Arnold Schwarzenegger
Governor

MAINTAIN, ENHANCE AND IMPROVE RELIABILITY OF CALIFORNIA'S ELECTRIC SYSTEM UNDER RESTRUCTURING

APPENDIX VI

Synchronized Phasor Measurement Information Server: Architecture and Data Model

PIER FINAL PROJECT REPORT

Prepared For:
California Energy Commission
Public Interest Energy Research Program

Prepared By:

Lawrence Berkeley National Laboratory

CERTS
CONSORTIUM FOR ELECTRIC RELIABILITY TECHNOLOGY SOLUTIONS

MARCH 2006
CEC-500-2006-035-APVI
LBNL-58939

Prepared By:
Lawrence Berkeley National Laboratory
Consortium for Electric Reliability Technology Solutions (CERTS)
Joseph H. Eto
Berkeley, CA
Contract No. 150-99-003

Prepared For:

California Energy Commission

Public Interest Energy Research (PIER) Program

Don Kondoleon,
Contract Manager

Mark Rawson, Bernard Treanton, Linda Kelly, Ron Hoffman, Don Kondoleon
Project Managers

Mark Rawson
Program Area Team Lead

Laurie ten Hope
Office Manager,
ENERGY SYSTESM INTEGRATION AND ENVIRONMENTAL RESEARCH OFFICE

Martha Krebs, Ph. D.
Deputy Director
ENERGY RESEARCH AND DEVELOPMENT DIVISION

B.B. Blevins
Executive Director

Consortium for

Electric

Reliability

Technology

Solutions

Synchronized
Phasor
Measurement
System

Synchronized Phasor Measurement Information Server Architecture and Data Model

Revision 1

July 20, 2001

Prepared by

Global Enterprise Managers, Inc.
1434 Malcolm Drive
Dresher, PA 19025
215-659-7004
www.global-enterprise-mgrs.com

Table of Contents

List of Exhibits

1. Introduction and Executive Summary

The Consortium for Electricity Reliability Technology Solutions (CERTS) is developing and demonstrating a series of modular, but integrated, computer-based real time grid reliability management applications to assist the operating authorities (e.g., ISOs, RTOs, and Security Coordinators) in their management of the grid. One of the specific CERTS initiatives for the 2000-2003 time period is to develop and demonstrate security monitoring applications, designed for dispatchers and operating engineers, using Synchronized Phasor Measurements (SPMs.)[1] Synchronized Phasor Measurements are accurate measurements of voltage and current values that can be precisely correlated in time across widely dispersed locations.

SPMs can be characterized as a distributed measurement set with very high data volume. These two characteristics pose considerable challenges to making information available to interested users where and when it is needed. In addition, over the several years that have elapsed since SPMs were first implemented, the proprietary nature and competitive value of the information that can be derived from these measurements has also been recognized, leading to increased data security concerns.

Previous research yielded recommendations for a very capable and general information management environment that could accommodate all of the types of information needed for wide area monitoring of the electric network.[2] The work summarized in the current document has the more modest goal of defining a prototype information management architecture that can be used to test and validate the SPM analysis tools while offering a migration path to a more comprehensive SPM information architecture.

In summary, the basic recommendations of this investigation are the following:

1) SPM data should initially continue to reside in file structures on servers physically close to the concentrating points for SPM data and under the control of the entity that owns the data. Migration of this file structure to database storage in the future is desirable but not essential.

2) CERTS should create a database with metadata that describes the locations, organization, structure, and quality of the distributed, file-based SPM data.

3) The SPM information server design should include readily configurable security rules that can restrict access by monitoring point, data collection time, or both.

4) An API (application programming interface) should be provided that permits access to the SPM data in two ways:

   a) As if it were simply in a set of files on an FTP server

   b) By specifying the data and time range of interest, independent of the physical location or file structures of the data.

5) Publish/subscribe technology should be used to propagate essential SPM data in near-

real time for dispatcher purposes.

The remainder of this document provides more details regarding these recommendations.

2. Requirements

2.1 Business Requirements

The security monitoring applications will support both operating engineers and dispatchers. These two groups of users have both shared and unique information server requirements.

The fundamental requirements that are shared across the two groups are:

- A simple and consistent approach to accessing data, regardless of where or when it was gathered
- Protection of data that has been identified as proprietary or of a competitive nature
- Confidence in the quality of the data.

