IEEE PES GM2016
Panel: Domain-Specific Big Data Analytics Tools in Power Systems
Chair: Prof. Hamed Mohsenian-Rad, UC Riverside

From data to actionable information: data curation, assimilation, and visualization

Zhenyu (Henry) Huang
Chief Engineer/Team Lead
Pacific Northwest National Laboratory
July 19, 2016

Pacific Northwest
NATIONAL LABORATORY

IEEE PES Power & Energy Society®

IEEE

Rich new data hold the promise to transform grid view and management

- Data sources are more diverse, with increased data volumes
  - SCADA → phasor
  - Market, weather/climate, cyber/communication, …
  - Simulated data
- Generic 4 "V's": capture the data evolution in power grid.

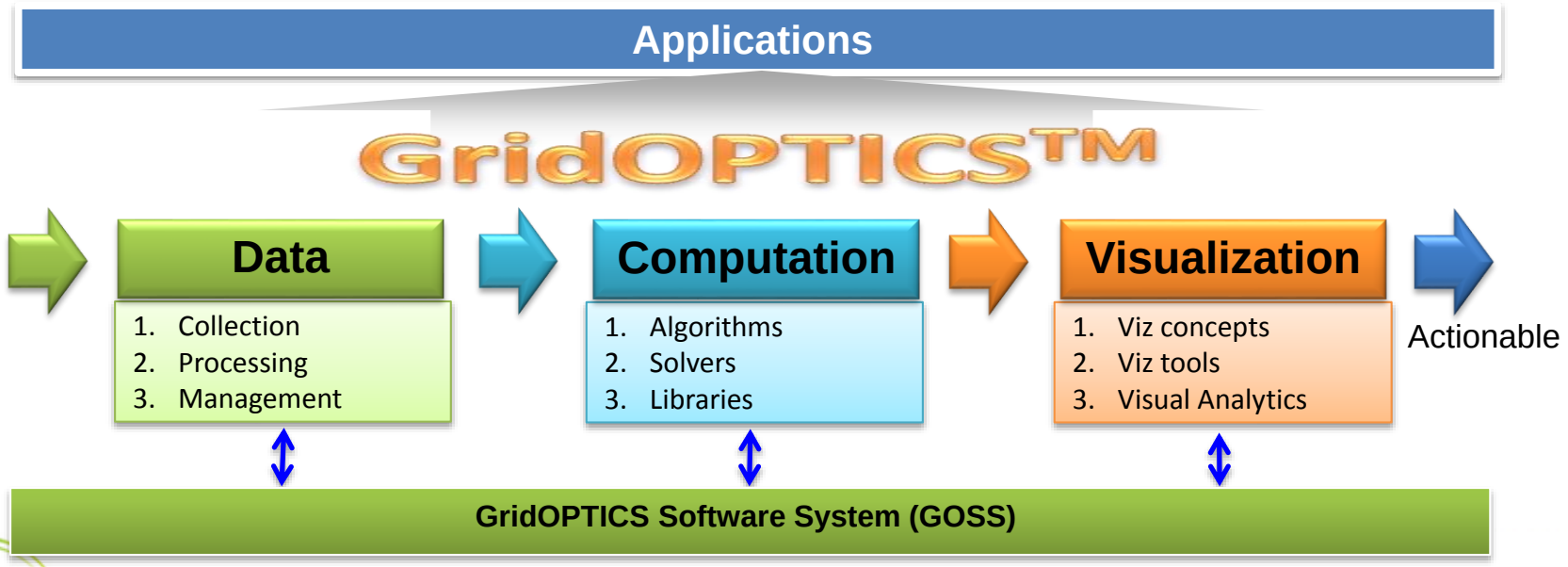
	Today – SCADA data	Emerging – phasor data	Improvement
Variety	voltage + current	+ phase angle, …	more information
Velocity	1 sample / 4 seconds	30-120 samples / second	~200x faster
Volume	8 terabytes / year	1.5 petabytes / year	~200x more data*
Veracity	unseen ms-oscillations	oscillations seen at 10ms	greater accuracy

* Transmission level only

- What to do with the data in domain-specific applications?

