Pilot Water Use Data Receiving and Storage System Project

April 5, 2018 - Status and Update to DMEC

Greg Gearheart, Deputy Director
Office of Information Management and Analysis
State Water Resources Control Board

“Water use” reporting

Terminology is important

- “Water use data” means data that is used to measure a quantity of water present, diverted, used, or returned (flow, volume, stage, level)
- “Discrete data” means individual data within a common category (e.g. “Dates” within the “Date of sample” column)
- “Time series data” means time-value pairs, often associated with other columns through primary key
- “Batch data” means aggregated or grouped data, generally by time period (e.g. Annual Report, monthly summary)
- “Bulk data” means multiple batch data uploads at one time.
- “Streaming data” means continually received, often with many channels (fixed unit) per stream

Assumptions

1. Collecting water use data electronically (discrete, batch, etc.) for one project or program is not that different than what is needed for other programs and projects. Current WB programs and “databases” in this market include eWRIMS, GEARS, CIWQS, CEDEN and Urban Water Supply (water conservation)
2. Building a team of users of all types (reporting users, regulatory users, technology advocate users, data scientists, etc.) to design and test this prototype will vastly improve the process of building in core user-centered design features to future projects related to this.
3. Keeping the use(s) of data in mind all the way through design of data system will better inform all parties.

Goals

This project will:

1. Establish proof of concept that collecting water use data in place can make that data available across program silos and satisfy user and business needs (authenticate who is the who, make data accessible with proper governance, etc.).
2. Test and improve data collection protocols like minimum data standards, valid values, data structure requirements and ability to use data in management questions that range in scope and scale typical to water use in CA.
3. Build better business requirements for the various current and near future information technology projects working their ways through the stage-gate process.

This project

This project WILL:

1. Be a proof of concept and pilot for standing up a secure, authenticated data system to receive, store and make accessible water use data across multiple scenarios and geographic scopes and scales in CA.

This project WILL NOT:

1. Provide a compliant path to report water data this year and probably not subsequent years, either.
2. Be the eventual system we use to water use data for purposes of complying with Water Board requirements.

Pieces to the puzzle

Data standards

- WaterML (depending on data type), condensed down to minimum fields (and optional fields)

Catcher’s mitt - the “system”

