WHITE PAPER

Hitachi Content Platform Anywhere Enterprise Global File System
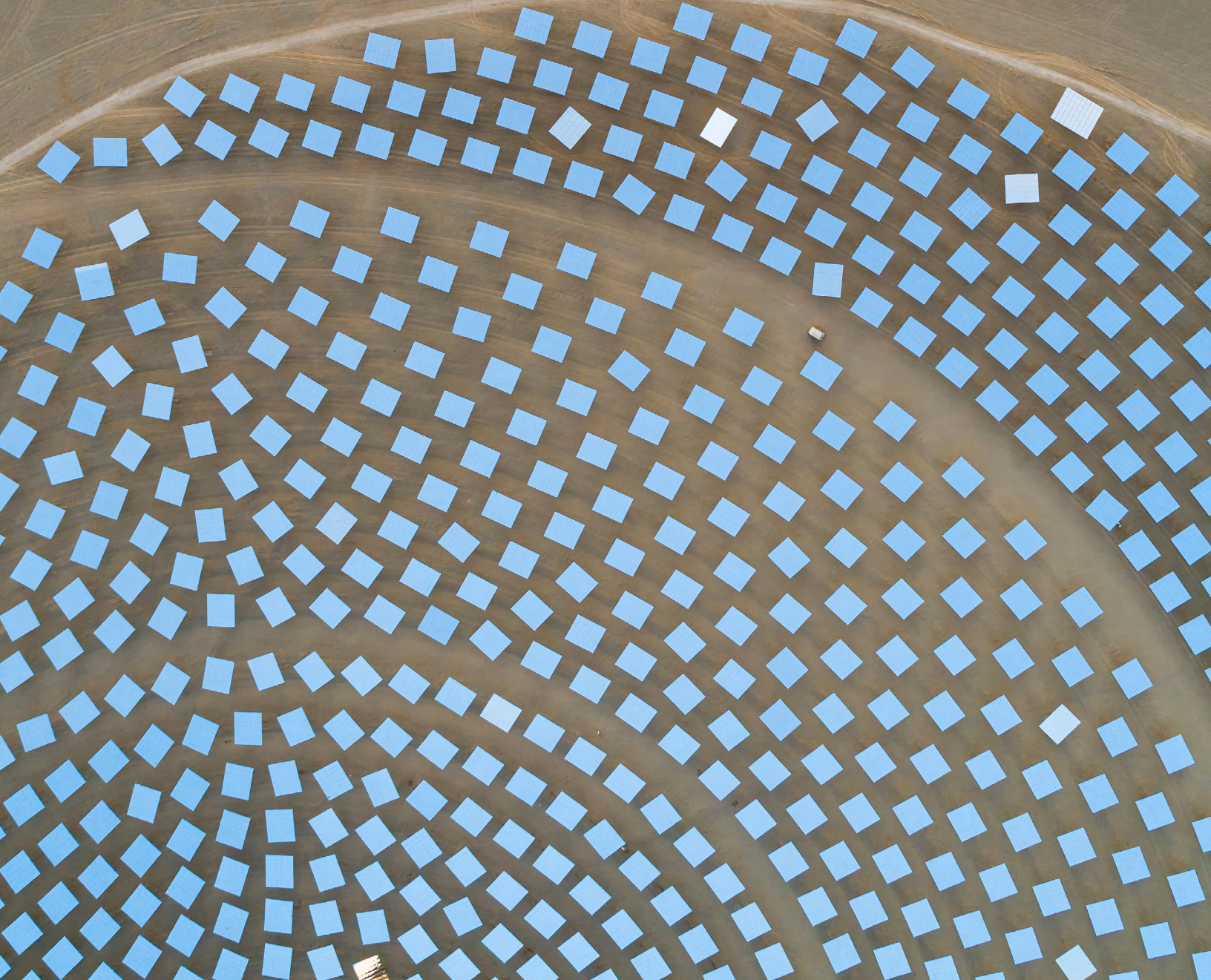
TABLE OF CONTENTS

Introduction

It's 7 p.m. on a Saturday night, and you are having dinner with your family. The phone rings. It's Jeff, the regional director of sales. This can't be good.

"Hey Tim, sorry to trouble you at home," Jeff says.

"I am at the office, and have been working all weekend on a proposal for a major deal that we need to close this quarter. It's due on Monday, and the VP of Sales must approve it first." You wait to hear the problem that led him to call you now.

"The thing is, the VP is on a business trip to Europe, and he needs access to the proposal," Jeff exclaims.

