Procurement of Information Technology (IT) Programming Services to Upgrade/Expand the Energy Consumption Data Management System (ECDMS)

Contact: Gary Occhiuzzo (654-4779) and Andrea Gough (654-4928)

Time Needed: 10 minutes

Action Requested: Commissioner support for a California Multiple Award Schedule (CMAS) agreement with Sun Sierra Software, Inc., for MS SQL and ASP.NET programming services in the amount of $350,000 to implement Stage 1 of a three-stage upgrade/expansion plan for the Quarterly Fuel and Energy Report's (QFER) ECDMS.

Business Meeting Date: April 11, 2012

Background:
Plan and Vendor Selection Process. As part of the FY-2011/12 workplan process, $350,000 was approved to upgrade/expand ECDMS, QFER's database that stores historic energy consumption information. At the Electricity and Natural Gas Meeting on November 9, 2011, Commissioner Peterman supported staff's proposed three-stage automation plan for ECDMS. Following the California Technology Agency (CaTA) approval of an IT Feasibility Study Report (FSR, discussed below), the Energy Commission released a CMAS/Request for Offer for Stage 1 on January 19, 2012. As a result, the evaluation team proposes the contract be awarded to Sun Sierra Software.

QFER Program. The purpose of the Energy Commission's QFER data collection and analysis program (CCR, Title 20, §§1301-1333) is to produce data that can (i) be used to support the Energy Commission's modeling methodologies, (ii) produce insights independent of any modeling methodology, and (iii) be collected and maintained in a searchable database that can be accessed to respond to inquiries from the public, academia, and governmental agencies. Accurate QFER data is a quintessential precursor to accurate Integrated Energy Planning Report (IEPR) electricity and natural gas demand forecasts (PRC § 25301/SB-1389).

Program Challenges. Over time a number of factors have adversely impacted QFER data quality. Examples include:

- Restructuring: Restructuring of California's utility industry circa 1995 had a non-trivial impact. Collecting and scrubbing QFER data submitted by utilities has been complicated by an increase in non-standard data submissions (e.g., introduction of energy service providers), difficulty in obtaining data since some utilities deemed certain data to be "competitively sensitive," and assertions by some utilities that restructuring relieved them of their obligation to report on non-bundled-service customers. At the same time the job of collecting QFER data became more

complicated and time consuming, resources devoted to supporting data collection declined.

- Expanded Use of IEPR Forecasts: Related data collection factors that further strained decreasing QFER resource levels included SB1389 (2002) that demanded an increased range in the analytic capabilities of the Energy Commission's traditional energy demand forecasting role. Additional demands came from the California Air Resources Board's legislative mandate (AB32) to effect significant reductions in California's greenhouse gas emissions, the California Public Utilities Commission's Long Term Procurement Planning process for determining California's long-term power generation needs, and AB2021 (2006) that mandated the Energy Commission's oversight of new energy efficiency requirement obligations of investor-owned utilities and publicly owned electric utilities. These new mandates placed new responsibilities upon the Energy Commission's IEPR energy demand forecast, which in turn called for a dramatic increase in the accuracy of that forecasting process and the data required to support it.

- Dampened Reporting Compliance Enforcement: The current QFER data processing system (i.e., ECDMS) is outdated, providing computer *assistance* rather than computer *automation* in key program areas such as reporting compliance assessment, communication tracking, data loading, data processing, data validation, and data publishing. ECDMS limitations, coupled with constrained QFER staffing levels, have dampened feedback to utilities regarding delinquent reporting practices and have precluded opportunities to educate utilities in how to remedy such shortcomings. Resulting judgment-based data adjustments have contributed to decreased forecasting transparency. Data compliance enforcement is an important component of the QFER data quality problem, a role staff can more effectively perform with support from a substantially expanded/upgraded QFER ECDMS.

