Cyber Infrastructure for the Power Grid

Dr. Anurag K. Srivastava, Dr. Carl Hauser, Dr. Dave Bakken

Computation Lecture 6: Grid Wide Area Measurement Systems
Towards Trustworthiness of Wide-Area Smart Grids

Prof. Dave Bakken
School of Electrical Engineering and Computer Science
Washington State University
Pullman, Washington, USA

Technische Universität Darmstadt
Darmstadt, Germany
September 22, 2014
IANAPP (power person): Computer Scientist
- Core background: fault-tolerant distributed computing
- Research lab experience with wide-area middleware with QoS, resilience, security, …. for DARPA/military
- WSU is #3 in power in USA
- Working with Anjan Bose since 1999 on wide-area data delivery issues appropriate for RAS and closed-loop – GridStat
Sources of Info (1)


  ▪ Best Paper Award for “Connectivity” track. This is the official communications/interoperability meeting for the pseudo-official “smart grid” community in the USA, namely DoE/GridWise and NIST/SmartGrid.

• Workshop on Trustworthiness of Smart Grids (ToSG): www.tosg-workshop.org
Sources of Info (2)

Outline

• Questions for Power Engineers & Researchers
• WAN Apps with Extreme Comms Requirements
• IT Guidelines for Achieving these Requirements
• GridStat: Industrial Internet for Electricity (IIE)
• GridStat Cyber-Physical Example
• Cyber-Security for Closed-Loop Apps
• Why Middleware is Needed for IIE
A Crucial and Wide Open Issue

• How do anomalies in ICT affect power “stability”?  

• Wide open: almost completely unstudied 😞  
  ▪ Lars Nordstöm of KTH best power researcher at this  
  ▪ Still a huge opportunity for research topics!

• Trying here to plant seeds to break chicken-egg  
  ▪ Power researchers can assume much “better” data delivery to come up with “better” apps  
  ▪ Computer scientists can come up with even better data delivery but need to know killer app requirements and acceptable tradeoffs (there are always tradeoffs!)  
  ▪ Data Analytics scientists can come up with better analytics given the tradeoffs and assumptions above
Comms Baseline: You Can Assume
• Data delivery over WAN can be (with GridStat etc):
  ▪ Very fast: less than ~1 msec added to the underlying network layers across an entire grid
  ▪ Very available: think in terms of up to 5+ 9s (multiple redundant paths, each with the low latency guarantees)
    – Even in the presence of failures!
  ▪ Very cyber-secure: for long-lived embedded devices and won’t add too much to the low latencies
    – E.g., RSA adds >>60 msec so not for RAS or closed-loop
    – Shared keys (61850-90-5): subscriber can spoof publisher 😞
    – GridStat solution not vulnerable and only adds 10msec
  ▪ Tightly managed for very strong guarantees (MPLS)
  ▪ Adaptive: can change pre-computed subscriptions ~INSTANTLY (and others FAST)
Questions to Ask Yourself

• How can power researchers exploit this better communications infrastructure?

• What rate and latency and data availability does my power app really need for remote data?
  ▪ Why fundamentally does it need that?
  ▪ How sensitive is it to occasional longer delays, periodic drops (maybe a few in a row), or data blackouts for longer periods of time?

• Can I formulate and test hypotheses for the above?
Beyond Steady-State-Only Thinking

• Previous is just for steady state: different in some contingency/mode situations?

• How important is my app \textit{in that given contingency/mode, compared to other apps}?
  ▪ E.g., simple “importance” number \([0,10]\)
  ▪ How much worse (latency, rate, availability) can I live with in steady state and in given contingencies?
    – But would still get strong guarantees at that lower quality
    – How much benefit do different levels really give me?
  ▪ Can I program my app to run at different rates, or is there a fundamental reason it has to run at one?

• What extra data feeds (or higher rates etc) could I use in a contingency/mode (could get in \(<< 1\text{ sec})\)
Bad Data

• How vulnerable is my power app to bad data?
  ▪ State estimation obviously has handed for many decades

• But
  ▪ How much bad data
  ▪ Does how much bad
  ▪ In what circumstances?

• E.g. can I specify assumptions about bad data?
  ▪ Number: absolute or (better) a function of the problem size
    – state/configuration/PMUs/etc
  ▪ Location: randomly distributed or worst case?
  ▪ Error degree: randomly off (what probability distribution) or worst case (from an adversary)?
Bad Manners

• How vulnerable is my power app or RAS scheme or stability assumptions to worst-case malicious behavior?