For operating engineers, it is expected and acceptable that the user will explicitly define the desired data, that a longer time delay for data retrieval may be acceptable, and that additional effort may be expended to interpret and validate the data. For dispatchers, it is essential that the information be available automatically in near-real time, that no manual intervention be required, and that any suspect data be identified as such without requiring further interpretation by the end user.

2.2 Performance Requirements

Specific performance requirements for dispatcher applications have not been developed as part of this investigation. Nonetheless, the key requirements are expected to be:

- Required time window for analysis (number of samples)
- Allowable latency
- Number of points required.

After these requirements are developed as part of the dispatcher SPM analysis algorithm design, the following characteristics of the SPM Information Server architecture can be used to help meet those requirements:

- High speed networking between the SPM Information Server and the dispatcher workstation
- Effective data compression algorithms for transmitted data
- Server-based event detection and publication of notifications to dispatchers.

Subscription, file, or query "cost", primarily in terms of processing load and bandwidth requirements should be estimated in advance. It should be possible to reject a subscription or access request if its estimated resource requirements exceed a threshold.

3. SPM Information Server Architecture

3.1 Technical Approaches

The current SPM system has a distributed information exchange capability for near-real-time data. But the usefulness of this capability is limited by the fact that a Phasor Data Concentrator (PDC) must be available to act as the data receiver for data gathered by other PDCs. The SPM Information Server Architecture described in the current work is intended to make SPM data more widely available without requiring the existence of special hardware at end user sites.

Today's wider availability of high-speed networks and the ubiquitous nature of Internet technology make it feasible and desirable to create a simple information transport approach based on commercial technology.

3.1.1 Measurement Data and Metadata

A measurement system such as the set of SPM measurements available in some area actually consists of two sets of information:

- Measurement values (voltage, current, time, data quality)
- Metadata (identification of measurement points)

In the current SPM system, both types of information are embedded in the measurement files, supplemented by associated initialization (INI) files. A limited amount of metadata-like information is also embedded in the data file names by means of file naming conventions.

A full set of metadata must be created in order to organize this distributed data environment. Metadata contents should include:

- Mapping of logical point names to physical file locations
- Data access permissions
- Overall data quality indication (detailed quality information for each sample should remain with the measurement values.)

For the purpose of the prototype, GEM recommends that the measurement values initially be left in files but that the metadata be migrated immediately to a database.

A set of consistent application programming interfaces (APIs) should be developed to allow access to the measurement values in two ways:

- As files

- Through an ad hoc query by point, time range, and granularity.

In the future, the data can also be migrated to a database, if desired. All APIs will use the concept of *encapsulation* to minimize future changes to applications if the data is migrated to a database. Encapsulation is simply the practice of hiding the details of data access and processing that takes place in response to a data request through the API.

3.1.2 Security

Security information should be defined within the metadata database. It should be possible to protect against access to any SPM data based on either or both of the following criteria:

- Point name
- Time range.

3.1.3 Data Processing

Modern, object-oriented software development techniques can be used to develop reusable components for construction of the data processing module. It is very straightforward to create objects in a language such as Java or C++ that perform all typical types of simple manipulations (filtering by security level, reformatting, data cleansing) and that are capable of interfacing external applications where required for more sophisticated analysis.

Processing will generally be performed *on demand*, i.e., data should generally be left in its raw form until requested by users. This provides the maximum flexibility, particularly for applying the essential security rules.

The data processing will be implemented as a distributed set of data processing capabilities that include automatic access to other data processing servers when needed to satisfy any information request. So, for example, a user may have a preferred SPM information server to which the user directs his or her request. (This should generally be the server that has the most information of interest to the user.) Nonetheless, if the request cannot be satisfied in whole or part by that server, the server will be able to request the needed information from other servers, based on the location and security metadata that is available in the Metadata Database.

3.1.4 Publish/Subscribe Technology

Publish/subscribe technology disseminates new or changed data from a data originator (publisher) to any systems that have indicated interest in that data (subscribers). The role of the SPM information server will be to:

- Provide the subscription service, i.e., a means to register interest in data
- Accept changed data from:
    - The PDCs, through their raw stream reader capabilities
    - Analysis programs that operate on the stream data and identify events of interest.

