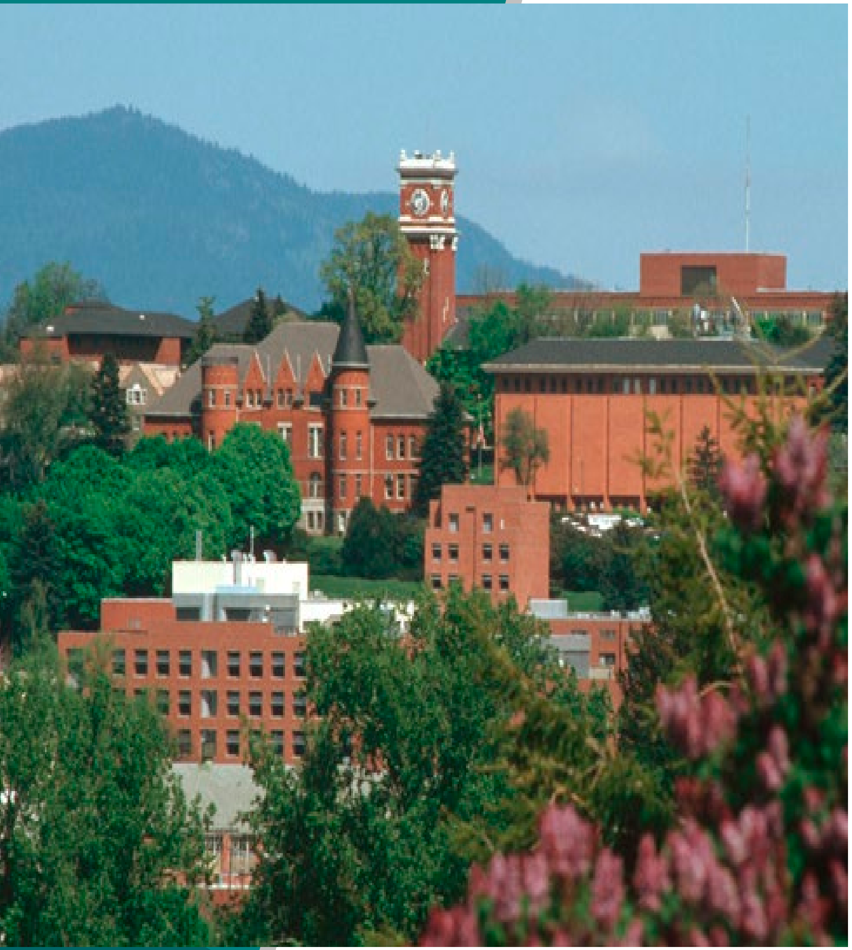
Washington State University
World Class. Face to Face.

College of Engineering and Architecture



The Increasing Data Streams in Power Grid Operation

Anjan Bose
Washington State University
Pullman, WA, USA

IEEE BDA Tutorial Series: Big Data & Analytics for Power Systems

January 24, 2022

Thoughts I want to leave you with

The data landscape for the power grid

- Static data, real-time data, historical data

The system operation data

- Markets: Energy, Capacity
- Transmission: EMS-SCADA, PMU
- Distribution: DMS-SCADA, OMS, AMI

The operation challenges for transforming grid

- Monitoring, operating distributed generation
- Resiliency (including recovery)

Data communication, management, application

Power Systems Data

Fixed Data (Assets)

- 7,500 generation plants
- 75,000 transmission substations
- 300,000 miles transmission (100,000 lines and transformers)
- 2.2 million miles distribution (1 million distribution feeders)
- 300 million customers

