Leveraging the Next-Generation Power Grid: Data Sharing and Associated Partnerships

Daniel Germanus, Ioanna Dionysiou, Harald Gjermundrød, Abdelmajid Khelil, Neeraj Suri, David E. Bakken, and Carl Hauser

***Abstract*—Data delivery in the power grid today is, for the most part, hard-coded, tedious to implement and change, and does not provide any real end-to-end guarantees. Applications have started to emerge that require real-time data delivery in order to provide a wide-area assessment of the health of the power grid. This paper presents two novel communication infrastructures that facilitate the delivery of power data to intended recipients, each based on a different communication paradigm. The necessity of forming and managing trusted partnerships in either framework is further discussed.**

***Index Terms*—power grid, communication models, trusted partnerships**

I. Introduction

One of a nation's overarching responsibilities is to protect its citizens by limiting the impact that malicious and accidental acts as well as natural disasters could have on the society. Therefore, the protection of critical infrastructures is elevated as one of the essential elements of ensuring a nation's security [1]. In order to do that, federal agencies, state and local departments, the private sector, and other foreign partners must collaborate and coordinate to share information without compromising it. The *need-to-share* principle [2] has found support, something that is not surprising: evolving from access to information necessary to perform one's official duty towards access to information for the greater common good is seen as a sign of the much-wanted collaboration.

The power grid falls into the category of critical infrastructures, and it is more complex and harder than the other infrastructures, mainly due to its geographical scale and real-time requirements. Consider the North American electric power grid, for instance, with nearly 3500 utility organizations [3]. These individually owned utility systems have been connected together to form interconnected power grids, which must be operated in a coordinated manner. There are many points of interactions among a variety of participants and a local change can have immediate impact everywhere. In order to detect disturbances that could escalate into cascading outages and take corrective actions, real-time information about the grid dynamics must be obtained to enhance the wide-area system observability, efficiency, and reliability. The next-generation power system is envisioned to be consisting of automated transmission and distribution systems that support efficient and reliable supply and delivery of power, with an appropriate underlying network infrastructure to support the framework's mechanisms for data retrieval from field equipment and issue control commands to power system equipment, among field devices, between field devices and systems located in substations, and between field devices and various systems including SCADA, utility control centers, engineering, and planning centers. Data can be real-time data, statistical data, or other calculated data and informational data from the power system to systems and applications that use the data.

However, this collaborative environment is nontrivial to establish and operate mainly because the current deployed power grid communications are inflexible and non-extensible [4], without provisions for managing wide-area communication channels among participants and facilitating secure, non-compromised and trusted data exchange subject to real-time constraints. This paper presents GridStat and INSPIRE projects [5], [6], [7], [8], two novel approaches to address the communication infrastructure deficiency in the power grid, based on different communication paradigms: GridStat is a specialized publish-subscribe middleware with support for QoS whereas INSPIRE is a managed group of interconnected Peer-to-Peer (P2P) overlay networks.

The remaining of the paper is organized as follows: Section II motivates the need for data sharing in power grid settings. Section III gives details on the GridStat and INSPIRE framework functionality, with a discussion on forming partnerships that could leverage the functionality of both frameworks given in section IV. Section V describes related work, with concluding remarks found in section VI.

II. Data Sharing in the Power Grid

Unfortunately, as of today, power utilities are reluctant to disclose information in order to protect themselves financially and legally. Sharing of data might jeopardize their business due to their inability to quantify the risk regarding interactions with other grid participants. For example, unrestricted access to a utility's data that are market-sensitive indicators could give a competitor an unfair advantage in adjusting its own contracts and prices. Similarly, a utility could distribute inaccurate data to mislead the other market participants. This trend is not only observed in the power industry but it rather concerns other markets as well; according to a CSI/FBI report [9], companies are reluctant to report security incidents because of the potential loss of revenue. Thus, even though a data dissemination communication medium was in place, still power grid participants would most likely have been objected to its usage unless they were obliged by legal contracts.

Ioanna Dionysiou and Harald Gjermundrod are with the Department of Computer Science, University of Nicosia, Cyprus

Daniel Germanus, Abdelmajid Khelil, and Neeraj Suri are with the Department of Computer Science, TU Darmstadt, Germany

David E. Bakken and Carl Hauser are with the Department of Electrical Engineering and Computer Science, Washington State University, WA, USA

