CERTS Database

Ernest Orlando Lawrence Berkeley National Laboratory

Chris Marnay *Afzal S. Siddiqui*

Jeremy V. Guttenplan *Mark Khavkin*

Ernest Orlando Lawrence Berkeley
National Laboratory
1 Cyclotron Road
Berkeley, CA 94720

August 2000

Contents

Introduction

To facilitate research in restructured electricity markets, the Consortium for Electric Reliability Technology Solutions (CERTS) has been collecting and maintaining a database of California electricity market information that covers the entire period that the California Power Exchange (CalPX) and California Independent System Operator (CAISO) have been operating, that is, since 1 April 1998. All data currently in place are in the public domain. To the best of our knowledge, this is the most complete data set freely available with up-to-date data from each California market.[1] In June 2000, with funding from the Department of Energy (DoE), Berkeley Lab initiated the process of building a database on a more robust and readily accessible platform, that is, an Oracle 8*i* database running on Solaris with web-enabled query capability. The maintenance of such a database enhances the ability of CERTS to conduct research on reliability of power systems in restructured markets, electricity market operation, and efficient design of competitive electricity markets. This document describes the database as it currently exists and plans for its improvement.

California Electricity Markets

The existing database includes information on the CalPX day and hour energy markets, and on the ten CAISO ancillary services (AS) markets plus its imbalance energy market (see Figure 1). All of these

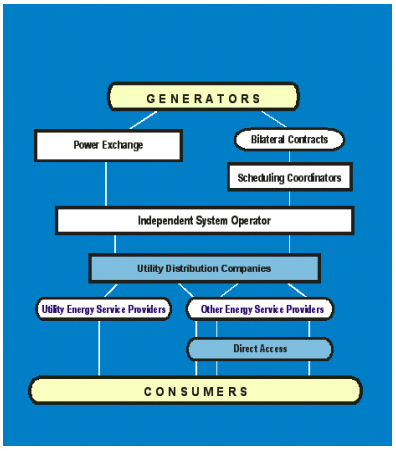


Figure 1 California Market Structure (Source: DoE)

markets are currently settled hourly, and simple prices and quantities are reported.[2] Data covering the AS markets are categorized by the AS, and by the twenty-four California electricity congestion zones (see Figure 2). The ten currently operating AS markets are the day and hour ahead markets for regulation up, regulation down, spinning reserve, non-spinning reserve, and replacement (see Appendix for definitions). For each

[1] Although the data are kept current, they have not yet been cleaned.

[2] CAISO markets will soon be settled in ten-minute intervals.

AS in each zone, the quantity and market-clearing price are given, along with information on the actual *ex post* generation. When no line congestion exists, CAISO buys these services from any source and reported prices are the same at every congestion zone. When congestion exists, CAISO begins buying AS separately for each congestion zone, and prices diverge.[3]

Database Structure and Capabilities

The CERTS database is relationsal and designed to store data from multiple ISOs in a flexible manner that reduces redundancies and facilitates updates. Indeed, administrative and economic changes such as the addition of new zones and markets, or changes in the way that information is reported will not disrupt the collection and maintenance of data. Examples of such data include disaggregated bid data of the generators that the Federal Energy Regulatory Commission (FERC) has already approved the CAISO to release. This decision means that beginning July 2000, bid data will be routinely released following a six month grace period. Another significant increase in the volume of CAISO data will also occur when the imbalance energy market is switched from hourly to ten-minute settlements which is expected to occur in August 2000. Data from other operating ISOs, notably New York, New England, and Pennsylvania-Jersey-Maryland, are