Power Flow data for transmission system ~ 2GB

Data for all equipment in PB

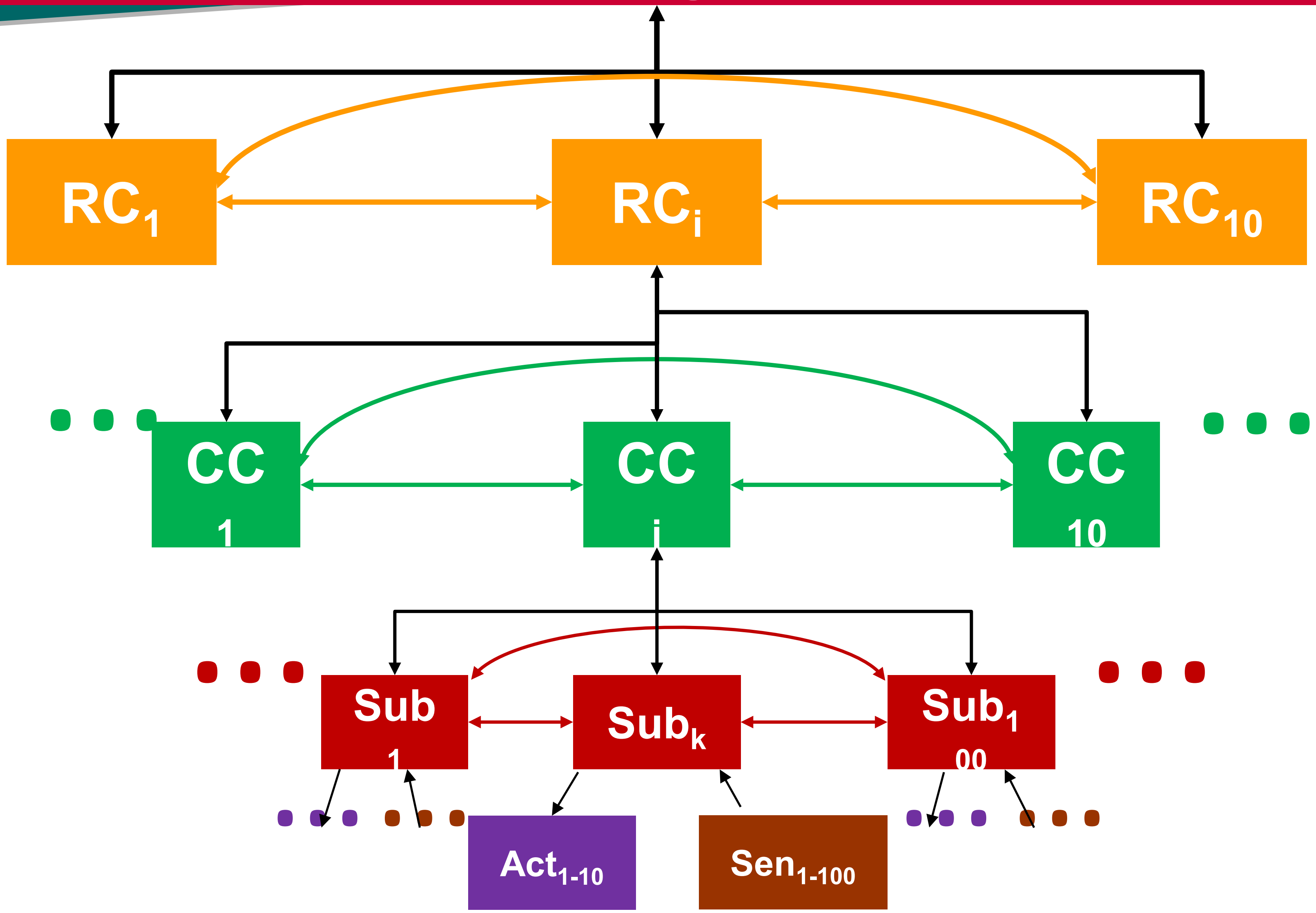
Applications for System Data

- System data for assets is fixed (changes slowly)
- Each engineering application requires a (small) subset of systems data
- Are there applications that require all data?
  - Asset Management
  - Inventory Control
  - Maintenance Records
  - Automated Mapping/Facilities Management
  - Etc.

Measurement Data (Variables)

- Power, Var, Voltage, Current, Frequency, etc.
- SCADA (EMS, DMS)
- PMU data (transmission)
- AMI data (customers)
- Substation, Generation, Microgrid (stored local)
- Measured at various frequencies
  - EMS SCADA at 2-10 seconds
  - DMS SCADA at 10-60 seconds
  - PMU 30-120 times per second
  - AMI 5-15 minutes
  - Substation/Plant data stored at various rates

Eastern Interconnection Control Monitoring Center



Average EMS Data Flows Today

- Average Reliability Coordinator has 10 Balancing Authorities (control centers)
- Average Control Center has 100 high voltage substations
- Average substation has 100 measurement points
- Average polling rate for real time data is 5 seconds

So

- Average data rate from each substation is 20/sec
- Average data rate to a control center is 2K/sec
- Average data rate to a RC is 20K/sec

