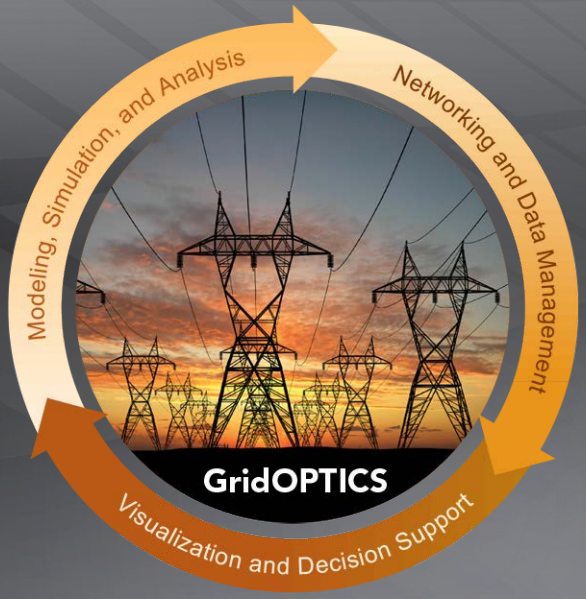


Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

Challenges of Cybersecurity Research in a Multi-User Testbed

THOMAS EDGAR, *THOMAS E. CARROLL*, AND DAVID MANZ

Pacific Northwest National Laboratory

Richland, WA

Cyber-Physical System are Complex

- CPS are large, complex distributed systems, comprising specialized, utilitarian equipment
- Heterogeneous equipment manufactured by multiple vendors
- Systems are expensive, difficult to configure, deploy, and maintain
- Requires expertise
  - Demand exceeds supply

How to support research in the CPS domain?

Challenges for CPS Research

- Equipment must available and accessible
- “Real” data needs to be available
- Researchers shouldn’t be forced to be operational experts
- Experiment in a “safe environment”
- Support for the scientific method
- Enable open science

Multi-user testbeds enable CPS research

Testbeds Support CPS Research

- A *testbed* is platform for experimentation (NSF 2002)
  - *Proof-of-Concept*: Purpose-built for demonstration
  - *Multi-User*: Shared resource pool

- CPS multi-user testbeds should:
  - Be dynamic, flexible, and remotely configurable
    - Researcher-friendly configuration
    - Libraries of scenarios, templates
  - Support concurrent experiments
  - Have broad and diverse pool of real world equipment
  - Be modular, extensible, and scalable
  - Support the research community and open science

Put the Researcher in Control

- World-wide accessibility
- Researcher-friendly interfaces to configure and initialize resources
  - GUI are adequate for small scale experiment
  - ...inefficient when experiments comprise hundreds of components
- Library of common designs, architecture, and designs
- Activation should be on the order of hours, not days
- Mechanisms to simulate “normal”
- … and to orchestrate events, processes, etc.
- Default instrumentation and visualization
- …and other mechanisms to inform and collect system state

Concurrent Experiments Demand Isolation

- Goal is to make efficient use of testbed resources
  - Concurrently running experiments
- Experiments should be isolated from one another
- Depending on constraints, minimize shared resources
  - Separate management from experiment
  - Support infrastructure, services duplicated per experiment
  - CPS equipment reserved for a single experiment
  - Virtual machine monitors per experiment
- Some resources must be shared
  - E.g., Network infrastructure
  - CPS devices, cannot separate the management from experiment
- Effects of sharing must be documented and quantified
- Method to reserve testbed for single researcher
- Resources returned to initialize state on experiment termination

Sensitive Data Must Be Protected

- Organizations demand we protect sensitive data
  - Architecture and design are often considered proprietary
  - Data often contains system state information
- If data is released, may harm or embarrass organization

- Testbed must enforce access controls on the data
- Obscure experiment designs
- Anonymize data employing a scientifically valid approach
  - Paul Ohm's law: "data can either be useful or perfectly anonymous but never both."
  - Several examples of anonymous data that were re-attributed

Testbeds Facilitate Reproducibility

- *Reproducibility* is the condition that allows a skeptic to independently verify results
- From a theory/model the researchers define a *system under test*
- Description of the system is the testbed configuration
  - What resources were used
  - Initial configuration
  - Connectivity between devices, characteristics of links
  - Operating system images, device firmware
  - Logs and serial, network traffic capture
  - Parameters for simulated components
- Unfortunately, uncertain what this means for physical processes…
- Provide mechanisms for researchers to share experimental designs, data, and documentation

Testbeds Can Enhance External Validity

- While scenarios/template greatly enhance external validity
- …fidelity, equipment, and scale are challenges
- Real always best, but not always possible
- Put the researcher in control of fidelity
- Combine the real with emulated and simulated
  - Procure broad and diverse set of equipment
  - Federate with other testbeds to gain access to additional resources
  - Emulate and simulate other components
- Simulation should be scientifically valid and researchers aware of shortcomings
- Some progress on simulated physical processes
- Bring everything together for experimenting on large-scale systems

powerNET Features

- Project-/program- based access controls
- Remote configuration/execution of experiments
  - Web application
  - Configure using GUI/declarative language
- Network emulation/simulation
  - DS, SONET, dial-up, wireless
- Phasors
  - 9 PMUs from multi-vendor/
    1 PMU development platform
  - 1 Hardware PDC/Many software PDCs possible

