Implementation of cloud-ready technology designed to address longstanding data interoperability and accessibility issues through the use of a comprehensive semi-automated seismic interpretation workflow

Marion Barré, Stuart Walley, Vianney Savajol, Eliis; Jie Deng, Christine Rhodes, AWS; Aurelien van Welden, Paul Endresen, Marius Storm-Olsen, Andy James, Bluware*

Summary

Energy companies maintain numerous subsurface applications with local data stores. This can lead to upstream workflow inefficiency, data silos, and unnecessary cost. As the E&P industry continues its digital transformation journey, step-change technologies are needed to improve workflows efficiency and reduce costs, particularly when migrating IT infrastructure to the cloud where data interoperability becomes a critical consideration. The purpose of this study is to test an alternative, cloud enabled technology that facilitates the access and management of subsurface data by petro-technical software, enabling a data-centric approach. By embedding an adaptive streaming technology into a semi-automated interpretation platform, data can seamlessly be accessed from a shared, cloud-based standard data repository. This innovative approach avoids the need for traditional time-consuming import and internal conversion steps. And using a standard seismic data format, data management and geoscientist teams can dramatically decrease the physical size of their projects and reduce data duplication. This technology improves upstream workflow efficiency and has a direct cost reduction impact for the benefit of operating companies and the E&P industry.

Introduction

Seismic reflection data enables exploration and development teams to assess reservoir storage potential, monitor reservoir fluid movements, appraise risks, and optimize well placement; all key exploration and reservoir management practices. Increasing volumes of pre-stack and post-stack seismic data are needed for processing and evaluation to support critical decision making.

Energy companies use specialized geophysical software applications for seismic processing, interpretation, inversion, and basin analysis. Typically, these applications use internal data formats in addition to the original SEG-Y format, resulting in data duplication. Furthermore, many companies differentiate themselves by developing custom workflows, combining advanced technologies with internal knowledge adapted to specific geological challenges and operational needs. In the seismic interpretation domain, we are witnessing increased use of platforms that offer automation and machine learning. Consequently, energy companies are burdened with the spiraling costs of petabytes of difficult to manage, duplicated subsurface data, siloed in proprietary formats that lack data interoperability, leading to additional storage, transfer and conversion costs. To meet these challenges, we need data-driven and intelligent technologies across the asset life cycle to improve legacy workflows and data management systems.

In response, leading organizations are participating in an industry wide initiative to address such challenges. The Open Group Open Subsurface Data Universe™ (OSDU) Data Platform has been developed to unlock the full potential of upstream data by standardizing how data types are formatted, organized, stored, and accessed. Combining this platform with compatible interpretive applications, makes it possible for technical specialists to screen large volumes of raw and interpreted data rapidly, further enhanced using cloud technologies. To optimize the accessibility and storage of seismic data, the OSDU Forum has recognized Volume Data Store™ (VDS), through the open-source OpenVDS API, as one of its standardized seismic data formats.

The following paper outlines an ongoing effort to advance E&P upstream workflows, by testing a cloud-based solution to ensure seismic interpretation platforms, can seamlessly access seismic data stored on the OSDU Data Platform. The approach described utilizes adaptive streaming technology, designed to mitigate storage and legacy latency issues.

Data Interoperability

The authors first tested using the OSDU Data Platform as a data repository to prove the interoperability of seismic data exchange with external data sources. Since its formation in 2018, members have contributed domain and technology expertise to iterate and refine the platform's design and functionality. Many improvements have focused on industry data interoperability, making the data platform a promising solution to securely store, catalogue, and exchange subsurface data.

The chosen workflow (Figure 1) tested the OSDU Data Platform's data exchange, where seismic data was stored and retrieved using the OpenVDS API. VDS files from external data sources were ingested following OSDU's canonical data definition. These files were grouped into logical seismic datasets and each group used standardized attributes to describe relevant properties, such as processing projects and seismic domain types. Once ingested, the VDS files were discovered by running high-performance text or range queries against those properties. The OSDU's canonical data model allowed applications, or organizations, when granted proper access, to consume the data easily and securely by following the platform's published data standards and retrieval methods.