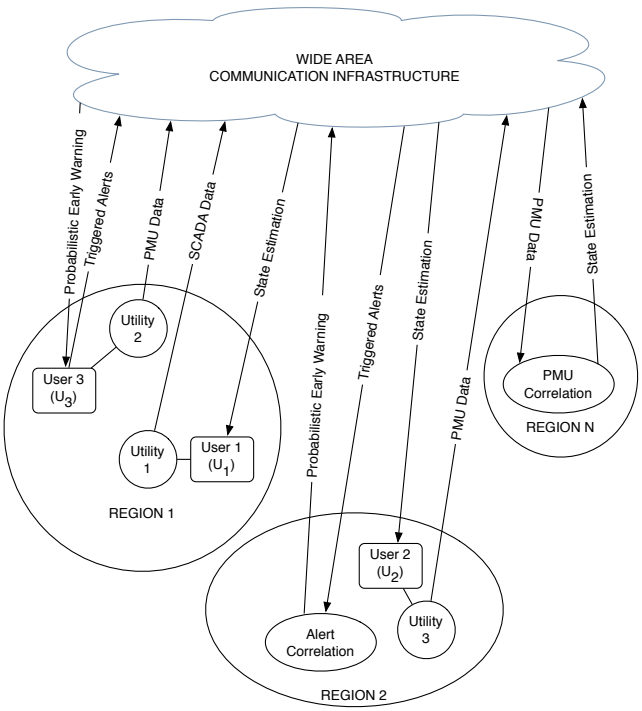

Fig. 1. Power Grid Data Exchange Facilitation

The "no sharing" policy could be relaxed under normal operating conditions if the risk of sharing were systematically contained, and could be relaxed much more if needed to help avert a looming crisis. The data exchange could be possible if trusted partnerships were formed among all parties involved, which would enable the entities that have knowledge about critical infrastructures to share confidential, proprietary, and business sensitive information with reliable partners.

Nevertheless, the provision of access to local and regional data is crucial to a wider-area monitoring and assessment of the stability and security of the grid. Sharing of current information beyond the scope of a single electric utility will help increase robustness of the power grid. In order to illustrate the usefulness of the dissemination of data, three power grid application suites are considered (Figure 1) along with concerns regarding the quality of the received data. The application domain setting involves utility organizations that generate data and users that receive computation results on that data. All applications span a number of regional utility districts with different administrative policies and ownership. The *Phasor Measurement Unit (PMU) Aggregation* application involves the dissemination and aggregation of PMU data, the *SCADA Data Dissemination* application deals with the distribution of local SCADA data for wide-area assessments, and the *Alert Correlation* application provides (probabilistic and imperfect) early warning of an impending power crisis.

Prior the discussion, we note a couple of simplifications made in the illustrated scenario. First, without loss of generality, user U_i could be an end-user, an administrator, a regional operator, etc. Second, the two applications disseminate the results of their processing to entities other than the original sources of data. In an actual deployment, feedback could (and likely would) be provided to the original utilities as well.

A. PMU Data Aggregation

The first application suite deals with a type of real-time electric power data, which is PMU data. PMUs are instruments that take measurements of voltages and currents and time-stamp these measurements with high precision. They are able to measure phase difference at different substations and have been implemented as a source of information to detect faults on transmission lines. These measurements are collected and aggregated at a central place in order to derive system state estimations. In order to preserve correlation of these readings, it is important to temporally synchronize the PMU measurements during aggregation [10], [11]. These estimations are disseminated to interested end users U_i, including entities that monitor or control the grid, such as Independent System Operators (ISOs), Regional Transmission Operators (RTOs) and, in the future, the Department of Homeland Security (DHS.) Currently, the Eastern Interconnection Phasor Project (EIPP) is deploying PMUs on the eastern U.S. grid [12].

A collaborative environment as illustrated in Figure 1 gives rise to new challenges involving the data quality of the aggregated state estimation. Suppose that the aggregated function $f(d_1, d_2, \cdots, d_M)$ takes as inputs PMU data d_i from utility $Utility_i$ and outputs state estimations. An accurate state of the power grid is based on the quality of the received d_i. Accidental or malicious faults observed at $Utility_i$'s sensors may affect its ability to generate correct data d_i. Assume that a faulty sensor at $Utility_2$ produces inconsistent PMU measurements. Including these d_2 readings in f gives an inaccurate view of the current state of the power grid. Users U_1 and U_2 are not able to detect the source of the inaccurate readings since they only receive the aggregated result. Thus, U_1 and U_2 entities must establish bilateral trust partnerships with *PMU Aggregation* entity that will allow them to rely on the received aggregated output.

Additionally, a data provider that produces and disseminates correct data does not necessarily imply that the received data is also correct. An unreliable data communication medium may also tamper with the transferred data, which will result in producing inaccurate state estimations. In order to manage the risk of producing inaccurate state estimations, the entity that performs PMU aggregation must make trust assessments of all interacting entities that collaboratively execute the task of generating and delivering PMU data.

B. SCADA Data Dissemination

Supervisory Control and Data Acquisition (SCADA) systems are a core component of many critical infrastructures and power grids are not an exception. Different techniques are applied to support SCADA systems during perturbations. These techniques include proactive and reactive measures through SCADA system monitoring and reaction.

Nowadays, interconnections among SCADA systems of different utilities and distributors are important to enable data sharing. In terms of the power grid, areas of large geographical

extent are covered by a conglomerate of different utilities for economical reasons and to improve the service quality to the customers. Especially, to safeguard customers against blackouts and to assure safe operation of power generating, transmitting, and distributing equipment. Interconnections promote economic efficiency on the one hand but also bear risks regarding the grid's integrity, since failures may be a consequence of cascading effects among different utilities. Furthermore, the attack surface of large scale SCADA networks increases and renders cyber attacks more likely.

