

Sensor Data and Analysis Framework (SDAF)

Don Landing

781-377-8312 • dlanding@mitre.org

Stephan Nadeau

781-377-7084 • snadeau@mitre.org



MITRE Sponsored Research

Problem

Changes in the sensor environment:

- Daily volumes of data are increasing from gigabytes to many terabytes.
- New types of information are being extracted from existing data types.
- Number and types of multi-function sensors are increasing.
- Critical response times for situational awareness decreasing.

Problem Statement: Current archive data and streaming data querying techniques are insufficient by themselves to harmonize sensor inputs from large volumes of data. These two distinct architectures (push versus pull) have yet to be combined together to meet the demands of a data centric world. The input of sensor streaming data from multiple sensor types further complicates the problem.

Background

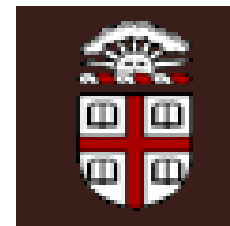
- A great deal of research in standard data streams using the stock market and network traffic analysis has been carried out.
- Sensor data streams research is still in its infancy.
 - Sensor data streams sample the entire area and are error prone.
 - Volumes of data from sensor streams fluctuate.
 - Little work has addressed the government problem space (e.g., pedigree, moving sensors, large volumes of data).
- Academic areas identified for further research:
 - Integration of stream processing and archive databases
 - Stream indexing, searching, and similarity matching

Objectives

- **Develop an integrated query capability that simultaneously accesses streaming and archive data sets from multiple sensor types.**
- **Design and test different techniques for incorporating data pedigree with integrated queries.**
- **Develop an approach that can scale and that addresses the need for quicker response times.**

