Cloud-based Well Log Database Prototype

Petroleum geoscience data management platform for data analytics

Chitra Viswanathan[1], Irina Emelayanova[1], Ben Clennell[1], Stacey Maslin[1]
[1] CSIRO

ENERGY BUSINESS UNIT
www.csiro.au

Geoscience data including seismic, well log, sensor and core measurements are fundamental for Petroleum exploration. Geoscience data volume has exploded due to recent advancements in sensor and computer technologies. This study explores the current approaches in geoscience data management and implements a well log database prototype for the target exploration geosciences group within CSIRO.

Introduction

Most industries including Geoscience are facing challenges posed by handling BIG DATA which is characterised by 3Vs, volume, variety and velocity.

Geoscience data are based on observations and measurements coming from in situ and remote-sensing data with ever-growing spatial, temporal, and radiometric resolution, requiring handling of big **volumes** (Peter Baumann, 2016).

They encompass both structured and unstructured data in the form of maps, images, spatial data, array based data (**variety**) and they are often real time data demanding fast processing (**velocity**)

This study addresses data management issues faced by the Exploration geosciences group within CSIRO. Lack of communication between the researchers often leads to no data sharing and multiple downloads of same log data from public websites (NOPIMS, WAPIMS) which increases total computing time and data processing.

The well log database (WLD) prototype would dramatically cut duplicate data downloads, provide easy access to the downloaded data and promote data sharing.

Design

Primary design criteria were:

- User friendly web based GUI to upload and share the sanitised well log data
- Storage for the bulky raw data from various logs
- An easy access to the raw data
- Ability to query the database as to what sort of data is available for each well.

Choice of architecture:

- DATABASE: HBase and MongoDB (A. B. M. Moniruzzaman and S. A. Hossain, 2013) which are NoSQL (Not Only SQL) databases were investigated as they are specifically designed to perform data analytics on Big Data. **PostgreSQL** platform was chosen for its ability to support advanced data types (arrays, JSON etc.), plug in to languages like Python, and its link to PostGIS, a spatial database extender
- RAW DATA STORAGE: **PAWSEY** supercomputing facility (a joint venture of CSIRO with Western Australian universities) was chosen as it offers cost free big volume data storage, with in-house data scientists' support. (Figure 2)
- WEB GUI: **ASP .NET** platform was used to develop the web portal. (Figures 3 to 6)

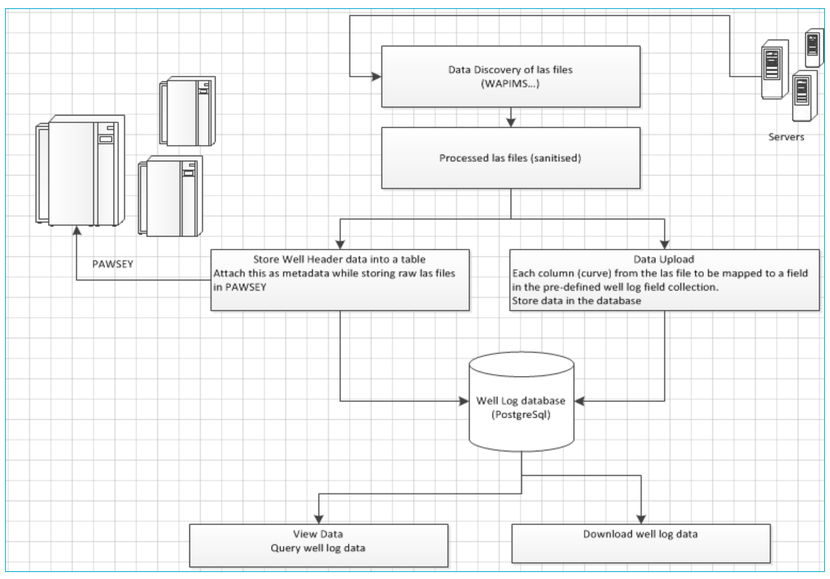

Figure 1: Schematic of the Well Log Prototype Implementation

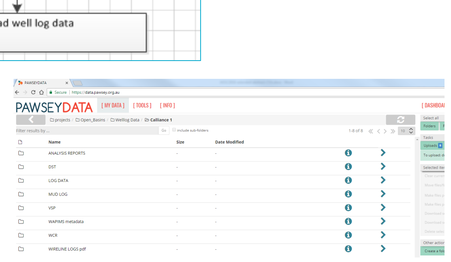

Figure 2: PAWSEY data portal

FIELD	UNIT	DESCRIPTION
Depth	m	Depth (drillers depth in meters below rig floor)
GR	API	Natural gamma ray activity
K	%	Formation potassium content from spectral gamma ray
Th	ppm	Formation thorium content from spectral gamma ray
U	ppm	Formation uranium content from spectral gamma ray
BS	in	Borehole size (smooth hole diameter produced by drill bit)
CAL	in	Caliper (hole diameter)
SP	mV	Spontaneous potential between ground and electrode at depth
Rxo	ohm.m	Resistivity of zone invaded by mud filtrate
Rs	ohm.m	Shallow Resistivity of formation close to borehole
Rm	ohm.m	Medium Resistivity for formation beyond shallow zone
Rt	ohm.m	Deep or true resistivity of formation
Rmf	ohm.m	Resistivity of mud filtrate
RHOB	g/cc	Formation density from gamma-gamma absorption lithodensity tool
ΔRHO	g/cc	Density measurement correction
PEF	B/e	Photoelectric factor (barnes per electron) from lithodensity tool
NPHI	v/v	Neutron porosity, limestone basis standard
DTC	us/ft	Delta-T Compressional (sonic travel time of compressional wave)
DTS	us/ft	Delta-T Shear Slowness (sonic travel time of shear wave)
DTST	us/ft	Sonic travel time of Stoneley wave

