BUILDING AN INTELLIGENT GRID
TO MANAGE UTILITY DATA
# Table of Contents

Introduction................................................................................................ 1  
The Problem Space ................................................................................... 1  
Industry Influencers....................................................................................2  
Time Series Data........................................................................................ 2  
Solution Components ............................................................................... 3  
Service Levels............................................................................................ 4  
  Deployment Scenario One—Basic Collection........................................... 4  
  Deployment Scenario Two—Buffered Collection ..................................... 6  
  Deployment Scenario Three—Decentralized Storage and Analytics...7  
  Deployment Scenario Four—Control Centers................................. 8  
Conclusion................................................................................................ 10  
For More Information................................................................................ 10
Introduction
In this paper, Cisco and OSIsoft™ present several networking and data collection topologies that will help utilities scale grid management in response to changing regulations, new technologies, adoption of renewables, and prosumer engagement. Many of these factors are leading utilities to adopt more intelligent data analytics to help manage the tsunami of data evolving from an increase in grid telemetry. An intelligent grid will help enable utilities to do more with less by distributing data analytics and control on top of the grid infrastructure. A distributed data model allows for faster remediation of instabilities in grid state, and a proactive monitoring scheme to alert operators before grid instabilities occur. The net results are more scalable, condition-based maintenance programs, reliable wide area measurement systems, improved visibility, and control of micro-grids.

The Problem Space
Typically more data is available for utilities than they are able to collect and consume. Substations contain numerous analog and digital devices with instrumentation and calculation logic, generating thousands of streams of time-series data. This data contains valuable information which, if available, could be strategically important to both the utility and their customers. Historically though, utilities have faced major impediments to the acquisition and application of useful substation data. Some of these include:
• Telecommunications availability—Substations, be they electrical, gas, or water, are often geographically remote and therefore large distances from existing telecommunication infrastructure.
• Legacy protocols—The source devices have been in place for so long that the protocols they use to communicate data are proprietary and becoming obsolete.
• Operational boundaries—In the past the need to share telemetry data from the grid with other parts of the organization was deemed unnecessary. As a result, there was less of a need to distribute data to multiple departments.
• Environmental challenges—Electrical noise is high. Temperature control is non-existent, and ambient environmental conditions range from desert to rain forest to arctic.
Industry Influencers

Growing population, increased per capita energy consumption, changes in regulation, renewable energy technologies, and an aging fleet are just a few of the forces changing the face of the utility industry. Utilities must be responsive to an ever growing list of demands, and in most cases, each of these brings with it a new required data set. Along with the challenges comes government’s recognition of the aging power grid and its role in overall economic stability.

Utility operations are responding to these industry challenges by implementing programs to improve situational awareness and respond to changes in grid state in near real-time. Some program examples include:

- Uptime—delivering uninterrupted service to their customer base
- Safety initiatives to maintain employee and customer safety, an operational priority for utilities
- Physical and Internet security monitoring of substations
- Use of synchronous phasor data (high speed, high resolution) for real-time grid stability monitoring in wide area monitoring systems
- Centralization of Remedial Action Schemes for grid protection and restoration
- Condition-based and predictive maintenance programs for substation equipment
- Residential advanced metering infrastructure and automated meter reading

Each of these efforts (and hundreds more like them) is in place due to a specific economic or regulatory pressure, and each of them carries a data acquisition requirement that is an order of magnitude greater than the traditional Supervisory Control And Data Acquisition (SCADA) control scenario.

Time Series Data

Regardless of source device, or the nature of the substation, the time-series data generated in the field has a set of homogeneous characteristics. These include:

- An enormous volume of data. Millions of data streams are available, some at high speeds. The engineering time to decide which streams are needed prior to acquisition is not available, incomplete, in error, or cost prohibitive; and so all of the data should be collected.
• Data needed by an unknown (but presumably large) number of users and applications, using an unknown set of tools and technologies and at remote and centralized locations simultaneously.
• Data that must be secure throughout its lifecycle. Proper authentication, authorization, and privacy controls must exist along the complete path from source device to consumer.
• The importance of latency of data transmission and timeliness of arrival, as in the case of power management unit (PMU) data or other event data. The infrastructure must provide options for both high performance communications as well as for edge data processing.
• Scalability of the infrastructure. Simply being able to build a large system is not sufficient. It must be possible to grow in a modular fashion from small research systems up to large production class installations as necessary.

