Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Analyzing Network of Networks at Scale

JOHN FEO AND VITO CASTELLANA
PACIFIC NORTHWEST NATIONAL LABORATORY

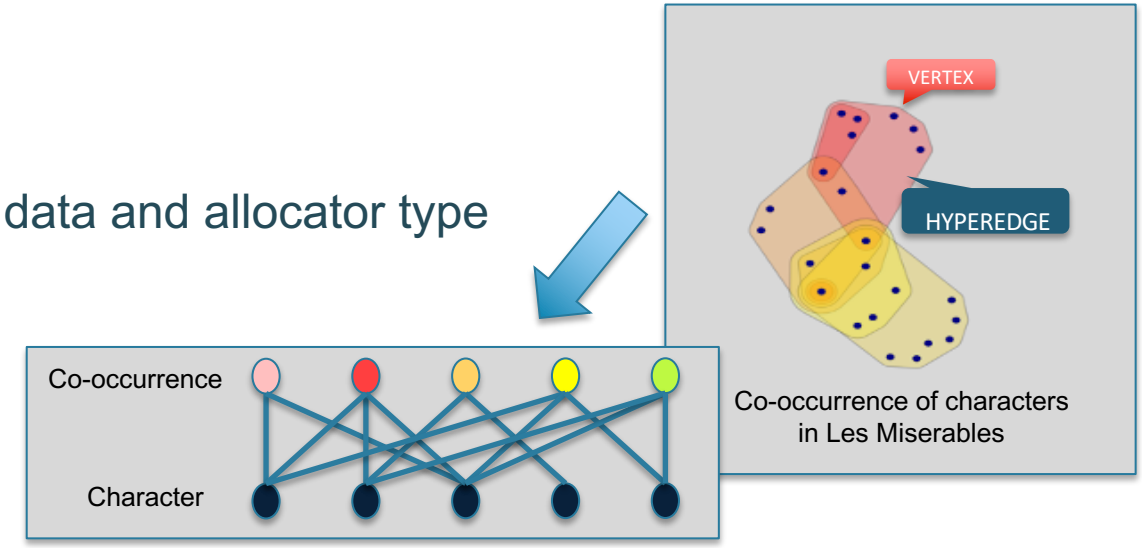
CLSAC
SEPTEMBER 2020

PNNL's platform supports scalable mixed analytics

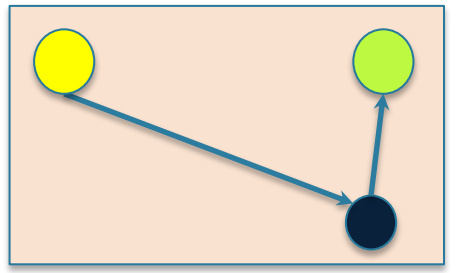
- Data views
  - Matrices – dense and sparse integer and floating point
  - Tables – heterogeneous data columns; tensors
  - Graphs – property graphs and hypergraphs
- Composable method libraries
  - Matrix and linear algebra routines
  - Relational database and machine learning methods
  - Graph methods for line and hypergraphs
    - Streaming
    - Fuzzy matching
- Scalability
  - Billions of vertices and edges
  - Shared memory and HPC cluster implementation

Recent enhancements

- Linear Algebra Library
  - Basic matrix operations and solvers
  - Matrix and vector data structures are templated by data and allocator type
  - Allocators are used to
    - modify data layout and distribution
    - target accelerators
    - support external libraries (SLATE, ScalaPack)



- Hypergraph Library
  - Represented as bipartite property graph
  - Collapse, BFS, Louvain, s-overlap, s-shortest path, s-components
    - Modify simple graph algorithms for “wedges”
    - Construct hyperedge – hyperedge graph OR vertex – vertex graph



A problem with hypergraphs

- Small data sets everything is fine … but
- Large data sets MAY result in an explosion of work or edges

Say you have a hypergraph with **10M hyperedges** and **1M vertices** and
100K hyperedges have 100 vertices and each of those vertices belong to 100K hyperedges, then

those just 100K hyperedges generate 100K * 100 * 100K wedges of work in the simple algorithm

OR

100K * 100 * 100K edges in the hyperedge – hyperedge graph

Collapse graph or filter out work while preserving accuracy

SHADes

- Client-server analytics platform, inspired by Arkouda
- Jupyter Notebook / Python frontend, SHAD backend
  - Front-end commands are mapped to SHAD functions
  - The SHAD server implements an interpreter for client commands
- Multiple clients can connect to the same server at the same time
  - Multiple users can explore or query same data images
- Clients can connect to multiple backends or separate data views

