDPR COMPLIANCE

Table 1 lists excerpts from GDPR articles that relate to digital processing, storage, and transmission of personal data and describes the FlashBlade capabilities that help users comply with them.

Table 1: Relationship of GDPR Provisions to FlashBlade Properties and Capabilities

Article	Provision Excerpt	Relationship to FlashBlade Systems
3.1	This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not.	FlashBlade systems have virtually no optional features. The systems' high availability, data security, and access controls are the same throughout the product line. Thus, no matter what model or where it is deployed, a FlashBlade system's role in GDPR compliance is as described in the body of this brief.
3.3	This Regulation applies to the processing of personal data by a controller not established in the Union, but in a place where Member State law applies by virtue of public international law.	
4(2)	'processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;	FlashBlade systems perform or participate in the operations highlighted in green.
4(12)	'personal data breach' means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed;	The body of this brief describes how FlashBlade's always-on data and metadata encryption, in conjunction with network encryption facilities for data in transit minimizes the possibility of data breaches. In addition, credential-based administrative access, together with rigorous logging of events and administrative actions minimizes the possibility that unauthorized processing or misappropriation will occur or go undetected. Finally, the systems' built-in resiliency against single and double component failures represent state of the art protection against accidental loss, destruction, or damage of personal data.
5.1(f)	processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality').	
6.4(e)	…the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia: … the existence of appropriate safeguards, which may include encryption or pseudonymisation.	
17.2	Where the controller has made the personal data public and is obliged pursuant to paragraph 1 to erase the personal data, the controller, taking account of available technology and the cost of implementation, shall take reasonable steps, including technical measures, to inform controllers which are processing the personal data that the data subject has requested the erasure by such controllers of any links to, or copy or replication of, those personal data.	From a storage system standpoint, this provision primarily applies to bulk erasure of entire data sets. FlashBlade system administrators can destroy and eradicate entire file systems at a stroke. It is incumbent upon applications to direct the systems to delete individual files and objects.

Article	Provision Excerpt	Relationship to FlashBlade Systems
24.1	Taking into account the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons, the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation. Those measures shall be reviewed and updated where necessary.	FlashBlade encryption uses the widely-accepted AES-256 algorithm (implemented in custom storage unit hardware) to keep stored data secure. The systems' high availability, administrative access control, and logging features all help processors comply with the GDPR data availability, security, and access control provisions.
24.2	Where proportionate in relation to processing activities, the measures referred to in paragraph 1 shall include the implementation of appropriate data protection policies by the controller.	
25.1	Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.	FlashBlade systems use the widely-accepted AES-256 algorithm to encrypt stored data. All data security, high availability, administrative access control`, and logging features are integral to every system. Taken together, they help data processors "design in" protections for personal data in digital form that comply with the GDPR's high availability, security, and access control provisions as they implement new processing systems.
29	The processor and any person acting under the authority of the controller or of the processor, who has access to personal data, shall not process those data except on instructions from the controller, unless required to do so by Union or Member State law.	
30.1	Each controller and, where applicable, the controller's representative, shall maintain a record of processing activities under its responsibility. *...several processing activities called out...*	FlashBlade systems log all administrative actions, including: • file system creations, resizings, protocol changes, and destructions, and recoveries • snapshot creations and destructions • client connections and disconnections. As such, they provide the raw information required to comply with this article.
32.1	Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate: (a) the pseudonymisation and encryption of personal data; (b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services; (c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident; (d) a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.	FlashBlade data encryption, high availability, administrative access control, and logging features all help processors comply with GDPR data availability, security, and access control provisions. However, full compliance also entails network encryption of data in transit, application trustworthiness (e.g., with regard to pseudonymisation), and rigorous policies for data and IT equipment access and handling by agents of the data processor.

Article	Provision Excerpt	Relationship to FlashBlade Systems
34.3(a)	The communication to the data subject referred to in paragraph 1 shall not be required if any of the following conditions are met: (a) the controller has implemented appropriate technical and organisational protection measures, and those measures were applied to the personal data affected by the personal data breach, in particular those that render the personal data unintelligible to any person who is not authorised to access it, such as encryption;	Paragraph 1 of this provision defines the data controller's obligation to notify data subjects of personal data breaches "without undue delay." FlashBlade encryption and access control are regarded as "appropriate technical protection measures," so this provision is expected to apply where FlashBlade systems are used for data storage.
35.1	Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks.	Through its system engineering organization, Pure Storage makes available a series of technical reports and briefs that controllers can use to help assess the impact of proposed operations on personal data protection. A complete assessment, however, would take into account all facets of a proposed operation, including data acquisition, processing, storage, transmission, and eventual destruction.

Users are advised to seek appropriate legal advice on GDPR compliance. Pure Storage customers are solely responsible for obtaining legal advice from competent counsel in appropriate jurisdictions on any actions they may need to take to be in compliance with any relevant laws and regulations.