By Andrew Long (Andrew.Long@pgs.com)

Real-Time Marine Seismic Acquisition and Processing

The SEG will host a post-convention workshop at the annual meeting in San Antonio, Texas (September 2019) entitled 'Real-Time Processing for Large-Scale Streaming Seismic Data'. Two relevant questions on the processing of seismic data streamed in real time from field acquisition are, 'How can real-time processing make a difference in operations and decisions and what are the current bottlenecks?' and 'How do data acquisition systems shape real-time processing workflows?' I address these issues in the context of towed streamer marine seismic operations where streaming is dependent upon available geostationary satellite bandwidth.

Scenarios below address marine seismic surveys acquired with towed streamer operations and processed in near-real time using office-based resources. Onboard computing and human resources are large enough to perform various testing flows, but are insufficient in scale to process the full data in real time. This rate is hypothetically achievable in the office—if the data could be streamed each day from the vessel to the office.

How Much Daily Data is Recorded by Modern Towed Streamer 3D Vessels?

For continuous recording of towed streamer seismic data over a 24 hour period with 100% productivity (no downtime), the number of sources towed and shot interval is irrelevant to the data volume being recorded—each 'shot event' simply representing the onset of a new burst of energy on the data recorded by each streamer. Modern day towed streamer acquisition is therefore always 'blended', as the overlapping energy from all the shots fired over the 24 hour interval is what constitutes each 24-hour data record from each streamer. The volume of raw data recorded is directly proportional to the number of receiver channels recorded x the number of samples recorded per channel. Therefore, in the example of 18 x 8 025 m streamers with 12.5 m receiver interval, and recording at 2 ms sample rate, there will be 642 x 18 x 43 200 000 ≈500 billion samples. Using 1 byte = 8 bits, it follows that 32-bit data recording (4 bytes per sample) will correspond to ~2 TB (terabytes) of data recorded each day, plus the relevant trace header and ancillary acquisition data.

If we assume that a marine seismic survey is unable to access any LTE-type broadband networks from offshore platforms or land-based base stations, we are dependent upon geostationary satellite networks to stream the seismic data recorded each day to the office for any near-real time data processing exercise.

Available Satellite Technology

Satellites have three basic kinds of orbits, depending on the satellite's position relative to Earth's surface:

- Geostationary orbits (also called geosynchronous or synchronous: see below) are orbits in which the satellite is always positioned over the same spot on Earth. Many geostationary satellites are above a band along the equator, about 35 800 km above the surface. Television, communications and weather satellites all use geostationary orbits.
- Asynchronous orbits average about 640 km in altitude, which means they pass overhead at different times of the day.
- Polar orbits, where the satellite generally flies at a low altitude and passes over the planet's poles on each revolution. The polar orbit remains fixed in space as Earth rotates inside the orbit. As a result, much of Earth passes under a satellite in a polar orbit. Because polar orbits achieve excellent coverage of the planet, they are often used for satellites that do mapping and photography.

Correspondingly, there are relevant styles of orbit for streaming marine seismic data:

- A geosynchronous orbit (sometimes abbreviated GSO) is one in which a satellite has an orbital period that matches the rotation of Earth on its axis (one sidereal day) of about 23 hours 56 minutes and 4 seconds. A geostationary orbit, geostationary Earth orbit or geosynchronous equatorial orbit (GEO) is a circular orbit about 35 800 km above the equator and following the direction of Earth's rotation (refer to Figure 1).

  - Note that GPS satellites are in another sweet spot known as semi-synchronous orbits that take 12 hours to complete and are about 20 200 km above the surface.

- A low Earth orbit (LEO) is an orbit around Earth with an altitude of about 2 000 km and an orbital period between about 84 and 127 minutes (this includes most satellites).