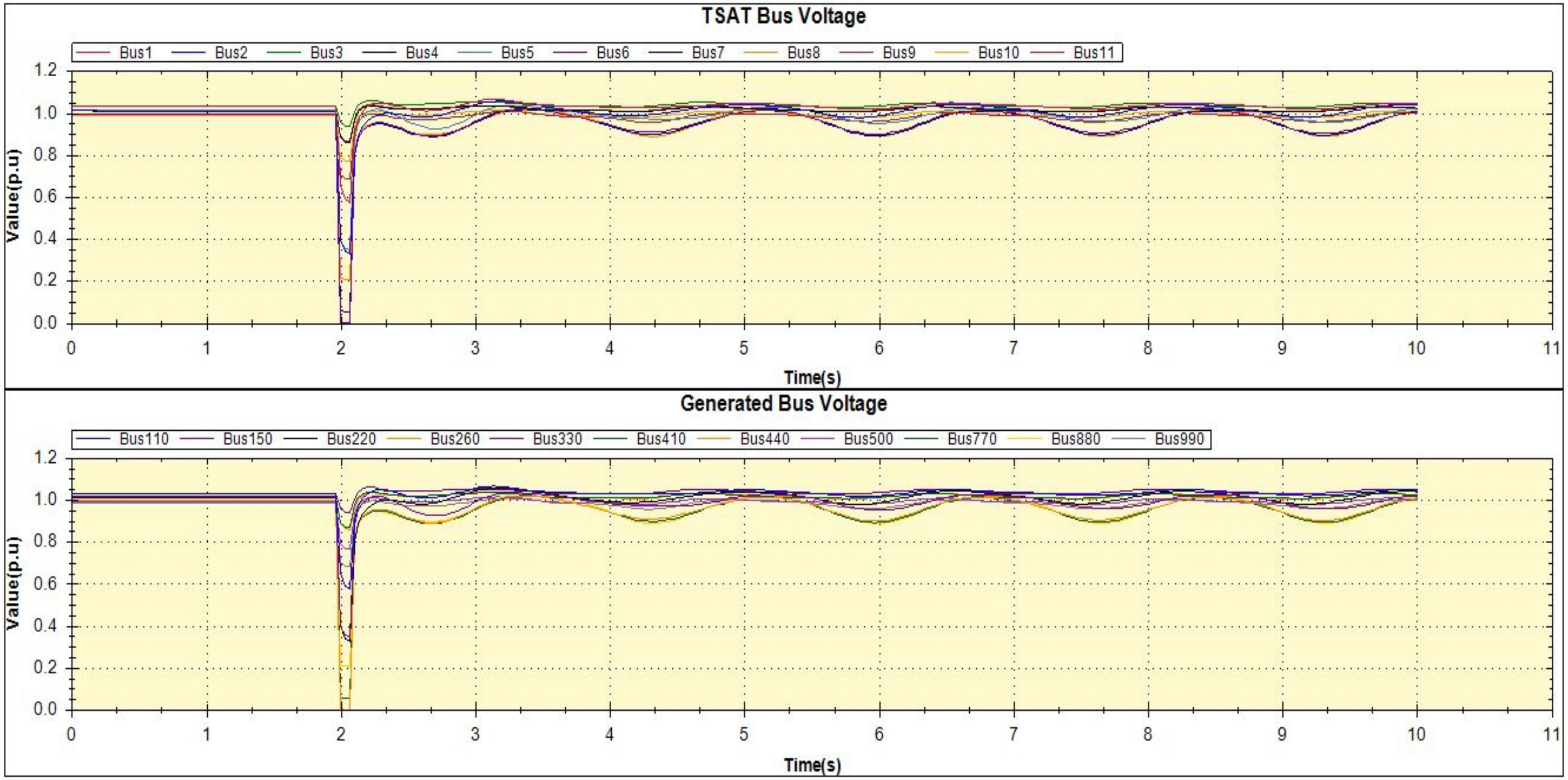
Data Collection by PMUs

- PMU sampling rates: 30-120 per second
- Assume 100 values per second

If we assume all 100 points in a sub are PMUs

- Average data rate per sub is 10K/sec
- Average data rate for the total of 100 subs in a BA is 1M/sec
- Average data rate for the RC is then 10M/sec

Simulated Bus Voltages by Powertech TSAT Generated PMU Measurements 33 msecs time steps