Obviously, addressing a list of challenges like this can be daunting. Doing it multiple times across multiple projects makes little sense, if it is even possible. The better approach is an enterprise-wide infrastructure which implements all of the services required for utilizing time-series data. In order to build such an infrastructure, highly scalable and secure components are needed for networking, compute platforms, and software systems.

**Solution Components**

These enterprise-class components needed to implement a time-series infrastructure are readily available from OSIsoft and Cisco Systems. OSIsoft provides the software components needed to collect and manage time-series data. Cisco provides the network communications, compute, and central data storage needed to deploy the software and manage connectivity between systems and locations in a secure and reliable fashion.

The OSIsoft software suite is called the PI System™. The PI System includes all of the software components necessary to collect, store, organize, analyze, and deliver time-series data. PI System Interfaces make the connections to proprietary as well as standards-based protocols. The PI Server™ provides high-performance data delivery mechanisms for real-time streams, and scalable storage for data archives. PI Analytics™ and PI System visualization tools give consumers of data easy access to raw and computed data in a consistent manner.
Cisco provides the compute and communications platform which connects enterprise data centers and operations centers with the grid. Beyond the enterprise, Cisco® Connected Grid platforms include industrial-grade routers and switches which form the foundation of the substation network. The Cisco 2010 Connected Grid Router (CGR 2010) provides intelligent routing to and from the substation. Using embedded security, the CGR 2010 secures critical Internet assets and prioritizes control data through the network. The Cisco 2520 Connected Grid Switches (CGS 2520) form the substation LAN. The CGS 2520 interfaces with Ethernet-enabled, Intelligent Electronic Devices (IEDs), relays, and Phasor Measurement Units (PMU) providing ultra-low latency switching for real-time grid communications inside the substation. The Cisco Connected Grid portfolio provides a highly secure, scalable network platform to manage the growth in data streams across substation networks.

Service Levels
At the most basic level, diverse forms of data need to be acquired from various sources and delivered to a set of destinations for processing and storage. End user visualization tools and transactional business systems represent common destinations, but this is certainly not a comprehensive list. In almost all cases, analytic calculations are required to be performed on the data either just after acquisition (edge processing) in transit, in bulk after storage, or both. Where along the process continuum these functions exist define the various tiers of capability, performance and reliability.

The infrastructure must be flexible enough to support many architectural permutations. Not all topologies are required to meet every use case. In addition, the infrastructure must be highly scalable—allowing the operator to start small and grow as more and more data streams are brought online and under analysis.

Deployment Scenario One—Basic Collection
The entry level of service involves establishing a connection between the source of data and the end users utilizing a minimum of components. Figure 1 displays what data acquisition in a substation and routed natively to a central store looks like.
In this configuration, Cisco Connected Grid platforms transport the substation data back to the control center and data center in a highly secure manner. Legacy protocols are encapsulated in a tunnel, but are exposed directly to PI System Interface programs running in the control center.

This architecture requires the smallest footprint in the substations. However, this is also the lowest possible level of service. Situational awareness is provided by exposing previously unavailable data streams, but minimal data processing is available.

At the utility control center, this architecture concentrates the data processing, analysis and storage in a central location. Cisco Unified Computing Systems® (UCS®) help enable the utility to start with a small number of the Cisco UCS C-series rack mount servers with integrated storage and the Cisco Nexus® family of data center switches. As more substations are brought online and the data volume and bandwidth requirements increase, additional servers and switches can be easily added.
Deployment Scenario Two—Buffered Collection

The next logical expansion of the architecture is to provision computing services at the substation. This allows for the deployment of PI System Interfaces with a local connection to their source devices. Figure 2 illustrates data acquisition at the substation.

Advantages to this service level are numerous:

- PI System Interface software is capable of buffering data streams destined for the control center. This protects the data from connectivity outages, storing data values temporarily when necessary and forwarding them (in order) as soon as possible.

- PI System Interfaces can intelligently analyze a stream of data and communicate only defined significant changes upstream. This is referred to as exception processing. In this way, significant bandwidth savings are realized with no loss of information or situational awareness.

