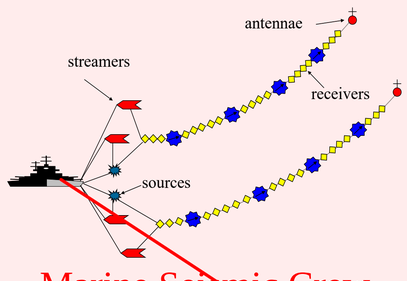
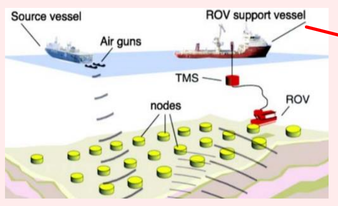
Data Ingestion Pipelines for Seismic Crew

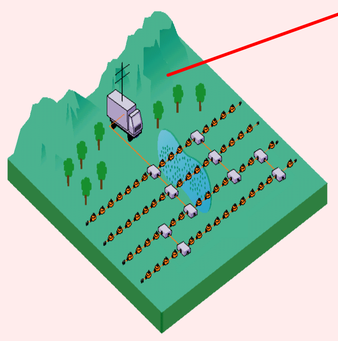
Mohit Lokwani, Vijaykumar Raj (vraj@argas.com)
QC - IT, Arabian Geophysical and Surveying Co. (ARGAS), King Fahad Road, 34232, Dhahran, Saudi Arabia



Marine Seismic Crew



Marine Automated Node Survey



Land Seismic Crew

1. Data Producer
1. Data Producer
1. Data Producer
2. Data Consumer - Polling
Field Data
5. Datasets
5. AI/ML
3. Field Data

Kafka
Topic Marine
Topic OBN
Topic Land

API
Dataset
Seismic Processing Software
API
AI/ML
SEGD
SEGY

Field

Processing Center in Cloud / On Premise DC

1. Data Ingestion with Kafka:

- Seismic data from field and marine sources is captured by a recorder, acting as a Kafka producer.
- The recording application publishes this data to specific Kafka topics, with each crew assigned a separate topic.
- Topics are divided into multiple partitions to enhance parallel processing and scalability.

2. Real-Time Data Access:

- A processing application, acting as a Kafka consumer, accesses real-time field data via APIs as soon as new data is ingested into Kafka.
- This enables immediate processing and reduces latency in data availability.

3. Data Storage and Direct Processing:

- Ingested data is stored for further processing; however, direct processing from Kafka topics is also possible, enabling the generation of datasets on-the-fly.

4. AI/ML Module Integration:

- The processing application dynamically loads necessary AI/ML/DL modules via APIs based on the field data received.
- This modular approach ensures the application is equipped with the right tools for each processing task.

5. Advanced Seismic Processing:

- AI/ML/DL modules are utilized to enhance seismic data processing, focusing on improving SNR, conducting velocity and attribute analysis, detecting faults, performing seismic inversion, and interpreting results.