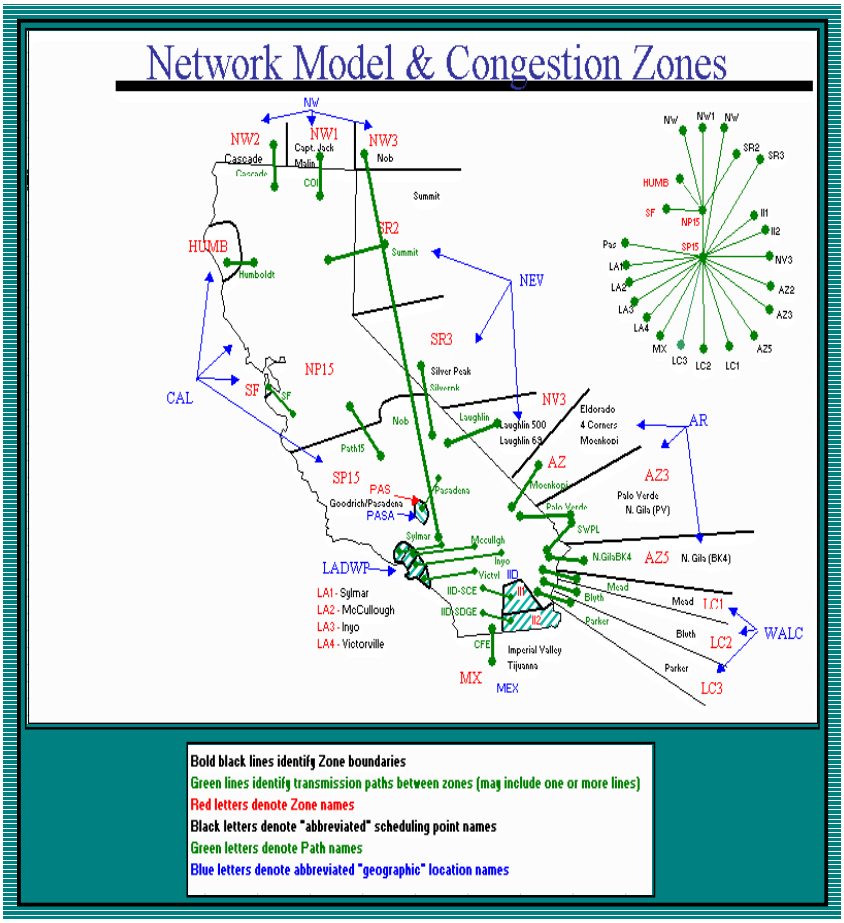


Figure 2 California Zonal Map (Source: CalPX)

available and could be easily added in a consistent format. The development of the CERTS database is summarized by Figure 3. The strength of the database's design is in its simplicity and flexibility. A look at the entity-relationship model (see Figure 4) illustrates that there are few tables, and the relationships among them are straightforward. Currently, data are available for only California, but data from other ISOs can be easily accomodated. The centerpiece of the database is the *Market_Equilibrium* table. This entity stores market-clearing prices and quantities for *all* CalPX and CAISO markets.[4] Not surprisingly, it is quite large (currently exceeding five million records). However, since its main purpose is to be queried (rather than to be opened), its large size is not a drawback. By comparison, the other tables are small. Currently, all tables in the database are operational except for the *Market_Bid*, *Market_ Demand*, and *Market_Supply* tables, which will contain bid data. The confidential *Market_Bid* table will not be accessible to the public, but the rest of the database can be queried through the internet. Researchers investigating the impact of deregulation will find this database to be an invaluable source of information. In fact, the CERTS database has already been used to provide empirical support for policy studies at Berkeley Lab and by CERTS research partners.

[3] Over the period 1 April 1998 to 31 September 1999, congestion existed for approximately 10% of the hours.

[4] To our recollection, the only missing data are market-clearing quantities for the CalPX constrained markets.