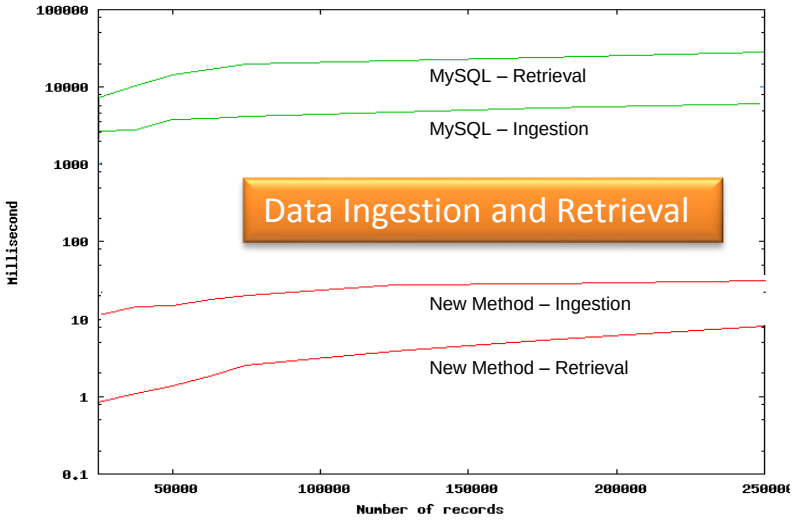
Power grid "Big Data" Challenge: making diverse data reliable, available and actionable

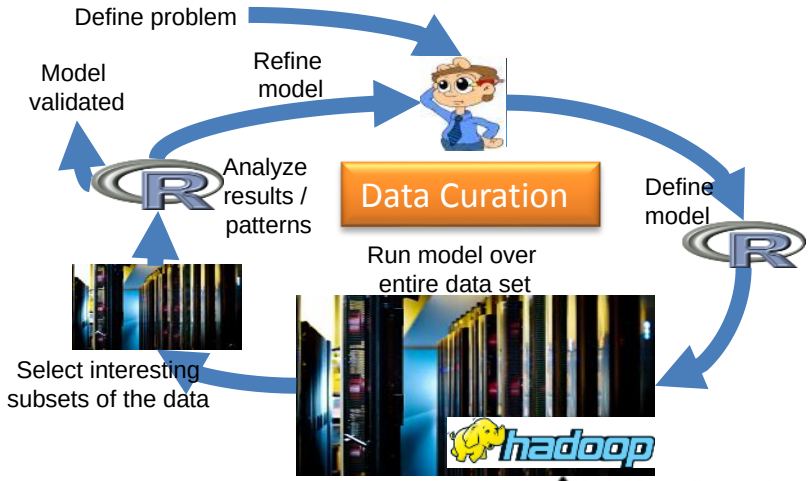
- GridOPTICS™: A suite of methodologies & software modules for accelerating the development and adoption of new **data analytical tools** for the power grid facing new **complexity**, **stochasticity**, and **dynamics**.