<u>ECDMS Upgrade/Expansion Plan</u>. Computer automation of ECDMS would permit staff to redirect focus from QFER administrative mechanics to (i) improve data quality, in part, through timely and more frequent interactions with utilities, and (ii) analyze data trends that could provide energy consumption insights independent of any modeling methodology as statutorily required. Owing to the nature and magnitude of this undertaking, a QFER staff prepared IT FSR guided by Rita Champlion was approved by the CaTA. FSRs are primarily concerned with assessing/reducing the risk that an IT investment could fail to achieve practical (or business) objectives. The FSR revealed three natural project stages whose sequential performance is more likely to achieve FSR-required business objectives than if all three stages were attempted at once under a single contract. Hence the procurement request of this agenda item is for Stage 1 only.

The following table presents the primary objectives and key elements of the Stage 1 $350K CMAS agreement. Also included are those for Stages 2 and 3 for context.

SUMMARY OF IT 3-STAGE ECDMS AUTOMATION PROJECT (Total Estimated Project Cost: $1,275,000, Distributed Over Six FYs)				
Stage	**Primary Objective**	**Key Elements**	**Funding (FY)**	**Contract Length**
#1	Maximize data management system automation of reporting compliance assessment, data loading, and communication and data tracking	• detect incomplete, incorrect, and omitted utility data reporting • correct select inaccuracies on-the-fly using a rule-based system • generate compliance reports for immediate feedback to delinquent utilities to improve data quality, in part, through timely and more frequent interactions with utilities • store, retrieve, and display all reported data as well as communication exchanges between staff and utilities via drag & drop and flexible retrieval user interface	$350,000 (2011/12)	19 months (4.30.12 to 12.31.13)
#2	Target computer automation of QFER data processing, validation, and in-house publishing of results	<u>Develop Data Processing Routines</u> • import data from old QFER forms; map SIC data to NAICS codes • distribute unclassified kWh amounts (when utilities have assigned non-meaningful NAICS codes); notify staff to resolve with utilities • aggregate energy consumption totals by specified geographic areas (e.g., planning area, county, climate zone) • load self-generation data <u>Develop Data Validation Routines</u> • perform trend and other statistical analyses on NAICS and non-NAICS QFER data; notify staff of irregularities • import and auto-compare QFER data to official alternate data sources (e.g., EIA, FERC) • compare data to benchmarks and/or adjust to official alternate data sources • import and auto-compare utility NAICS audit reports to verify Title 20 NAICS accuracy requirements • assist in filling missing historical data cells using rule-based protocols <u>Develop In-House Data Publication Capabilities</u> • develop automated report-generating algorithms to accelerate staff response to data requests from in-house (e.g., IEPR forecast), public, academia, and government agencies	$450,000 (2013/14)	27 months
#3	Web-base appropriate segments of Stages 1 and 2 for further automation gains and resulting data quality improvements	• web-base data loading and reporting compliance assessment for direct utility use via secure internet link • web-base data publishing: automate report upload procedures for posting scrubbed QFER data to the Energy Commission's public website; expand public website data manipulation tools	$475,000 (2016/17)	27 months
All Stages	• Employ an IT California Multiple Award Schedule (CMAS) competitive contract for MS SQL and ASP.NET programming services • Ensure software development fully complies with the Energy Commission's IT standards • Document all software development including user's and programmer's guides for QFER staff and an administrator's guide for the IT Services Branch (ITSB) • Train QFER and ITSB staff in the use and maintenance of all software upgrades/expansions • Provide post-implementation debugging and maintenance support for a specified time interval			

Justification for Action Requested:
IEPR electricity and natural gas forecasts are not possible without QFER data, and *accurate* forecasts are not possible without *accurate* QFER data. Toward that end, implementation of Stage 1 of the ECDMS automation plan would offer labor savings that can be invested in several key areas to improve QFER data quality.