Therefore, grid operators need a holistic view to ensure safe and appropriate operation, e.g., metering and load balancing has to be performed in realtime or close-to-realtime to prevent failures. Moreover, protection mechanisms need to bear intentional misbehavior in mind by addressing SCADA cyber security aspects. Apart from overarching protective measures, the interconnected grid's data sharing is of importance for business departments, e.g., to detect excess capacities or for billing purposes.

C. Power Grid Alert Aggregation and Early-Warning System

The third application family demonstrates the potentials of sharing of more operational information between electric utilities. The first scenario deals with the situation when problems start happening in one part of the power grid, many alerts can fire in a short period of time, and utilities can get buried in the (largely redundant) barrage of such alerts. Alert correlation allows for transformations of a series of lower-level alerts into higher-level alerts which have a much lower false alarm rate and much richer semantics, and are thus a much more useful indicator of trouble. Such transformations should be based on taxonomies of power grid devices, and be policy-programmable, since different configurations of a given device will have different thresholds for operational reasons. However, alert correlation must be complemented by similar data quality assessment techniques as the one described in the earlier example since incorrect data may suggest a catastrophic situation that does not occur.

The second situation is that some data are market sensitive, meaning if a competitor has the reading of some key data (for example, the output of a utility's generators) it can, over time, deduce the company's production and pricing strategies. As an example of this problem, instead of sharing market sensitive data directly, derived values such as the instantaneous rate of change (or moving averages thereof) can be shared. Thresholds for particular kinds of devices can be monitored, and alerts generated if they exceed a certain threshold. Since there are currently no means to quantify the risk of sharing sensitive data or derived indicators, the next best alternative is to restrict their access to non-competitors. Based on observed behavioral trends, a utility can decide whether or not access to sensitive data should be granted or denied.

III. Communication Infrastructures for the Electric Power Grid

This section presents two communication frameworks that provide a wide-area communication infrastructure for critical infrastructures. The first framework is GridStat developed at Washington State University, while the second is the INSPIRE project developed by a consortium of European academic institutes. The two research teams have made different design assumptions and impose different requirements on the physical infrastructure needed to provide the ability for wide-area communication. However, both GridStat and INSPIRE provide an abstraction to the application layer for delivering status information for a critical infrastructure.

A. GridStat: A Status Dissemination Middleware

The GridStat project has been in development since 2001, and funded by various US research grants. A prototype of the project has been field-tested by Avista Utilities and Pacific Northwest National Lab (PNNL). Insights from the GridStat project have been used in the formation of the NASPInet [13] specifications.

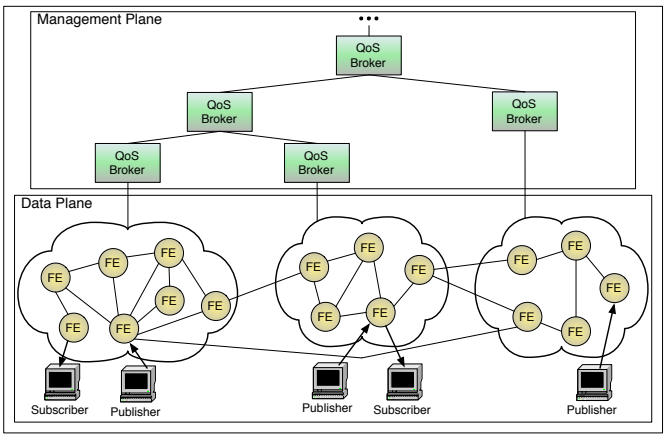

Fig. 2. GridStat Architecture

1) Design assumptions: The GridStat framework [14], [6] was architected and designed to deliver status updates with very high reliability, and a very small latency. The major communication requirements were to provide end-to-end reliability and timeliness for event streams scalable into the wide-area setting. In order to achieve the aforementioned requirements, a number of assumptions on the underlying infrastructure was specified:

- Closed and dedicated network infrastructure: In order to provide high reliability and tight control, it is assumed that a GridStat deployment will work in a closed, dedicated environment. In other words, the network will be completely controlled by GridStat, i.e. it will be a backbone network for communication for a critical infrastructure. This is costly but we argue justifiable, as the seamless operation of a nation's critical infrastructure is important for its citizens and its competitiveness. To alleviate the cost originating from this dedicated network, the communication infrastructure could potentially be shared by all the critical infrastructures like water supply, food supply, transportation, etc.
- Mostly static information flows: In order to provide small end-to-end latency in the wide-area setting a number of

optimizations must be done. An important observation for critical infrastructures like the electric power grid, is that information flows (control and monitoring) stay fairly static. i.e. a new power plant doesn't get built overnight and similarly, it takes time to put up a new high voltage power line. Thus, prior to become operational, the information flows are known in advance, as well as the sources of the information, and their respective requirements as far as control and monitoring are concerned. The forwarding logic of the communication infrastructure takes advantage of the static feature and devises forwarding algorithms that exploit it in favor of minimizing the end-to-end latency.

- Accurate timestamp: Each end-point that is connected to this communication infrastructure has access to a very accurate time source, like a GPS clock. Within the the electric power grid domain, this requirement is already fulfilled as PMU devices rely on accurate clocks. The GridStat framework depends on the existing technology for time stamping to optimize the forwarding logic and provide the end-applications with a consistent snapshot of the current state of the grid.