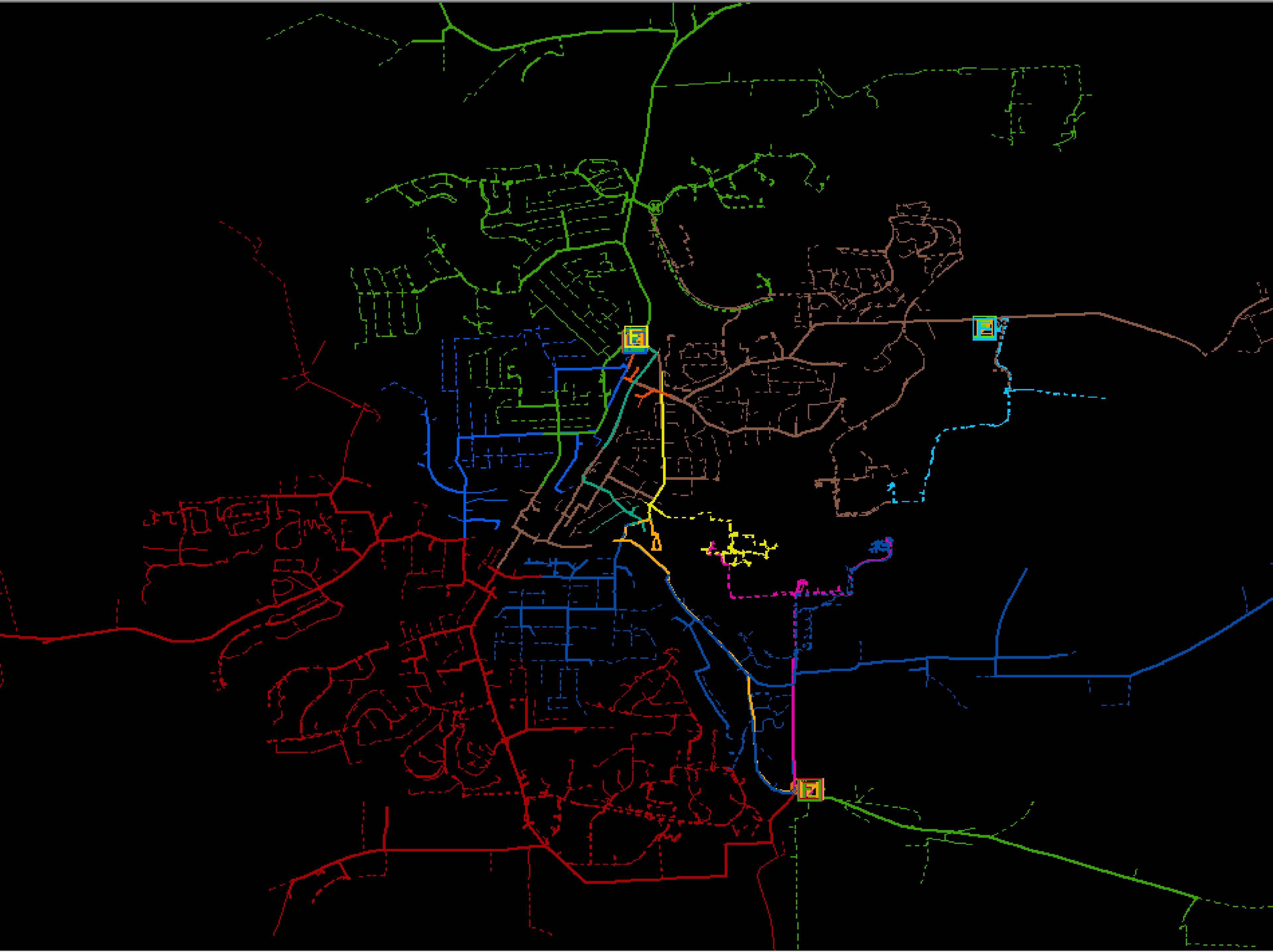


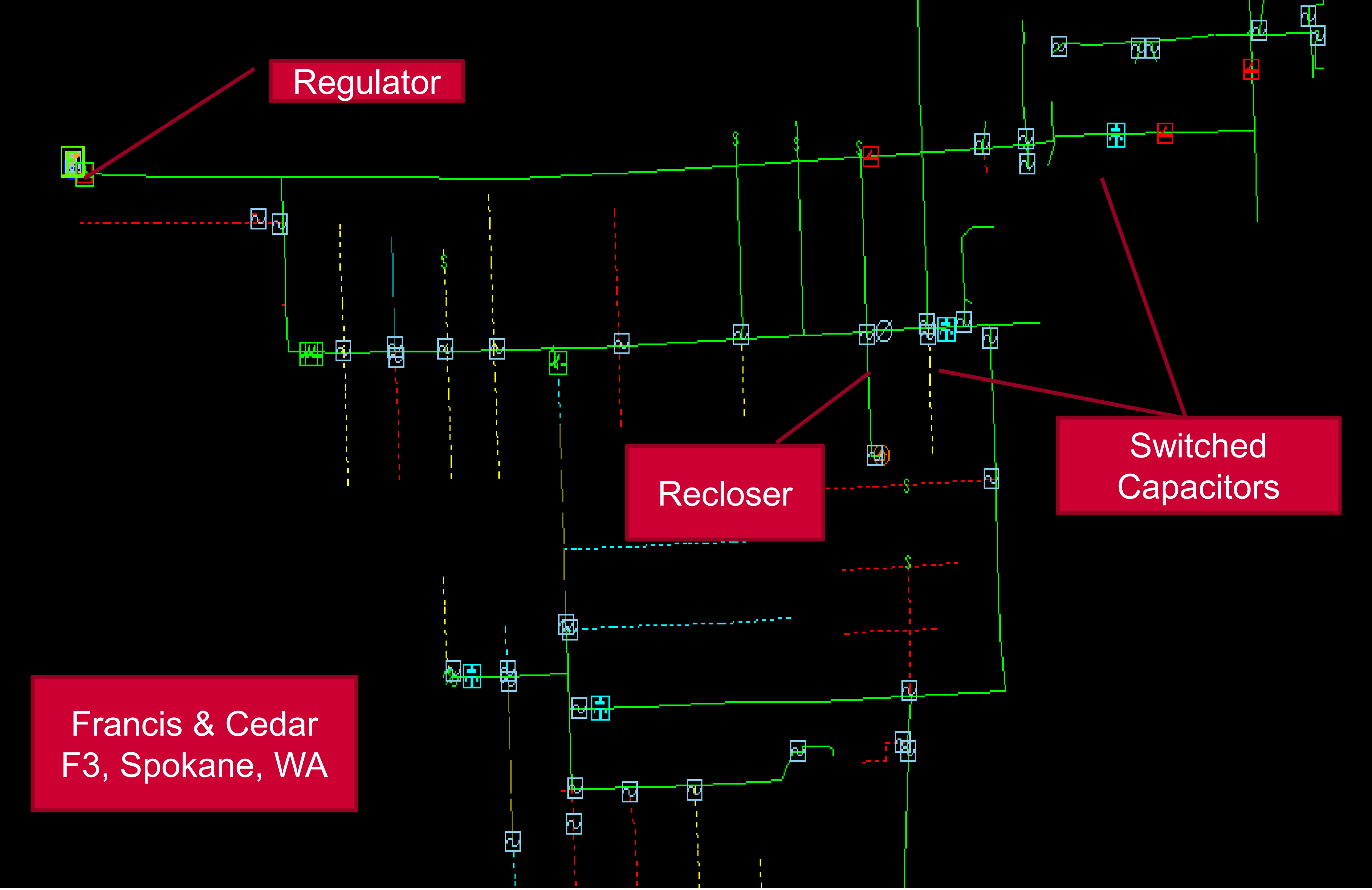
ADMS Data

- DMS SCADA data
  - P, Q, regulator at substation
  - P, Q, status at intelligent breakers (few)
- Outage Management System (OMS)
  - Customer status
  - Estimated status of breakers/sectionalizers
- AMI (customer) data
  - P, Q
  - Usually not available in real time
- No data from primary, secondary transformers

Distributed Generation & Microgrids

- Generation connected to distribution feeders
- Storage connected to distribution feeders
- Generation on secondary transformers (customer side like rooftop solar)
- Storage on secondary transformers (customer side like EV or HVAC)
- Microgrids: grid side or customer side