QFER Labor Savings: Targeted labor savings objectives cited in the FSR are presented and discussed below:

- Reporting Compliance Assessment: 25% of QFER staff time (3.8 PY) is now dedicated to assessing reporting compliance. This activity involves manual review of up to eight QFER reporting forms from approximately 60 utilities per quarter to determine if (a) the required forms have been filed, (b) all required fields of the filed forms have been completed correctly, and (c) the reported data are accurate with respect to time and cross-form comparisons. At present, this process occupies approximately 16 weeks per quarter. The goal is a 75% reduction in completion time, or 4 weeks to complete reporting compliance assessment.

- Data Loading: 20% of staff time is dedicated to key entry of the data. Most of the data is electronically delivered (e.g., XLS or TXT files) as e-mailed attachments. The information must be key-entered via a fixed-format data entry screen. Data volume exceeds several million field entries per year, with per quarter load-time exceeding 20 weeks. The goal is an 80% reduction in completion time, or 4 weeks to complete data loading.

- Data and Communication Tracking: 15% of staff time is applied to distributing the loaded data to internal tables that compose the QFER database. Once there, the tasks of data scrubbing, compilation, and validity checks can commence. All communications (e-mail, telephone, post-mail) between staff and utility contacts are individually archived in a manual folder structure. These activities currently occupy in excess of 20 weeks. The goal here is a reduction in completion time of 80%, or 4 weeks to complete data and communication tracking.

- Note: The remaining available staff time (40%) is presently allocated to scrubbing, compiling, and using the data (i.e., responding to data requests), and will be the focus of Stages 2 and 3.

QFER Data Quality Investment Benefits From Labor Savings:

- Reporting Timeliness: Increased staff interaction with reporting utilities can improve timely delivery of data. A 3rd quarter 2010 manual audit revealed that fully one-third of approximately 60 utilities were delinquent in meeting QFER reporting deadlines. Untimely utility deliveries have hampered staff from discovering and acquiring the latest consumption data revisions for use in staff's IEPR forecasts. Utility accounting units prepare QFER data reports that can subsequently be revised by utility forecasting units. Such revised data often arrive courtesy of Title 20 Forms and

Instructions only months prior to IEPR forecast workshops and too late for QFER staff to incorporate into ECDMS for use by Energy Commission forecasters. ECDMS automation of administrative mechanics would provide potential labor savings needed for QFER staff to work with utilities to resolve such issues.

- Data Accuracy Improvements: The North American Industrial Classification System (NAICS) permits energy consumption to be codified by the primary economic activity of the industry or business type at a given location. Without NAICS the Energy Commission's end-use electricity forecasts would not be possible. NAICS-based end-use models assist in revealing patterns in energy demand, identifying energy intensive industries, and estimating how energy demand shortages can affect California.
  (1) Unclassified NAICS: Unfortunately, other than satisfying QFER reporting requirements, most utilities have no intrinsic operational needs to classify energy sales by NAICS and are therefore predisposed to report energy use as "unclassified." Regardless of how sophisticated an algorithmic approach employed by QFER staff may be to approximate an end-use distribution of these unclassified "clumps," IEPR forecasting accuracy, reliability, and transparency is unavoidably compromised. Anticipated QFER labor savings would provide time to work with utility marketing departments to (a) showcase the inherent marketing benefits of a NAICS-classified customer data base for conducting surveys and performing other market-based research analysis, and (b) develop codification standards to correctly classify unclassified customer clumps.
  (2) Misclassified NAICS: Of those customer accounts for which utilities assign NAICS codes, non-trivial amounts of misclassification are indicated. Analysis of the Energy Commission's 2004 Commercial End Use Survey (CEUS) revealed errant customer classification assignments for 44% of small offices, 25% of food/liquor stores, 57% of unrefrigerated warehouses, and 26% of colleges. NAICS accuracy, crucial for supporting end-use modeling detail required by the ever expanding uses of the IEPR energy demand forecast, is further exacerbated by the magnitude of monthly turn-over in utility customer accounts -- 10,000 for PG&E alone, and approximately 25,000 statewide. Anticipated QFER labor savings would provide time to work with utilities to properly classify customer accounts to reduce NAICS misclassifications.
  (3) NAICS Accuracy Audits: Title 20, Ch. 3, Article 1, §1303, Subsection (m) indicates that unclassified consumption cannot exceed one-percent of non-residential customer accounts at the sector level. However, 2009 annual totals indicate that level was exceeded by a factor of 3 to 1. The Executive Director has authority to notify a utility that the accuracy of its NAICS data is in question. From date of notification, a utility has 3 months to deliver an audit to the Executive Director, 6 months to deliver a plan of correction, and 12 months to implement the corrections. Anticipated QFER labor savings would provide time for staff to prepare NAICS audit requests and to work with utilities to avoid inaccurate assignment of NAICS codes.
  (4) Legal Enforcement:
  QFER staff efforts unrewarded by utility improvements in NAICS assignments as