• E.g. not just false data (which may be able to be detected) but taking over command of a relay or other devices
  ▪ How many of these, and of what kind, could cause problems?

• Thinking cyber-physical here
  ▪ What are some worst case combinations of a physical attack (rifle, chaff, modifying sensors, ..) and a cyber attack (colluding customer meters, taking over relays, DDOS to throttle delivery of sensor data and commands,
  ▪ And worst case under what situations?
Bad Manners (cont.)

• “The event I fear most is a physical attack in a successful cyber-attack on the reconjunction with responders' 911 system or on the power grid,”

A Cloudy Forecast

• What could I do with cloud computing, assuming it is made mission critical, i.e.:
  ▪ Keeps same fast throughput
  ▪ Does not allow deliberate “inconsistencies”
    – e.g., a replica does a state update never received by others
  ▪ Is much more predictable with CPU perf., ramp-up time, …
  ▪ (BTW, ARPA-E GridCloud proj. w/Cornell+WSU doing for >2 years)
  ▪ Not all CPUs in datacenter, some (managed) in substations…

• How could I use
  ▪ Tens/Hundreds of processors in steady state
  ▪ >>Thousands when approaching/reaching contingencies
  ▪ Data from ALL participants in a grid enabled quickly when approaching a crisis

• Backup slides on killer cloud apps
CIP-Managed **Compute+Comms+Security**

- Computations + communications + security *can* be
  - Mission critical to power grid specs
    - Closed-loop WAN app requirements **WAY harder** than air traffic control, railways, military, ...
  - Changed rapidly in a coordinated manner
    - Providing app developers *much higher-level building blocks*
  - Managed in a network operations center 24x7
    - Much like a power control center
    - Needed if power grid stability really does depend on comms and computation and cyber-security
Outline

• Questions for Power Engineers & Researchers

• **WAN Apps with Extreme Comms Requirements**

• IT Guidelines for Achieving these Requirements

• GridStat: Industrial Internet for Electricity (IIE)

• GridStat Cyber-Physical Example

• Cyber-Security for Closed-Loop Apps

• Why Middleware is Needed for IIE
Wide Range of QoS+ Requirements

• QoS+:
  ▪ network/middleware “QoS” (latency, rate), availability/criticality
  ▪ Also things an implementer/deployer of WAMS-DD needs to know: geographic scope, quantity.

• Comparing Apples and Apples:
  ▪ Normalize each from 1 (very easy) to 5 (very hard)

• Wide ranges
  ▪ Across application families
  ▪ Sometimes within them (each configuration is different)
<table>
<thead>
<tr>
<th>Difficulty (5 is hardest)</th>
<th>Latency (ms)</th>
<th>Rate (Hz)</th>
<th>Criticality/Availability</th>
<th>Quantity</th>
<th>Geography</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5-20</td>
<td>&gt;240</td>
<td>Ultra</td>
<td>Very High</td>
<td>Across grid or multiple ISOs/RTOs</td>
</tr>
<tr>
<td>4</td>
<td>20-50</td>
<td>120-240</td>
<td>Very High</td>
<td>High</td>
<td>With an ISO/RTO</td>
</tr>
<tr>
<td>3</td>
<td>50-100</td>
<td>30-120</td>
<td>High</td>
<td>Medium</td>
<td>Between a few utilities</td>
</tr>
<tr>
<td>2</td>
<td>100-1000</td>
<td>1-30</td>
<td>-</td>
<td>Low</td>
<td>Within a utility</td>
</tr>
<tr>
<td>1</td>
<td>&gt;1000</td>
<td>-</td>
<td>-</td>
<td>Very Low</td>
<td>Within sub.</td>
</tr>
</tbody>
</table>
## Diversity of Extreme Apps