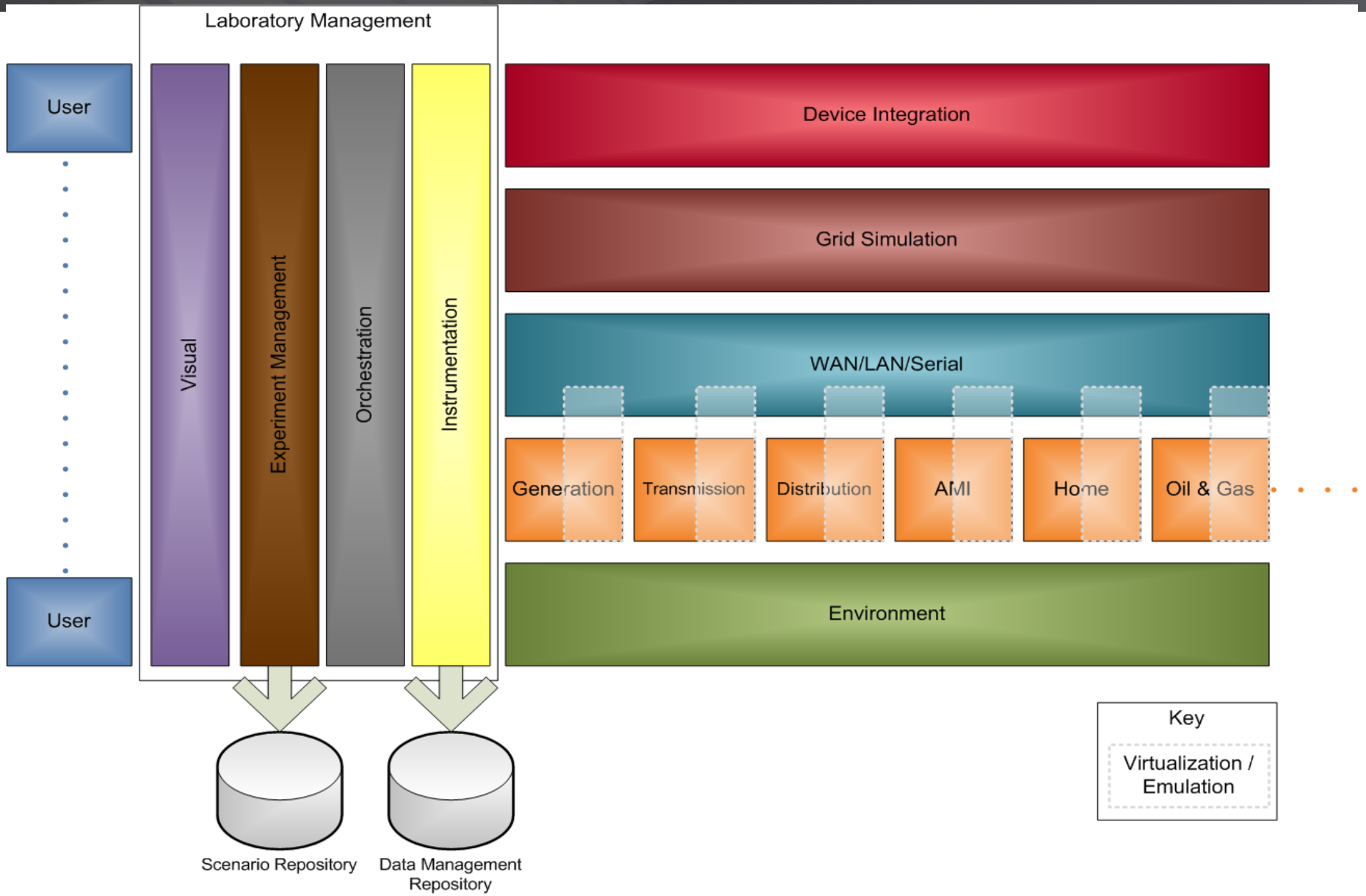
powerNET Features (cont)

- More than 250 virtual nodes possible
- Energy management system (in progress)
- Advanced metering infrastructure (in progress)
- Compute Cluster
  - 3 nodes with SSDs and Infiniband interconnects
  - Scale experiments to thousands of nodes
- 64 TB high-speed shared storage