2) Architecture and Design: GridStat is built on the publish-subscribe paradigm and hence, its architecture is appropriately composed by components such as publishers, subscribers, and forwarding engines. The GridStat architecture is depicted in Figure 2. It is separated into two planes: the *management plane* and the *data plane*. The role of the management plane is to control the resources in the data plane so as to provide QoS guarantees and scale to the wide-area setting. The role of the data plane is to forward events from the publishers to the subscribers, as efficiently as possible, while providing end-to-end QoS.

- Management Plane: The management plane provides services that involve primarily registration (and unregistration) of publications and subscriptions, path allocation decisions as well as secondary services, such as placement of condensation functions (see mechanisms below). The lowest level of the management hierarchy is composed of *leaf QoS brokers*, which manage a pre-defined domain consisting of a set of *forwarding engines*, publishers and subscribers; that's called a *cloud*. Each leaf QoS broker has complete knowledge of its own administrative domain and its resources. These resources include *event channels*, forwarding engines, and any other computational resources. The responsibility of the leaf QoS broker is to allocate and deallocate *subscription paths* from the publishers to the subscribers, subject to the QoS requirements. The non-leaf QoS brokers may manage multiple clouds and are responsible for allocating and deallocating inter-cloud subscriptions.
- Data Plane: The data plane, on the other hand, provides a virtual message bus used by publishers and subscribers to send and receive event streams. A forwarding engine is in effect a router with additional functionality to provide forwarding of status events when subscribed to and at the right rate (*rate filtering*). The management plane controls the content of the routing tables in the forwarding engines, and leaf QoS brokers inform forwarding engines to add, remove or modify the content corresponding to a subscription allocation or deallocation request.

3) Mechanisms and Services: The GridStat framework supports a number of built-in mechanisms, as listed below. The list is not exhaustive, as its purpose is to demonstrate the usability of the framework via providing services to the application layer or directly to the end-user. GridStat is extensible and could be customized to provide other mechanisms and/or services for the specific application domain.

- Multicast with predictable rate-filtering: Without rate filtering, many status updates from PMUs (and other sources) would be wasting significant bandwidth. It would be most unfortunate, however, if the rate filtering were to filter update streams from different sources differently; i.e. delivering, say, 4 updates per second from one stream timestamped at 0, 250, 500, and 750ms past the second while delivering updates from another stream timestamped at 125, 375, 725, and 875ms past the second. This of course would not provide a consistent global snapshot of the grid's conditions. GridStat's rate filtering is designed so that subscriptions with identical rate requirements and with compatible publication intervals result in identical timestamps on the delivered updates for all subscriptions. The filtering is only performed when the subscribers request to receive fewer updates than are generated. In the case where a subscriber subscribes at the same rate as the publication rate, then all events will be delivered to the subscriber.
  Both multicast and rate filtering are implemented by a single forwarding mechanism that works as follows. The publication ID and (GPS-accurate) timestamp are extracted from each incoming packet. The ID is used as a key to look up the outgoing links with active subscriptions for that ID. The lookup yields for each link a list of subscription intervals. A calculation based on the interval and timestamp yields a forward or do not forward decision for that link. If any of the subscriptions on a link produce the decision forward then the packet is sent on the link; otherwise, it is dropped.
- Hierarchical Operational Mode: The process of establishing individual subscriptions is a resource-intensive operation which, if done during run-time, may result in unsatisfactory subscription delays. GridStat enables subscription bundles to be allocated and pre-loaded into the forwarding engines' routing tables where operating modes control which routing tables the forwarding engine network will utilize [15]. A mode contains the necessary forwarding rules for a set of subscriptions and allows the forwarding engine network to quickly switch between bundles of subscriptions. This mechanism which allows the system to quickly adapt at run-time to pre-configured subscription set will help the monitoring application of critical infrastructures to plan what subscriptions are needed for various contingencies both in the application domain as well as in the IT-infrastructure.

- Middleware layer Event Patterns: Commonality in application logic could be provided as a middleware service for a number of domains, with a number of benefits. The most obvious benefit is the reuse of application logic. Another benefit is computational resources reduction. Consider the case where multiple users are performing the same application level computation. If the middleware layer handles this computation, it can perform it once and share the result with all the interested parties.
  GridStat provides the condensation function mechanism [16] that is used to migrate application logic, such as event patterns, in the middleware layer. The pattern is specified using a tool at the application level and then forwarded to the management plane which placed it in the data plane. Once deployed the output of the condensation function is publish just like any other publication and it is transparent to the subscribers that this is a derived value compared to a raw measurement.
- Security Management System: As a GridStat deployment would be a long lifetime system, i.e. the system cannot be shutdown for maintenance once a year, it provides a security management system [17] for long lifetime operation. The system allows for security modules to be securely added and replaced at runtime in an non intrusive way. The security management system consists of a set of modules that provides tradeoffs between performance and security. These can be combined in order to provide high-performance secure multi-cast supporting confidentiality, integrity, authentication, and obfuscation.
- Reliable RPC: GridStat's one-way delivery mechanisms are sufficient for delivering updated values from remote sensors. However, they are inadequate for a round-trip remote procedure call (RPC), for example to an actuator in the power grid or between control centers. Therefore a timely and reliable wide-area RPC mechanism over a QoS-managed, one-way publish-subscribe mechanism [18] are provided within the framework. The RPC mechanism supports three distinctive techniques for redundancy, offering tradeoffs between worst-case deadlines, use of network resources and resiliency towards a variety of network failures. Applications are allowed fine control of redundancy semantics. In addition, the RPC mechanism has integrated pre- and post-condition into its call semantic.