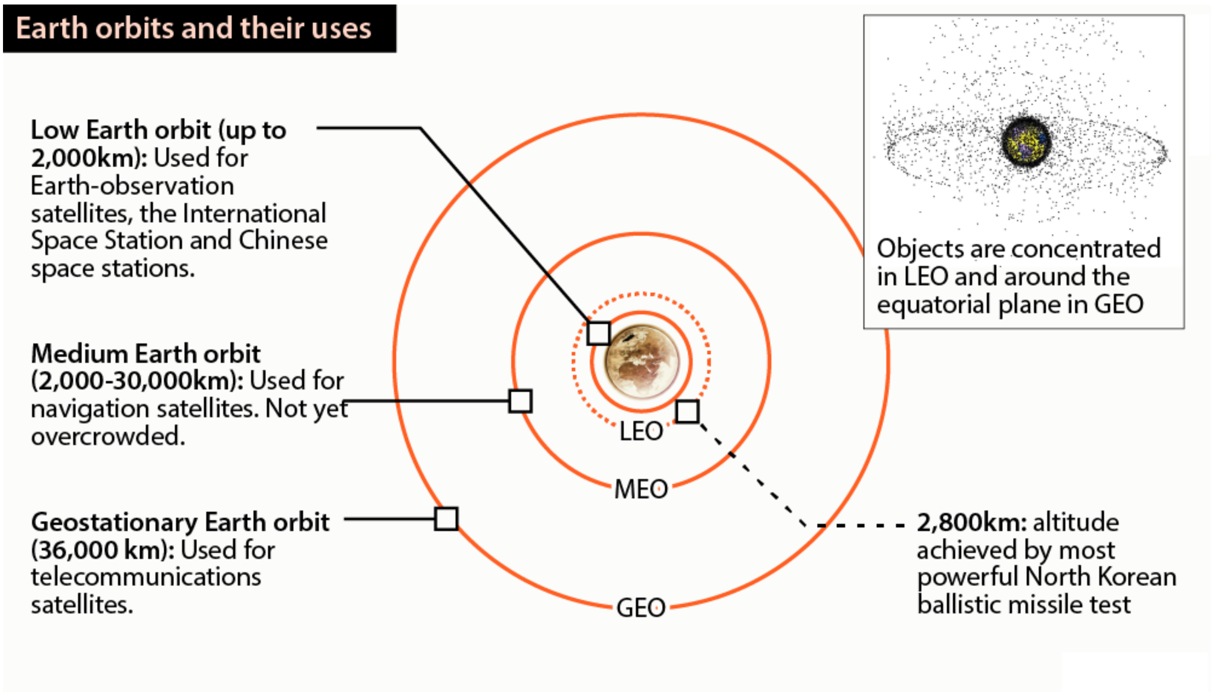

Figure 1. Schematic illustration of various satellite orbits. Source: NASA.

Geosynchronous satellites operate on the Ku-band (12-18 GHz), some with a Ka-band overlay (25.5-40 GHz). The higher the frequency the more bandwidth you can squeeze out of the system, however, the Ka band is more susceptible to weather variations. As illustrated with the inset panel in Figure 1, LEO satellite networks require many satellites to maintain coverage at any ground location. SpaceX recently launched the first 60 satellites of what will be 2 200 satellites deployed over the next five years to form the 'Starlink' network.

Each satellite weighs 227 kg, has multiple high-throughput antennas and a single solar array, and are equipped with electric propulsion—a system that expels electrically charged atoms of krypton to provide thrust. The engine is necessary to lift a Starlink from its drop-off altitude of 440 km to its operational height of 550 km. The propulsion system will also act to maintain the satellite's correct position in the sky, and to bring it down at the end of its service life. The newly launched Starlinks are an iterative design operating only in the Ku-band, but later platforms will have a higher specification, featuring for example inter-satellite links.

In the meantime, marine seismic data streaming is limited to GEO satellites that suffer from high latency due to their orbit height (about 600-800 ms). This latency makes remote interactive office-vessel sessions such as velocity picking and model building impractical, but is otherwise irrelevant for remote batch processing operations.