- Notify the subscribers, and ensure successful delivery to the subscribers.

The same security rules for historical data access should be used for the publish/subscribe service.

Numerous commercial middleware tools exist to provide publish/subscribe capabilities. Most are very powerful and general tools designed to solve a variety of communications and integration needs, and unfortunately they are priced accordingly. It should be adequate to create a simplified publish/subscribe scheme for the initial prototype and then migrate to a commercial publish/subscribe middleware tool in the future if necessary when full deployment of the SPM application occurs.

3.2 Phased Implementation

Exhibits 3.2-1 through 3.2-3 illustrate a phased implementation of the recommended architecture.

Currently, a typical approach for analysts to share data is based on FTP transfer of files as shown in Exhibit 3.2-1. Utility programs currently exist for connecting to an FTP server, transferring the files, and converting them to various formats.

In addition to the need for the user to be familiar with the locations and naming standards of the data files on remote computers, the current approach cannot conveniently enforce security except through the mechanism of enabling or disabling access to entire directories or files. This is a highly manual operation, which will not foster the wide dissemination of information.

Exhibit 3.2-2 shows the recommended first step toward the SPM information architecture. A Metadata Database will be established and will serve as the central repository for security information and data location information. The Filtering and Processing function, which will reside at each PDC location, will be designed as an extensible framework of SPM processing modules. The initial implementation will incorporate filtering by security attributes and an external API (the FTP Proxy Plug-In) that provides a subset of FTP commands needed to download files. The net result of this initial architecture is that existing end user applications that use FTP can continue to operate without change, but the security policies for data access will be enforced.

A web interface to the Metadata Database for determining data location will also be implemented during this phase.

Exhibit 3.2-3 shows the complete architecture to support both operating engineer and dispatcher users. The architecture has the following characteristics:

- The FTP Proxy Plug-In will continue to operate as described in Exhibit 3.2-2. Its functionality may be extended, if desired, by creating virtual files that span data resident on multiple SPM information servers.

- The Query Plug-In will allow submission of a data query by points and time range. Information will be returned in simple data sets or structured files of either a fixed format or a tagged format, such as XML. This data exchange can be either an actual file-based exchange or a direct application-to-application data object exchange (or both), based on a more detailed assessment of user requirements.

- The Publish/Subscribe Plug-In will allow clients to register interest in particular data streams and/or predefined event types.

- The Analysis/Notification Plug-In can be used to perform monitoring of critical data points (either directly or by interfacing to other analysis engines) for predefined patterns. The results of this type of monitoring would normally provide the triggering mechanism for publishing specific events using the Publish/Subscribe Plug-In.

- The Extender Plug-In creates the true distributed information server environment. When a query or subscription request is received that cannot be satisfied by a server, the server references the Metadata Database and retrieves or subscribes to the needed information from the server where the information resides. This information is merged with locally available information to create the required data set.

Although individual connections are shown in the conceptual diagram of Exhibit 3.2-3 between the FTP Client, the Query Client, and the Subscription Client and their respective information server plug-ins, the actual connection and transport approach will actually be a high-speed TCP/IP connection, either via the Internet or a private wide area network.

The information server architecture described here also serves as an adapter approach, providing a means of merging data from various types of phasor measurement systems into a consistent information architecture.