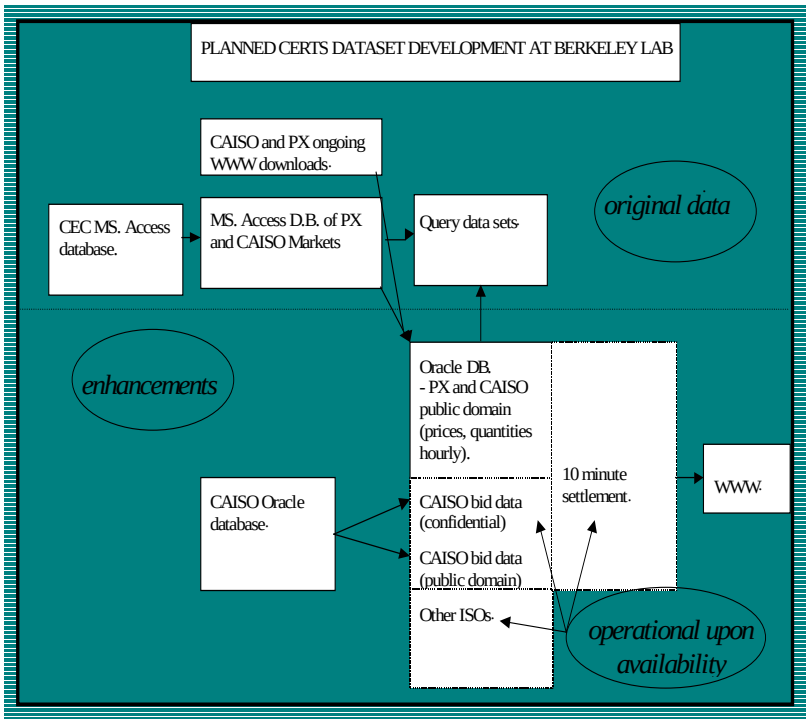


Figure 3 CERTS Database Development

Examples of the potential of the database can be seen from Figures Figure 5, Figure 6, and Figure 7. By using structured query language (SQL) queries, data can be extracted to yield information about market conditions. This information can then be represented in graphical terms to shed light on complex market behavior. In Figures Figure 5 and Figure 6, the regulation reserve price is plotted from the inception of the market until 31 March 2000 along with notes on changes in market rules. As a result, the effects of market reform on prices are simple to identify, and the data can be used by researchers to test theoretical claims. Figure 7 shows how the cost of purchasing regulation reserves (as a percentage of the CalPX day-ahead market value) has evolved during the first two years of the market's operation. While the hourly percentages (in blue and not shown for values above 100%) have not evolved with any predictable pattern, the *overall* percentage (in pink) has declined

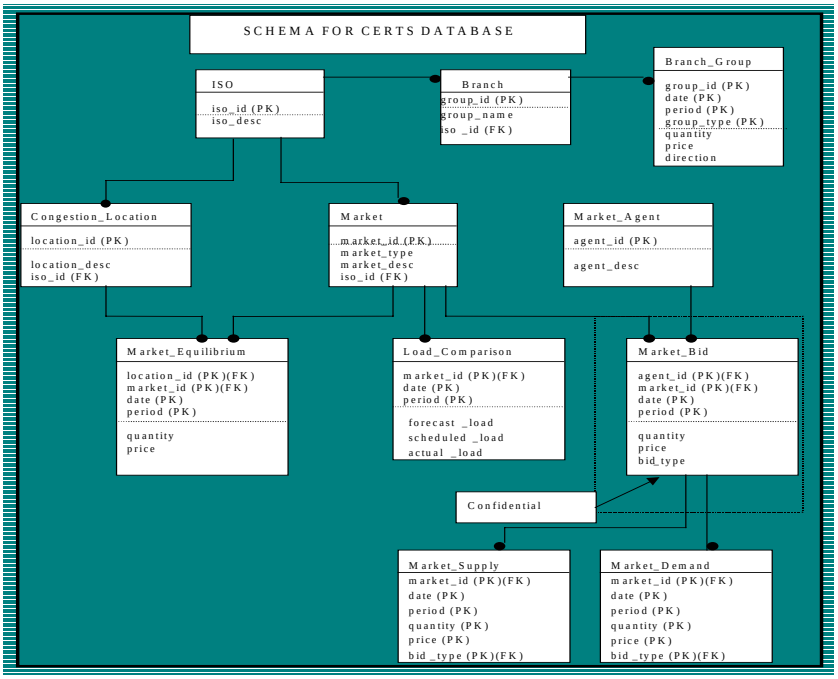
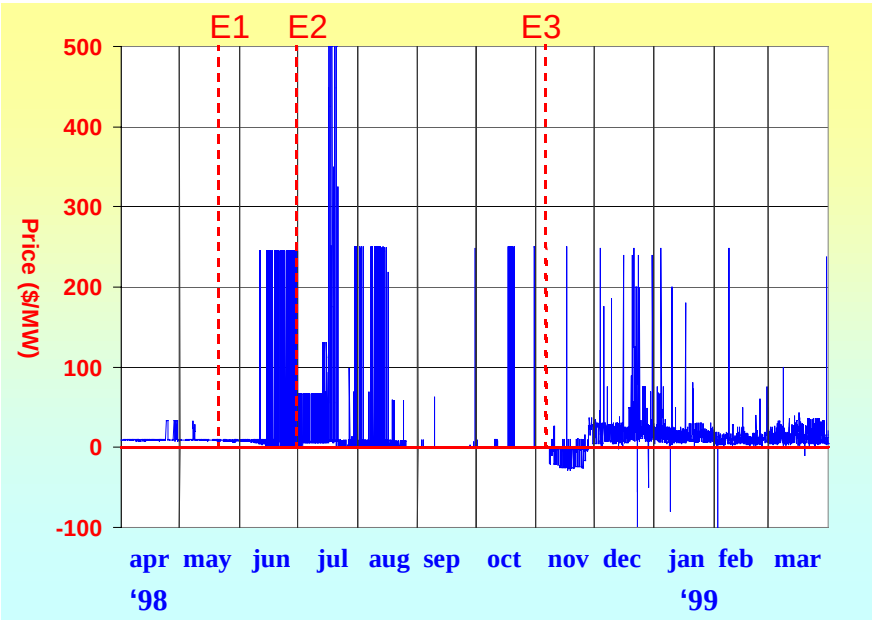