One advantage of LEO satellites is the much lower latency of about 30-50 ms, however, satellites in LEO have a small momentary field of view; only able to observe and communicate with a fraction of the planet at a time, meaning a network of satellites is required to in order to provide continuous coverage. Satellites in lower regions of LEO also suffer from fast orbital decay, requiring either periodic re-boosting to maintain a stable orbit, or launching replacement satellites when old ones re-enter.

Using a common default GEO transmission bandwidth of 512 kbps (kilobits per second) downlink (to shore), where 'b' = bit, 'B' = byte, and 1 byte = 8 bits (i.e. about 64 kBps downlink), 2 TB of daily seismic data would take about one year to stream to the office. This reduces to about three days by using a 64 mbps link. Clearly, we need to consider one of the following options, or a combination thereof:

- Use a much higher transmission rate (at least 256 mbps).
- Use data compression to reduce the file size.
- Only transmit 'strategic' data volumes to the office by satellite, and wait for periodic 'data drops' of the physical data (on hard disk drives) to deliver all the data to the office.

Data Compression: What is Acceptable?

Data compression algorithms are referred to as either lossless or lossy. The common 'zip' compression is a lossless format as the uncompressed file recovered after compression is perfectly reconstructed. Data compression is all around us, and most of the time we do not notice that the compressed data is often vastly smaller in size than the uncompressed version. For example, we use JPEG file format to store our photos, which is a lower fidelity version of the RAW format available on most cameras. How we achieve the compression is of course the key issue, and it is noted that the higher the compression ratio (and associated distortion level), the greater the amount of data loss.

Most seismic data compression algorithms work in a similar manner:

- A two-dimensional wavelet transform of some type is applied to the seismic data (not the headers).
- Quantization (where the precision of the wavelet-transformed data is reduced and thus data loss occurs) is achieved by setting coefficients (or 'wavelet sub-bands') to zero.
- A lossless rearrangement of the header values to reduce file size.

Note that standard SEGY seismic data is a combination of EBCDIC and binary trace header information, plus floating point seismic data records.

The compression ratio offered by various solutions can vary from 4:3 (lossless) to 100:1 (highly lossy), and the solution chosen is typically determined by the desired data bandwidth and fidelity. Stacked data can be more highly compressed without perceptible distortion than a shot gather.

The distortion level is defined as follows:

$$Distortion\ Level = \frac{\sum^{samples}\left(x_{original} - x_{decompressed}\right)^2}{\left(\sum^{samples} x_{original}\right)^2} \times 100.00$$

where $x_{original}$ are the uncompressed data values, and $x_{decompressed}$ are decompressed data values, and the distortion level is traditionally preferred to be in the range of 0.1% to 5.0%.

High compression rates would enable daily transmission of all the data from the vessel to office at reasonable cost, such that all testing could be pursued in near-real time on all the survey data while acquisition continues—without having to wait for the physical data drops. So, when the physical data from the completed survey arrives in the office, much of the testing could already be complete. Note this scenario involves some of the data being processed twice—first the compressed/uncompressed data is used for testing, and then the raw data is processed through the production flow. Historically, final seismic data deliverables have not included any form of loss data compression/decompression, meaning that all final deliverables have been derived from uncompressed physical data from the vessel.

What effect does compression have on data fidelity? Figure 2 compares shot gathers before and after 10:1 compression. Even when the difference result is amplified several times, the signal distortion is imperceptible.

Pragamatic Workflows for Real-Time Processing

Most pre-imaging processing steps can use heavily compressed/decompressed data without any corruption of parameter testing, QC and decision-making. Therefore, highly compressed data could hypothetically be streamed off the vessel at close to the rate of acquisition; enabling testing and flow parameterization to proceed in the office, and in parallel with acquisition.