"That shouldn't be a problem," you respond, suggesting that Jeff sends the proposal via email.
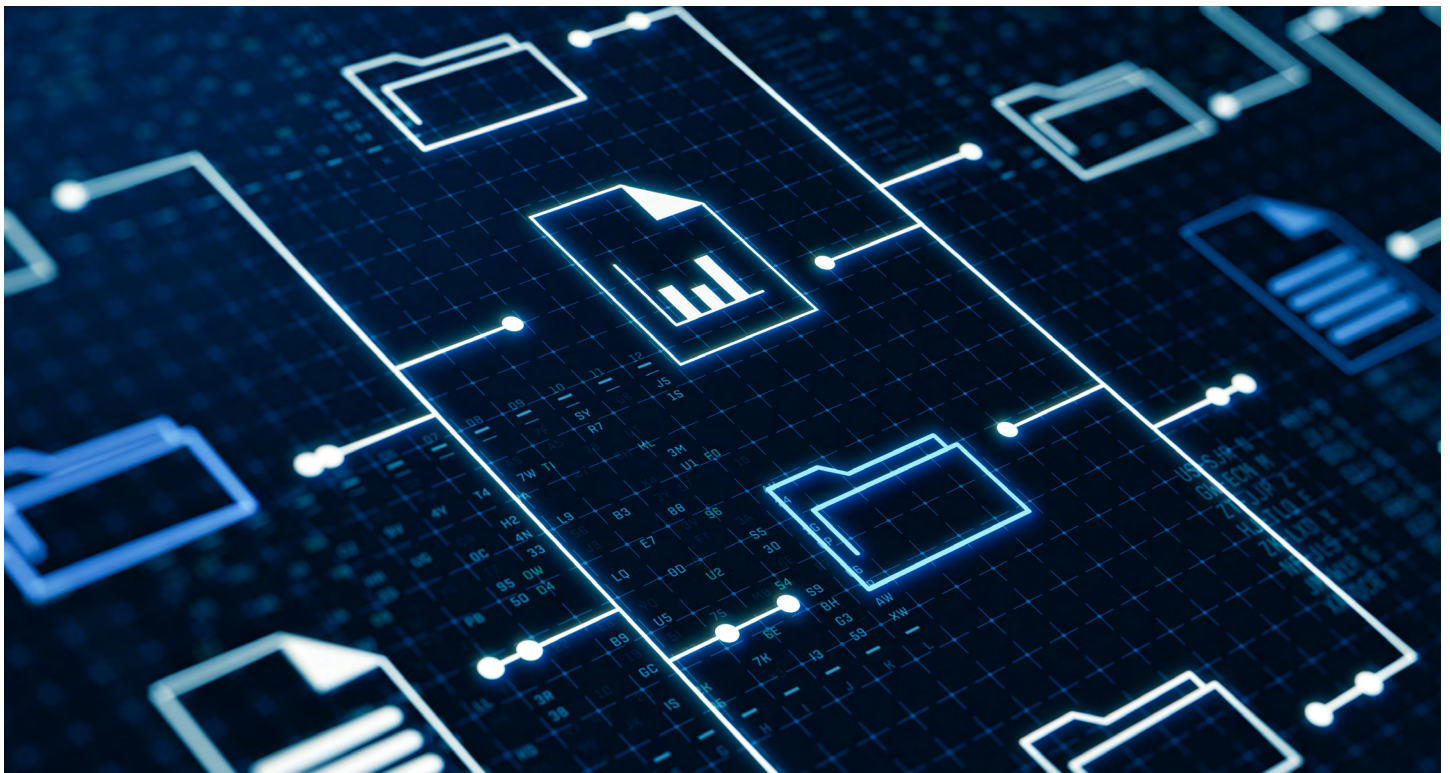
"You don't understand, Tim!" Jeff interrupts. *"The proposal consists of a dozen large files with images and drawings, plus references to other sales documents. I can't email* You've heard complaints before from the European team about how impossibly slow it is for them to access files stored at the company HQ. Solving that problem was on your list of things to do - you just hadn't gotten around to it yet.

Luckily, Jeff works at the company HQ just as you do, and not in one of the dozens of offices elsewhere. But you know what it means: you'll need to get into your car and meet him at the office. The only option you have is to log into the European file server, create a replica of the sales directory, and copy everything over. This should let the VP access any sales content he needs. Unfortunately, it could take hours to successfully complete the copy, so there goes your Saturday night dinner.

Does this scenario sound even remotely familiar?

Are you responsible for delivering and managing file services for users throughout your enterprise?

If so, a global file system (GFS) may be the right solution for you.

What is a Global File System (GFS)?

Files are the "lifeblood" of modern enterprises. Documents, spreadsheets, presentations, images and videos clips are but few examples of the content generated by today's knowledge workers.

Files are continuously created, modified and shared by workers, and each file must be stored, cataloged and retrieved upon request. Traditionally, most files have been maintained on storage mediums such as general purpose file servers, or on dedicated network attached storage (NAS) devices. Managing these files has been the role of a 'File System' – a service used for storing, organizing and retrieving files from a storage medium. File systems have been primarily 'local', managing file storage and retrieval on a single storage medium, accessible by users at a certain location.

Things have changed, though. Enterprise have gone global, and employees have become mobile. Geographical distribution created the need to access and share files between and across different physical locations (e.g. branch offices). And employee mobility called for accessing files via mobile devices (e.g. laptops, smartphones, tablets) – including from home and on the road.

Storage technology didn't stand still either. The advent of cloud computing hasn't skipped the storage world. Cloud-based object storage became ubiquitous, offering cost effective and scalable, virtually unlimited storage space for files. Rather than storing files on individual file servers or NAS, enterprises can take advantage of object storage in public-cloud, private-cloud or any combination of the two.