B. INSPIRE: INcreasing Security and Protection through Infrastructure REsilience

In order to enable interconnections among SCADA systems, the trend goes towards using components-off-the-shelf (COTS) and to utilize public networks like the Internet. The major drawback of COTS and public networks is that cyber attacks are more likely to emerge than in isolated and proprietary systems. Therefore, protection of SCADA systems is crucial in order to protect the underlying critical infrastructure for the sake of public safety & security and business objectives.

The INSPIRE research project [19] aims at increasing the security and protection through infrastructure resilience. The techniques that will be presented in this section have been evaluated using simulations and shown good results in terms of dependability and security gain [20]. Currently a prototype implementation is developed to serve as a proof of concept study.

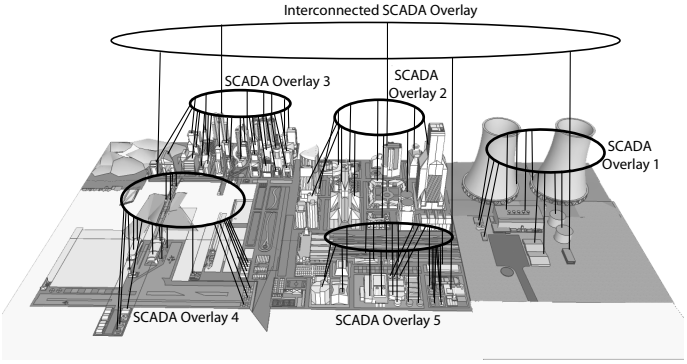

Fig. 3. Two Layered Overlay Architecture for Multiple Interconnected SCADA Systems

1) Architecture and Design: : As shown in Figure 3, different facilities coexist, e.g., power generation, distribution and consumption. Each SCADA system of the aforementioned facilities spans its own overlay network for data storage and replication. In addition, a global overlay network includes peers from all interconnected and interdependent SCADA systems. This two layered overlay hierarchy provides data sharing among all facilities using the interconnected SCADA overlay. Each facility may individually define a subset of its data that is replicated to both hierarchy layers. Legal, privacy, or performance aspects may be a basis for specific decision on the subset extent.

2) SCADA Protection Requirements: A SCADA protection mechanism should fulfill the following crucial requirements. First, the protection efforts should be beneficial with respect to mitigating lasting perturbation and ensuring that no new threats are introduced to the system. Further requirements include support for legacy systems and heterogeneous environments as well as scalability in large scale and interconnected SCADA systems.

This can be realized by monitoring critical functionalities such as (i) data sharing among remote terminal units (RTUs) and master terminal units (MTUs) or data sharing among different SCADA systems on the interconnected level, (ii) detection of perturbations that might harm the system, and (iii) mitigation of effects that are caused by perturbations. The INSPIRE project considers two perturbations, namely node crashes and illicit data modification on behalf of an attacker.

3) P2P-based SCADA Protection Layer: P2P techniques represent an excellent platform for data sharing. Different P2P protocol families have been qualitatively evaluated regarding their appropriateness for SCADA system protection enhancements in previous studies [21]. In particular, structured P2P protocols are well suited for SCADA protection enhancements; this protocol family implements distributed hash tables (DHTs) which provide efficient data storage and replication mechanisms. The P2P-based data sharing simplifies perturbation

detection and mitigation by exploiting the distribution of peers across the network and their independent operation. In particular, the protection layer supports the following:

- Concordance to Requirements: P2P overlays satisfy the previously defined requirements and constitute a key technology used in the INSPIRE project. P2P's middleware nature allows to mask the heterogeneity of large scale topologies which are made up of a plethora of node types. Structured P2P protocols [22], [23] scale well, since they provide efficient lookup and routing strategies and are thereby able to manage millions of peers. Legacy systems may be integrated because P2P protocols demand few resources in terms of computation cycles and data storage.
- Protective P2P Mechanisms: The protection enhancement using P2P overlays is based upon path redundancy and data replication which are inherent characteristics of structured P2P protocols. Path redundancy increases the overlay's robustness during node crashes and retain the overlay intact. Data replication denotes data storage on several nodes. Thereby, data remains available during node crashes and also timeliness of data requests may be improved, since a requesting peer might accept the first returned answer.
- Mitigation Strategies: To overcome perturbations, two mechanisms are implemented in the P2P-based protection layer. Communication between RTUs and MTUs is intercepted by the P2P middleware. Then, SCADA message payloads are extracted and stored in the overlay. sThis is helpful in case the traditional SCADA communication flow is disturbed. It is assumed that multiple paths exist between RTUs and MTUs, in order to exclude single points of failure in the underlay network layer. In case of a node crash, e.g., of a non-edge router, the underlay network's reconfiguration process would exceed SCADA timeliness requirements. Therefore, communication outages need to be bridged by requesting missing or inaccessible SCADA data via the overlay network. In case of illicit data modification, it is assumed that an attacker controls a router and is able to modify data in transit. Since the P2P middleware on each RTU intercepts and replicates the data beforehand, and due to the P2P concept of overlay path redundancy, changes are high that an attacker that controls only one or few nodes cannot intercept and modify all replicas. Now, either proactively or reactively, the P2P-based protection layer requests each SCADA message that has been received through the traditional channel also via the overlay. The originally received message and the copies from the overlay are subsequently compared to decide upon the trustworthiness of the messages.