Exhibit 3.2-1
Current Data Exchange Technique for Analysis

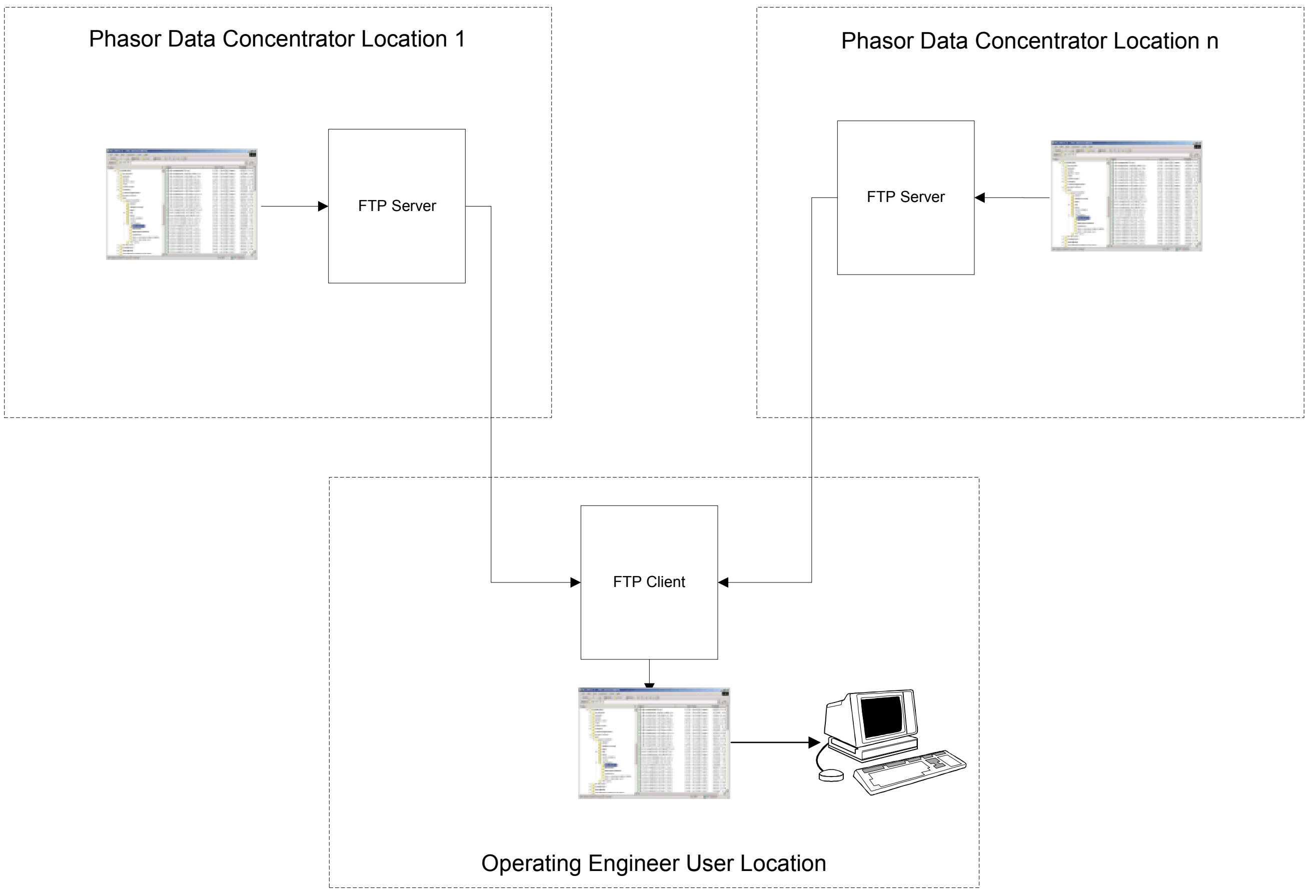


Exhibit 3.2-2