Centralizing file storage in the cloud has many benefits. However, a large number of employees work at remote sites and branch offices. Those workers are located at the "edge" of the enterprise wide area network (WAN), and therefore experience delays caused by latency and bandwidth constraints when trying to access centralized resources. Cloud storage solutions must therefore support fast and secure file access from the network edge.

The expanded file access patterns, coupled with the growing use of cloud storage, drove the evolution from 'local' file systems into 'global' file systems. Any global file system (GFS) should handle geographically distributed access to files, take into account network bandwidth and latency limitations, support mobile access, and take full advantage of cloud storage.

Key Requirements

With all the changes described above, it should come as no surprise that a global file system must address a diverse set of requirements. Let's look at the key ones:

- **Distributed access:** with employees increasingly spread around the globe, a global file system solution must offer reliable access to files – wherever the employees are. Techniques such as WAN optimization and caching should be employed to enable fast and secure file access throughout the distributed enterprise. Providing seamless remote access may involve deployment of 'access points' (aka 'cloud storage gateways') to facilitate a site access to the cloud.

- **Mobile access:** many enterprises have adopted a 'bring your own device' (BYOD) policy, allowing employees to access corporate assets using mobile devices, including laptops, smartphones and tablets. A GFS should enable employees to access files using a device of their choice, and to do so from home or on the road. This may require installation of software apps on users' laptops or mobile devices.

- **Security:** with files created and accessed by a geographically distributed organization, and often stored on externally hosted cloud infrastructure, security is a paramount requirement. A GFS must encrypt files both 'at rest' – wherever they are stored, and 'in transit' – while sent across a corporate network or the internet. Users who wish to access files must be first authenticated via corporate directory services (e.g. AD, LDAP) and granted access according to defined access control lists (ACL).

- **Flexible cloud storage:** a GFS should support any storage combinations enterprises choose to use, including "private cloud" hosted within the data center, a public cloud, or a dedicated virtual private cloud (VPC) hosted by one or more service providers. Furthermore, it should allow seamless transition from one storage target to another, avoiding any vendor lock-in. The chosen storage targets should not impact how users view and access their files via the GFS.

- **Central scalability:** demand for storage in general, and file storage in particular, continues to grow rapidly. A global file system should support seamless addition of storage space, taking full advantage of either private or public cloud storage resources.

- **Edge scalability:** The growth of the central storage pool should not impact users at the edge. They should be able to access any file, and to see all files available to them, regardless of their device's storage capacity.

- **Global name space:** Users should have the ability to access any file from any location – subject to their access rights. Therefore, each file should have a unique "global name", and a GFS should decouple physical location from logical file names. Each file should have a unique "global name," and be accessible to users (subject to access rights) from anywhere.
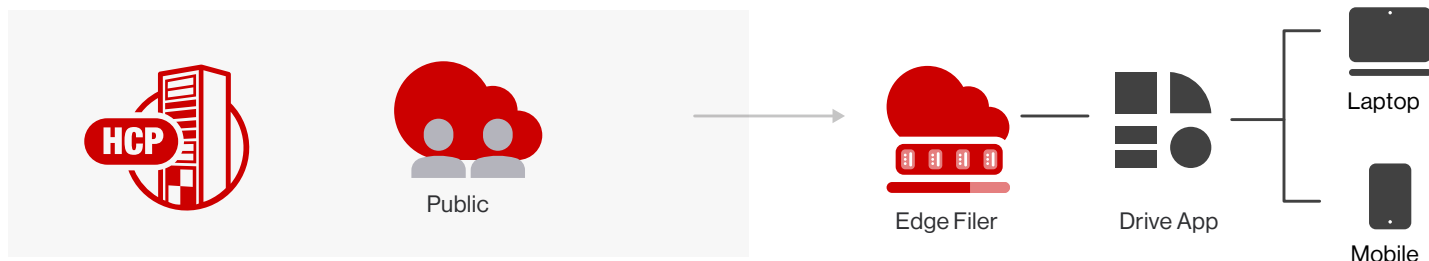
- **Collaboration:** employees often collaborate with others, creating and sharing files on a regular basis. Traditional methods for file-based collaboration rely on storing a file on a shared server, and then enabling access to others. This method typically requires involvement of a system administrator to set the proper ACL for designated files. A more modern collaboration method is file sync & share (FSS), which allows users to set desired collaboration attributes to specific files or folders, and then invite others to share them. A GFS should support both traditional and modern file sharing methods.

- **Data protection:** files embody significant portions of knowledge workers' output and corporate intellectual property. Protecting data stored in files is therefore a top priority. Data protection is an umbrella term that covers a variety of processes, such as backup and restore tools that protect files from external disasters or user errors and anti-virus (A/V) utilities that protect files from malware. A GFS solution should offer data protection tools both built-in and integrated via third parties.

- **Central management:** managing a global file system involves operations such as storage pools allocation, security policies, configuring and monitoring remote access points, etc. These operations are best handled through a central management platform, with well-defined role-based access.

- **Compliance:** demand for regulatory compliance is on the rise, with regulations such as Sarbanes-Oxley (SOX), HIPAA, GDPR forcing companies to carefully track their use and processing of data. A global file system should support detailed logging and monitoring of file related operations, allowing companies to meet the most stringent auditing requirements.