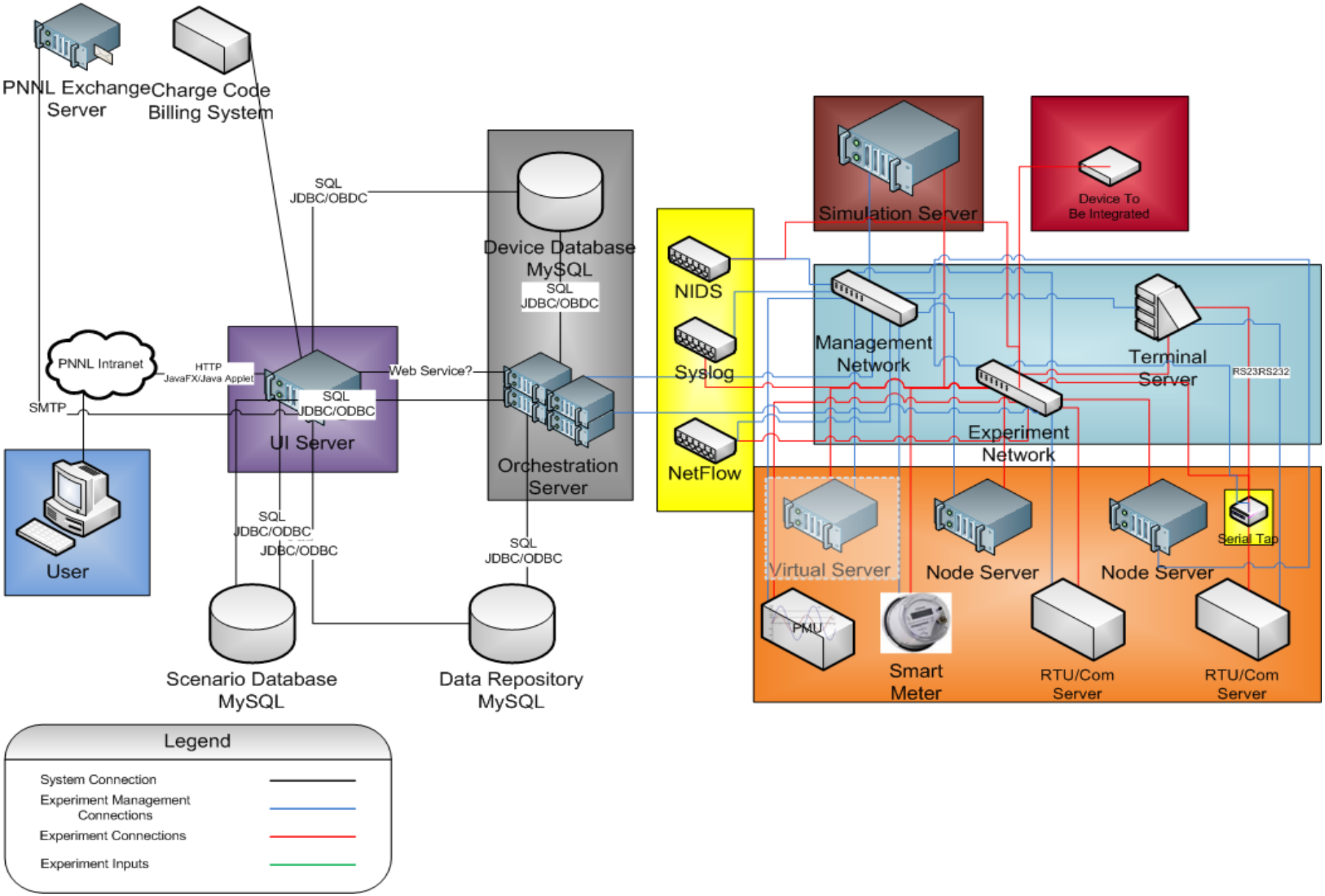
Use Cases

- Validation and verification
- Technology assessment and prototyping
- Simulation and modeling
- Training and education
- Demonstration