<table>
<thead>
<tr>
<th>Inputs</th>
<th>System Analysis</th>
<th>Situational Awareness</th>
<th>Stability Assessment</th>
<th>Wide-Area Control</th>
<th>System Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lat.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3-5</td>
</tr>
<tr>
<td></td>
<td>Rate</td>
<td>3-4</td>
<td>2-3</td>
<td>2-5</td>
<td>3-4</td>
</tr>
<tr>
<td></td>
<td>Crit.</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Quan.</td>
<td>5</td>
<td>3-4</td>
<td>4</td>
<td>2-4</td>
</tr>
<tr>
<td></td>
<td>Geog.</td>
<td>2-5</td>
<td>2-5</td>
<td>2-5</td>
<td>1-5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outputs</th>
<th>System Analysis</th>
<th>Situational Awareness</th>
<th>Stability Assessment</th>
<th>Wide-Area Control</th>
<th>System Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lat.</td>
<td></td>
<td>1</td>
<td></td>
<td>3-5</td>
<td>5</td>
</tr>
<tr>
<td>Rate</td>
<td></td>
<td></td>
<td></td>
<td>2-4</td>
<td></td>
</tr>
<tr>
<td>Crit.</td>
<td></td>
<td></td>
<td>4-5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Quan.</td>
<td></td>
<td>1-2</td>
<td></td>
<td>1-3</td>
<td>1-3</td>
</tr>
<tr>
<td>Geog.</td>
<td></td>
<td>2-5</td>
<td></td>
<td>1-5</td>
<td>1-3</td>
</tr>
</tbody>
</table>
Outline

- Questions for Power Engineers & Researchers
- WAN Apps with Extreme Comms Requirements
- IT Guidelines for Achieving these Requirements
- GridStat: Industrial Internet for Electricity (IIE)
- GridStat Cyber-Physical Example
- Cyber-Security for Closed-Loop Apps
- Why Middleware is Needed for IIE
## Internet vs. NASPInet/WAMS-DD

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Internet</th>
<th>NASPInet/WAMS-DD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Size</td>
<td>~10⁹ hosts worldwide</td>
<td>10⁵ hosts on a grid, 10³-⁴ “routers”</td>
</tr>
<tr>
<td>Per-Flow State</td>
<td>Death (RIP RSVP)</td>
<td>Very feasible</td>
</tr>
<tr>
<td>Network Design Goal</td>
<td>Best-effort delivery for any user and purpose</td>
<td>Guaranteed QoS in several dimensions for specific users and purposes</td>
</tr>
<tr>
<td>Admission Cntl Perimeter</td>
<td>None</td>
<td>Complete</td>
</tr>
<tr>
<td>Fraction of Managed Traffic</td>
<td>None/Very Little</td>
<td>Almost all. All traffic subject to policing. &gt;&gt;90% periodic.</td>
</tr>
<tr>
<td>Central Knowledge of Topology</td>
<td>Not attempted: large scale and dynamicity</td>
<td>Feasible: small scale and relatively slow changes</td>
</tr>
<tr>
<td>Topology changes (w/o failures)</td>
<td>Often &amp; without warning</td>
<td>Not often &amp; virtually always with warning</td>
</tr>
<tr>
<td>Frequency of Route Changes</td>
<td>Frequent; can converge slowly</td>
<td>Infrequent ➔ use static topo.</td>
</tr>
<tr>
<td>Characteristic</td>
<td>Internet</td>
<td>NASPIInet/WAMS-DD</td>
</tr>
<tr>
<td>----------------------------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Latency Achievable</td>
<td>Slow to Medium</td>
<td>Very fast</td>
</tr>
<tr>
<td>Latency Predictability</td>
<td>Poor</td>
<td>Very Good to Excellent</td>
</tr>
<tr>
<td>Recovery Delay after dropped packet (with “reliable” delivery)</td>
<td>High (timeout waiting for data or acknowledgement; then two one-way msgs)</td>
<td>Zero (redundant copies sent over disjoint paths arrives virtually at same time) ➔ <strong>DO NOT USE post-error recovery, be proactive!</strong></td>
</tr>
<tr>
<td>Forwarding Unit</td>
<td>Uninterpreted packet</td>
<td>Update of a sensor variable</td>
</tr>
<tr>
<td>Traffic Predictability</td>
<td>Low</td>
<td>Very High</td>
</tr>
<tr>
<td>Elasticity of QoS requirements</td>
<td>None/Low</td>
<td>Potentially Medium-High (if power apps programmed per previous questions)</td>
</tr>
</tbody>
</table>
Other IT Guidelines for WAMS-DD

• Don’t depend on priority-based “guarantees” for hard real-time
• Exploit a priori knowledge of traffic
• Optimize for rate-based sensor updates
• Use static, not dynamic, routing
  ▪ Can still have adaptations to update routing tables
• Reject unauthorized messages quickly & locally
• Provide per-subscriber QoS+
• Don’t over-design consistency and (re)ordering
• Questions for Power Engineers & Researchers
• WAN Apps with Extreme Comms Requirements
• IT Guidelines for Achieving these Requirements
• **GridStat: Industrial Internet for Electricity (IIE)**
  • GridStat Cyber-Physical Example
  • Cyber-Security for Closed-Loop Apps
• Why Middleware is Needed for IIE
What is GridStat?