Figure 4 CERTS Database Schema

steadily. This might imply that the reforms have had the intended effects. At the very least, this one-of-a-kind database will provide researchers with a means to investigate the impact of deregulation on electricity reliability markets.

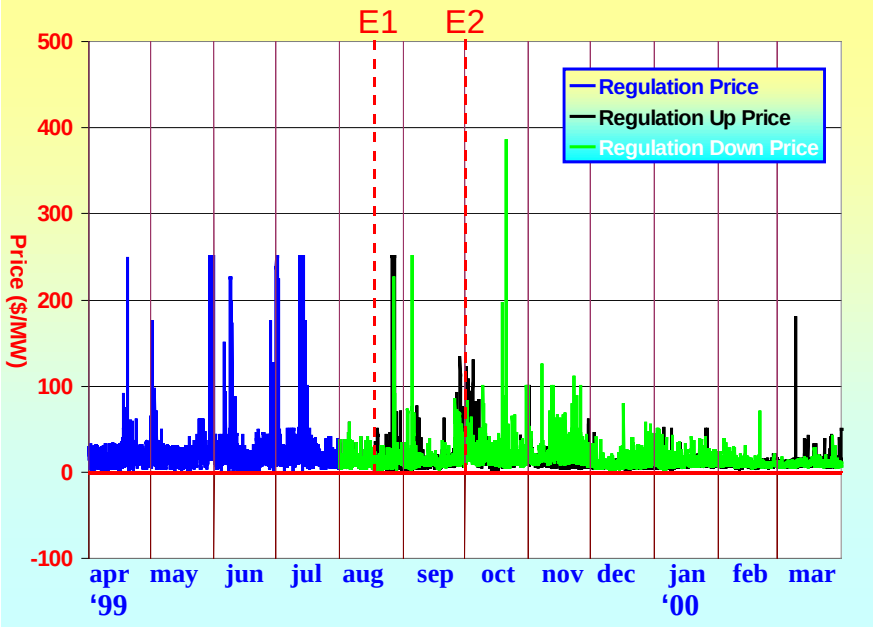


E1. 21 may 1998 - REPA

E2. 30 june 1998 - market based rates approval began

E3. 6 november 1998 - ISO reminds market participants of negative bids

Figure 5 First Year Regulation Reserve Price (Source: CAISO)



E1. 18 august 1999 - separation of regulation up and down

E2. 1 october 1999 - price cap is raised to $750/MW

Figure 6 Second Year Regulation Reserve Price (Source: CAISO)