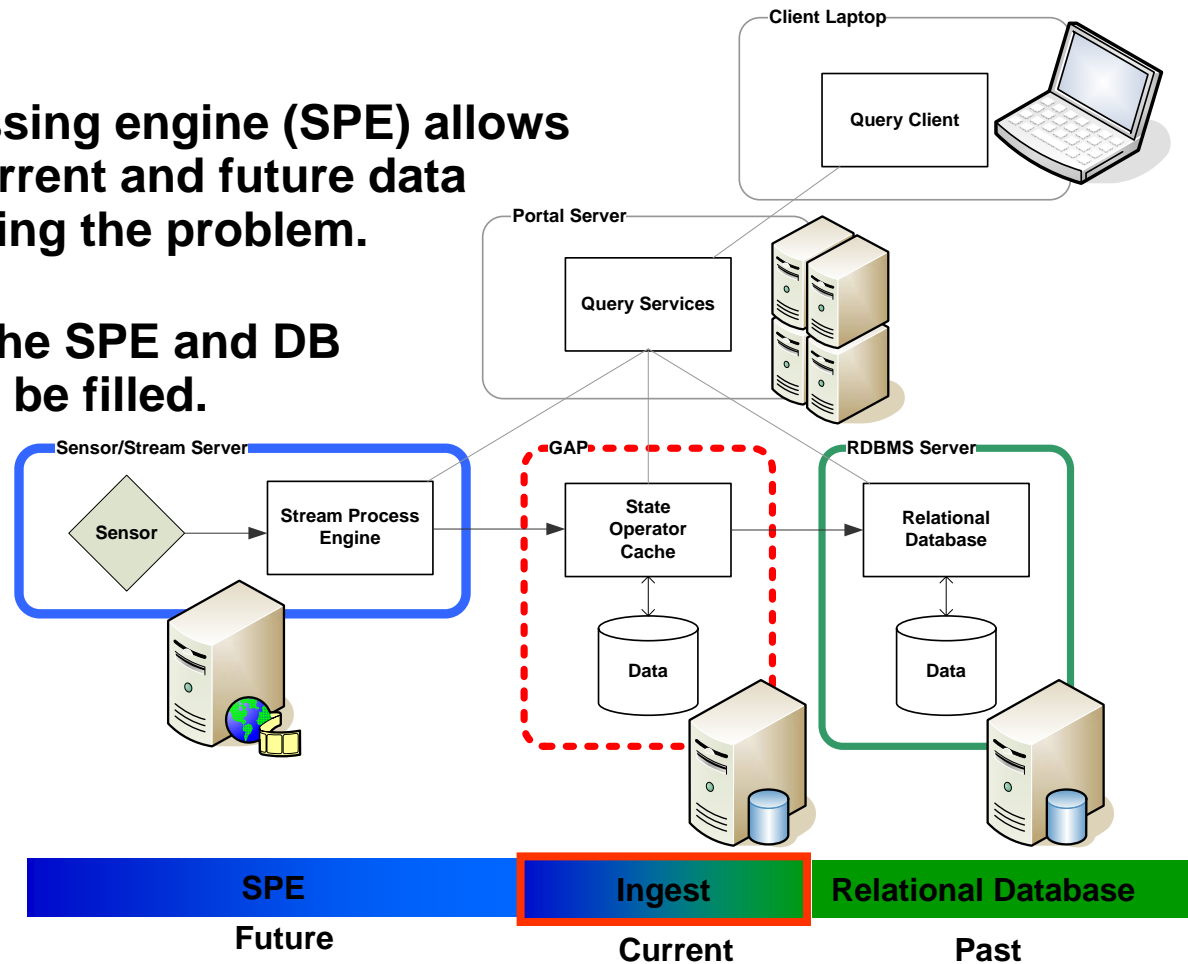
Activities

- Collaborating on research with academia:
 - Borealis Open Source Community (MIT, Brandeis, Brown)
 - Boston University
 - Washington State University
 - University of Dartmouth
- Processing sensor data sets using a stream processing engine:
 - Moving Target Indicator (MTI) data
 - Cursor on Target (CoT) data
- Maturing an analysis framework
- Analyzing a stream processing engine
- Collaborating with ISR Forensics – an Internal MITRE project
 - ISR Forensics provided a GMTI data warehouse as a baseline for developing an archived data source to feed the framework and capture pedigree.



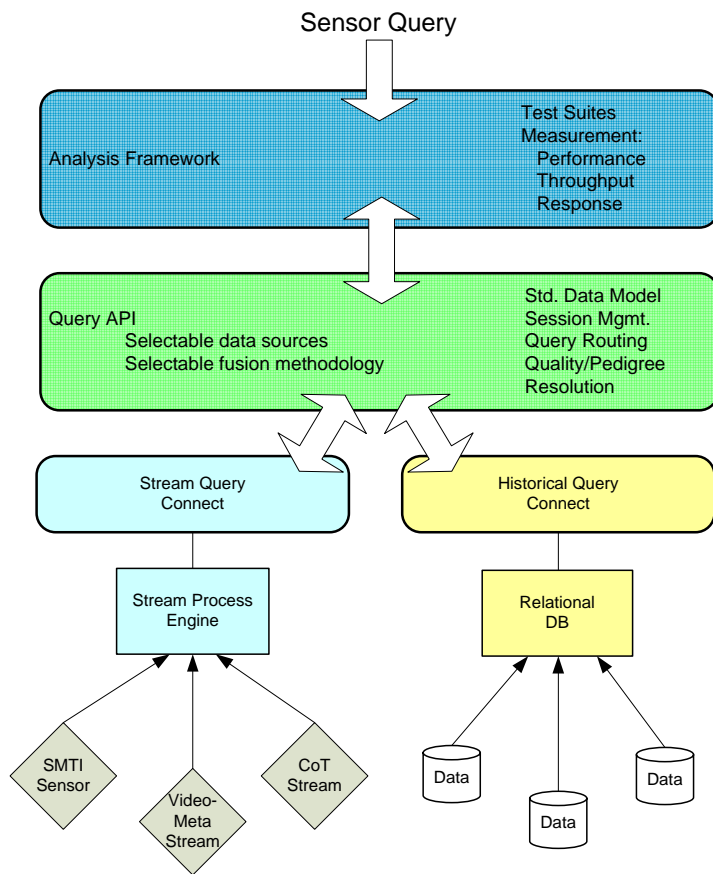
Highlight

- New sensor data in the relational database (DB) is not available during the ingest period.
- The stream processing engine (SPE) allows queries against current and future data only, partially solving the problem.
- The gap between the SPE and DB technologies must be filled.



Highlight

Sensor Data Analysis Framework Layers



3D Visualization Engine Front End



Stream Processing Engine with Data Operators



Impacts

- **Knowledge transfer:**

- Streams research in academia is transferred to MITRE.
- MITRE/government problem space influences academia.

- **Positive results from the SPE using current datasets:**

- SPEs allow for large amounts of sensor data to be filtered in near-real time.
- Additional datasets can be dynamically combined to improve result sets.

- **Academia:**

- Research from academia is now focusing more on dynamic sensors collecting dynamic data sets.
- Additional research areas have been developed from our discussions (University Washington).

Future Plans

- Continue work with academia to address integration of dynamic stream and archive queries.
- Enhance framework to quantify performance and improve scalability.
- Optimize the archive structures to include pedigree and work with the SPE.
- Use more realistic and diverse data sets to assess SPE and framework performance.