MULTIARCHIVE MANAGEMENT

Gain efficient data collection with HPE DRAGON
Information is increasingly important and needs to be captured from various sources using different types of storage. It's crucial to find a solution that treats these different archives as if they are on a single system—for heavy-duty data management and datastore.

Get a handle on your massive data—in an efficient, secure, and cost-effective manner—with several state-of-the-art solutions for data archiving and management, to transfer it into useful information.

Collect and consolidate data from different sources to analyze and generate required information using HPE Data Retention and Guardian Online (DRAGON) solution. This multiarchive collection of solutions treats all of your data as if it was on a single system. Central to the HPE Investigation solutions, DRAGON is enhanced to accelerate query speed and performance across multiple databases and improve system security. Other improvements include full integration of different database technologies from the standard RDBMS, to columnar and to the Hadoop ones.

TABLE 1. DRAGON features

Following table summarizes product features for each of the DRAGON software layer component, while a more detailed description is provided in the next sections.

1. DRAGON manager: Secured web user interface

   DRAGON provides its own Warranty/Request management solution featuring:
   - A single, unified front-end for managing request fulfillment for lawful intercept and retrieval of retained data
   - Increased workflows reduce the effort and timeframes for processing requests
   - Flexible deployment models enable localized deployment at communications service providers, or remote deployment for self-service operations at law enforcement and intelligence agencies
   - Document management capabilities support warrant and request documentation
   - Different integration layers enable external provisioning mechanism in compliance with different international bodies like ETSI RDHI and country specific

2. DRAGON server: Multiarchive Database Management

   DRAGON enables the possibility to query heterogeneous archives, ranging from fast relational databases to big column-based ones, and grid systems to simple XML-based or file system-based archives for data capture. Its federated server architecture is based on the concept of data source. And the federated database is a collection of databases that are treated as one entity and viewed through a single user interface offered by DRAGON Manager WebGUI. DRAGON software offers its federated query engine, which enables parallel processing of queries that involve data from several different archives at the same time with the ability to perform aggregated functions on it. DRAGON can also use data from specified database technologies including HPE Ezmeral Data Fabric, Oracle®, MySQL, Teradata, and Micro Focus Vertica. Other key functionalities offered:
   - Processing pipelines enable the same request to be executed using different processing chains in order to fulfill customer’s needs
   - Cross archive search has the possibility to spread a query over different archives and merge the results together
   - Flexible query capabilities can exploit native archive’s features for maximum speed
   - Large queries are decomposed into sub queries with parallel processing and priority management over multiple nodes and search types
   - Plug-in architecture covers many aspects of the processing chain and can be altered through dedicated plug-ins
   - Cluster support enables adding more servers to distribute query workload, while providing a fault tolerant mechanism that moves queries from faulty nodes to working ones.
3. DRAGON server: Data loading

The DRAGON Loader is a general-purpose graph computation engine that allows to process data as it moves from one source to several sinks. Each computation is called a process and is defined by a topology, where several building blocks (sources, processors, and sinks) are linked together.

The Server’s plug-in architecture is leveraged to provide several source types (local and remote files, jdbcdatabases, etc.), processor types (normalization, projection, encryption, etc.) and sink types (local and remote files, jdbcdatabases, etc.).

The transformations provided by the Loader are not SQL based but plug-in based, and this allows to implement hard tasks like metadata extraction from unstructured data and analytical model creation.

Besides the ingestion phase, the Loader also takes care of the data lifecycle, automatically purging data that goes out of the retention period, and data optimization, compacting small daily files into larger ones in order to speedup queries.

In addition, the implemented backward data lineage allows to track data as it moves through topology’s stages. This simplifies tracing errors back to their sources.

4. Security

HPE DRAGON’s security and privacy is woven throughout the system to enable auditable user access with enhanced security. This tight security enables information to be used as legal evidence and provides privacy to citizens’ communications activities. Core features include:

- Role-based access controls provide strong security with ease of management.
- Robust security and encryption protect data privacy and nonrepudiation.
- User action logging and audit controls validate security.
- Integration is optional with external security authentication systems.
- Integration is optional with external security systems for handover (PGP/GPG).
- Secured communication protocols with mutual authentication enforcements are available for intra and extra modules.

5a. Data exposure layer capabilities: REST API

This layer provides a REST interface to clients that want to access data archives. A set of built-in services provides search capabilities and system maintenance functionalities; custom services can be added on demand to fulfill special requirements.

A command layer takes care of exposing these services to clients, enforcing the proper authentication and authorization policies. Every service access is tracked by the audit component, which can write the log to a disk or different system and is analyzed by a statistics component that updates some usage counters.