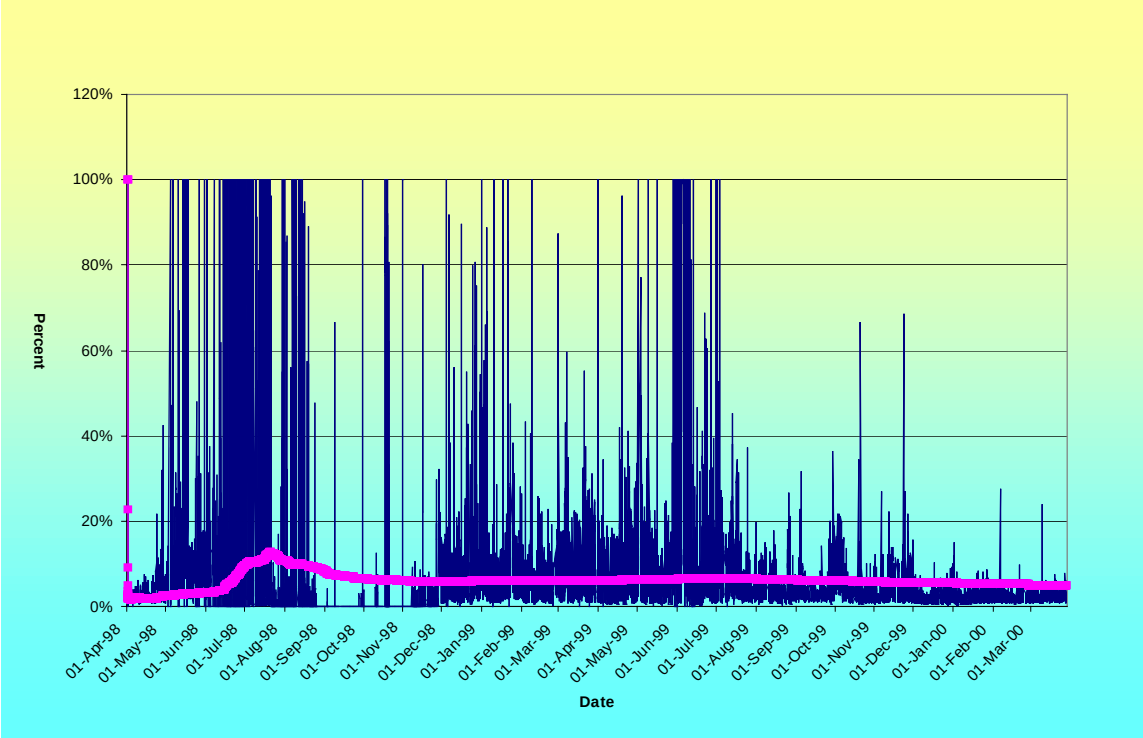


Figure 7 Regulation Reserve Expenditures as Percentage of CalPX Energy Value (Sources: CAISO and CalPX)

Appendix

In this section, we provide some technical specifications of the CERTS database.

First, the various ancillary services traded competitively in California are defined:

1. Regulation Up: increased use of generation equipped with governors and automatic generation control to maintain minute-to-minute generation/load balance within the control area to meet NERC control-performance standards

2. Regulation Down : decreased use of generation equipped with governors and automatic generation control to maintain minute-to-minute generation/load balance within the control area to meet NERC control-performance standards

3. Spinning reserve : the provision of generating capacity (usually with governors and automatic-generation control) that is synchronized to the grid and is unloaded that can respond immediately to correct for generation/load imbalances caused by generation and transmission outages and that is fully available within 10 minutes and is able to maintain output for 120 minutes

4. Non-spinning: the provision of generating capacity (usually with governors and automatic-generation control) that is not synchronized to the grid and is unloaded that can respond immediately to correct for generation/load imbalances caused by generation and transmission outages and that is fully available within 10 minutes and is able to maintain for 120 minutes

5. Replacement: generation available within two hours to compensate for the transmission system or other generator outages.

