Large-Scale Network Data Analytics

S&T Campaign: Computational Sciences
Data Intensive Sciences

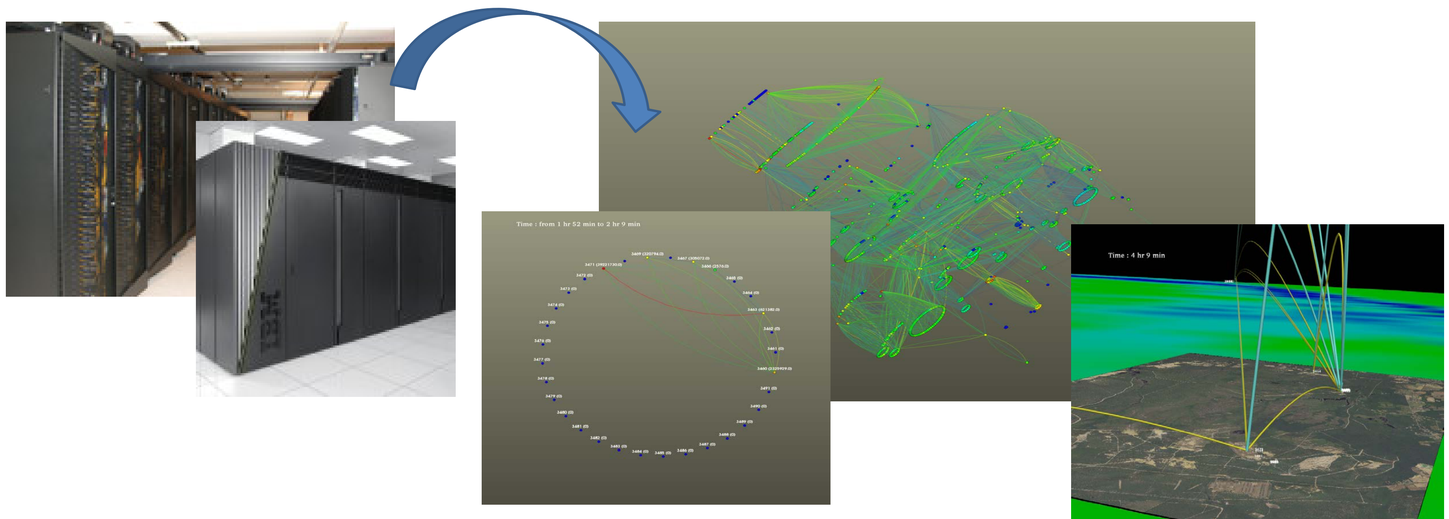
Brian Panneton
(410) 278-5345
brian.c.panneton.civ@mail.mill

Research Objective

- Investigate non-real time and real-time data analytics for C4 and automotive data
- Generate performance metrics from the data
- Apply stream processing to extend analysis and evaluation capabilities of the Army
- Enable additional analysis of data through advanced algorithms and data visualization

Challenges

- Decomposition of data for parallel execution
- Formatting requirements for dynamic analysis
- Mapping of real-time analysis to existing HPC platforms



Transformation of raw data into higher dimensional data models.

Complementary Expertise/ Facilities/ Capabilities Sought in Collaboration

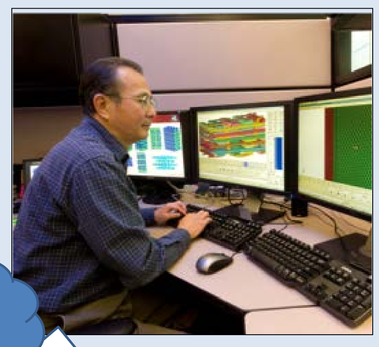
- High-speed distributed data storage to support visualization and analysis
- Scalable visualization tools for network test data
- Innovative methods for robust streams analysis and evaluation in presence of network disruptions
- Dedicated HPC hardware for static and real-time network data capture, reduction, visualization and analysis

ARL Facilities and Capabilities Available to Support Collaborative Research

- HPCMP/ARL-DSRC supercomputer facility DSP (Dedicated Service Partition) provides computation and storage
- Streams software including IBM InfoSphere Streams developers licenses
- Unclassified and Classified computing environments depending on sensitivity of data being analyzed
- Internally developed HPC data reduction framework and modules for C4 data







DREN

- Data collected via NetADMAS on various platforms
  - Network packet data
  - GPS location
  - GPS Time – clock correction
- Harvested every 12 – 48 hours
- Transferred to ATC

DREN

- Pre-processing of data
  - Generation of context
  - Indexing and “view” selection
- **Transfer data to ARL HPC**
- **Data Reduction**
- Populate data model

- Delivery of data model to evaluators
- Service requests for additional information
  - Reach back to raw or intermediate data for clarification
- Conduct analysis and take action

Data Flow for C4 and Automotive Analytics