• Bottom-up re-thinking of how and why the power grid’s real-time data delivery monitoring services need to be

• Comprehensive, ambitious WAMS-DD middleware for power grid in coding since 2001
  • Rate-based pub-sub middleware with
    • Predictably low latency
    • Predictably high availability
    • Predictable adaptation
  • Different subscribers to same variable can get different QoS+ {rate, latency, #paths}

• Influencing NASPInet effort
Overview of GridStat Implementation & Perf.

- Coding started 2001, demo 2002, real data 2003, inter-lab demo 2007-8
  - But power industry moves very, very slowly……
    - “Utilities are trying hard to be first to be second” Jeff Dagle, PNNL
    - “Utilities are quite willing to use the latest technology, so long as every other utility has used it for 30 years” unknown
  - And NASPI is pretty dysfunctional in a number of dimensions

- Implementations
  - Java: < 0.05 msec/forward, 500k+ forwards/sec
  - Network processor: 2003 HW ~.01 msec/forward, >1M fwds/sec
    - Current network processors are ~10x better, and you can use >1 …
  - Future: FPGA/ASIC
    - Should be competitive w/IP routers in scale: doing much less, on purpose!

- Note: no need to use IP for core …… (ssshhhhhhh!): less jitter and likely more bullet-proof (no IP vulnerabilities)
What is GridStat? (cont.)

• GridStat at two layers
  • APIs & services (including management, monitoring, ...) at edges (e.g., last DNMTT comment)
  • I.e., Middleware overlay with mechanisms only at edges (P2P)
  • Augmented with core software defined network (SDN) utilizing rate-based, in-network router-like Layer-3 forwarding engines (FEs)
    • Also then richer management that exploits them

• Even with only 10% penetration of FEs have much more control over data delivery
Outline

- Questions for Power Engineers & Researchers
- WAN Apps with Extreme Comms Requirements
- IT Guidelines for Achieving these Requirements
- GridStat: Industrial Internet for Electricity (IIE)
- **GridStat Cyber-Physical Example**
- Cyber-Security for Closed-Loop Apps
- Why Middleware is Needed for IIE
GridStat Modes

- Observation
  - Path allocation algorithms complex, not for a crisis \(10^3\) +
  - But power grid adaptations planned way ahead of time

- GridStat supports **operational modes**
  - Can switch (preloaded) forwarding tables very fast
  - Avoids overloading subscription service in a crisis
  - Can have modes for different contingencies/modes/etc
Data Load Shedding

• Electric Utilities can do load shedding (I call power load shedding) in a crisis (but can really hurt/annoy customers)

• GridStat enables Data Load Shedding
  • Subscriber’s desired & worst-acceptable QoS (rate, latency, redundancy) are already captured; can easily extend to add priorities (e.g., [0,10] above)
  • In a crisis, can shed data load: move most subscribers from their desired QoS to worst case they can tolerate (based on priority, and eventually maybe also the kind of disturbance)
  • Works very well using GridStat’s operational modes
  • Note: this can prevent data blackouts, and also does not irritate subscribers

• Example research needed: systematic study of data load shedding possibilities in order to prevent data blackouts in contingencies and disturbances, including what priorities different power apps can/should have…

• Enables critical infrastructures: adapt data comms infrastructure to benign IT failures, cyber-attacks, power anomalies, changing req, …
Multi-Level Contingency Planning & Adapting

- Electricity example: Applied R&D on coordinated
  1. Power dynamics contingency planning
  2. Switching modes to get new data for contingency
  3. New visualization window specific for the contingency involving contingencies with
    A. Power anomalies
    B. IT failures
    C. Cyber-attacks

- State of art and practice today: 1 & A only, offline
- Very possible: \{1,2,3\} X \{A,B,C\} and online
Outline

• Questions for Power Engineers & Researchers
• WAN Apps with Extreme Comms Requirements
• IT Guidelines for Achieving these Requirements
• GridStat: Industrial Internet for Electricity (IIE)
• GridStat Cyber-Physical Example
• Cyber-Security for Closed-Loop Apps
• Why Middleware is Needed for IIE
A Note on Closed-Loop Cyber-Security