- **Protocol compatibility:** migrating from a local to a global file system should not disrupt existing workflows. Traditional protocols (e.g. SMB/CIFS, NFS, S3) used for accessing files on file servers and NAS should continue to operate in the presence of a GFS. In addition, access policies (e.g. ACL) previously assigned to files should be preserved.

Hitachi Content Platform Anywhere Enterprise Global File System

Hitachi Vantara allows organizations to launch and manage a modern global file system that delivers a wide variety of edge-to-cloud file services.

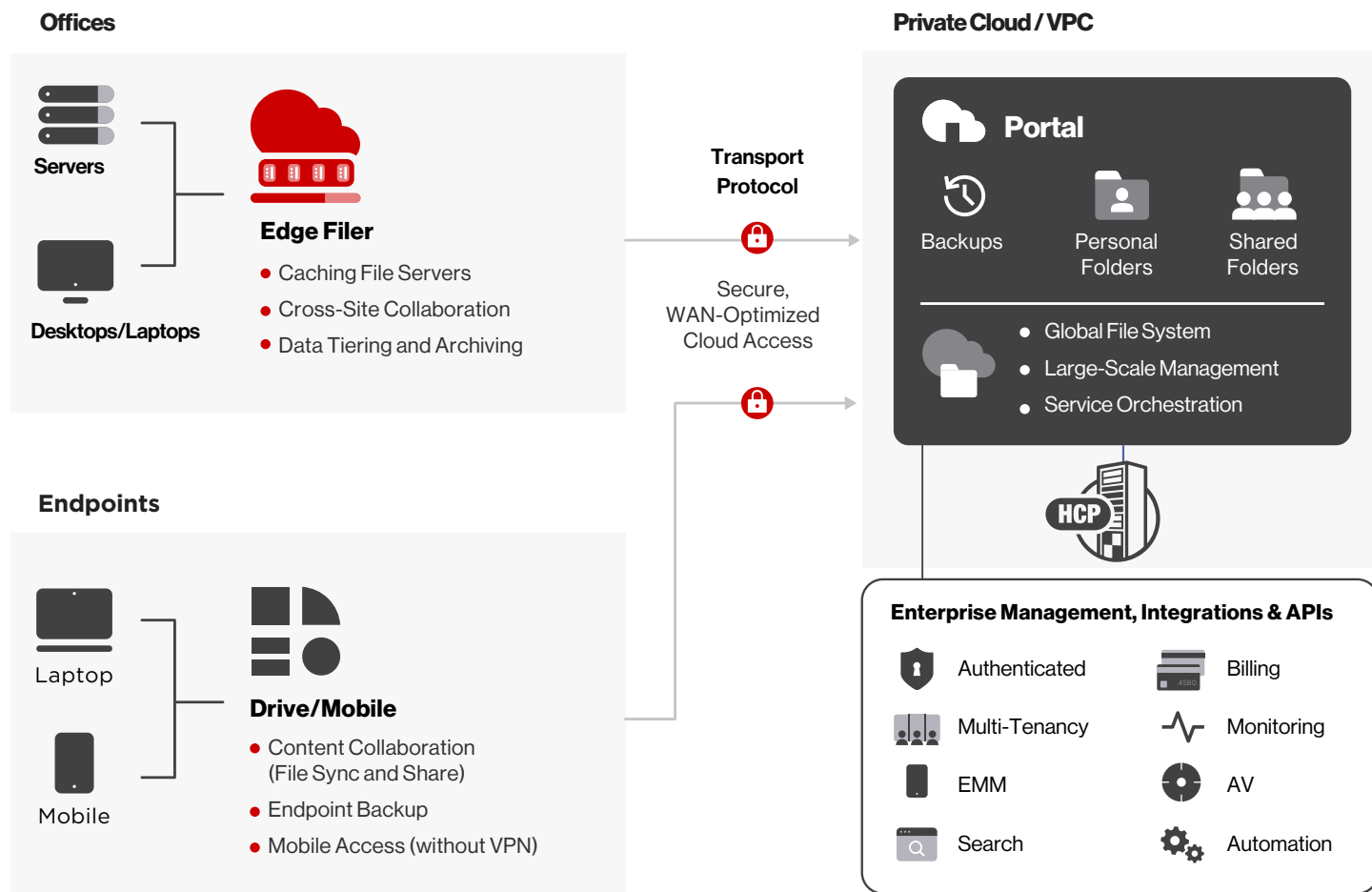**The Hitachi Enterprise File Services Platform includes the following components:**

- **Portal:** facilitates access to cloud storage services; handles data protection and file sync & share services; used for provisioning and monitoring global file services.
- **Edge Filers:** a line of physical and virtual appliances that function as 'cloud storage gateways' to streamline cloud storage access for remote sites
- **Drive:** Endpoint clients that support accelerated remote file access, data protection and file sharing, for users' workstations (desktop, laptops), as well as mobile apps (smartphones, tablets).

Together, these components allow Hitachi to offer true global file services: files are centrally stored and protected, while users can easily access them everywhere; Edge Filers and Drive clients guarantee fast and secure file access for remote sites and mobile users; and modern content collaboration services allow users to freely sync and share files. These services are offered with total security, featuring military-grade encryption and full control over data residency.