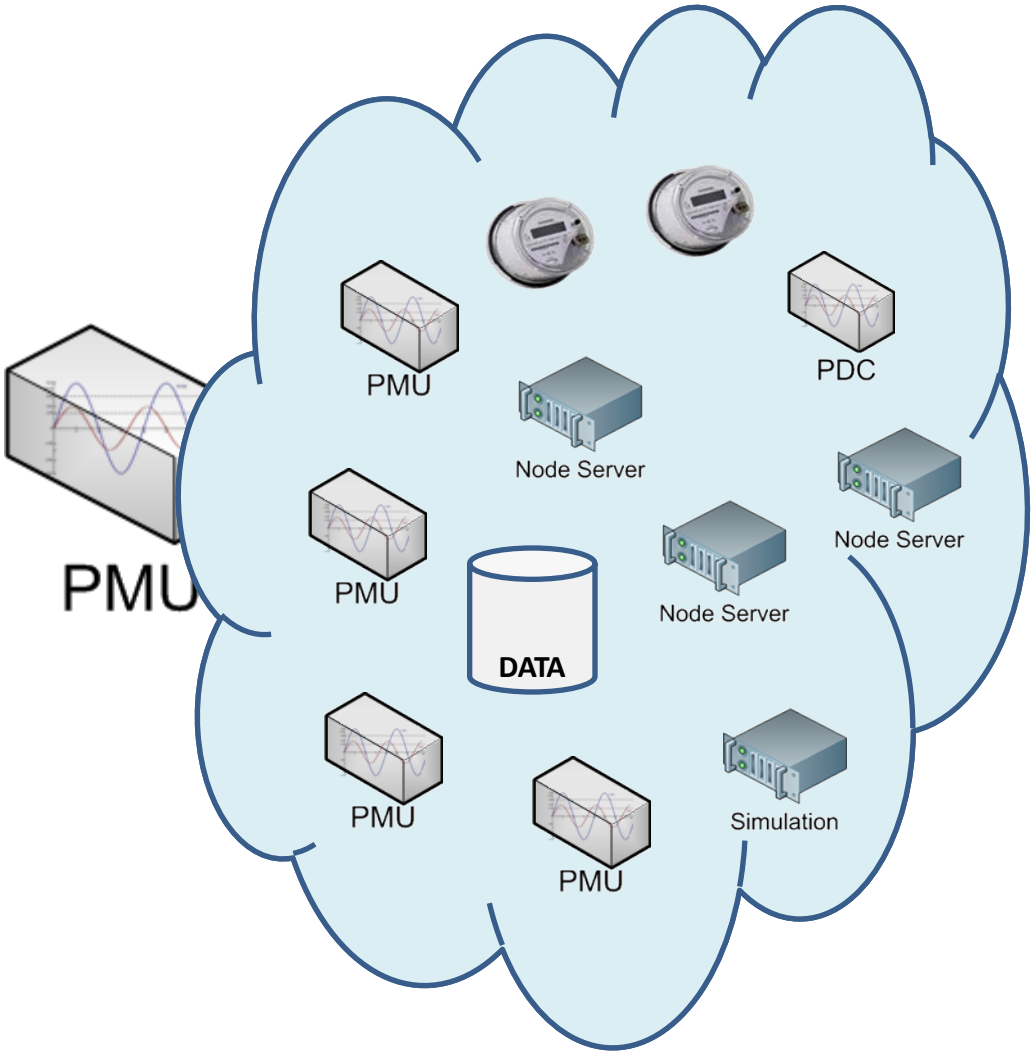
Logical Description



Technical Architecture

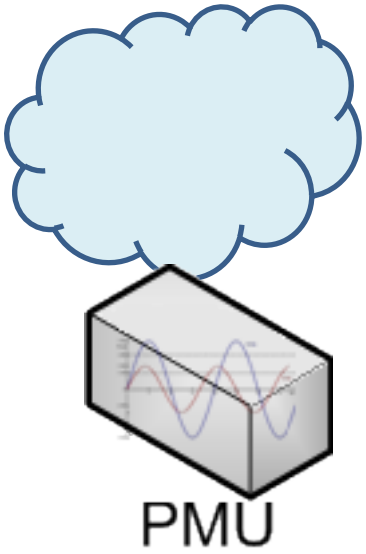


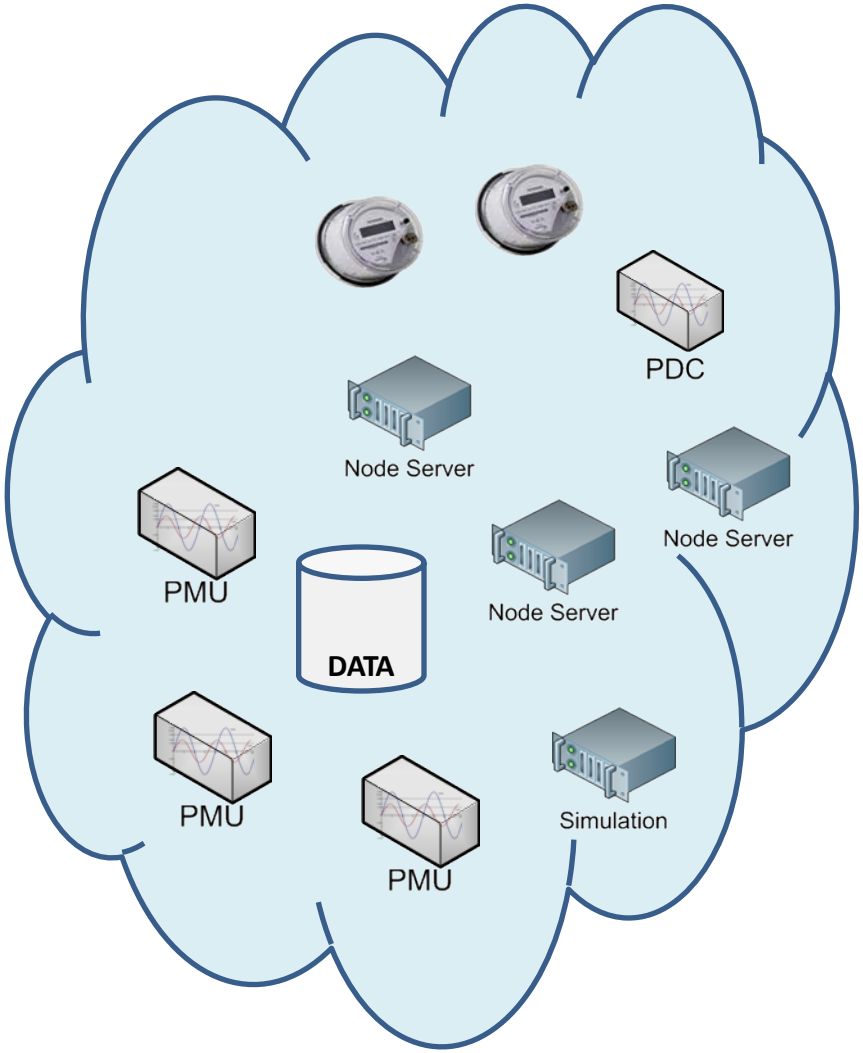
Testbed Operation



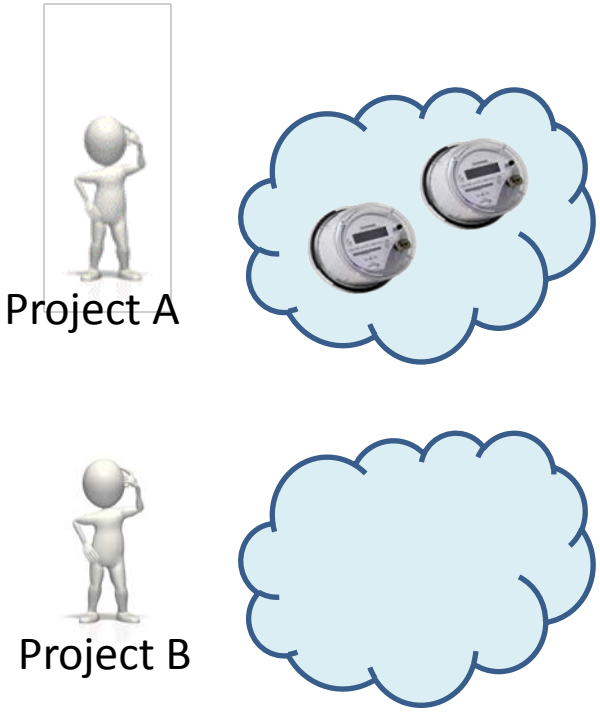
Testbed Operation

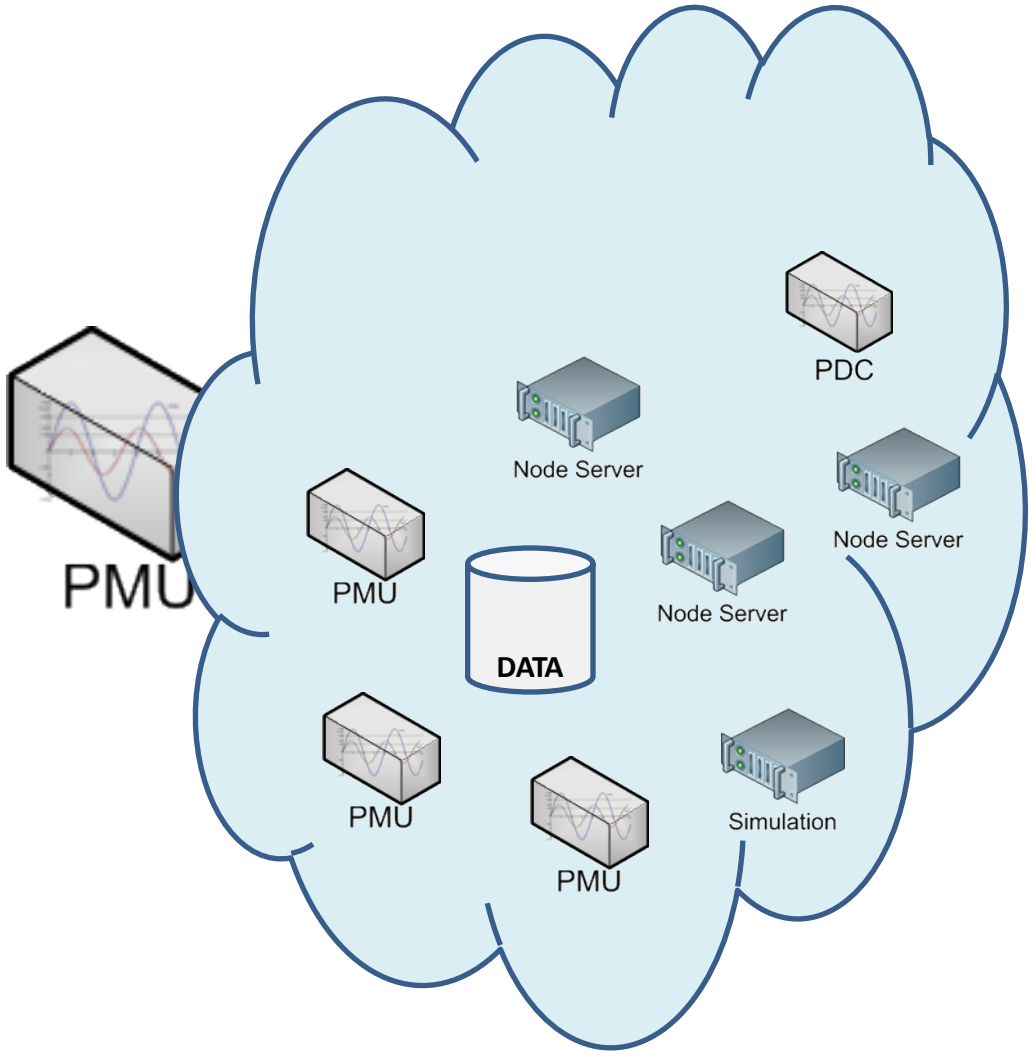
Project A



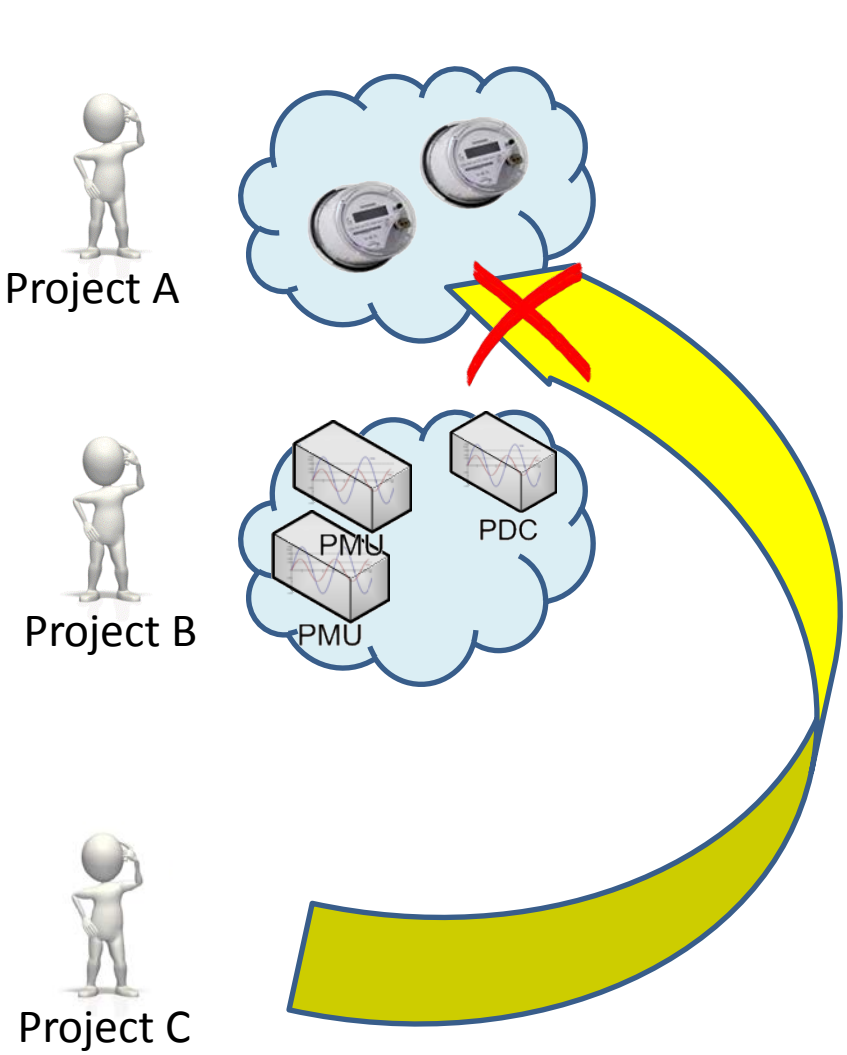


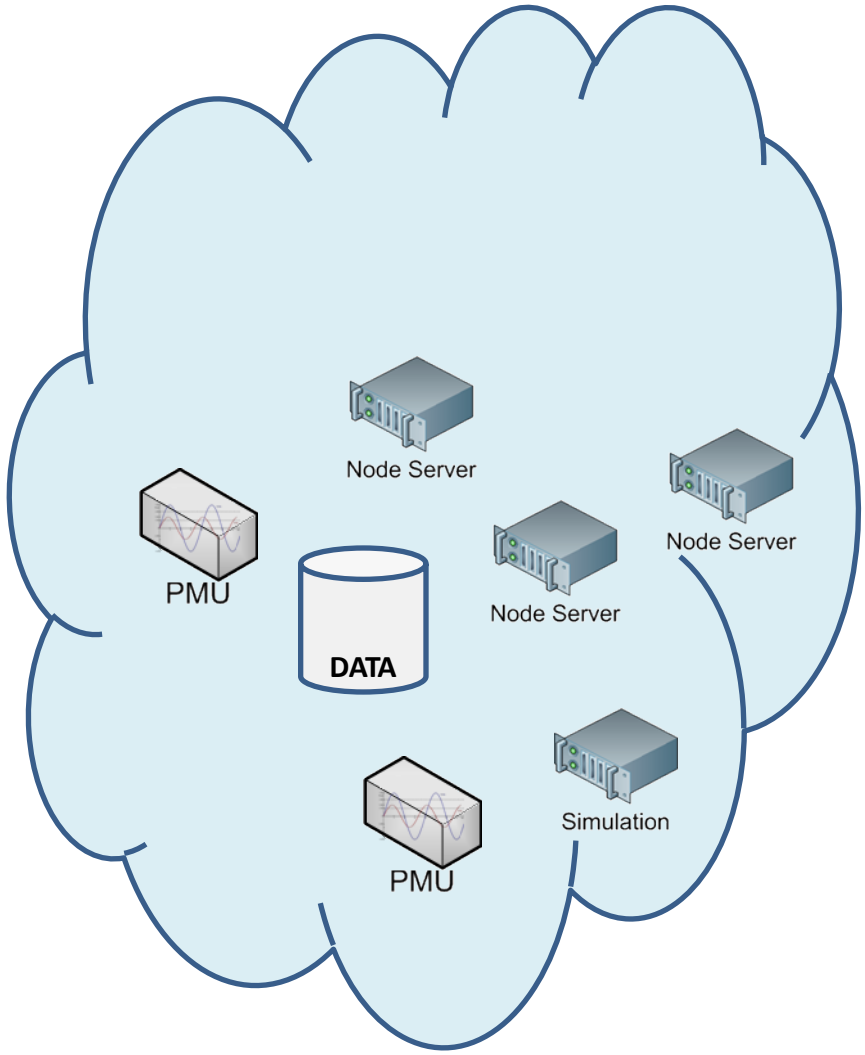
Testbed Operation





Testbed Operation

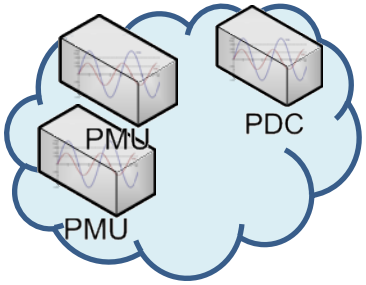




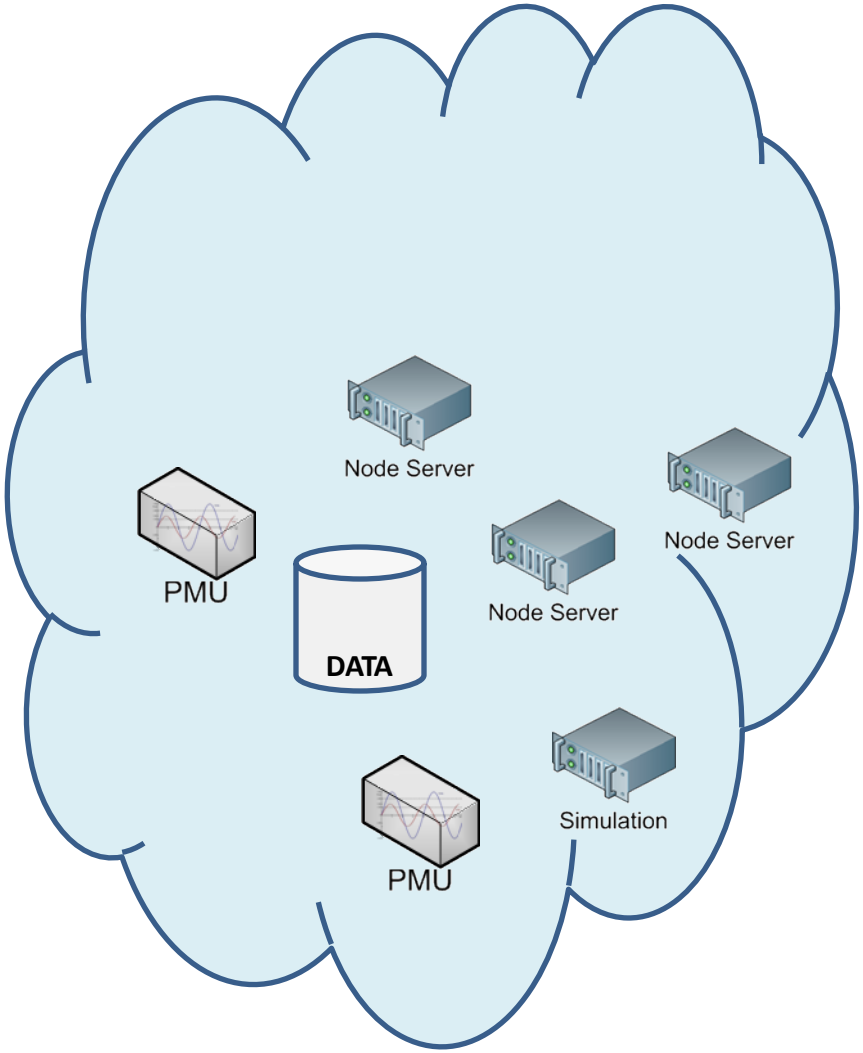
Testbed Operation

Project A