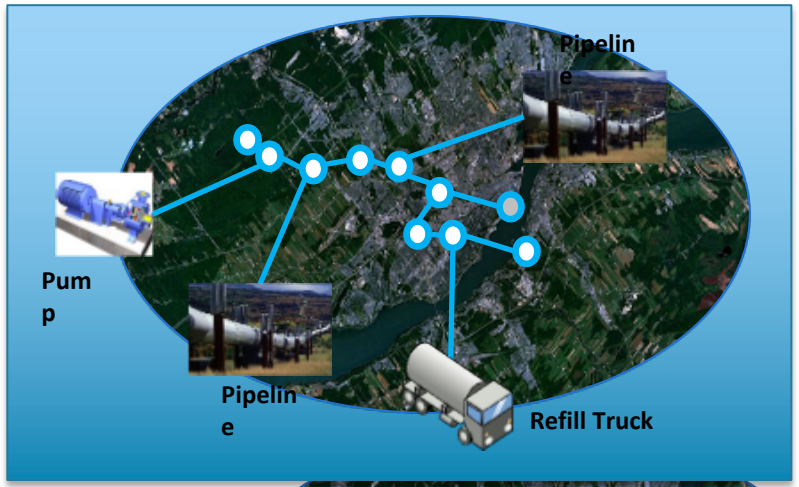
https://github.com/pnnl/SHADes

Platform for system of systems analysis

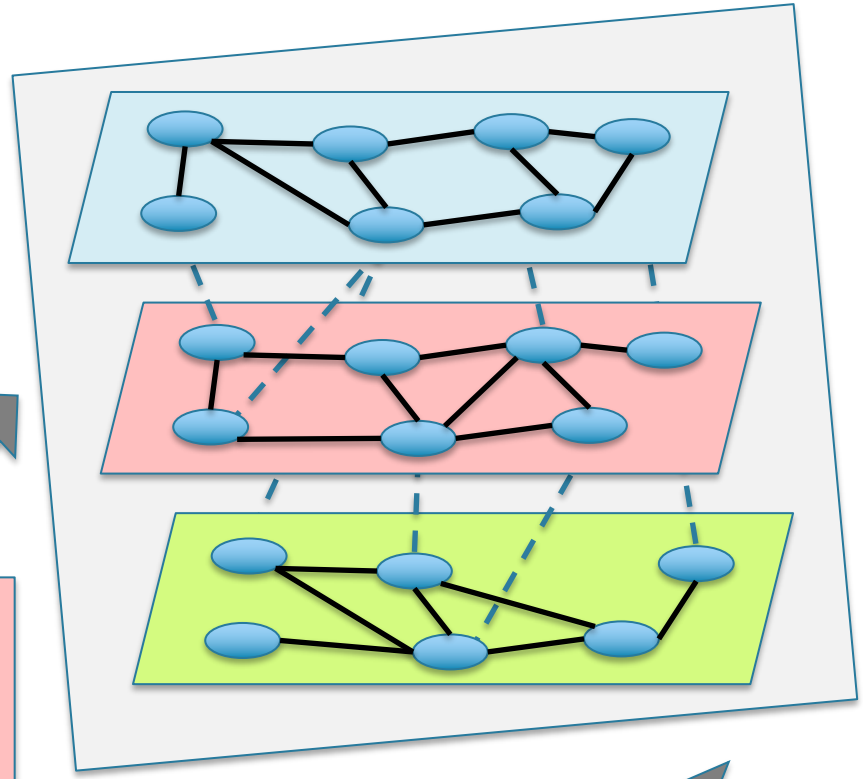
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

PNNL's Analytic Platform models multiple systems as one enabling cross system analysis, consequences, and decision support