Solution Architecture

Let's take a closer look at the underlying architecture for the global file system platform:

The Portal is a software service deployed as a virtual appliance. It can run in an enterprise data center (i.e. private cloud) or on IaaS (infrastructure as a service) cloud providers. When using a public cloud service, the Portal can run in a virtual private cloud (VPC), leaving control over security in customers' hands. Hitachi supports a wide range of on-premises and cloud storage infrastructure providers, including Amazon Web Services (AWS), Microsoft Azure, IBM Cloud, HPE, and others. Customers are free to pick their cloud service of choice.
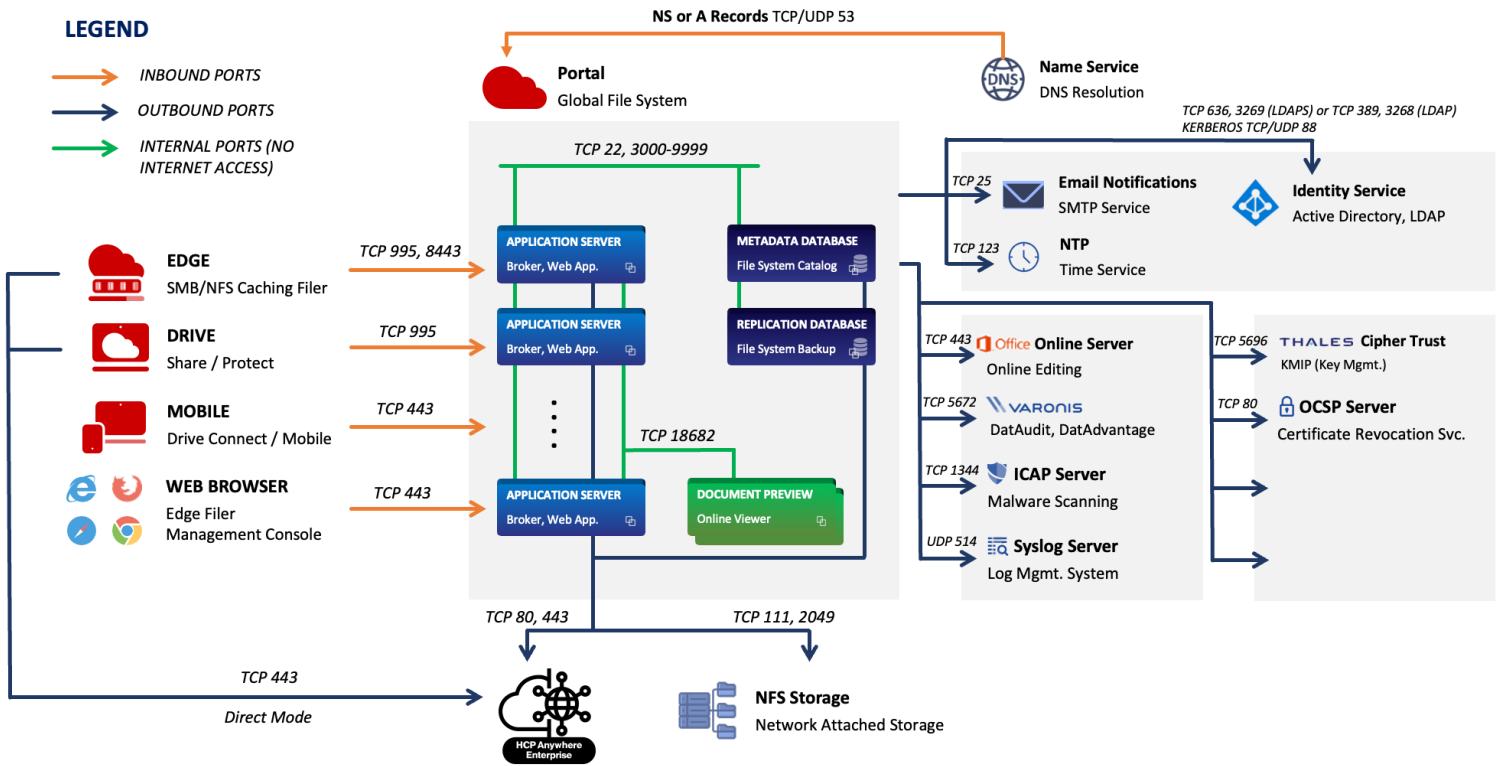
Portal software has two main roles: a) it offers a broad range of data management services, such as file storage/retrieval, data protection, sharing, etc., and b) it handles service orchestration tasks, such as assigning storage quotas, configuring remote devices/agents, monitoring, etc. The Portal comes with built-in integrations with a wide range of third party services, including Active Directory, and various mobile device management (MDM), and antivirus (AV) platforms. Furthermore, the Portal has a rich set of documented APIs allowing customers to integrate with any other third party service.