Project B

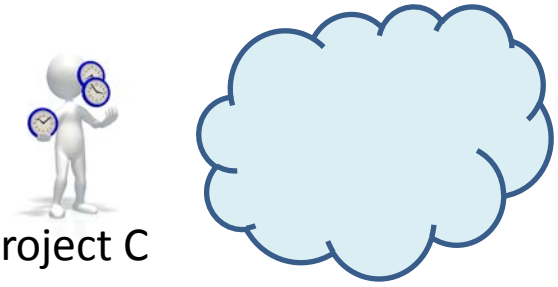


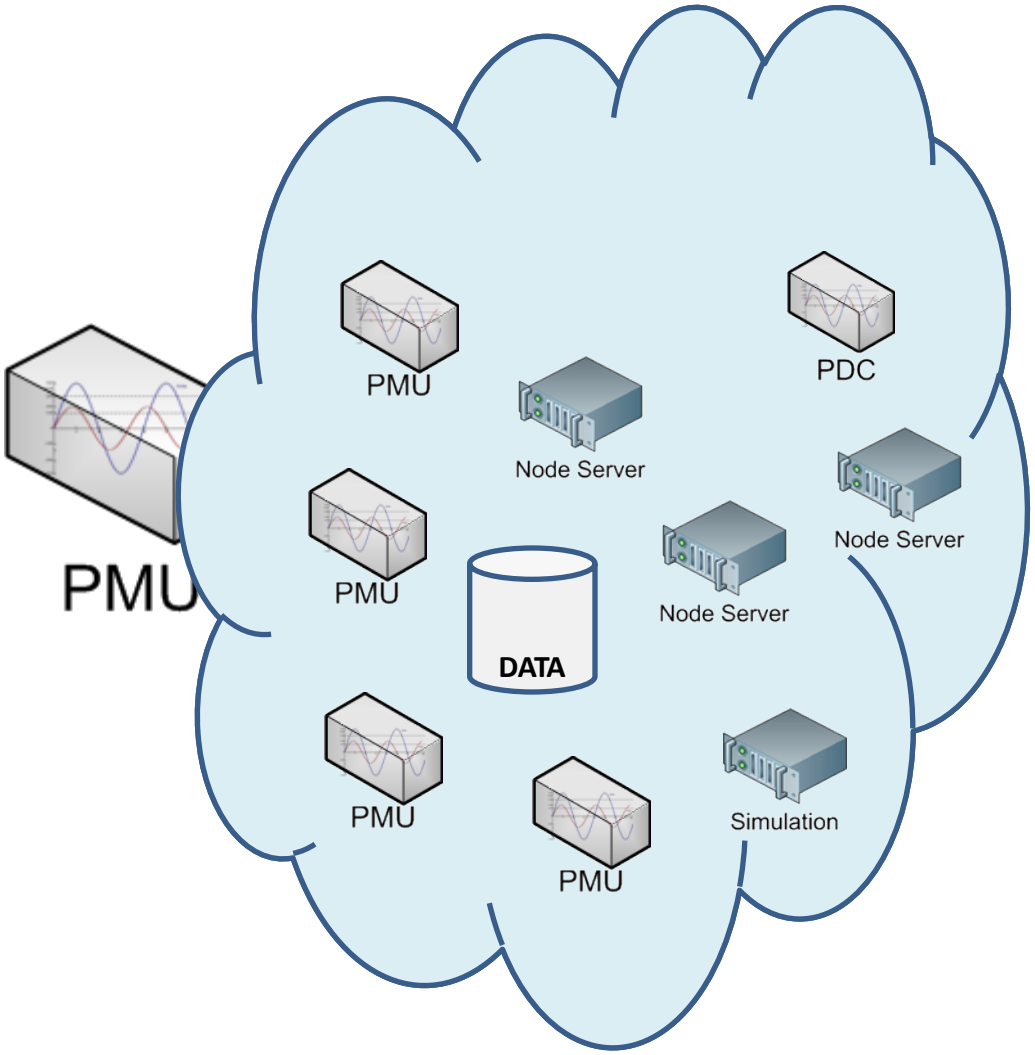
Project C



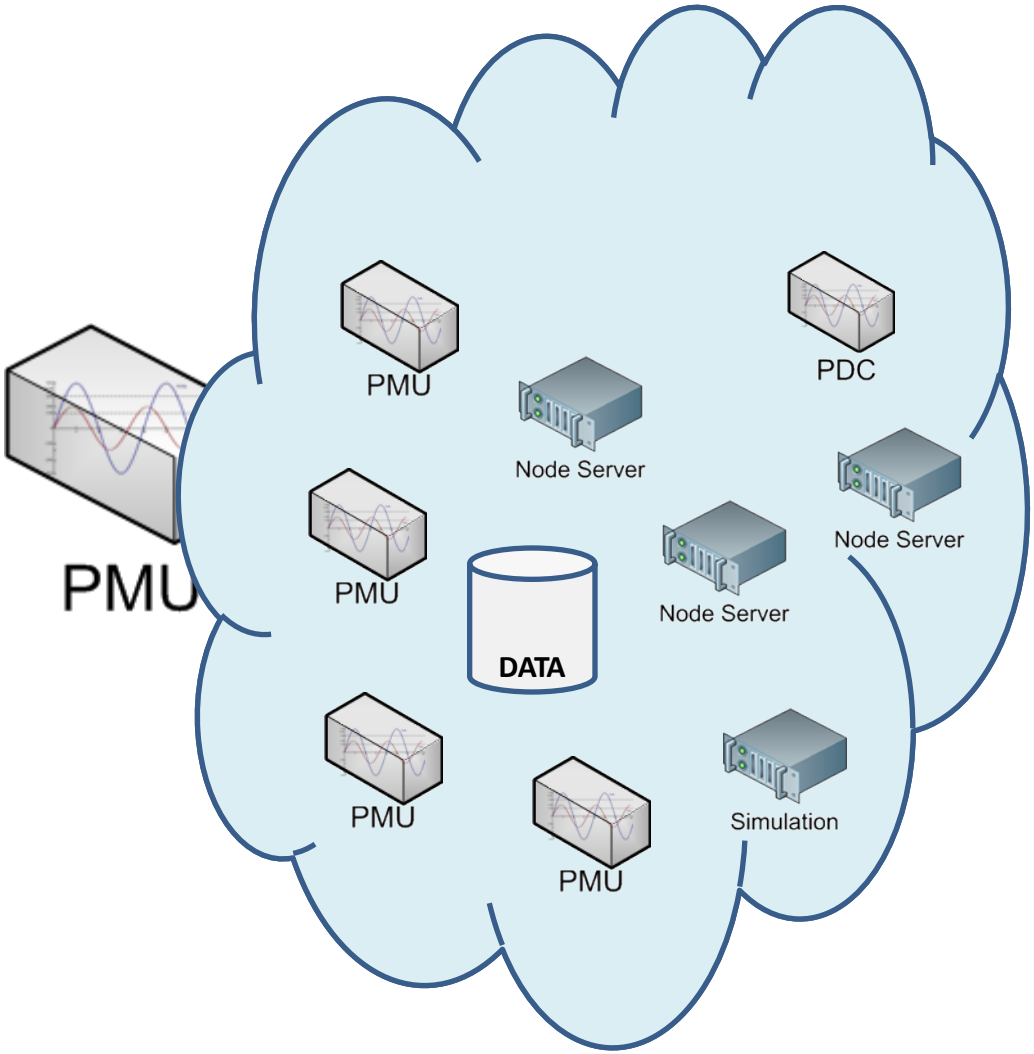
Testbed Operation

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*





Testbed Operation



Project C

powerNET: Researcher Driven Control

- Researchers remote connect to facility
- Can configure through a GUI or a descriptive language
  - Initial device configuration and impediments
  - Provide templates and scenarios
- A subsystem for event orchestration

powerNET: Isolation