Secure Extraction of Data

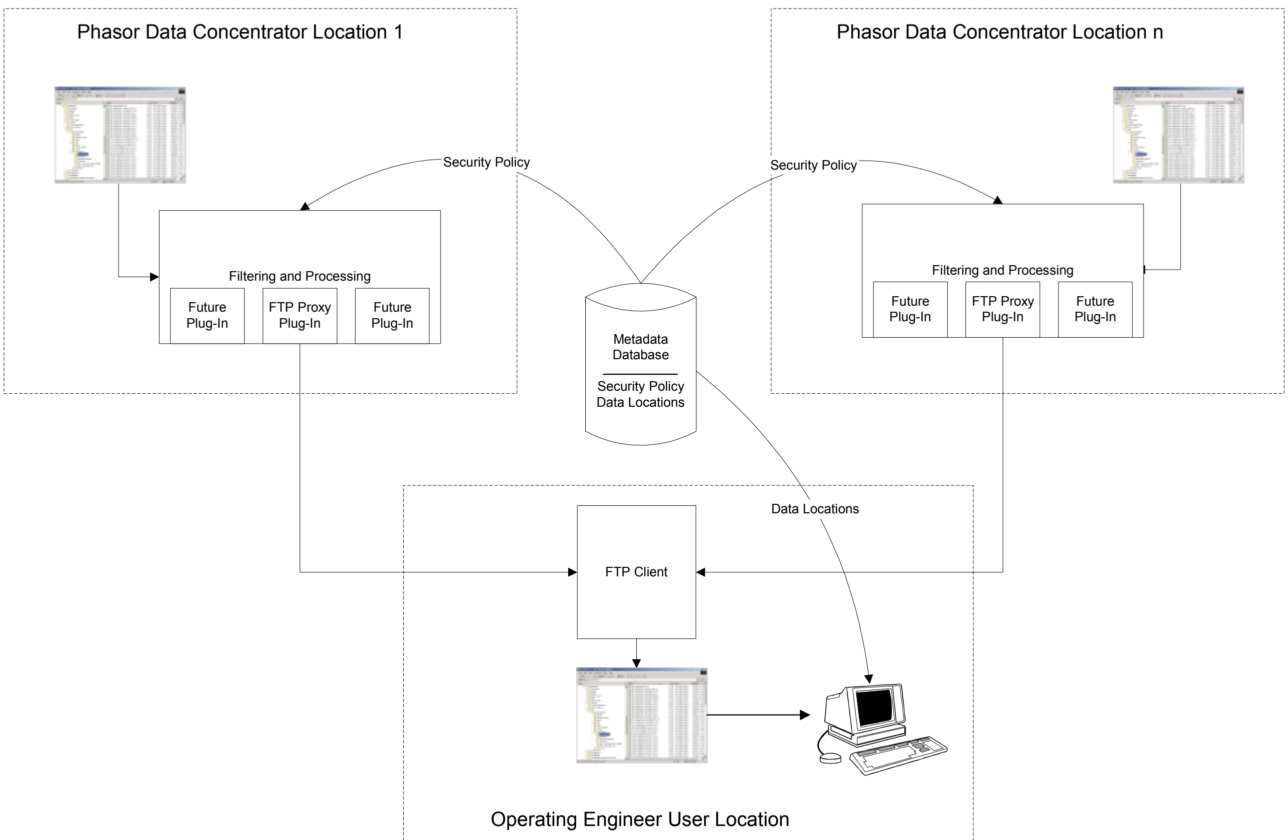
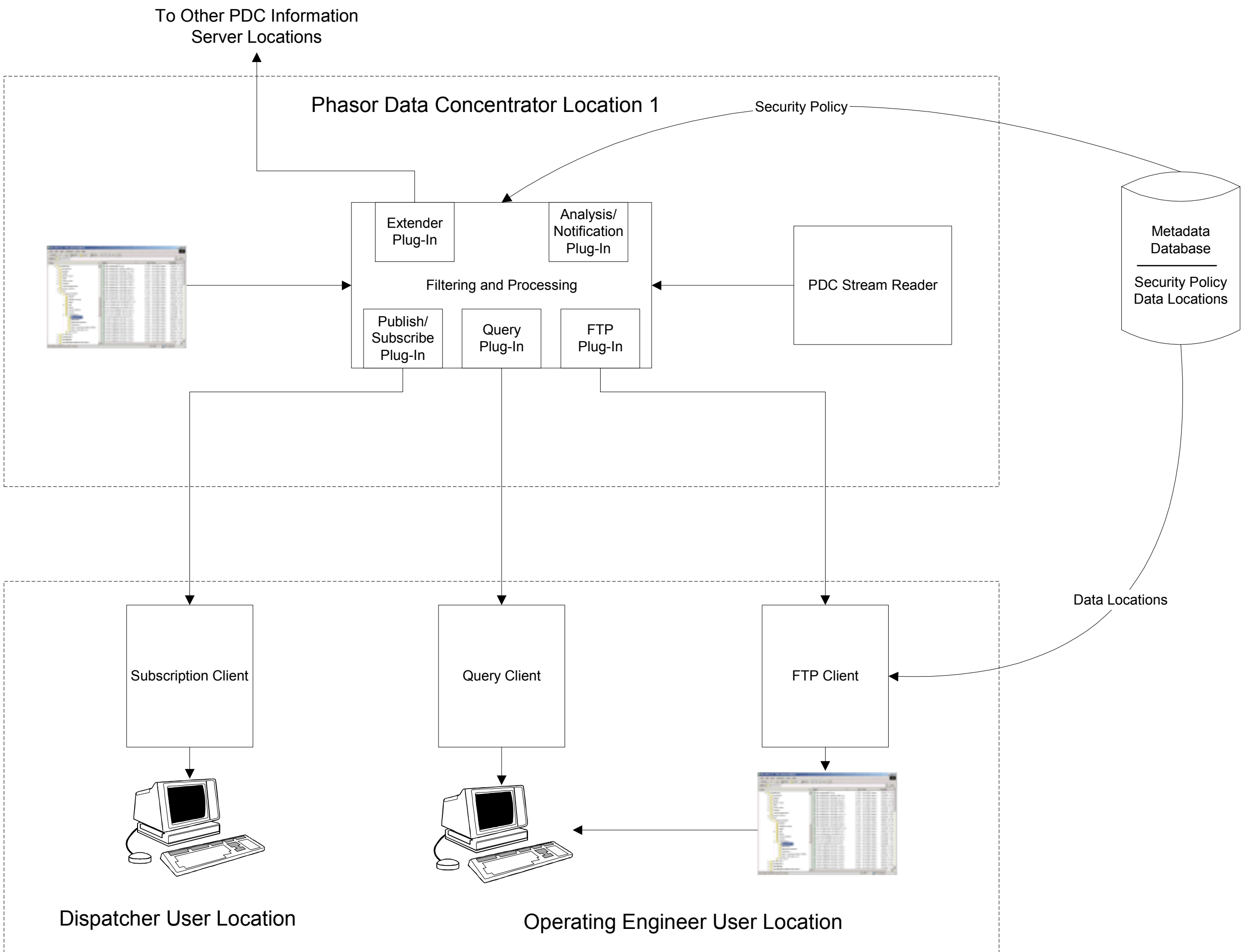