This opens the door for several potential scenarios:

- Delivery of fast-track seismic images soon after acquisition concludes—using data that was decompressed and processed in the office.
- If most parameters are already established (for example, in 'mature' 4D programs with several monitor surveys already acquired and processed), then the compressed/decompressed data could be used to deliver a full-integrity seismic product (no abbreviation to the production processing flow) soon after acquisition concludes.
- Parameterization of several processing steps in the office using compressed/decompressed data, and then processing of the uncompressed data in an accelerated manner when the physical data are received off the vessel.
- Parameterization of several processing steps in the office using compressed/decompressed data, and these parameter choices used by on the vessel using onboard computing resources. This frees the onboard personnel to work on, for example, velocity picking and model building (an exercise that can be impractical using remote sessions due to the aforementioned GEO satellite latency issues). Several 'production' processing steps can be applied before the data leaves the vessel.
- Transfer of interim interpretation products from the vessel during acquisition, e.g., angle stack and angle domain common image gathers (ADCIGs), or first-pass velocity models.

The options chosen depend upon a variety of factors, including the availability of vessel-based human and computing resources, the availability of office-based resources, the duration of the acquisition phase (i.e. how much data processing can be achieved during that time interval), the available satellite bandwidth, the project priorities (whether accelerated processing is required, and if so, involving what delivery schedule), and the appetite for data compression.

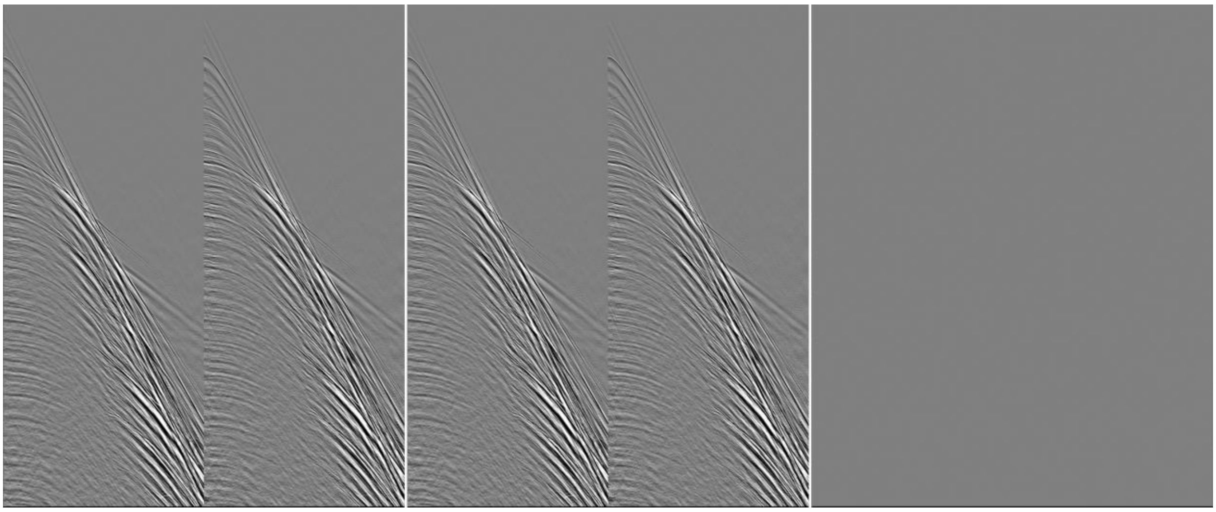

Figure 2. (left) Uncompressed shot gathers; (middle) after 10:1 compression and decompression; and (right) difference.

Summary

After discussing the satellite transmission logistics for modern towed streamer seismic surveys, I addressed the complementary benefits that lossy data compression can provide if the processing flow is strategically designed to meet the various data delivery schedule requirements.

Overall, it is clear that data compression needs to be considered wherever satellite or LTE-type networks cannot cost-effectively transfer large data volumes in near-real time from vessels. LEO satellite networks may significantly improve transmission rates in the future, as well as enable interactive remote sessions due to the reduced latency, but in the foreseeable next few years we will continue working with GEO satellites and high latency.