- Program/Project access controls
- Resources are reserved by and dedicated to experiments
  - …including virtual machine monitors
- Resources are wiped and re-initialized to a known good state between uses
- Separate control/management traffic from experiment traffic
  - Leverage multiple NICs in devices
- Experiments are isolated from one another using VLANs
- Authentication/authorization resources are duplicated
- Devices cannot communicate directly with one another on the control network
- Data access controls mapped to NFSv4 acls and data confidentiality/integrity provided by NFSv4 and CIFS

Isolation: Reproducibility

- Researchers are free to export their data from the facility
  - Try hard to store data in standard formats
  - Sometimes restrictions on images/firmware
- Community portal/wiki that assists communication between researchers
- Provide archive in support of open science
  - Storage experimental designs, configurations, and data

powerNET: External Validity

- A current focus on PMUs, PDCs
- …in talks with other equipment vendors
- We are federated with other testbeds within PNNL
- …and are in the process of federating with DETER and UIUC

Conclusion

- Cybersecurity research in CPS has high barrier of entry
- Testbeds can ease the burden by providing access and enhancing reproducibility and external validity
- Testbeds create new challenges such as isolation and data protection

- As a community, we need
  - Scientifically valid approaches for simulating devices and physical processes, synthesizing normal activity and data
  - Access to real data
  - …Scientifically valid approach, with acceptable risk, for anonymizing data