Clients can easily connect to DRAGON using provided SDK libraries. This simplifies access to built-in services and new custom ones.

5b. Data exposure layer capabilities: XQuery

DRAGON real-time XQuery to SQL engine enables telecom operators to extract and expose customer information from multiple Big Data archives in a standard, efficient, and multitenant way. This layer lets external partners query DRAGON or SPS archives using a subset of the XQuery language (specifically, XQuery Core). This subset includes FLWR expressions (for, let, where, return), if then else conditions, axis navigation (self, child, attribute), set operations (union) and relational and arithmetic and logical expressions.

Operators can define general views of their models and specialize them for each partner—adding removing or changing information. The engine also supports per partner access control lists (ACL), which are lists of XML paths—in the resulting XML—that a partner can or cannot see. Due to the engine’s design, ACL can be applied without impacting SQL query performance. The ACL application prunes the resulting XML structure causing a possible reduction on the number of SQL queries executed and a possible performance improvement. After the XQuery to SQL transformation is complete, the calculated SQL queries are executed in parallel, and their results are combined together in real time. The resulting XML is built on the fly from this tuple stream and returned as a data stream to the partner.

6. Data Lake integration

Extend the DRAGON product marketing offering, with its own Communication Data Records Data Lake from raw data loading to the full DRAGON data lifecycle management along its embedded security and compliance capabilities. Embracing HPE Ezmeral Data Fabric is one of the preferred choice for implementing a DRAGON Data Lake, leveraging all those product capabilities like data tiering and data replica from the edge to the cloud.

Of course, also the main Hadoop distributions are supported by DRAGON including Cloudera and Hortonworks and the latest Cloudera Data Platform. In this context DRAGON offers a completely new DRAGON loader tailored specifically for it, achieving extraordinary performance benchmarks.
COLLECT, MANAGE, AND ARCHIVE DATA

DRAGON manages the end-to-end lifecycle of data through collection, management, and archiving processes. See Figure 1, which outlines these steps.

COLLECT DATA

Collecting the necessary data in a valuable format and understanding the information sitting deeply inside data flows are the winning strategy. And it can be powered by HPE DRAGON IP probe Deep Packet Inspection (DPI) solution. DRAGON DPI enables monitoring, extracting, and filtering of data directly from real-time networks, offering passive probes that generate no overhead on network devices, and leave no impact on network operations. Knowing every company has different needs, our main focus is to customize DRAGON DPI to monitor exactly the same level of data flow insight you need to collect exactly the data you require.

The data collection layer could be extended with the adoption of the HPE eIUM mediation software. Such option will provide file collection, file conversion from binary to ASCII and correlation capabilities to prepare the data before to be moved to DRAGON Loader.

HPE enhanced Interactive Unified Mediation (eIUM) is a telecommunication platform designed and built to help telecommunication service providers in their digital transformation journey. Even though microservices paradigm was not around at the time when HPE eIUM was born, it was specifically architected to build lightweight applications, which can process big volumes of data coming from network core, running in a distributed computing environment.

Following some key product differentiators introducing this data collection layer within DRAGON architecture:

• Leverages HPE eIUM strength in telco mediation
• High performance, modular architecture is very scalable for multiple data flows and high xDR data rates
• Incoming data is cleansed, correlated and normalized prior to loading into the data repository which optimizes the data stored
• Correlate data with external sources such as CRM, directory servers, etc.
• Templates for common data flows are included to reduce implementation time and cost
• Configuration flexibility to ease the introduction of new data sources
• Providing a distributed platform for massive parallel processing with high performance and low latency
• No coding, configuration-only without sacrificing processing performance; easy to introduce new data formats and protocols with configuration-only
• 50+ protocols available out of the box
• 60+ rules to quickly support new use case, new workflow, and build new API
MANAGE DATA

Data retention or archiving are not just about having data storage and storing data. There are a number of important aspects to be considered—authentication, accessibility, performance, management, and maintenance of databases. As companies use various types of databases to store data, it becomes imperative to synchronize the data among databases and manage them as a single point of management. All these requirements are addressed by DRAGON—the right solution for Big Data management.

ARCHIVE DATA

Every day, companies must store and manage a large amount of information used for everything from marketing, billing, capacity planning, legal, and compliance. In fact, it's one of the biggest corporate topics, making data management no longer an IT problem but a corporate strategy consideration.

Moreover, different data needs different archival, retrieval, and security capabilities. They can be heterogeneous—a mix of text and multimedia, for instance. Additionally, some of it needs strong security and encryption.

With DRAGON, you use the right archival technology per data type to reduce system size and improve data archival efficiency while realizing a cost-saving management strategy. DRAGON also implements federated database management to manage various types of databases as single points of management screen (DRAGON GUI) and creates relations between different database types.