Real-time data ingestion, retrieval, curation from a distributed sensor network

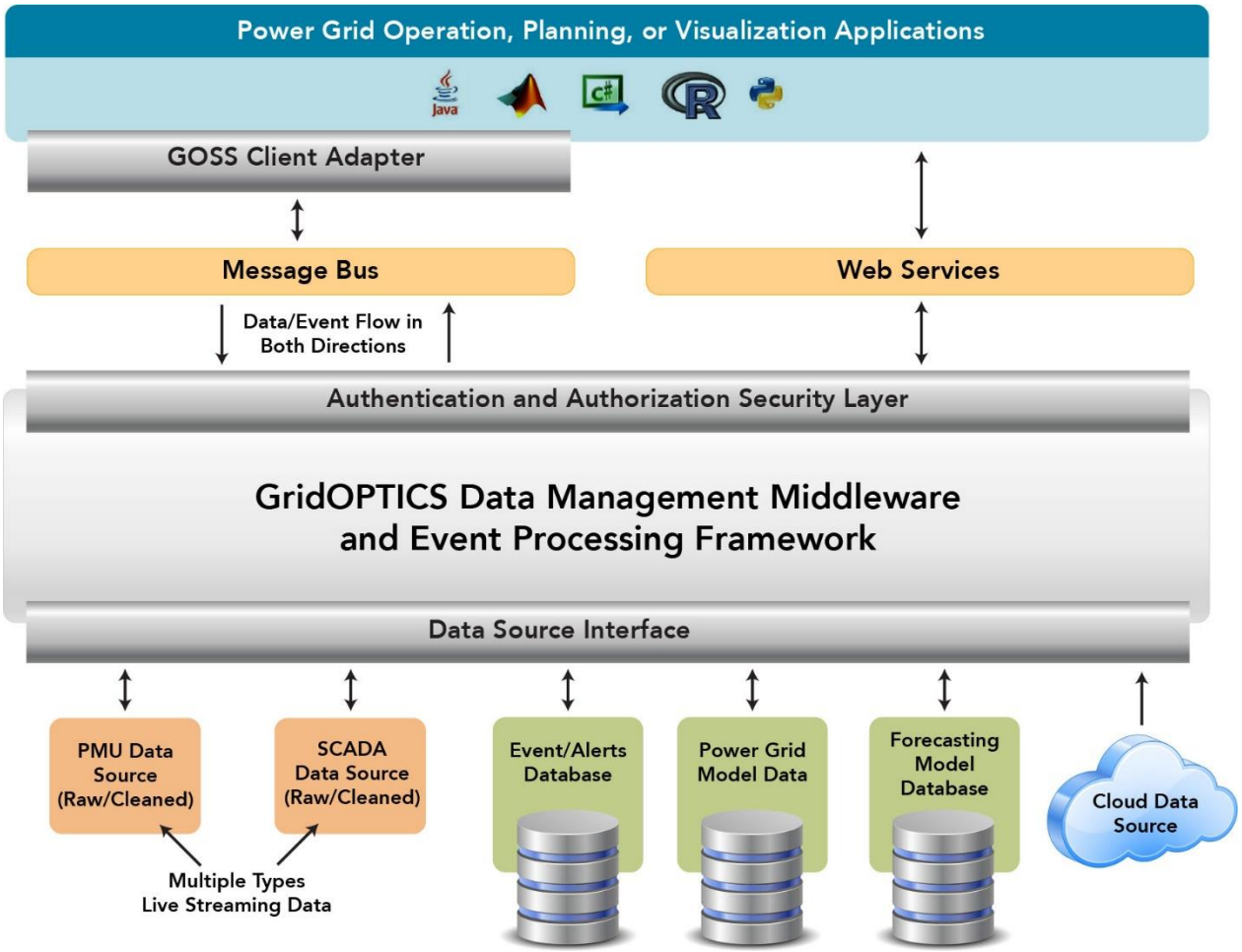
- **Requirements**
  - Cyber-secure sensor network
  - Data provenance and privacy
  - Real-time processing
- **Solution: scalable, flexible middleware and R/Hadoop statistical analysis capabilities**
  - Data ingestion is 10^3 times faster than MySQL
  - Linearly scales to many nodes
  - Data curation cleans data and detect events with confidence in real time





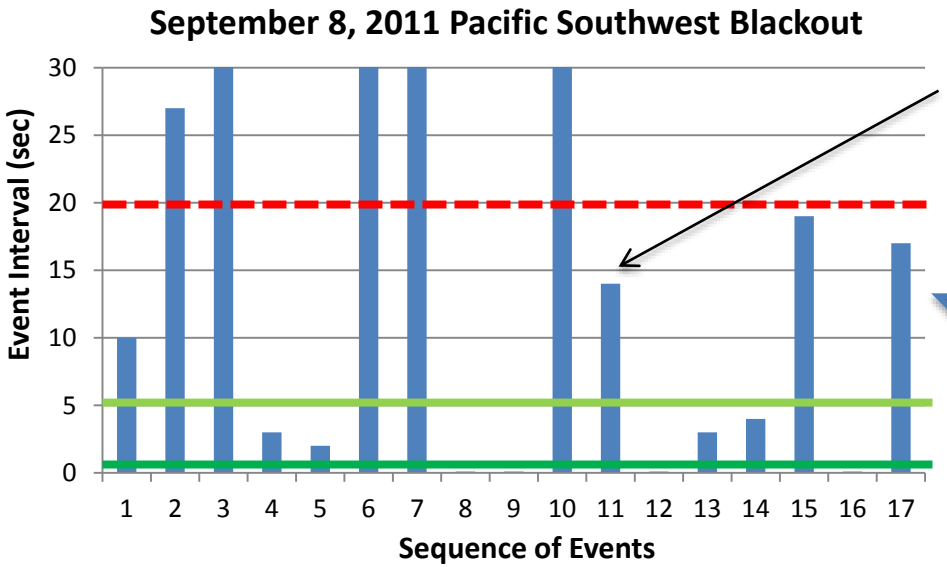
GOSS™: link data to applications

https://github.com/GridOPTICS/GOSS



Data assimilation: State Estimation (SCADA + power flow model)