In defining the roles of the database tables, the “PK” notation denotes the primary key, i.e., all of the

attributes (or fields) of an entity (or table) that uniquely identify a record. Similarly, "FK" signifies a foreign key, i.e., an attribute that refers to those in another entity and is used to specify relationships between entities. The roles of the entities in the database and their relationships are summarized as follows:

1. ISO: has the ID number and description of the ISO's
2. Branch: a child of ISO; this identifies and describes a group of transmission lines
3. Branch_Group: a child of Branch; this identifies for each period in time and each group of lines the direction, quantity, and price of congestion
4. Congestion_Location: a child of ISO; this simply has all of the congestion locations (be they zones or nodes) within the relevant ISO's jurisdiction
5. Market: another child of ISO; it contains all of the markets operated within the ISO's domain (including markets not operated by the ISO, such the PX in California)
6. Market_Agent: a list of all the players that bid into markets; this is not a child of ISO because market participants are not tied down to an ISO
7. Market_Equilibrium: for each time period, this lists the market-clearing quantity and price for a given location and market; this is a child of both Market and Congestion_ Location
8. Market_Bid: a child of both Market and Market_Agent, this table stores the quantity and price bid into a specific market by a given agent; consequently, this is the only confidential data table in the entire database
9. Market_Demand: a child of Market; this table aggregates the demand bids from Market to make a segment of the demand curve for the given market at a point in time
10. Market_Supply: a child of Market; this table aggregates the supply bids from Market to make a segment of the supply curve for the given market at a point in time
11. Load_Comparison: a child of Market; this simply lists the forecasted, scheduled, and actual loads for each market at each point in time

The attributes of these entities are also defined carefully to prevent redundancies or confusion. For example, the "date" field is set according to standards specified by the International Standards Organization, i.e., the date 24 December 1999 is represented as 1999/12/24. Also, to refer to a time in the day when an event occurs, we use the field called "period", instead of "hour." This is to keep the design as flexible as possible since the CAISO will change hourly reporting of market-clearing prices to ten-minute increments and data from other ISO's or futures markets may not be hourly. The full specifications for the attributes are as follows:

1. actual_load (Load_Comparison)

double

ex-post load in a market at a given time

2. agent_desc (Market_Agent)

text

describes the market agent (name, location, etc.)

3. agent_id (Market_Agent, Market_Bid)

integer

uniquely identifies the market agent

4. bid_type (Market_Bid, Market_Supply, Market_Demand)

character

specifies whether the bid is for a demand (D) or supply (S)

5. date (Branch_Group, Market_Equilibrium, Load_Comparison, Market_Bid, Market_Supply, Market_Demand)

custom specification: yyyy/mm/dd

6. direction (Branch_Group)

character

the direction of congestion: "N" for north, "S" for south, etc.

7. forecast_load (Load_Comparison)

double

load forecasted for a market at a given time

8. group_id (Branch, Branch_Group)

integer

uniquely identifies the branch group, i.e., the group of transmission branches

9. group_name (Branch)

text

describes the group of transmission lines (name, location, etc.)

10. group_type (Branch_Group)

character

specifies the type of congestion being reported: "S" for same day, "D" for day ahead

11. iso_desc (ISO)

text

describes the ISO (name, location, etc.)

12. iso_id (ISO, Branch, Congestion_Location, Market)

integer

uniquely identifies the ISO

13. location_desc (Congestion_Location)

text

describes the congestion location

14. location_id (Congestion_Location, Market_Equilibrium)

integer

uniquely identifies the location (e.g., zone or node) of congestion

15. market_desc (Market)

text

describes the market (name, location, etc.)

16. market_id (Market, Market_Equilibrium, Load_Comparison, Market_Bid, Market_Demand, Market_Supply)

integer

uniquely identifies the market for electrical energy or reserves

17. market_type (Market)

character

specifies the type of market, e.g., "S" for same day, "D" for day ahead, etc.

18. period (Branch_Group, Market_Equilibrium, Load_Comparison, Market_Bid, Market_Demand, Market_Supply)

double

this is the time of the given day during which an event occurs

19. price (Branch_Group, Market_Equilibrium, Market_Bid, Market_Demand, Market_Supply)

double

the price in dollars per unit for electricity being traded or bid

20. quantity (Branch_Group, Market_Equilibrium, Market_Bid, Market_Demand, Market_Supply)

double

the amount of electricity being transacted or bid (either in MW or MWh)

21. scheduled_load (Load_Comparison)

double

load scheduled for delivery