Table 1: Well log data fields in the sanitized well log table

Implementation

- Data fields finalized in consultation with Petro physicist. (Refer to Table1)
- Sanitized well log data*, associated metadata, original mnemonics, and a list of additional logs with storage links to PAWSEY data storage stored in the PostgreSQL DB
- Raw data along with the associated metadata from various logs are stored in PAWSEY facility
- Web GUI is used to download, view, append and share the sanitised well log data

*Unit – Converted into one set of common units and stored in the database

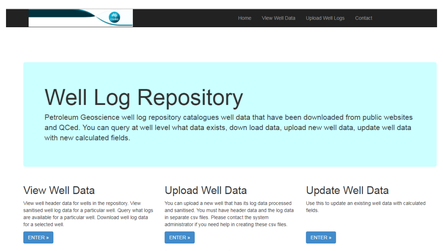

Figure 3: WLD web portal Home page

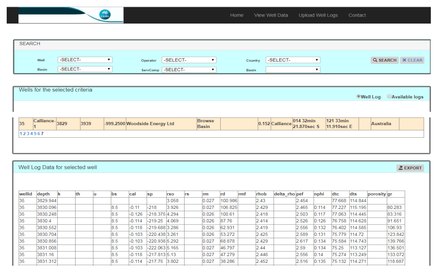

Figure 4: WLD web portal – View data & Download

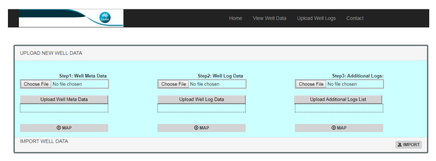

Figure 5: WLD web portal – Upload data

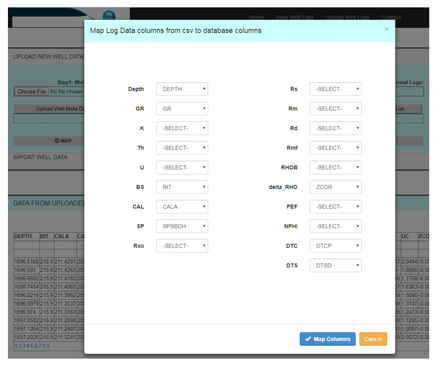

Figure 6: WLD web portal – Map data columns

Conclusions

Development of the well log database (WLD) prototype using the latest data management technologies has addressed the data integrity issues in the target Petroleum Geoscience research group within CSIRO. It has curtailed duplicate data downloads from public websites. Selection of the design features was heavily influenced by the available hardware / software infrastructure within CSIRO and affiliated facilities such as PAWSEY supercomputing facility.

In the next phase of the project other types of geoscience data such as seismic will be included and PAWSEY's cloud services will be used to perform data analytics and visualization. The data management technologies used in this WLD prototype to address the data integrity issues in a research organisation are easily transferable to manage the needs of industry as well.

FOR FURTHER INFORMATION
Chitra Viswanathan
e chitra.viswanathan@csiro.au
w www.csiro.au/Energy

REFERENCES

A. B. M. Moniruzzaman and S. A. Hossain, "NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison," ArXiv13070191 Cs, Jun. 2013.

[Hsu, L., Martin, R. L., McElroy, B., Litwin-Miller, K., & Kim, W. (2015). Data management, sharing, and reuse in experimental geomorphology: Challenges, strategies, and scientific opportunities. Geomorphology, 244, 180-189, doi:https://doi.org/10.1016/j.geomorph.2015.03.039.

REFERENCES (CONTINUED)

Iwanaga, T., El Sawah, S., & Jakeman, A. (2013). Design and implementation of a web-based groundwater data management system. Mathematics and Computers in Simulation, 93, 164-174, doi:http://doi.org/10.1016/j.matcom.2012.11.009.
Kingdon, A., Nayembil, M. L., Richardson, A. E., & Smith, A. G. (2016). A geodata warehouse: Using denormalisation techniques as a tool for delivering spatially enabled integrated geological information to geologists. Computers & Geosciences, 96, 87-97, doi:https://doi.org/10.1016/j.cageo.2016.07.016.
Li, Z., Yang, C., Jin, B., Yu, M., Liu, K., Sun, M., & Zhan, M. (2015). Enabling Big Geoscience Data Analytics with a Cloud-Based, MapReduce-Enabled and Service-Oriented Workflow Framework. PLOS ONE, 10(3), e0116781, doi:10.1371/journal.pone.0116781.