Edge Filers enable sites with bandwidth and/or latency limitations to accelerate and secure remote access to cloud storage. Edge Filers utilize caching and optimization techniques to deliver the best user experience. They act like a local NAS, supporting traditional file protocols, while serving as a "gateway" to cloud storage. Edge Filers eliminate the need to deploy, configure and manage a local file server. Each Edge Filer synchronizes its file content to the cloud, adding the benefit of offsite cloud storage. The result is a perfect solution for remote site data protection and disaster recovery.

Drive is installed on user workstations, enabling file sync & share (FSS), and endpoint backup. When installed on mobile devices (iOS, Android), they offer simple access to centrally managed files, and FSS services. The built-in integration with Microsoft Office365 allow users to create a file on their desktop, modify it using their smartphone, and then access it via a browser. This enables true anytime, anywhere, any device workflow.

See next page for a more detailed architecture.

HCP Anywhere Enterprise Reference Architecture

LEGEND

→ INBOUND PORTS
→ OUTBOUND PORTS
→ INTERNAL PORTS (NO INTERNET ACCESS)

NS or A Records TCP/UDP 53

Portal
Global File System

Name Service
DNS Resolution

TCP 636, 3269 (LDAPS) or TCP 389, 3268 (LDAP)
KERBEROS TCP/UDP 88

TCP 22, 3000-9999

APPLICATION SERVER
Broker, Web App.

METADATA DATABASE
File System Catalog

APPLICATION SERVER
Broker, Web App.

REPLICATION DATABASE
File System Backup

APPLICATION SERVER
Broker, Web App.

DOCUMENT PREVIEW
Online Viewer

TCP 18682

EDGE
SMB/NFS Caching Filer
TCP 995, 8443

DRIVE
Share / Protect
TCP 995

MOBILE
Drive Connect / Mobile
TCP 443

WEB BROWSER
Edge Filer
Management Console
TCP 443

TCP 25 **Email Notifications**
SMTP Service

TCP 123 **NTP**
Time Service

Identity Service
Active Directory, LDAP

TCP 443 **Office Online Server**
Online Editing

TCP 5672 **VARONIS**
DatAudit, DatAdvantage

TCP 1344 **ICAP Server**
Malware Scanning

UDP 514 **Syslog Server**
Log Mgmt. System

TCP 5696 **THALES Cipher Trust**
KMIP (Key Mgmt.)

TCP 80 **OCSP Server**
Certificate Revocation Svc.

TCP 80, 443

TCP 111, 2049

TCP 443

Direct Mode

HCP Anywhere Enterprise

NFS Storage
Network Attached Storage

Aside from offering "all-in-one" global file services, the Enterprise File Services Platform architecture has several unique attributes:

- **Military-grade security:** A private and secure architecture powered by end-to-end encryption, advanced authentication, anti-virus, and behind-the-firewall deployment.

- **Global deduplication:** most modern storage solutions apply deduplication only to centrally stored files. Hitachi has taken deduplication to the next level, applying the algorithms at both cloud and edge. Not only does the Portal support global deduplication, but Edge Filers and Drive clients offer source-based deduplication, greatly reducing the size of files being sent to the cloud and lowering storage costs substantially.

- **WAN optimization:** to overcome bandwidth and latency limitations, a slew of optimization techniques are used in order to reduce file sizes and transfer times to/from any access point.

- **Intelligent caching:** every Edge Filer and Drive application comes with a built-in file cache. Caching accelerates remote access, plus it enables access points to "view" the full file storage space, and have on-demand access to every available file.

- **Cloud neutrality:** The platform is cloud agnostic. It works with multiple cloud providers, so customers can choose any cloud service, combine multiple services, or migrate between them. The platform offers full flexibility with no vendor lock-in.

- **Automation and orchestration:** managing a large global file system, with thousands of access points and tens of thousands of users, can be quite challenging. To simplify the process and support scale, the Portal comes with advanced management tools, including template-based automation.

- **Central monitoring:** The Portal comes with rich activity dashboards and analytics, allowing administrators to observe, monitor and troubleshoot every aspect of their global file system.

- **Multi-tenancy:** file services can be easily partitioned into multiple tenants - each with dedicated storage and service settings. Service providers with multiple external customers, or IT organizations that serve different departments (each viewed as internal customer), can provision and manage file services per each "customer."

- **Horizontal Scalability:** the platform easily scales to tens of thousands of users and sites, and is actively used in extremely large deployments.

- **IT-as-a-Service:** Hitachi offers capabilities such as user quotas, chargebacks and billing APIs that support the financial procedures of both service providers, and enterprises that implement cross-departmental IT charges.

A "Security-First" Approach

Files may contain sensitive enterprise data, such as intellectual property, know-how, or customer information. As such, they must be protected from theft, leakage and loss — globally. Founded by security veterans, Hitachi has adopted a "security first" approach for its global file system platform.

Here are some of the highlights:

- **100% In-Firewall installation option:** The Portal, the heart of its global file services, can be installed either on-premises, or at a virtual private cloud (VPC), allowing customers to maintain full control over all data, security, and access rights.

- **Source-based encryption:** files are encrypted by the global file system right at their source. Every file is fully encrypted — at rest — wherever it is stored (AES 256), or when transmitted — in transit — across the network (TLS/SSL)

- **Private Key Management:** encryption keys are created and managed by customers, and never leave their control.