- Fast State Estimation captures real-time changes and offers an opportunity to stop cascading



When the event interval is less than the ability to respond, there is a cascading effect. This means that the region of impact from the disturbance is expanding.

Today's view, >20 sec

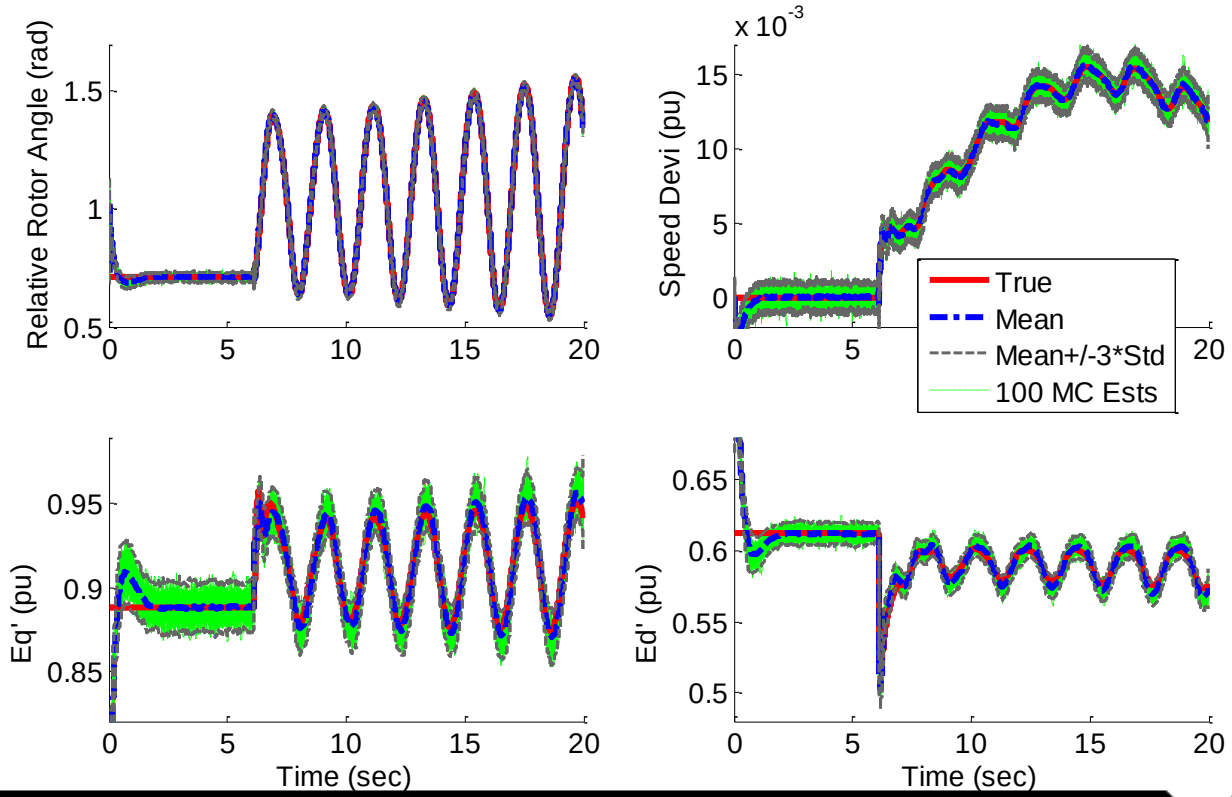
Need for speed improvement

