EEAC Database SubCommittee Minutes:
Tuesday October 29th – 2:30pm – 4:30pm
Mass DEP conference rooms 2A&B

Attendees:
Leo Steidel, Mike Myser – Energy Platforms
Christina Halfpenny, Ian Finlayson, Lawrence Masland – MA DOER
Matt Saunders – AG
Sharon Webber – MA DEP
Bob Rio – AIM
Christina Dietrich – ENE
John Howat – LEAN
Brad Swing – City of Boston
Eric Winkler – ISO-NE
Jeff Schlegel – Optimal Consulting
Shaela Collins – Rich May
Lisa Shea – NU
Lyn Westerlind – NGrid
Monica Cohen, Columbia Gas
Whitney Brougher - NGrid

The meeting opened with an introduction from Ian Finlayson, covering the proposed plan for this and the next three meetings of the Database subcommittee and how they coincide with the separate and more technical Database working group.
The meeting goal was to review what was covered in the first database working group meeting early that day, focusing on understanding the technical possibilities, and the current way that data is managed at a high level. Detailed discussion of some of the database design decision topics were collected and listed in order to ensure that they will be addressed more fully in the next (November) meeting.

Leo Steidel then led an accelerated 'Data Bootcamp' exercise to develop a common vocabulary for discussion of the Database design phase. Specifically discussing how data can be viewed in three primary categories of Source, Derived, & Copied data, and to map out how source data is transformed within the efficiency programs to get to the broad set of outputs required and requested by stakeholders.
Some of the key points covered were:
The value of automation of source data transformations where possible to reduce error points – with an example of how McDonalds manages sales in realtime.
The challenges of relying on copied data given that source data inputs change over time.
The benefits of versioning / version control in order to address source data changes over

Following this overview, Leo led a discussion of the high level data flow currently used by PA programs. This illustrated that copied rather than source data is currently used at several points along the path to final reporting, and it particular the Technical Resource Manual (TRM) is not automated.

Brad Swing asked the group what primary problem the statewide database should address. There was a group discussion to define the problem, focusing on energy efficiency savings, benefits and costs.

Comments from the PAs included:

National Grid – has a largely automated tracking system, which is also their work management system. It has a significant degree of versioning and automation because they viewed this is necessary as they operate different but related programs in different states. With this approach the customer incentive can be calculated directly from the tracking system
NSTAR – Their tracking system has savings calculations but not benefits.

Columbia Gas and Leo noted that there is a delay in any official reporting because the financial system for a PA has to be closed before the EE reporting can be run and savings can be 'counted' in the appropriate quarter.

It was acknowledged that a number of the smaller PAs use sophisticated excel tools as their tracking system, and that this approach is appropriate particularly where only a small number of people working on the spreadsheet. Leo suggested that a statewide database could be designed to provide tools and dashboard services to the PAs to enhance their ability to track program progress.

Discussion returned to the possibility of automating the TRM calculations that are now run in multiple PA spreadsheets. The TRM expert is recognized as being Kim Crossman. Within the TRM – simple unitary measures require relatively simple calculations based on deemed savings values. At the next level in the TRM – some measure calculations need additional data inputs such as climate values etc for the screening of some measures. Concern was raised about the limited ability to automate custom projects which often depend on a custom site specific engineering study.
If you can build more complex calculations, you can get more custom measures moving into measured savings.

It was noted that there are two versions of the TRM maintained a Plan version – based on the DPU plan and a Report version – based on program implementation.

Discussion moved to Performance verification and privacy issues.
For Optimal or others to verify the performance of custom measures, PAs typically provide a project file with study results but with redacted data to protect customer privacy.

Leo presented some database design options to address concerns around privacy:
It was a common expectation that all data will reside in the cloud, rather than on any entities specific server. Rather than one commonly shared database, systems can be built with clearly separate spaces with controlled access, for example a PA specific space, a common PA space for merging statewide data, then linked to a 'public' space with

varying degrees of stakeholder access to finalized or published data. Published data can be further obfuscated as appropriate to conceal address or other customer specific non-energy data

Columbia Gas raised concerns around access to customer consumption data as typically being a higher bar than savings and benefits data. It was agreed that the degree of aggregation of consumption data should be addressed at the next meeting.
John Howat suggested that in some cases access to measure level savings could be used conceivably to market inappropriate things to customers.
Jeff Schlegel requested that consumption data for individual C&I customer's be released but masked, only for participants and once a year so that the Optimal team can analyze the savings from custom programs – which would be required for appropriately evaluating performance incentives.

The meeting was adjourned at 4:20pm