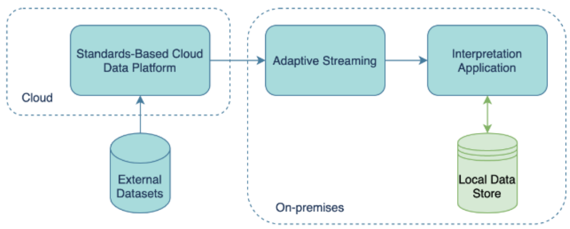

Figure 1: Architecture diagram representing how an on-premises application seamlessly accesses data supplied by external data sources through the use of a cloud data platform and adaptive streaming technology.

The OSDU Data Platform offers additional features to solve the challenges of past data exchange methods including implementation of legal and entitlement services to govern the right of data use, and data lineage preservation by tracking data change history as part of a centrally managed data repository. This workflow also demonstrated how organizations, moving to a cloud-based OSDU instances can extend desktop applications to leverage cost-efficient and highly scalable object storage, such as Amazon Simple Storage Service (S3), to store raw seismic files.

However, legacy geoscience applications that use normal file systems cannot directly integrate with cloud-based object stores. A lift-and-shift strategy, in which companies deploy traditional HPC file systems in the cloud, is expensive and does not easily accelerate innovation using machine learning tools. This is why adaptive streaming technology is critical to the success of this workflow.

Adaptive streaming

Cloud-native formats enable on-demand data delivery from a central source into applications. By spreading data into thousands of objects, access speeds are drastically accelerated and enable data streaming and data compression. Streaming data from the cloud is an established technology that is used by video streaming services. The same approach has been applied to seismic data using compression with adaptive streaming to deliver data directly into geophysical applications.

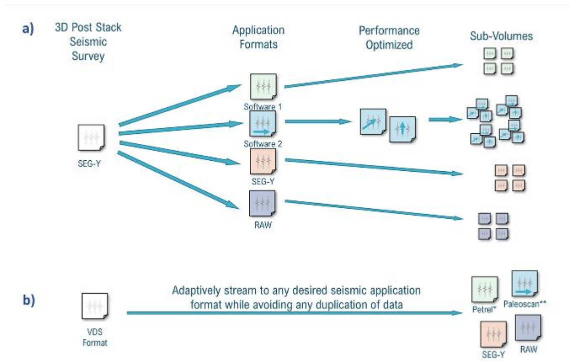

Figure 2: a) Traditional method of using seismic data across multiple applications whereas method b) highlights how an adaptive streaming technology eliminates the duplicative datasets.

Contributed to the OSDU Forum in 2020, OpenVDS is an open-source API for fast, random access to seismic data in VDS format, to provide the energy industry with a modern and cost-effective approach to efficiently store and retrieve seismic data in the cloud. Data stored in VDS can be converted back to any format (Figure 2B). The adaptive streaming capability allows the data signal quality to be catered for each workflow, minimizing data storage, eliminating data duplication, and optimizing usage in the cloud or on-premises.

Application of Adaptive Streaming to a semi-automated seismic interpretation platform

Conventionally, users initiate a seismic interpretation workflow by importing seismic data, either using a direct import application, or through data connectors that bridge applications. Often seismic volumes are converted from industry standard SEG-Y into an application native file format. As an example, in PaleoScan™ software, the custom format (PS) currently used, cannot be compressed, and must store the 'no_value' area, often increasing seismic volume stored internally, when compared to the original SEG-Y data. This is especially true when the seismic survey coverage does not have a typical cuboid shape. Once created, a PS volume is siloed in the software's project repository, away from its original SEG-Y archive. For performance optimization, PS volumes are typically stored close to the application, to ensure the maximum data I/O access speed and making its accessibility to other applications onerous and requiring data connectors. The main advantage of this native format is apparent early in the workflow, normally during data reconnaissance, as it allows the user to seamlessly cross-navigate through three different data dimensions (inline, crossline and horizontal slice).