Exhibit 3.2-3
Full Information Server



4. Implementation Timetable

If work begins immediately on the SPM information server architecture, it should be possible to meet the following implementation timetable for a prototype SPM information server that provides the capabilities described in this document:

- 3Q 2001 – System Design
- 4Q 2001 – Implement secure FTP-syntax proxy access for operating engineer SPM workstation
- 1Q 2002 – Implement secure publish/subscribe and extender plug-in to support dispatcher SPM workstation
- 2Q 2002 – Implement query access for operating engineer SPM workstation

5. References

1) *Draft - Functional Specification for Security Monitoring Prototype Workstations for Dispatchers and Operating Engineers using Synchronized Phasor Measurements*, Transmission Reliability Program, Office of Power Technologies, U.S. Department of Energy, Washington, DC: October, 2000

2) *A Dynamic Information Manager for Network Monitoring of Large Power Systems*, EPRI, Palo Alto, CA: 1999. TR-112031.

3) *Full PDC Manual*, Bonneville Power Administration, Portland, OR: July, 2000.

Appendix A SPM Information Server Data Model

Appendix A
SPM Information Server Data Model

Exhibit A-1 shows a preliminary, high-level SPM Information Server data model. The primary emphasis of this data model is the metadata needed to enforce security, represent the locations of data throughout the distributed SPM data storage environment, and support operation of the information server. This information would be stored in a centralized Metadata Database, which could be replicated for reliability purposes. Such replication would allow operation of the distributed environment even if the main Metadata Database server were unavailable.

A sample table for actually storing data points is also shown. Similar to file storage, GEM recommends that database storage of measurements be distributed, most likely at the same locations as the Phasor Data Concentrators. Proper indexing of these tables, which will become extremely large, will be critical to achieving acceptable performance. Database management systems, such as Oracle with time series enhancements, are also able to segment large sets of temporal data across multiple tables to reduce data access time; prior to committing to a database implementation for phasor measurement storage, performance tests of this type of approach should be conducted.

The security schema of this data model assigns users to access groups and in turn grants permission to access data on a coarse level (the PDC) and a fine level (individual measurement point.) Point permission includes an access level. For the SPM Information Server functions, Read access or No access levels will be sufficient, but it may also be useful to define Create, Update, and Delete permission levels so that the same security tools can be used for maintenance of SPM data. The *historical* and *recent* attributes of the PointPermission entry control access along the time dimension.

Additional database tables for maintaining security policies (e.g., password expiration) are not shown in the data model, because implementation of these policies is not unique to this type of system.

Most database entities in this model include versioning, allowing the configuration of PMUs and PDCs to evolve over time while still allowing access to historical data.

Exhibit A-1
SPM Information Server Data Model