To survive the data explosion means governing all the different aspects:

• Reduce available storage dramatically
• Use the right archival technology for each data type
• Ease system and data center usability
• Require data reliability on a wide range of devices used in different settings
• Ensure compliance and secure data management

With DRAGON, you use the right archival technology per data type to reduce system size and improve data archival efficiency while realizing a cost-saving management strategy. DRAGON also implements federated database management to manage various types of databases as single points of management screen (DRAGON GUI) and creates relations between different database types.

FIGURE 4. HPE DRAGON Manager user interface

FIGURE 3. HPE DRAGON Server (Query) component
COMBINE DIFFERENT DATABASE TECHNOLOGIES

It is possible to combine different databases with the DRAGON multiarchive capability. For example, DRAGON Large Event Archive is used for transactional raw data, including call data records (CDRs), more storage, and other databases like HPE Ezmeral suite for analysis purposes.

- DRAGON is the choice for multipetabyte deployments
- Based on HPE Reference architecture (both HPE Apollo and HPE ProLiant DL based)
- Data tiering allows the movement of data between different storage tiers, which allows an organization to ensure that the appropriate data resides on the appropriate storage technology based on different factors such as frequency of access, security needs and cost considerations.
- Different storage formats are adopted by DRAGON: using parquet format leveraging the target storage systems integrated by DRAGON from HDFS to S3 and to CloudStorage offering a good compression factor (typically above 4x)
- Data is queried by DRAGON using different engines like Drill, Impala, Athena and others, designed for high-performance, agile processing of information at massive scale.
- DRAGON query engine ensures that performance is freed up to meet their SLAs for the data that needs the fastest read response times. As a result, end users and apps can experience unmatched performance.
- For the Hadoop datalake, the map/reduce framework is not used by DRAGON and the Hive and HBase components are not required, as well, it implies a better disks utilization for the benefit of HDFS storage capacity.
INFORMATION ARCHIVE SOLUTION WITH HPE DRAGON EDGE TO CLOUD, ON-PREMISES, HYBRID

With our HPE DRAGON product, the information archive solution could be deployed in a Data Center (from central, to regional and to the edge) and move towards a private (e.g., HPE GreenLake) or public cloud platform (e.g., Google Cloud Platform™, Microsoft Azure or AWS-Amazon Web Services). Leveraging the right archive technologies that HPE DRAGON offers, it could be chosen with the proper and required technologies without any vendor lock-in in each implementation layer.

The following are some deployment examples supported by HPE DRAGON:

- On-cloud information archive: The cloud deployment solution depicted in Figure 8 is to deploy HPE DRAGON on Amazon Cloud infrastructure and use ‘Amazon AWS S3’ as information archive and ‘Amazon Athena’ for analyzing and searching the data.
  - The value-added to the HPE customers is an easy scale-in and scale-out of infrastructure based on data traffic, a significant drop in deployment cost, and minimal CAPEX and OPEX cost based on usage.
  - The complete solution deployment is automated which drastically reduces the deployment time and also go live time.

---

HPE Financial Services: Creating investment capacity to accelerate digital transformation.

HPE Financial Services helps organizations create the investment capacity they need for digital transformation, in an innovative and sustainable way. HPEFS partners with customers to develop a playbook for their entire IT asset portfolio (from edge to cloud to end user), one that is unique to their aspirations and size. Our financial and asset management solutions are anchored by best-in-class tech upcycling services. For more information, visit: hpe.com/us/en/services/financial-services.html.

---

FIGURE 7. DRAGON on Hadoop ecosystem in a real customer environment with a storage capacity of more than 7 PB of raw data stored in two different sites

FIGURE 8. HPE DRAGON on AWS Cloud Information Archive Solution
• Hybrid information archive: Information archives on-premises vendors are increasingly losing ground to cloud-based, opensource solutions leveraging hybrid implementation approaches.
  – The hybrid deployment depicted in Figure 9 provides advantages provided by both on-premises and cloud storage as information archives.
  – This solution is to use "Amazon AWS S3" / "Google Cloud Storage™" as information archive on cloud and Data Fabric as on-premises Information Archive. "Amazon Athena" / "Google™ Big Query" / Drill for analyzing and searching the Archives. The value to HPE customers is a significant drop in deployment cost, minimal CAPEX and OPEX cost, based on usage. This solution aims at deploying HPE DRAGON on-premises.