FY13, 5 sec view

FY14, 0.5 sec view

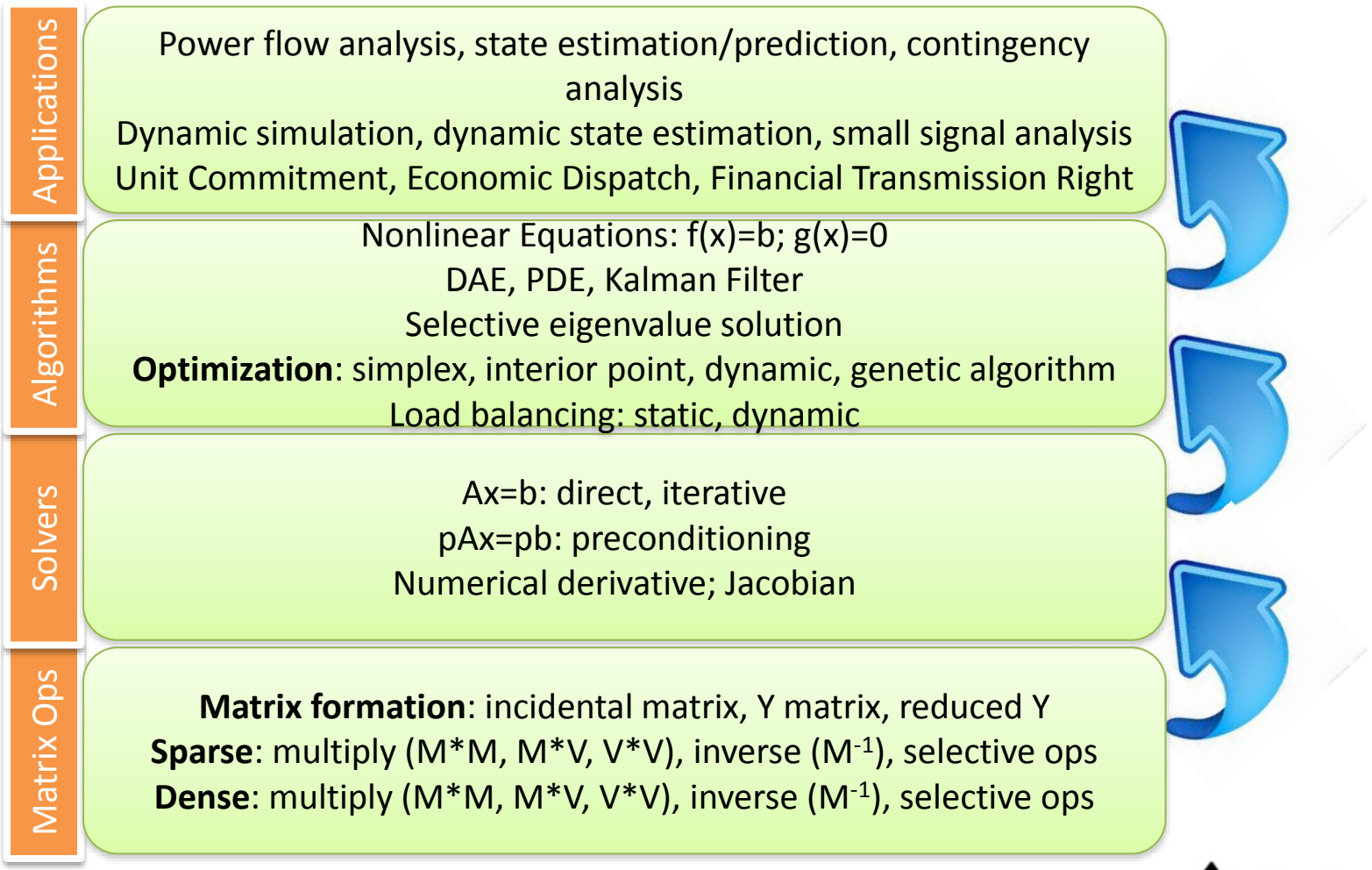
Data assimilation: Dynamic State Estimation (Phasor + DAE model)

- Estimating power system dynamics states (and parameters) in real time. Excellent tracking using Kalman filter with imperfect model and realistic conditions. Scalable to 1000s cores.