This advanced seismic interpretation platform relies on a comprehensive semi-automated process (Pauget et al, 2009). A model-grid is computed by extracting a myriad of polarity-consistent micro horizons. To facilitate the auto-tracking of these elements using a cost-minimization function, the input data is pre-conditioned, using structural or 3D Gaussian smoothing, which generates a new stored copy. Thousands of horizons are then auto tracked, allowing the interpreters to focus on critical decisions relating to geological interpretation process. Once validated through QC, a Relative Geological Time (RGT) model is computed from the interpolation of the auto-tracked and user-refined horizons within the model-grid. During the interpretation, cross navigation allows the interpreter to access the data in three dimensions, which is crucial to assess the depositional environments whilst investigating exploration prospects or performing detailed reservoir analysis. From basin-scale evaluation to prospect maturation, this technology enables the user to drastically decrease the E&P cycle time, reduces subsurface uncertainty, allowing subsurface teams to make informed decisions regarding company investments. The model-grid algorithm, combined with the cross-navigation functionality, was successfully tested in various geological contexts (Gupta et al., 2008; Lemaire et al., 2010; Lacaze et al, 2011, Schmidt et al., 2013). However, this functionality requires the creation of additional volumes and becomes storage space intensive.

The move toward data-centric infrastructure, coupled with increasing need to reduce data storage is driving the integration of modern technology. This part of the experiment embedded adaptive streaming technology into this seismic interpretation platform to assess if the application could directly and efficiently access seismic data stored from the OSDU Data Platform.

To test cloud-based adaptive streaming and assess the benefits of utilizing object-based storage for seismic data in the data platform, the authors used a ‘real’ publicly available seismic dataset from the Exmouth Sub-Basin. The source SEG-Y seismic volume has a physical size of 3.4 Gb and covers an area of over 900 km². A comparison was performed between the import-based (Figure 3a) and OpenVDS adaptive streaming (Figure 3b) workflows, from data access, through pre-conditioning and the seismic interpretation process to the creation of the model-grid and the ultimate RGT model. The import-based method uses 41.1 Gb of storage due to the native format conversion and the creation of volumes in all directions. In contrast, the adaptive streaming technology, used from the start of an interpretation workflow, allowed the user to seamlessly access the initial raw data directly from a shared data repository which halved the physical size of final project at 23.4 Gb while preserving data quality and the cross-navigation functionalities.

Conclusion

This paper has shown how data can be accessed and delivered directly to an application from the OSDU Data Platform. The original seismic data, stored in VDS format either as a file system or object storage, can be retrieved from the cloud platform and delivered to a seismic interpretation software instantly without needing to locally import and