C. Revisiting Data Sharing in the Power Grid

Even though it is imperative to portray in detail the two communication paradigms, still it is also useful to revisit the data sharing scenarios mentioned in section II and examine the applicability of the two data delivery frameworks in those situations. In all three scenarios, the minimum set of requirements on the communication infrastructure are scalability, reliability, timeliness, and security.

The aforementioned requirements are fullfilled by the two systems in the following manner:

- Scalability in size: GridStat eliminates the case of isolated islands of grids, and supports wide-area connectivity within the entire grid, i.e. each entity is able to communicate with any other entity in their operational grid. This is accomplished via its publish-subscribe underlying communication paradigm that delivers data to any interested subscriber, subject to proper authorization. Similarly, structured P2P overlay networks are evaluated in the INSPIRE project which scale up to the range of millions of peers and are therefore suitable for large-scale and interconnected infrastructures.
- Scalability in data: Heavy data load can be eased by optimization in data delivery. GridStat deploys multicasting techniques in order for any datum to only traverse a communication link at most once. The P2P technology employed in the INSPIRE project distributes data across the overlay network using DHTs. Caching algorithms for past data retrieval requests increase efficiency as well as protocol extensions for appropriate replica positioning.
- Reliability: Situational awareness counts on information delivered not only in a timely manner, but without any loss during the transmission process. GridStat and INSPIRE's P2P technology provide fault tolerance mechanisms (e.g. redundant paths) to assure data delivery within the specified fault model.
- Timeliness: Data timeliness is addressed by GridStat as it provides end-to-end QoS guarantees, including latency requirements. A subscription would not accepted unless the infrastructure can provide the end-to-end requested latency. P2P protocol extensions exist for continuous overlay network adaption with respect to bandwidth provisions in the underlying network medium. These can be applied to overcome network bottlenecks and thereby improve data transmission latencies.
- Security: Basic security services such as message confidentiality, message integrity, user authentication are added-on features in GridStat. The P2P-enhancement in the INSPIRE project provides secure admission of peers. Only SCADA hosts which are known a priori are allowed to join the network to preserve confidentiality, availability, and integrity of data. Also, P2P protocol extensions exist to enable secure DHT operations and encrypted or digitally signed message exchange.

IV. Leveraging the Communication Frameworks: Forming Partnerships

Section III discussed two communication infrastructures that could be deployed in a wide-area setting to facilitate the distribution of power grid data. However, the need-to-share direction has an optimistic approach in the sense that all participants will willingly come to the realization that information sharing is beneficial and as a result coalitions will be formed to exchange data for the common good. But,

critical infrastructures, such as the electric power grid, exhibit both non-competitive and competitive characteristics due to the involvement of both the private and public sector. The voluntary sound of the need-to-share directive could lead to data sanitization with little or no access to data, i.e. need-to-share could be exploited by practicing unilateral sharing and restricting access to information for a number of reasons including risk of compromising work and other technological, political, economical, and cultural reasons.

However, as it was demonstrated in section II, the data sharing is vital for a wide-area situation awareness. In order to to address the above concerns and at the same time provide the much-needed data exchange, a compromised solution is proposed where data sharing is realized through two different kinds of partnerships:

- Legislative Partnerships: data exchange is enforced in a non-voluntary basis and bounded by regulatory contracts; could be unilateral
- Trusted Partnerships: data exchange is based on negotiable trust relationships on voluntary basis; could be either bilateral or unilateral

Both legislative and trusted partnerships impose strong assumptions on the underlying communication infrastructure. Among others, the communication medium is expected to be reliable and secure during the dissemination of data to its destination, with accountability mechanisms in place to hold entities liable for misconduct and contract violations.

A. Legislative Partnerships

There is a minimum set of data that must be available to grid operators to maintain a stable and healthy grid. Given that such set exists, grid participants should make provisions to supply it to the intended recipients, using the deployed communication infrastructure, regardless of any corporate policies. On the contrary, legislation policies mandate the data access and availability.

Experimenting and exploring the information landscape is essential to derive the above set. Given that data from both public and private sector is available, research efforts could be directed on deriving models that perform data mining on complete and incomplete sets. Needless to say, it will be challenging to enforce mandatory sharing based on those findings. Regardless, some challenges that need to be taken into consideration are the following:

- What data must be shared in order to predict something useful? Sharing only partial information could potentially have negative impact on decision- making.
- What are the economical, technological, political implications of sharing?
- One of the fundamental obstacles in data sharing is the diversity in the matter of data ownership. Who owns the data? The one who generates it? The one who handles it? The one who receives it? And to makes things even more complicated, privacy laws that govern personal data disclosure differ substantially from country to country.
- In the case of an incident that disturbs the normal operations of the grid, what can be learned by post-mortem analysis on data collected prior, during, and after?