GridPACK™: building blocks for scalable power grid computing

https://www.gridpack.org/



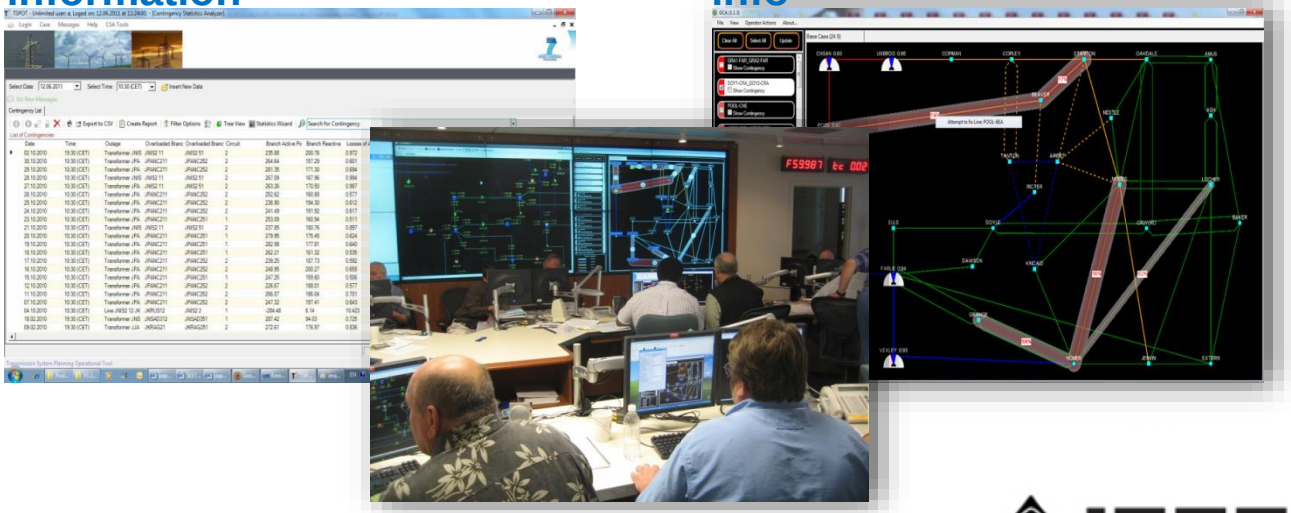
Visual analytics of massive Contingency Analyses for real-time decision support

Contingency Analysis	Number of scenarios	Serial computing on 1 processor	Parallel computing on 512 processors	Parallel computing on 10,000 processors
WECC N-1 (full)	20,000	4 hours	~30 seconds 469x speed up	
WECC N-2 (partial)	153,600	26 hours	~3 minutes 492x speed up	~12 seconds 7877x speed up

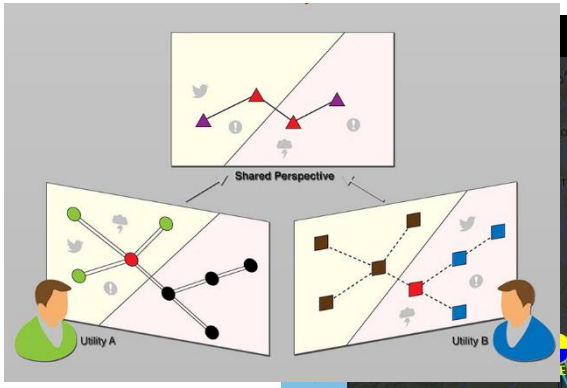
- Easy-to-interpret visualization of contingency analysis data
- Prioritized areas of concern and recommended corrective actions
- Operators reported 30% improvement in emergency response