- **Strong Authentication:** Hitachi supports the most advanced user authentication and authorization schemes, including: Active Directory (AD), two-factor authentication, and smartcards/ military common access cards (CAC).

- **Full ACL support:** file access rights have been traditionally enforced by access control lists (ACL). Hitachi handles the migration and support of previously defined Windows NT ACLs, and continues to enforce them at both the filer and browser levels.

- **File sharing policies:** Hitachi supports fine-grained policies for file sharing, based on file names, users or user groups.

- **Centralized anti-virus (AV):** the Portal can be configured to work with third party anti-virus utilities, allowing easy quarantining of infected files.

- **Remote Wipe:** the platform supports global file access, including via mobile devices. Drive can be instructed to delete file content in the event of a user losing his phone or laptop.

- **Compliance:** HCP Anywhere Enterprise is designed to meet the requirements of a broad range of data privacy and data security regulations (e.g. FIPS 140-2, HIPAA, GDPR and more).
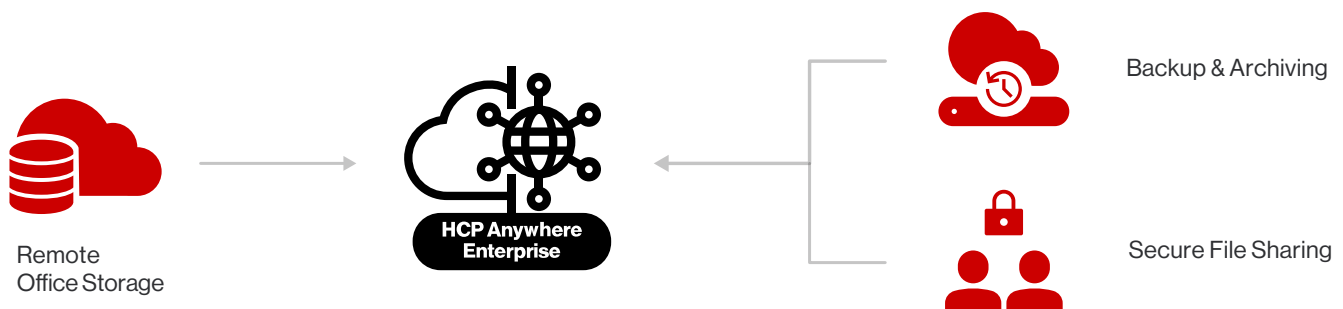
"Caching Everywhere"

HCP Anywhere Enterprise has fully embraced caching technology and applies it throughout its GFS architecture. In particular, both the Edge Filers and Drive applications have built-in caching technology. Whether you work at a remote site, use your laptop on the go, or access files via your smartphone – caching comes to your aid.

The first obvious benefit of caching is reduced response time. Retrieving a file from the cloud often involves pulling its content over a long network link. Bandwidth and latency limitations may cause 'read' operations to be quite lengthy. Keeping a "local copy" of frequently accessed files eliminates the need to retrieve the same data over and over again, thus accelerating 'read' operations. Caching also streamlines 'write' operations – clients can quickly update the local cache, while letting the system handle in the background the lengthier process of uploading the new data to the cloud.

HCP Anywhere Enterprise effectively handles file consistency: it ensures that whenever a file is read by a client, its latest version is indeed served. And when a client updates a file locally, its cloud source is quickly updated. In the event of simultaneous updates by multiple clients, the platform automatically creates file versions, allowing users to reconcile the changes if needed.

Caching has another important benefit: it allows clients to access virtually unlimited file storage space. Edge Filers and Drive applications provide full access to directory and file attributes information (aka 'metadata'). The amount of metadata is relatively small in size, even for extremely large file repositories. Metadata enables users to easily retrieve any file visible to them – based on permissions. The local cache stores only the most frequently accessed files, or files "pinned" by users, thus enabling access to extremely large file repositories while requiring a modest amount of local storage. This capability relieves administrators from the need to regularly upgrade storage capacity at remote sites or on users' workstations. The cloud repository size may grow exponentially over time, but the amount of storage used by the cache remains fairly constant.

Suppose you invested in deploying a global file system. How would your life and your users' lives change? To answer this question, let's examine a few use cases for the platform:

Remote Office Storage → HCP Anywhere Enterprise ← Backup & Archiving, Secure File Sharing

- **Remote-Office Filers:** in a globally distributed enterprise, a large number of users work at remote offices. Traditionally, addressing their file storage needs meant deploying a file server or a NAS at each office. Those devices had to be continuously maintained, and regularly upgraded. To further complicate matters, all the data stored on them had to be backed-up and hauled offsite. With the solution, simply deploy an Edge Filer at a remote site. Users will have access to a familiar looking NAS, but all data will be automatically synced to the cloud. No need to worry about complicated backup processes or specialized DR plans.

- **Cross-Site Collaboration:** sharing files between users has traditionally been a complex/inefficient process, a security challenge, or both. Setting up a file server with ever changing access rights is a pain. Exchanging files through emails is inefficient and leakage prone, and using external file sharing services comes with significant security risks.

  Hitachi offers an intuitive and secure file sync & share (FSS) service. The service is built into its Drive apps and Edge Filers. It is therefore available to any user, at any site, using any device. Simple, fine-grained sharing controls guarantee information security, without sacrificing ease of use.