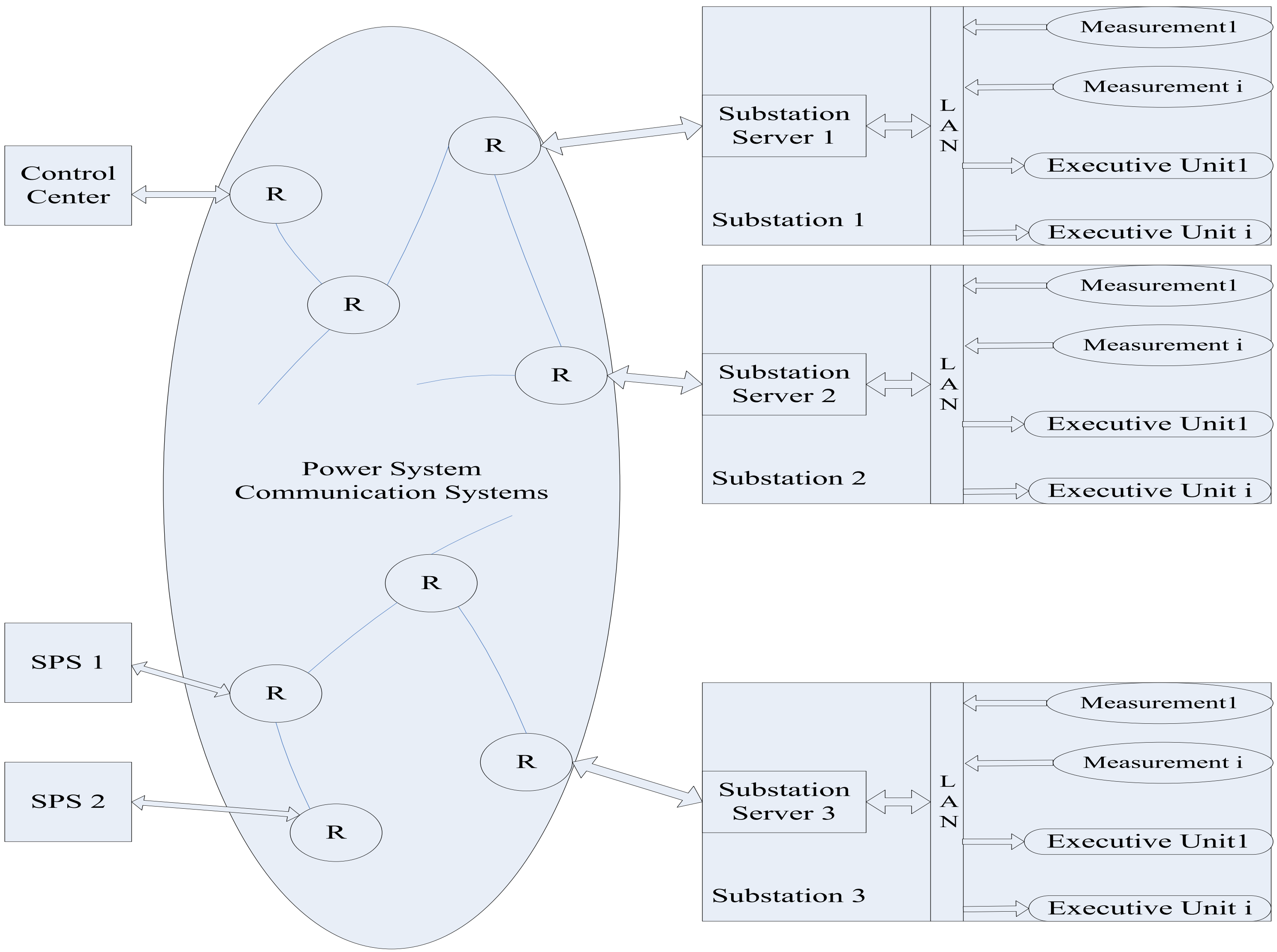

Regulator
Recloser
Switched Capacitors
Francis & Cedar F3, Spokane, WA

Data Exchange Issues

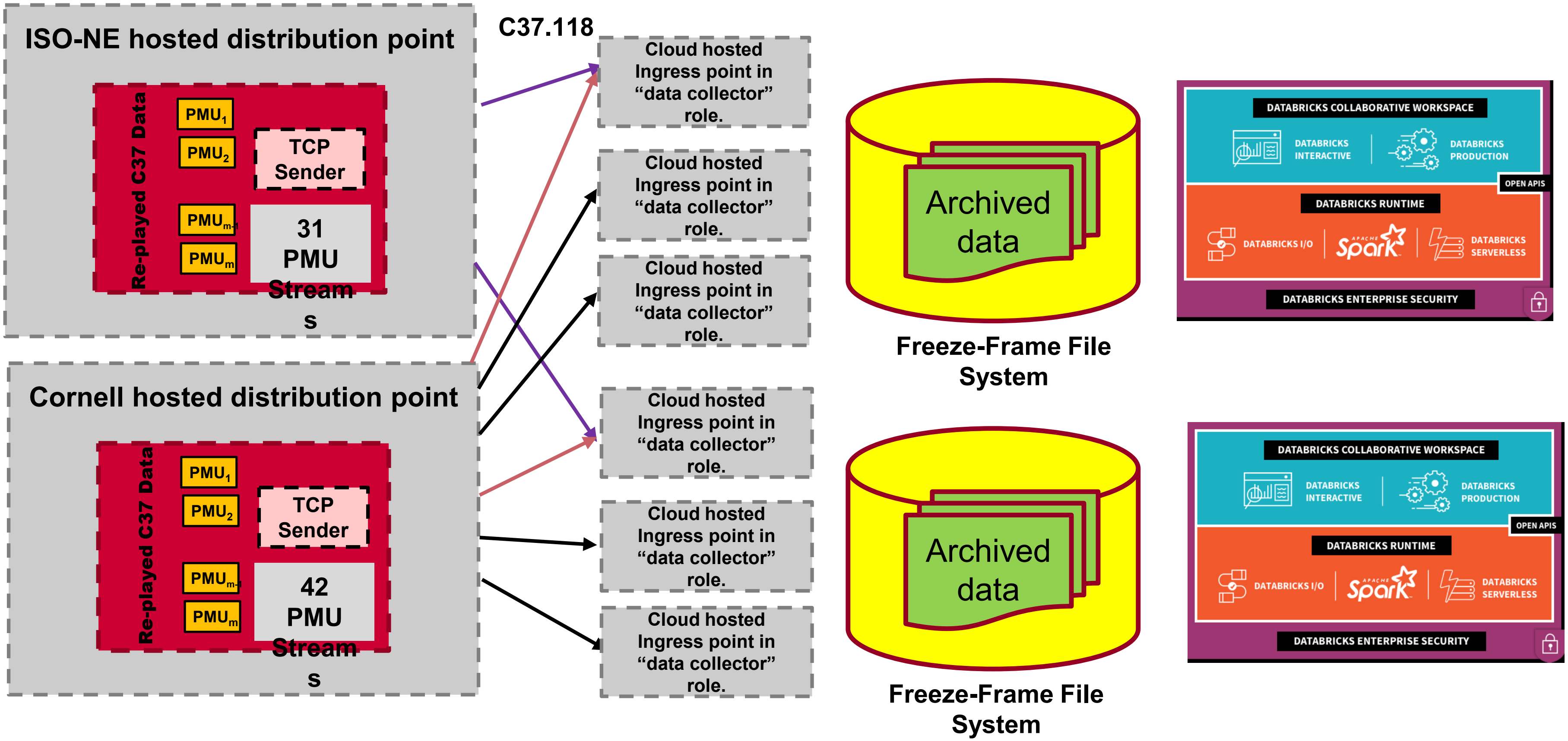
- Within one organization
  - Data movement between EMS, BMS, DMS, OMS, AMI, etc. is non-trivial
- Within one hierarchy
  - Several TOs to ISO
  - Several BAs to RC
- Laterally between neighbors
  - Bilateral agreements too many to be manageable
- Bandwidth, volume, latency