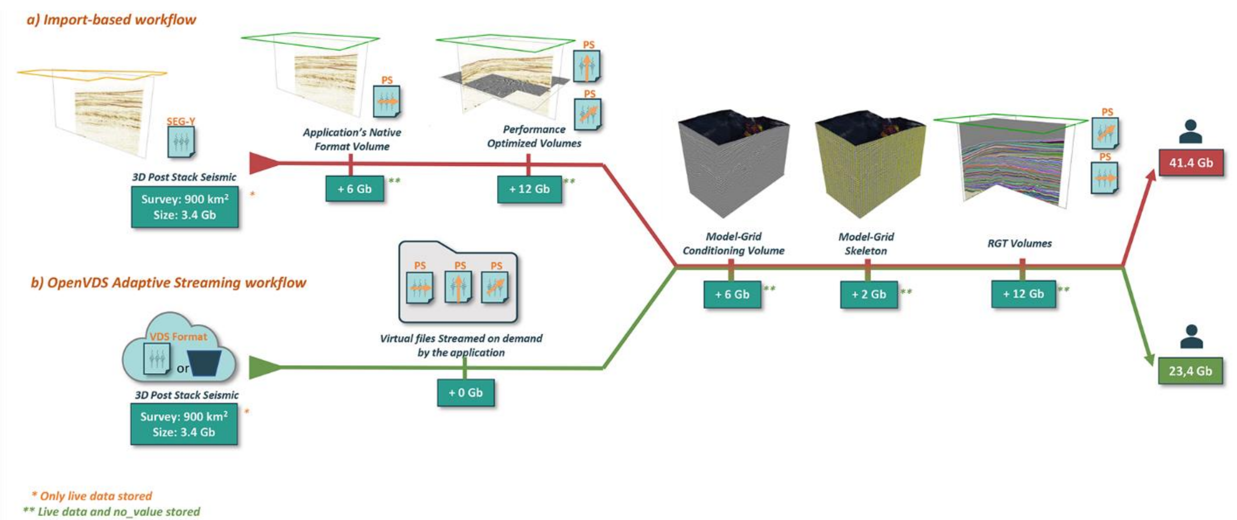

Figure 3: Comparison of two workflows from data access to comprehensive seismic interpretation approach. A) Import-based workflow: the initial SEG-Y seismic data is imported and converted into the application’s native format PS. B) VDS adaptive streaming workflow: the initial seismic data is stored on the cloud either as VDS file or VDS object storage.

convert the data. Streaming, rather than copying data, removes local storage needs on traditional file systems and eliminates additional storage costs. It also centralizes the data access by an organization decreasing seismic interpretation cycle time by reducing the time and effort required for data transfer and management. Also, the object storage system is designed to compress and handle massive libraries of data in an affordable and highly scalable way. Nevertheless, in the presented approach, all subsequent volumes generated throughout the seismic interpretation workflow, are stored in the application's native format. To solve completely the data interoperability challenge, petro-technical software would first require exporting key interpretation results toward this common data platform using a seismic standard format such as VDS. VDS format, being a bricked format, makes possible the access almost instantaneously of any slice of a volume. Therefore, over a second phase, redesigning the core architecture of these platforms to read and write natively the VDS format would standardize all intermediate results and increase drastically the performance of applications, reducing the overall interpretation time cycle.

Acknowledgments

The authors would like to thank Geoscience Australia for their permission to use the Exmouth seismic data of the block HCA2000A.

References

Gupta, R., T. Cheret, F. Pauget, and S. Lacaze, 2008, Automated geomodeling a Nigeria case study: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: https://doi.org/10.3997/2214-4609.20147601.

Lacaze, S., F. Pauget, M. Mangue, M. Lopez, and A. Gay, 2011, Seismic Interpretation from a geological model, a North Sea case study: 81st SEG Conference & Exhibition San Antonio, doi: https://doi.org/10.1190/1.3627403.

Lemaire, R., F. Pauget, S. Lacaze, T. Cheret, M. Mangue, and C. Horno Kort, 2010, A multi scale approach on large seismic volumes – Tunisia case study: 72nd EAGE Conference & Exhibition, EAGE Expanded Abstracts, doi: https://doi.org/10.3997/2214-4609.201400729.

Pauget, F., S. Lacaze, and T. Valding, 2009, A global approach in seismic interpretation based on cost function minimization: 79th Annual International Meeting, SEG, Expanded Abstracts, 2592–2596, doi: https://doi.org/10.1190/1.3255384.

Schmidt, I., S. Lacaze, and G. Paton, 2013, Spectral decomposition combined with geomodel interpretation: Creating new workflows by integrating advanced technologies for seismic imaging and interpretation: 75th Annual International Conference and Exhibition, EAGE, Extended Abstracts, 1309–1313, doi: https://doi.org/10.3997/2214-4609.20130567.

10.1190/image2022-3743148.1