B. Trusted Partnerships

Sharing should not have any implications on the right to data privacy, something that is fundamental in competitive markets. For example, one power generating company may not want to provide too many details of its contracts to another competitor which is part of a joint venture, but provide the whole data set to a subsidiary company. In this case, sharing of non-vital data goes over and beyond the data distributed through legislative partnerships, requiring entities to build their own bilateral partnerships, at their own risk. Some issues related to the establishment of these partnerships are:

- Build partnerships among organizations using trust relationships, which have to be specified, managed, monitor, and recover in case of a breach of trust.
- Evaluate the impact of the specific partnership on the company's decision-making process. For example, what is the risk for sharing particular data sets? Could the partial sharing/using yield the opposite results?

V. Related Work

One of the first initiatives that explored the possibilities of data exchange among grid participants was undertaken by the Electric Power Research Institute (EPRI), a consortium of academic and industrial experts that work collaboratively on approaches to electric power challenges. The IntelliGrid project [24] had two objectives: first, identify the business needs of power systems and their functional requirements, and second, devise a conceptual architecture that fulfills these requirements and used as a basis for the next generation power system. IntelliGrid assumes, but does not provide, an appropriate QoS managed middleware that is used to support the framework's mechanisms. However, to the best of our knowledge, there is no actual deployment of this technology.

On the other hand, an ongoing and promising effort is the North American SynchroPhasor Initiative (NASPI) [13], which aims in developing a secure, distributed, and flexible data communications infrastructure to support synchrophasor applications in North America.

Another research effort is the CRUTIAL project [25], which aims at new networked information and communication technology (ICT) systems for power grid management. The focus is on the interconnection of physical process controlling systems with information architectures like SCADA systems that are connected via multi purpose and public networks. Interdependencies in interconnected infrastructures are analyzed with respect to their dimensions having the goal to derive new architectural patterns to provide resilience during accidental failures and malicious attacks. The main objectives in CRUTIAL are threefold, namely to (i) investigate models and architectures that meet openness, heterogeneity and evolvability requirements of critical infrastructures, (ii) analyze critical scenarios in the information infrastructure which may lead to serious impacts in the critical infrastructure, and (iii) consideration of distributed architectures which promote dependable monitoring and control of critical infrastructures.

Finally, the IRRIIS project [26] aims at protecting critical infrastructures, like telecommunication networks or energy

supply. As a consequence of critical infrastructure interconnection and resulting interdependencies, critical infrastructures (CIs) become more vulnerable and IRRIIS addresses dependability, survivability and resilience of their underlying information infrastructures. The main objectives are to (i) find relevant requirements for both, public and private sector CIs, (ii) develop a middleware improved technology with focus on recovery actions and service stability, and (iii) to build a simulation environment for interdependencies in CIs.

Furthermore, the VIKING project [27] discusses security of SCADA systems. Its main objectives are to investigate the vulnerability of SCADA systems, the effect of cyber attacks on societies, and the proposal of test and mitigation strategies for SCADA systems.

VI. Conclusions

The communication infrastructure for today's electric power grid is based on old technology and piecemeal together. This prevents the operators of the grid to have full situation-awareness of the current state of the grid and limits the use of new control and monitoring techniques. Advances in distributed computing technology can be used to offer a new communication infrastructure for critical infrastructures that will be flexible, extensible, reliable, and adaptable.

This paper presented two such communication infrastructures developed in the US and EU. Even though they are based on different communication paradigms, they both facilitate the communication needs of the next generation power grid. The GridStat project is a managed publish-subscribe middleware for delivering streams of status events subject to QoS. The second project, INSPIRE, uses overlay P2P technologies for reliable and scalable communication between SCADA devices.

Regardless of the underlying infrastructure, it was demonstrated that it is essential to form partnerships, either legislative or trusted, to ensure proper dissemination of information.

Acknowledgment

The GridStat research has been supported in part by Grants CNS 05-24695 (CT-CS: Trustworthy Cyber Infrastructure for the Power Grid(TCIP)) and CCR-0326006 from the US NSF. The INSPIRE research is supported in part by EU INSPIRE, CASED (www.cased.de), and EU CoMiFin.

References

[1] "Protected critical infrastructure information (pcii) program," Department of Homeland Security, 2006, www.dhs.gov.

[2] "A Scientific Research and Development Approach to Cyber Security. Submitted to the Department of Energy on behalf of the Research and Development Community, December 2008," http://www.er.doe.gov/ascr/ProgramDocuments/ProgDocs.html.

[3] U. C. P. S. O. T. Force, *Final report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*, March 2004, https://reports.energy.gov/BlackoutFinal-Web.pdf.

[4] C. Hauser, D. Bakken, and A. Bose, "A failure to communicate: next generation communication requirements, technologies, and architecture for the electric power grid," *Power and Energy Magazine, IEEE*, vol. 3, no. 2, pp. 47 – 55, march-april 2005.