Standardization is the key

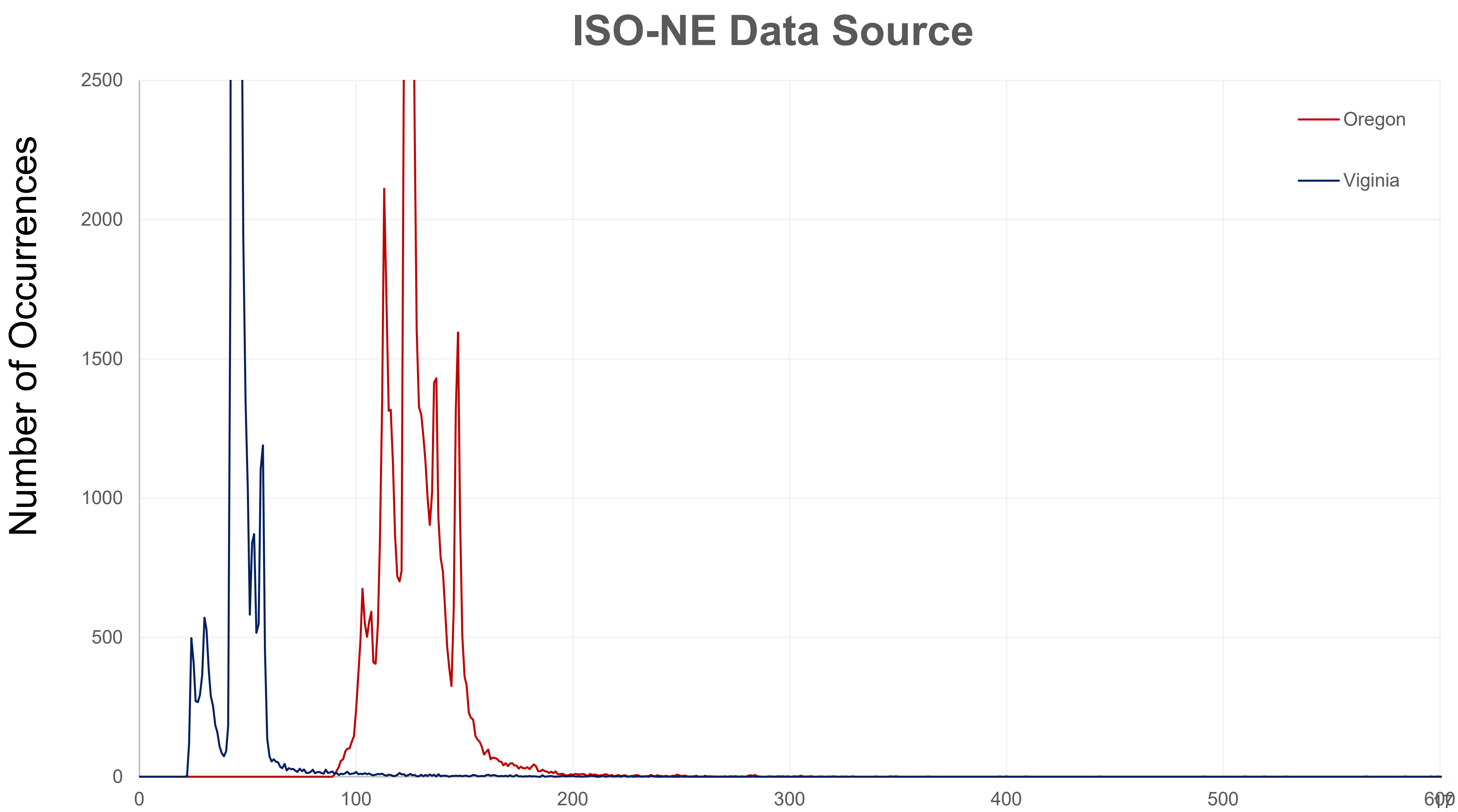
Networked Communications



Cloud WAMS Deployment: Data Archive + Analytics



Histogram: L3 Raw Data Round Trip Latencies



TSO-DSO Data Exchange

Today there is no EMS-DMS data exchange

- EMS and DMS get the same SCADA data from the substation step-down transformer
- There is voice communication between TSO-DSO

When large fraction of generation is on distribution

- More DMS data will be needed in the EMS
- Generation, storage, active loads, microgrids

Data exchange depends on generation mix

Data Base Issues

- **Real time data base must be distributed**
  - **Large amounts of calculated data must be part of this data base**
- **Static data base must be distributed**
- **Historical data base will require still another design**
- **Substation data bases and system level data bases have to be coordinated**
- **All data bases in the same interconnection will have to be coordinated**
- **Standards will be key (CIM – IEC61970)**

Applications

There are dozens of applications needed in the control centers for system operations

- New technologies promise new applications
- Some apps may be distributed (not centralized)
- Data must be stored/moved to support apps

Let us choose three apps to illustrate

- Volt-VAr control
- Power balance/frequency control
- Resiliency/restoration

Volt-VAr Control

- Voltage control requires controlling VAr resources
- VAr resources should be nearby