> FIGURE 9. HPE DRAGON Hybrid Information Archive Solution

• On-premises information archive: The solution is to deploy HPE DRAGON and HPE Ezmeral Data Fabric as information archive on-premises. Loader ingest data to Data Fabric using HDFS or NFS interface in parquet storage format. Query uses Drill as query engine to search data from the Data Fabric. The value to the HPE customer is faster, more-dependable data management along with lower TCO, complete data protection and no single points of failure.
GET THESE BENEFITS

HPE DRAGON delivers multidimensional scalability of performance, capacity, and functionality required to fulfill communications data requests. Flexible DRAGON supports many business models including mobile virtual network operator (MVNO), hosted multiple operating units, multiple geographies, and others.

The proposal for the above challenges is to implement an information archive solution with HPE DRAGON. HPE DRAGON enables the adoption of different archive technologies, supporting several implementation options. The key characteristics of the HPE DRAGON solution are:

• DRAGON is ideal for clustered, scalable, hybrid deployments leveraging both HPE Ezmeral Data Fabric and AWS
• As compared with the competition, DRAGON provides more features, is more secure, performant and standards compliant, and is easier to install, manage, and operate than ever.
• DRAGON implements unified agnostic database management supporting different archive technologies, with ease of scale
• Telco & security compliant with data retention regulations

DRAGON significantly reduces the operational cost and TCO and enables information archive corporate strategies.

HPE DRAGON is the secured application for loading and protecting the information archive, leveraging different archive technologies from Hadoop, HPE Ezmeral Data Fabric, AWS S3 and Google Cloud Storage.

HPE DRAGON handles massive data in an efficient, secure, cost-effective manner with several state-of-the-art solutions for data archive and management, ability to transfer data into useful information.

Main benefits of DRAGON include:

• A complete solution for integrated data collection, data retention, and archiving—offering highly scalable and flexible solutions to ease access to critical operations data
• Scalable solution for data archiving with the provision of several types of database, like HPE Ezmeral Data Fabric, Oracle, MySQL, RainStor, Micro Focus Vertica, and Hadoop as per data usage requirements to improve TCO
• Federated database management to treat databases as one entity and view it through a single user interface offered by DRAGON GUI
• Multiple complex integrated solutions
• Existing compliance environment simplified
• Industry standards for configuration, integration, and data exchange (ETSI, XML)
• Support for new IP services and network
• Compliance costs reduced
• New capabilities

Benefits through integration:

• Supports all of common integration standards and technologies including ETSI, CALEA, 3GPP, web services, and XML
• Reduces operating cost through a common integration approach for external and internal systems

Security benefits:

• Comprehensive data security features, including encryption and certificates that protect privacy and integrity
• Robust user security mechanisms, such as user access controls, auditing and reporting, and optional integration to external authentication systems
• Confidence in communications through secure messaging channels
SEE HPE IN THE TELECOMMUNICATIONS INDUSTRY

To grow in a fast-changing market, communications service providers (CSPs) must meet the huge demand for new services, and streamline internal operations. Hewlett Packard Enterprise (HPE) helps the world’s CSPs transform the way they do business. We are unmatched in our ability to help them drive transformation—more than 30 years of telecom experience, global IT leadership, expertise in entertainment, and leadership in consumer devices.

HPE offers a truly end-to-end portfolio from core network to handheld devices, including a suite of convergent IT and telecom solutions, such as software (OSS/BSS), service delivery platform (SDP), subscriber data management (SDM), and cloud. Our professional services include business consulting, integration, and managed services/outsourcing, joint go-to-market programs, carrier-grade servers, storage, printers, tablets, and smartphones.

COMMUNICATIONS AND MEDIA SOLUTIONS, HEWLETT PACKARD ENTERPRISE

HPE Communications and Media Solutions is dedicated to creating vertical solutions for the communications and media industry. With over 30 years of experience in the industry, we have over 50 solutions, over 1500 active contracts, and more than 300 telco customers in 160 countries. We provide software and services that enable your digital transformation, automate your operations, and help you grow your business with innovative cloud-native network solutions and digital, 5G-ready services.

ABOUT HEWLETT PACKARD ENTERPRISE

Hewlett Packard Enterprise is the edge-to-cloud company that helps organizations accelerate outcomes by unlocking value from all of their data, everywhere. Built on decades of reimagining the future and innovating to advance the way people live and work, HPE delivers unique, open, and intelligent technology solutions, with a consistent experience across all clouds and edges, to help customers develop new business models, engage in new ways, and increase operational performance.

LEARN MORE AT
hpe.com/telco

© Copyright 2022 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein. Google, Google Cloud Platform, and Google Cloud Storage are registered trademarks of Google LLC. Azure and Microsoft are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. Oracle is a registered trademark of Oracle and/or its affiliates. All third-party marks are property of their respective owners.

4AA4-4801ENW, Rev. 4