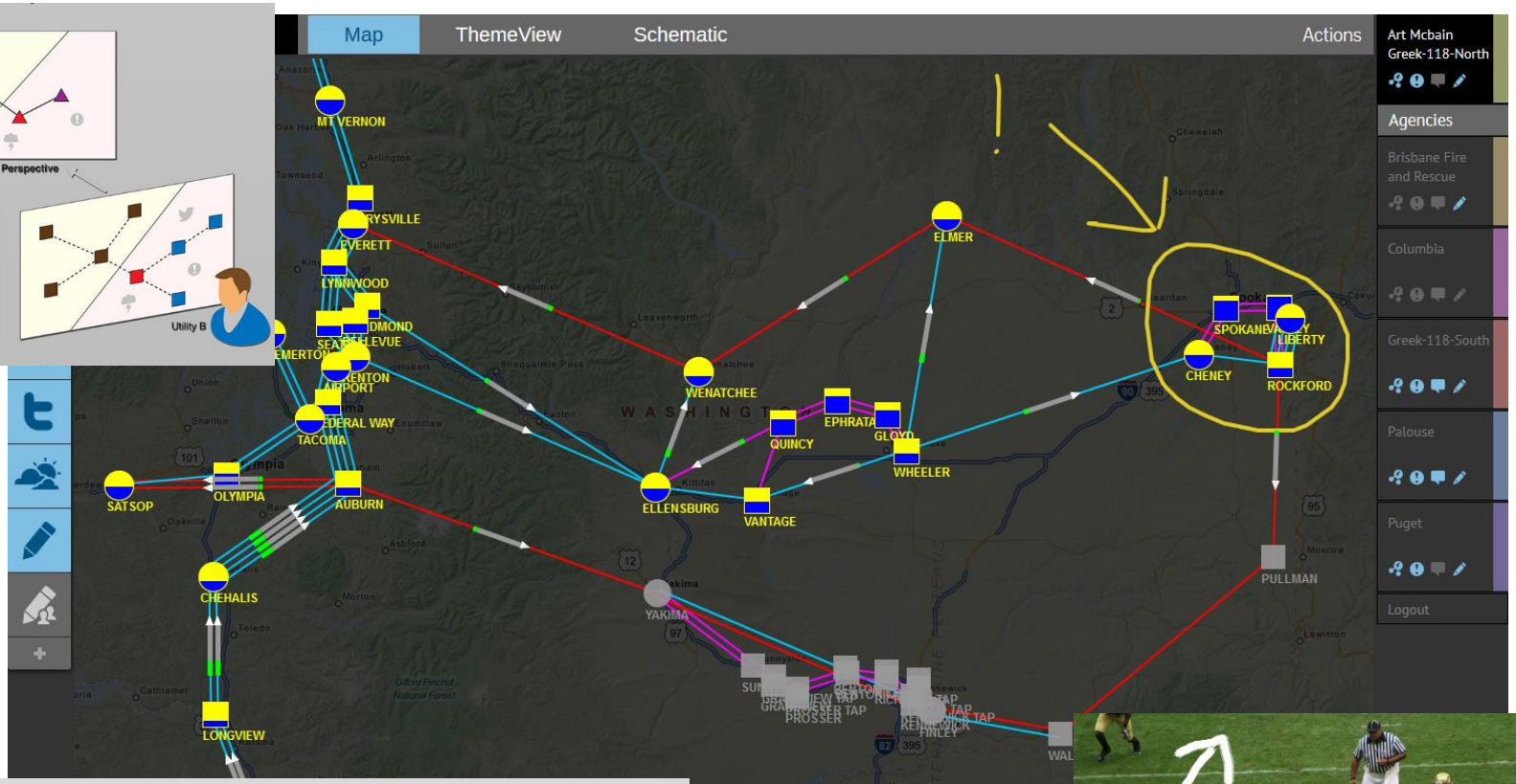
Current tabular format presents data, not information