- The network enforces policy to secure communications and route data to qualified sources in the control center. In addition, PI System Interfaces effectively serve as a demilitarized zone (DMZ), or perimeter network, between the control center and the field devices.
• Server presence in the substation enables controlled remote access to substation management packages which provides enhanced data acquisition beyond what is possible when the PI System is centralized in the control center.

• This design adds significant processing capacity to the architecture as a whole. Distributed processing of the interface software enables essentially unlimited scalability, while insulating the processing of each individual substation’s data.

Deployment Scenario two is generally considered the entry point for a fully functional architecture. While scenario one deployments are sometimes all that can be justified (and should be used in these cases), the benefits of a scenario two deployment generally justify any associated costs.

Deployment Scenario Three—Decentralized Storage and Analytics
Decentralizing data storage and analytic capabilities at the substation further enhances the capabilities of the architecture. Decentralized storage allows processing at the substation for lower latency applications. As more distributed energy resources come online, the substation evolves into a distributed data center for managing variable energy sources. These mini-data centers report real-time grid state based on data processed from distribution feeders, substation assets, and distributed generation sources. Figure 3 outlines a network diagram for a decentralized topology.

Figure 3. Data Acquisition and Storage at the Substation and Sent to Central Store
With local processing and storage, a full range of applications can be run in the substation itself. This is important when:

- A substation must be self-sufficient, even in an islanding condition.
- The real-time analytics to be performed cannot tolerate the latency of a wide area connection.
- Substations staffed with personnel are candidates for tier three deployment. Personnel need responsive access to data and are often consuming data in ad-hoc ways.
- Distribution of analytics is needed for scale. Complex, statistical or model based analytics are often better executed on more numerous platforms in a substation rather than on centralized super-computers.
- Telecommunications do not enable transmission of all data continuously. In the event that the telecom system can only handle a subset of the data, even after exception processing, then local storage will be required.

**Deployment Scenario Four—Control Centers**

At the top of the service stack is the control or data center, which houses the bulk of data storage and processing applications. Control centers generally house the bulk of data consumers, especially human users. The data center is also where connections to other lines of business systems most commonly take place. Figure 4 highlights a control center data usage architecture.

*Figure 4. Control Center Data Usage Architecture*
As explained above, the control center architecture may start small with one or more Cisco UCS C-series rack mount servers with integrated storage. VMware’s vSphere virtualization offers flexibility to support multiple operating system (OS) environments for application hosting on the Cisco UCS, reducing the initial investment to bring up multiple applications. In addition, the Cisco UCS Manager (UCSM) offers ease of deployment, reconfiguration and scaling using hardware virtualization and service profile management, and OS and application virtualization using vSphere. The benefits to the utility are reduced management and administration expenses, increased agility and reduced capital costs from server consolidation.

Architectural specifics are harder to generalize at Deployment Scenario four, as requirements begin to vary by customer installation. However, certain aspects are common among good scenario four designs. These include:

• **Modularity.** Scenario four architectures need to be expandable to a much greater extent than other scenarios. New uses for data are always being developed, and additional data streams arrive as the field installations grow.

• **Redundancy.** A scenario four system is always in use and so all maintenance must happen on the system while it is operational. Likewise, “five 9’s” or better availability is commonly required, and so the system must be fully fault tolerant.

• **A streamlined and centrally managed security model.** Scenario four has the largest number of components and the largest number of users. To ensure that systems of this size are adequately secured and that security is maintainable, a centralized directory-based approach is commonly mandated.

• **Connectivity.** User and business systems will need various quantities of data, with varying performance needs, from multiple technology platforms. Both the networking and the software stack need to allow for these different types of connectivity.

Deployment Scenario four systems frequently span multiple architecture tiers, with systems on operational networks, business networks, and DMZ networks. As these systems grow the ability to remain flexible and yet manageable becomes more and more critical to long term success.
Conclusion

In order to move from project-based solutions to a true enterprise-wide infrastructure for real-time data, a combination of advanced networking, scalable server architectures, and software is required. Cisco and OSIsoft provide the complementary technologies that enable universal access to data.

For More Information

For more information on the Cisco Connected Grid portfolio of products please visit http://www.cisco.com/web/strategy/energy/external_utilities.html.

For more information on the OSIsoft PI System please visit www.osisoft.com.