- **Home Folders:** supporting users' personal home folders has been a nightmare for administrators. Each home folder has to be backed-up, a process that often failed to complete, and the ever-increasing storage use forced upgrades to users' workstations. And following the Bring Your Own Device (BYOD) movement, users now demand access to their home folders everywhere.

  Hitachi lets users maintain their personal home folders from anywhere while they are automatically protected in the cloud. Users can seamlessly access their files via a Drive app installed on their workstation or mobile device, or via their office Edge Filer. The FSS service offers access to personal files anywhere, and the ability to share them with others when needed. Gone are all the complexities associated with protecting and managing users' home folders!

- **Caching:** Leverage a Edge Filer or Drive to ingest local file data and tier it to lower-cost object or cloud storage. The platform also supports file data archiving, providing a local view of cloud archival data (via stubbing) that can be accessed whenever needed.

- **Endpoint Backup:** With so much sensitive data residing on users' workstations and personal devices, it's no wonder administrators are worried. Multiple attempts to deploy and operate backup solutions on users' endpoints have proven futile. Backup cycles often fail to complete, and large amounts of data remain unprotected.

  Hitachi's global file system helps solve the endpoint backup challenge in several ways. First, existing endpoint backup utilities can point to a Edge Filer as their 'storage target.' The Edge Filer automatically uploads the backup data to the cloud, where it is safely stored. Second, endpoint backups can be sent over the WAN to the cloud for a direct-to-cloud, or disk-to-cloud, option that removes the hybrid appliance from the equation. Or, in a different manner, a user can take advantage of the GFS and directly place their home folders and shared folders in the cloud. With this approach, there is no longer a need to explicitly backup the endpoints. Those have become "access points" to the master copy of the data that resides in the cloud. Changes made to files by users on an endpoint device, are quickly synchronized to the cloud, where they are stored and protected.

- **Content Distribution:** Enterprises often need to make file-based content available to users globally. Examples include project documents, enterprise information broadcasting, marketing materials, etc. Replicating this content to every remote site is a time and resource consuming process.

  With the global file system, all that's needed is to place the desired content in a shared folder. Users throughout the distributed enterprise immediately "see" the new content via their Drive app or Edge Filer. There's even no need to "push" the content to remote locations ahead of time. The built-in syncing and caching mechanism can easily handle on-demand access to new content.

- **Ransom Protect:** Hitachi's global file system features Ransom Protect, a state-of-the-art AI-driven ransomware defense mechanism that identifies and halts ransomware attacks in real time. Furthermore, should a ransomware attack occur, organizations have the ability to virtually instantly recover the 'locked' data from immutable snapshots. Hitachi Ransom Protect employs advanced machine learning algorithms to swiftly identify and block suspicious file activities. With an incident management dashboard, administrators can monitor attacks in real time. The system securely stores extensive incident evidence and logs, aiding in post-attack forensics. Integrated into the Hitachi Edge Filer, Ransom Protect is easily activated with a single click for streamlined deployment.

- **Vault:** Hitachi Vault aids enterprises in guaranteeing the preservation and tamper-proofing of their data, while also ensuring compliance with data regulations through strict enforcement. Hitachi Vault provides the flexibility and granularity to create Write Once Read Many (WORM) Cloud Folders with customized retention modes for specific periods of time to fit any organization's regulatory or compliance requirements. The Hitachi Portal provides centralized control and management of the policies that are enforced at every remote Hitachi Edge Filer.

Summary

Back to Jeff, our desperate director of sales who needs his VP to approve the proposal, and Tim, the IT person committed to help him. Things have been quite different since the platform was installed.

"Calm down Jeff, I've got your back,"
you say, and log into the Portal.

You check out Jeff's home folders, and it takes just a few seconds for you to find the one named "Sales Proposals." With a few mouse clicks you share the folder with the VP. *"I've got good news and bad news,"* you tease Jeff. *"Start with the good news,"* he begs. *"The good news is that I have just shared your proposal folder with the VP."* You explain that not only the proposal, but the full sales directory is now visible to the VP in Europe. Thanks to the Edge Filer you installed there, he should have fast access to anything he needs.
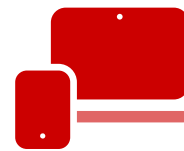
"So what's the bad news?" Jeff asks. You pause and chuckle: *"The bad news is that you could have easily done this yourself, if you just bothered to read the update I sent about the new global file system. So you now owe me a beer for my lost dinner time!"* Jeff can't hold his excitement. *"This is fantastic news! I've been on the phone with my boss the past hour, wracking our brains how to do this – and you just saved our lives. Thanks bro, I owe you more than just a beer!"* You wish Jeff good luck and return to the dinner table.

Global File System

HCP Anywhere Enterprise

Edge
Caching Filer

Drive
Endpoint Agent

VDI
Virtual Workspaces

This is but a simple example of how the global file system can impact your enterprise. Files are indeed the lifeblood of modern enterprises, and Hitachi helps manage, protect and share them effectively and securely anywhere.

Hitachi Vantara
A Hitachi Group Company