[5] D. E. Bakken, A. Bose, C. H. Hauser, E. O. Schweitzer III, D. E. Whitehead, and G. Zweigle, "Smart Generation and Transmission with Coherent, Real-Time Data," Washington State University, Tech. Rep. TR-GS-015, August 2010. [Online]. Available: http://gridstat.net/publications/TR-GS-015.pdf

[6] H. Gjermundrød, D. E. Bakken, C. H. Hauser, and A. Bose, "GridStat: A Flexible QoS-Managed Data Dissemination Framework for the Power Grid," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 136–143, January 2009.

[7] S. D'Antonio, L. Romano, A. Khelil, and N. Suri, "INcreasing Security and Protection through Infrastructure REsilience: the INSPIRE Project," in *Proceedings of The 3rd International Workshop on Critical Information Infrastructures Security (CRITIS'08)*, October 2008.

[8] S. D'Antonio, A. Khelil, L. Romano, and N. Suri, "Increasing Security and Protection of SCADA Systems through Infrastructure Resilience," *International Journal of System of Systems Engineering (IJSSE)*, vol. 1, no. 4, pp. 401–413, 2009.

[9] *Computer Crime and Security Survey*, CSI/FBI, 2005.

[10] A. G. Phadke, "Synchronized phasor measurements in power systems," *IEEE Computer Applications in Power*, vol. 6, no. 2, pp. 10–15, April 1993.

[11] R. A. Johnston, C. Hauser, K. H. Gjermundrød, and D. Bakken, "Distributing time-synchronous phasor measurement data using the gridstat communication infrastructure," in *Proceedings of 39th Annual Hawaii International Conference on System Sciences (CD/ROM)*, Kauai, Hawaii, January 2006.

[12] "Eastern interconnection phasor project," www.phasors.pnl.gov.

[13] "North American SynchroPhasor Initiative (NASPI)," http://www.naspi.org/naspinet.stm.

[14] K. H. Gjermundrød, "Flexible QoS-managed status dissemination middleware framework for the electric power grid," Ph.D. dissertation, Washington State University, July 2006.

[15] S. F. Abelsen, H. Gjermundrød, D. E. Bakken, and C. H. Hauser, "Adaptive data stream mechanism for control and monitoring applications," in *Proceedings of 1st International Conference on Adaptive and Self-adaptive Systems and Applications (ADAPTIVE'09)*, Athens, Greece, November 2009, pp. 86–91.

[16] H. Gjermundrød, D. E. Bakken, and C. H. Hauser, "Integrating an event pattern mechanism in a status dissemination middleware," in *Proceedings of 1st International Conferences on Pervasive Patterns and Applications (PATTERN'09)*, Athens, Greece, November 2009, pp. 259–264.

[17] E. Solum, C. Hauser, R. Chakravarthy, and D. Bakken, "Modular over-the-wire configurable security for long-lived critical infrastructure monitoring systems," in *DEBS '09: Proceedings of the Third ACM International Conference on Distributed Event-Based Systems*. New York, NY, USA: ACM, 2009, pp. 1–9.

[18] E. S. Viddal, D. E. Bakken, H. Gjermundrød, and C. H. Hauser, "Wide-Area Actuator RPC over GridStat with Timeliness, Redundancy, and Safety," in *4th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS'10)*. Washington, DC, USA: IEEE Computer Society, February 2010, pp. 17–24.

[19] "INcreasing Security and Protection through Infrastructure REsilience," http://www.inspire-strep.eu/.

[20] D. Germanus, A. Khelil, and N. Suri, "Increasing the Resilience of Critical SCADA Systems Using Peer-to-Peer Overlays," in *ISARCS 2010, 1st International Symposium on Architecting Critical Systems*, ser. LNCS, no. 6150. Springer, June 2010, pp. 161–178.

[21] A. Khelil, S. Jeckel, D. Germanus, and N. Suri, "Towards Benchmarking of P2P Technologies from a SCADA Systems Protection Perspective," in *Proceedings of The 2nd International Conference on Mobile Lightweight Wireless Systems (MOBILIGHT)*, 2010.

[22] P. Maymounkov and D. Mazières, "Kademlia: A peer-to-peer information system based on the xor metric," in *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*. London, UK: Springer-Verlag, 2002, pp. 53–65.

[23] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *SIGCOMM '01: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications*. New York, NY, USA: ACM, 2001, pp. 149–160.

[24] *The Integrated energy and communication systems architecture*, EPRI, Palo Alto and Electricity Innovation Institute, Palo Alto, 2004, www.epri.com/IntelliGrid.

[25] "Critical Utility Infrastructure Resilience (CRUTIAL)," http://crutial.erse-web.it./.

[26] "Integrated Risk Reduction of Information-based Infrastructure Systems (IRRIIS)," http://www.irriis.org/.

[27] A. Giani, S. Sastry, K. Johansson, and H. Sandberg, "The VIKING project: An initiative on resilient control of power networks," in *The 2nd International Symposium on Resilient Control Systems (ISRCS '09)*, August 2009, pp. 31 –35.