New visualization tool converts data to actionable info



Shared Perspectives enable real-time collaborative decision making

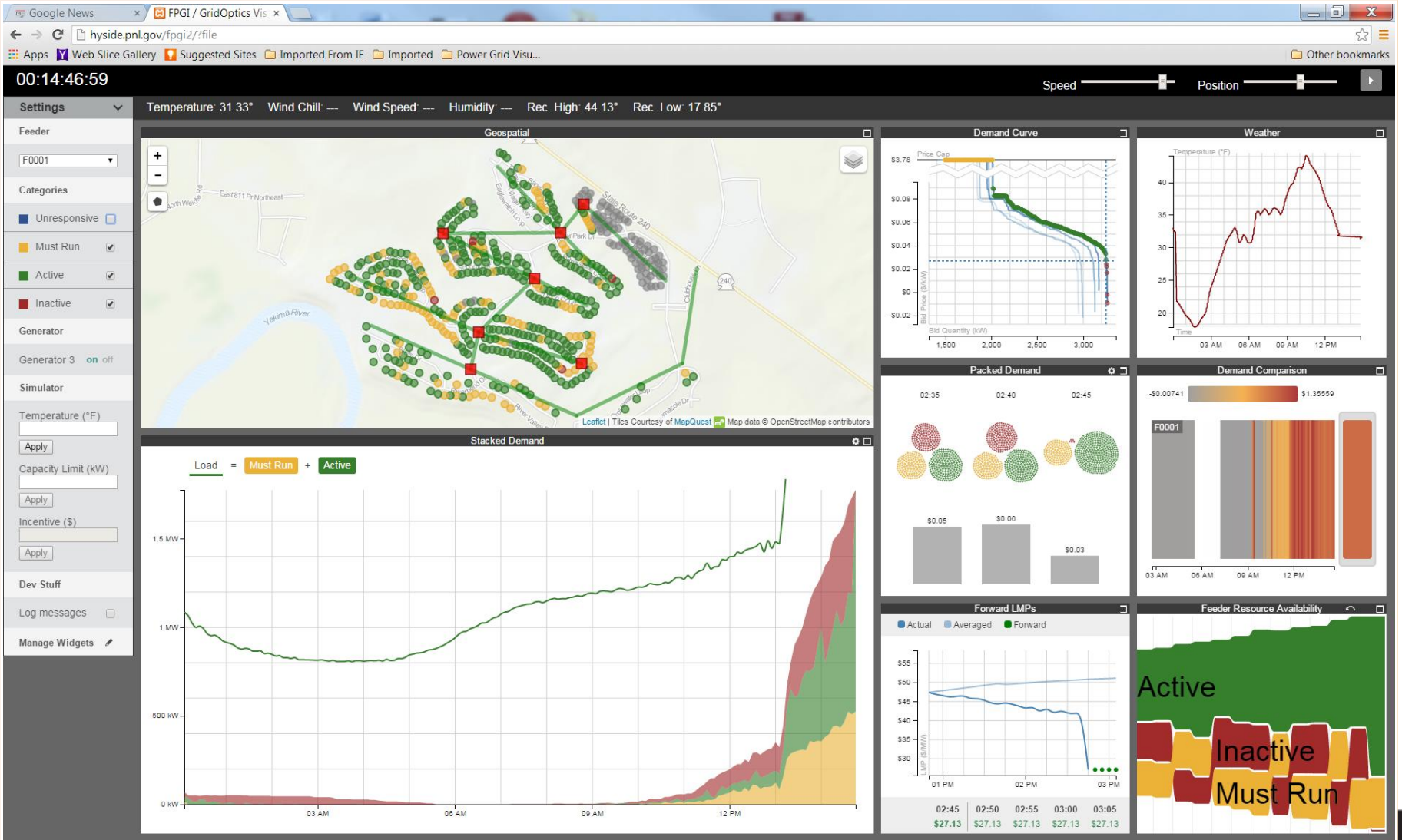




Visual Interface – multiple users communicate in a visual interface

- Hand Annotation – the **telestrator**
- Synchronized View – ensure participants are playing from the same information base

Advanced modular visualization for easy exploration of large-scale data



Summary

- The increasing dependency of grid on data calls for an analytical architecture for converting big data into actionable information.
- GridOPTICS™ is an implementation of this analytical architecture, with building blocks (such as GOSS, GridPACK, and visualization modules) available for application development.
- Data curation, assimilation, and visualization are essential functionality supported by GridOPTICS, achieving high performance.

Acknowledgement

- PNNL Researchers: *(Data and Computing)* Bora Akyol, Poorva Sharma, Yin Jian, Steve Elbert, Shuangshuang Jin, Bruce Palmer, George Chin; *(Power Engineering)* Ruisheng Diao, Yousu Chen, Mark Rice, Jeff Dagle
- Former PNNL Researchers: Terrence Critchlow, Ning Zhou

Questions?

Further Information:
GridOPTICS: http://gridoptics.pnnl.gov/
GridOPTICS™ Software System (GOSS): https://github.com/GridOPTICS/GOSS
GridPACK™ (open-source HPC library): https://www.gridpack.org/
Interactive Visualization and Demo Center: http://vis.pnnl.gov/

Zhenyu (Henry) Huang
Chief Engineer/Team Lead
Pacific Northwest National Laboratory
zhenyu.huang@pnnl.gov

Pacific Northwest
NATIONAL LABORATORY

IEEE PES Power & Energy Society®

IEEE