• Extra low added latency crucial for closed-loop apps
  • Otherwise delivery latency too high to work
• RSA is too slow (~60 msec on 2012-era PC)
• Using shared key (publisher and its subscribers) has a serious security vulnerability discovered by WSU
  • Nothing prevents a subscriber from spoofing a publisher
  • Shared keys (and RSA) are the choices for IEC 61850-9005
    • Example from 61850 workshop: presenter ducked out of next session after my pointing above out
• Real-time multicast authentication is an extremely hard problem
Overview of GridStat Cyber-Security

• GridStat has been a founding member of TCIP and TCIPG centers for cyber-security for the grid, 2005+.

• Stackable and changeable security modules at pubs and subs (2007)
  • Long-lived required ability to change modules as crypto technology evolves
  • Modules for encryption & authentication & obfuscation of data

• Authentication of management plane entities pairwise (2009, 2011+)

• Node security protecting data in management plane nodes (2012)
  • Secure key storage (quorum based, Byzantine fault-tolerant, …) ProFokus
GridStat Multicast Authentication

- Researchers: Prof. Carl Hauser and Kelsey Cairns (PhD expected early 2015)
- Started with TV-OTS (Illinois), an experimental data authentication protocol that appears to fit PMU data delivery needs:
  - Safe with Multicast (one-to-many) environments Low Latency
  - Independently verifiable signatures
  - Optimized for rate based communication (but not nearly fast enough for closed-loop apps with original TV-OTS)
- Achieved by probabilistic security (not perfect security)
GridStat Multicast Authentication (cont.)

- DETERLab experiments:
  - Measure signing and verification latency
  - Test achieved security with Security Analyzer node
- Low latency signing, lower verifying
- Lossy network has negligible effect
- Given message forgery probability ranging from $1.2 \times 10^{-7}$ to $9.6 \times 10^{-17}$
  - Faster cases less secure
- Generic latency adds only 10 msec to delivery
- In tightly-controlled WAMS-DD: $\sim1$ msec added
  - Easily fast enough for closed-loop apps
  - Ergo GridStat+authentication+encryption = $\sim2$ msec over underlying network layers
Final Thoughts on GridStat Security

• Would be very hard for an adversary to inject unauthorized traffic anyway even without authentication
  • Has to know exactly where to insert
  • Almost always would get dropped by next FE because not in forwarding table
• Recent MS thesis on rate-based monitoring can make adversary’s job even harder
Outline

- Questions for Power Engineers & Researchers
- WAN Apps with Extreme Comms Requirements
- IT Guidelines for Achieving these Requirements
- GridStat: Industrial Internet for Electricity (IIE)
- GridStat Cyber-Physical Example
- Cyber-Security for Closed-Loop Apps
- Why Middleware is Needed for IIE
Middleware (MW) Integrating Legacy (Sub)Systems

GridStat Management Plane

Publisher 1 → Common MW API → GridStat FEs → Nimbra, ATM → Subscriber 1
Publisher 2 → Common MW API → GridStat FEs → Nimbra, ATM → Subscriber M
Publisher N → Common MW API → GridStat FEs → Nimbra, ATM → Subscriber M

MPLS
OpenFlow/SDN
Microwave, Satellite
GridStat FEs

= Cntl Wrapper
= Data Wrapper
The “Smart Grid” Community Agrees

  - Argued that, by the stated goals of the US “smart grid” leadership (GridWise and NIST), middleware was needed for WAMS-DD
  - Argued that if you care about QoS, needed even more
  - **Best Paper Award for “Connectivity” track.** This is the official communications/interoperability meeting for the pseudo-official “smart grid” community in the USA, namely DoE/GridWise and NIST/SmartGrid.
  - Paper also describes what middleware is and how standardized
Other Reasons Middleware is Needed

• Very hard problem for computer scientists to get right: combining mechanisms for network QoS and security and other QoS+ properties
  ▪ Almost impossible for non-expert to get right
  ▪ Without middleware, each application has to rewrite
  ▪ Much better to let middleware developer do this, and apps reuse

• Future proofing: not locked into given mech.

• Only way to do monitoring/policing on a per-update per-sensor basis

• .... see Bakken book chapter on GridStat for 5 pages on this!