For Transmission Lines

- VAr sources at same sub (gens, shunts, trans)

For Distribution Feeders

- Main control at source sub
- Boost control with switch caps

If Distribution has gen, microgrid, active load

- More controllability but need coordination
- Transmission & Distribution can interchange VArs

Power Balance (AGC)

- Balancing Authority (BA) calculates ACE for Area
- BA EMS apportions ACE to trans owners (TO)
- Each TO EMS apportions their ACE to their gens

As higher fraction of generation becomes distributed to lower distribution voltages

- Those gen measures come to DMS
- The AGC signals have to be apportioned by DMS
- The DMS-EMS data exchange becomes crucial

Restoration

Restoration of customer service falls mainly on Distribution Operators

- Assessment of damage
- Availability of replacement inventory
- Crew dispatch

Resiliency will depend more on distributed generators/customers/microgrids

- Data availability and communications
- Imitate black start protocols

Control Centers

- The next generation of control centers will have a more flexible (decentralized) architecture
- The boundaries between various XMS functions (including protection and local controls) will fade
- The automatic coordination between entities that are interconnected will increase
- This will require large movement of data both hierarchically and laterally
- This cannot be done without wide adoption of standards across the interconnection

Large Data Applications

Historical Data (Measurement Data Only)

- **Data Science has many usable tools**
- **Identify measurement anomalies (model-free)**
- **Identify trends (loads, renewable generation, outages – equipment failures, control operations)**
- **Identify patterns (correlations between loads, solar, wind, maintenance, external events)**

Large Data Applications

Historical Data (Measurement + Model)

- Event analysis
- Identify measurement anomaly (SE bad data)
- Simulation management
- Planning scenario development
- Training scenario development
- Cross infrastructure analysis
- Controller design

WASHINGTON STATE UNIVERSITY
World Class. Face to Face.

College of Engineering and Architecture

Model Validation using CIM Tool

Fig. Dashboard for the CIM-based Framework

- **Simulation**: The tool populates all the customers (~500) in the GridLAB-D model with appropriate AMI data from PI server.
- **Comparison**: Simulation results are compared with actual SCADA measurements for the corresponding day.
- **Deviation:** Mainly due to inconsistent AMI data
- **Functionality** : Facilitates emulating real-network scenarios

November 2020

December 2020

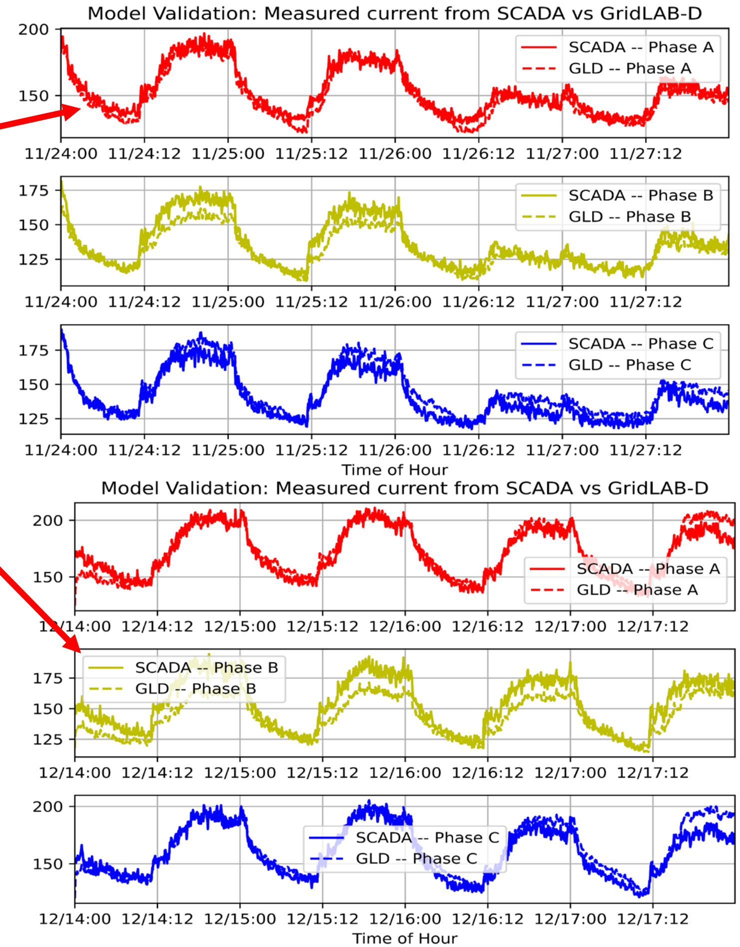


Fig. Model Validation: AMI populated GridLAB-D vs SCADA

Concluding Remarks

- Data science is the use of big data in NEW ways (NOT the use of existing power applications with bigger data)
- Are there uses of data that does not require physical models?
- Are there new applications that take advantage of both big data and physical models?
- Are there on-line applications that can use big data to help operators make decisions?