discussed above then would provide solid basis for imposing Title 20 enforcement penalties. Per Caryn Holmes, penalties contained in the Warren-Alquist Act may be applied using Title 20, §1230 (Complaints and Investigations) via a staff-initiated complaint to Legal alleging a violation of the QFER reporting regulations. The Energy Commission may then decide to serve a complaint to which the utility must respond in 21 days, typically followed by a hearing. Title 20 pecuniary penalties range as high as $2,000 per data category violation per day. Ultimately, lack of compliance with a Commission enforcement decision is grounds for requesting a court order. Utilities' legal staff are typically motivated to resolve issues well before a judicial proceeding is implemented. Sometimes, the filing of a complaint, or the noticing of a Committee or Commission hearing on data collection enforcement is sufficient to resolve the non-compliance. Once staff's efforts to resolve a compliance issue have been adequately documented, this approach can represent a very powerful compliance enforcement technique. Anticipated QFER labor savings would provide time to pursue this more extreme course of action as required.

- Data Insights Independent of Any Modeling Methodology (CCR, Title 20, §§1301-1333): Computer automation of ECDMS would permit staff to redirect focus from QFER administrative mechanics to analyzing data trends that could provide energy consumption insights independent of any modeling methodology as statutorily required. As opposed to causality-based modeling analysis performed by Demand Analysis Office's forecasting unit, the QFER program is statutorily charged with producing insights independent of any modeling methodology. Statistically or non-statistically based data trending analysis could alert QFER staff to data issues requiring utility resolution. Moreover, such analysis would highlight consumption pattern changes that among other things could contribute to more accurate IEPR electricity and natural gas forecasts.

Pros and Cons:

- Stage 1 ECDMS Upgrade/Expansion Plan
  - -- Pros:
    - (1) Anticipated labor savings would permit staff to redirect focus from QFER administrative mechanics to (i) improving data quality, in part, through timely and more frequent interactions with utilities, and (ii) analyzing data trends that could provide energy consumption insights independent of any modeling methodology as statutorily required.
    - (2) Automated compliance assessment would permit faster staff feedback to utilities upon receipt of incorrect filings (or no filings) signaling the importance of timely and accurate QFER data reporting to the Energy Commission.
    - (3) This proposal provides an opportunity for QFER staff to work with utility staff to correct reporting problems, especially unclassified and misclassified NAICS. Utilities are the best source for accurate consumption data, rather than approximation fixes currently employed to repair missing or incorrectly reported data. Such fixes can lead to staff vs. utility differences in underlying historical consumption data and to forecast modeling errors.

(4) Anticipated labor savings would support QFER data analysis independent of any modeling methodologies. In turn this would highlight consumption pattern changes that among other things could contribute to more accurate IEPR electricity and natural gas forecasts.

-- Cons: Utility IT and human resource efforts to fully comply with Title 20 QFER data reporting requirements could result in moderate utility cost increases.

What Happens Next: Commissioner support would advance staff's proposed CMAS agreement with Sun Sierra Software for $350,000 to the Energy Commission Business Meeting of April 11, 2012. Anticipated contract start-up is April 30, 2012, software implementation is slated for June 2013, and documentation, training, and final fixes would complete by late December 2013.