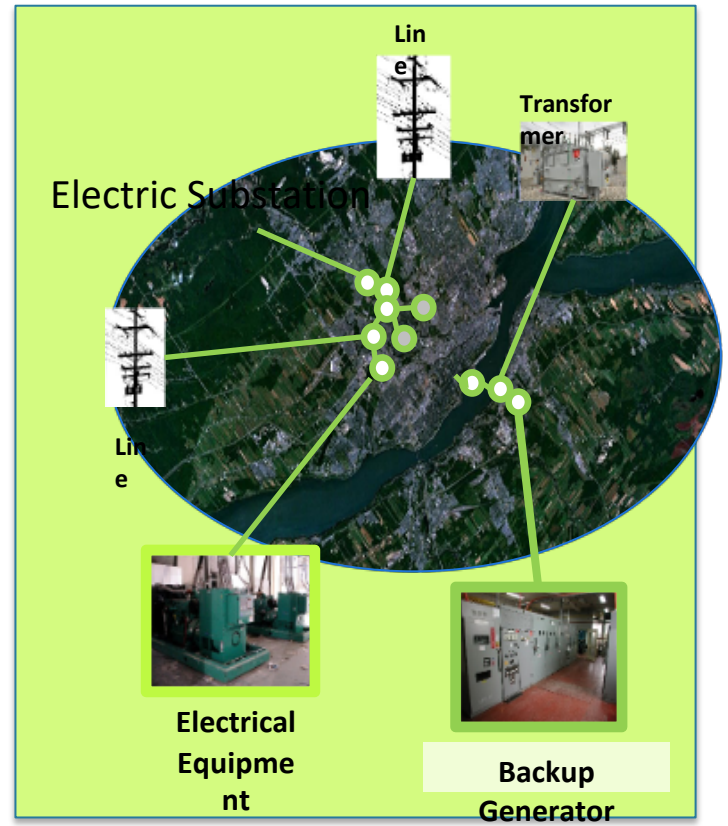
NATURAL GAS



COMMUNICATIONS

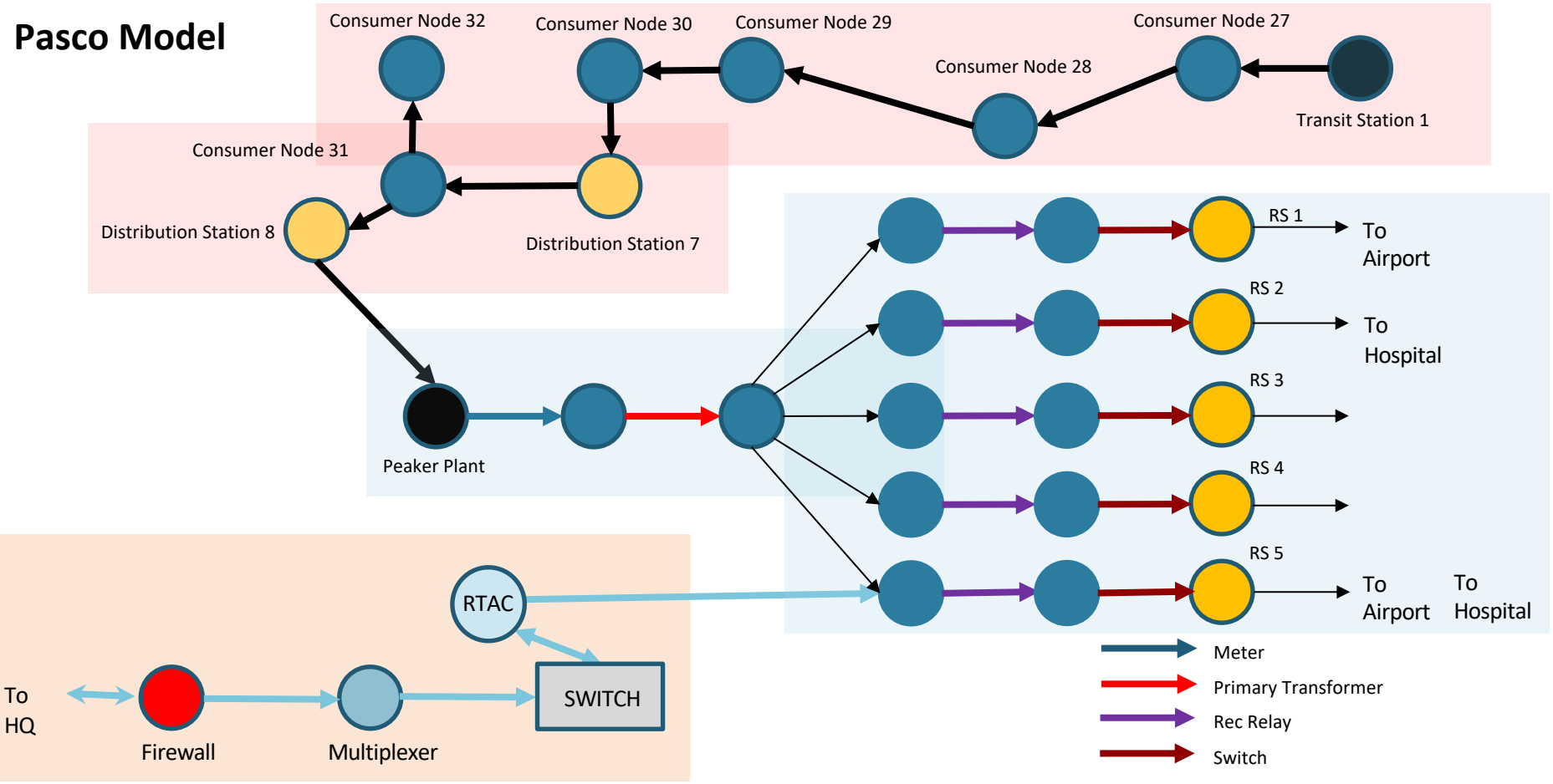


POWER GRID



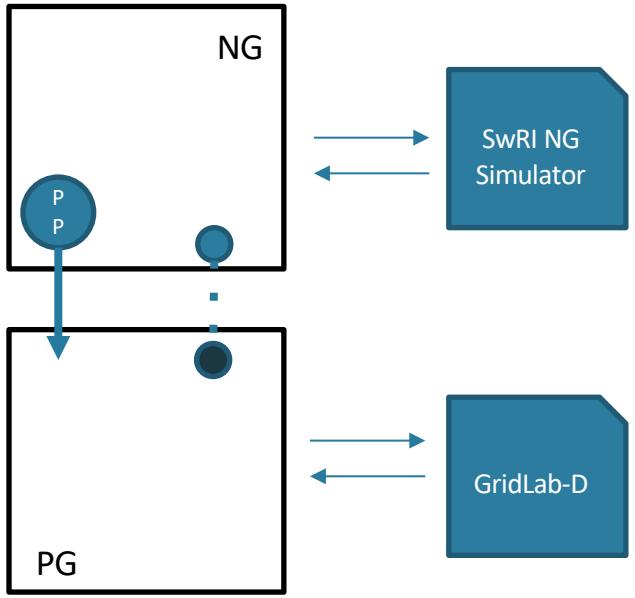
Predictive Consequence Modeling

- Model consequence of an action or event on one network on all networks



Action Graph = Properties + Methods

- Vertex and edge type definitions now include METHODS
  - Methods can take local action
  - Send messages to neighbors causing them to take actions
  - Invoke global functions

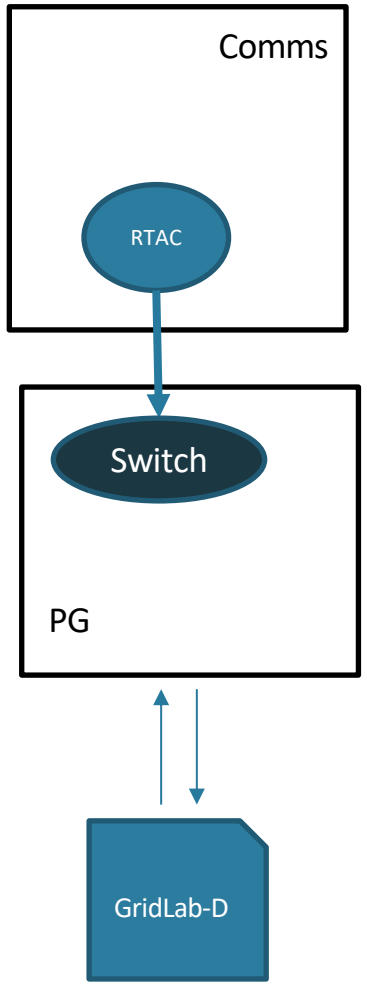


What is the impact to the Pasco hospital if I destroy the transformer that feeds the airport?

- Destroy transformer and collocated distribution station
- Status change launches PG and NG simulation with Δt time step
- Change in gas reaching Peaker Plant reduces power output change in output launches PG simulation
- Trace the electrical power reaching the hospital over time

Cybersecurity modeling

Edges are now asynchronous message queues between vertices



Coupled Rule-based / Scientific simulation

- Launch cyber attack to open switch

  Firewall/Router –
  - Message arrives at Firewall
  - Firewall accepts message
  - Message calls BOT subroutine.
  - BOT subroutine sends remote procedure request to Multiplexer

  Multiplexer –
  - Message arrives from Firewall
  - Multiplex accepts message and executes OpenSwitch

  OpenSwitch –
  - Open SSH session on RTAC
  - Send SEL protocol command to RTAC
  - Receive list of RecRelay that use SEL protocol (say RecRelay 4 is on the list)
  - Send command to send a message to RecRelay 4 to open switch
  - Send Logout command

  RTAC –
  - Listen for SSH session
  - Opens connection
  - Accept and execute command until Logout command

- Open switch and launch PG simulation
- Run until convergence or commander's decision window