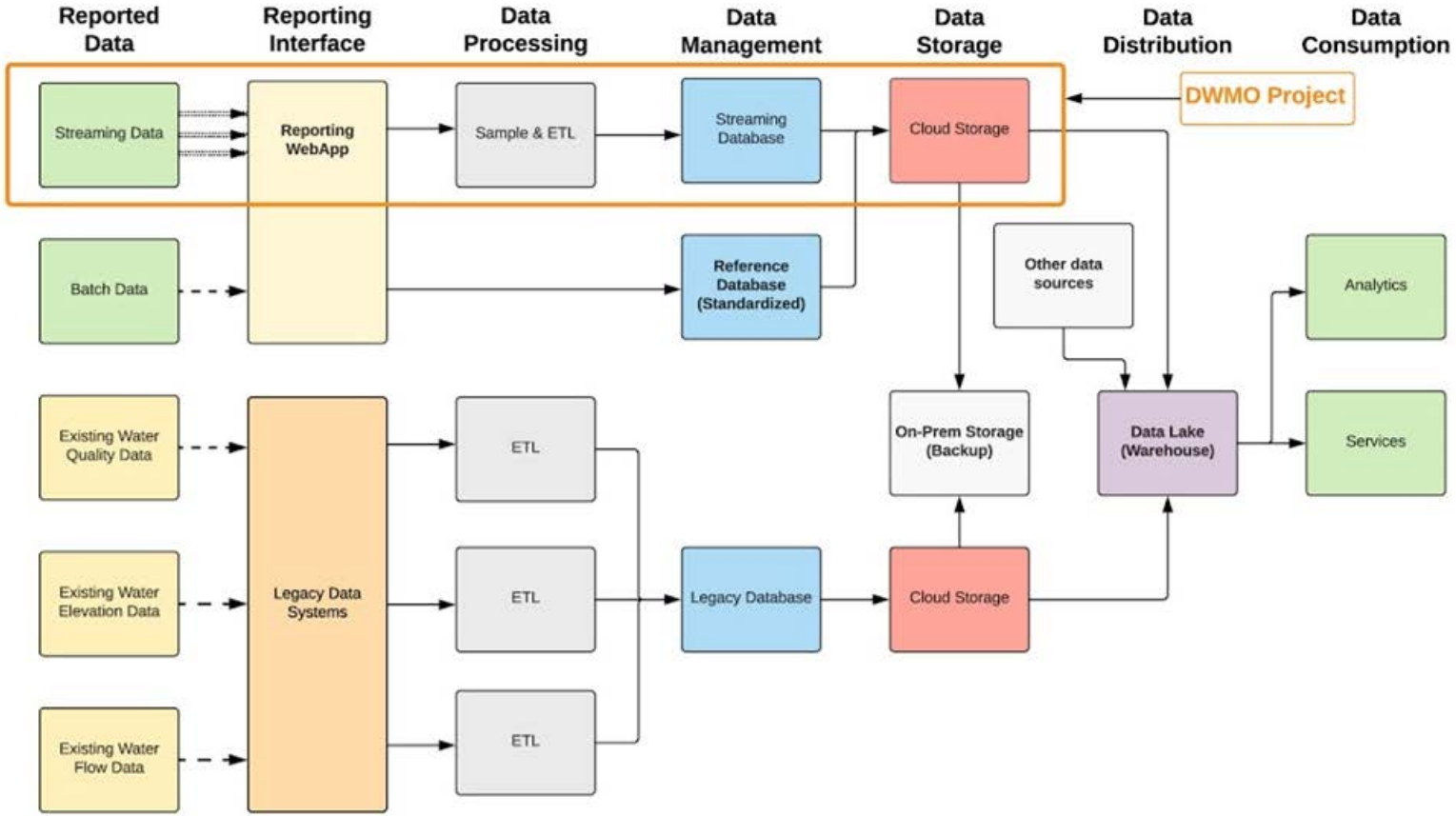
- Stand up a pilot in Amazon Web Services (AWS) space to test out the data ecosystem (data: web reporting, ETL, storage, access, use)

Pilot project stakeholders - the team

- Key water use reporting entities, geospatially diverse in CA, some tech reps, willing to help with coding, funding (or both), and testing use of data requirements

COMPREHENSIVE DATA ARCHITECTURE OVERVIEW - SWRCB Complaince Reporting

DRAFT



User roles

Data Reporter:

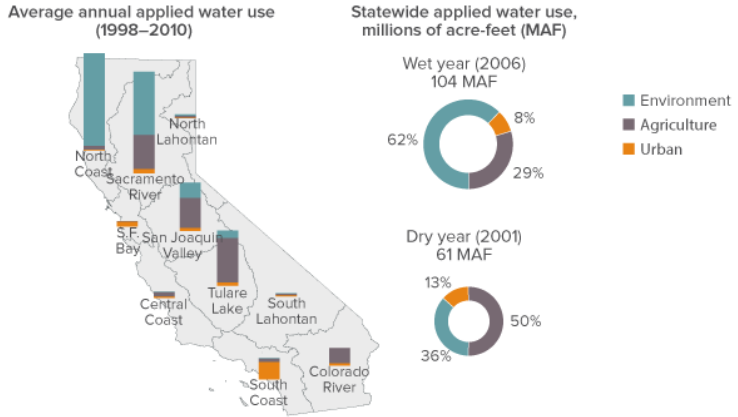
1. Ease of reporting
2. Reliable record keeping
3. Privacy

Data Consumer:

1. Ease of access (to data)
2. Quality data
3. System reliability (service uptime)

Data Manager:

1. Ease of DB query & report creation
2. Extensibility for new reporting needs
3. Adaptable QA/QC protocols
4. Metadata



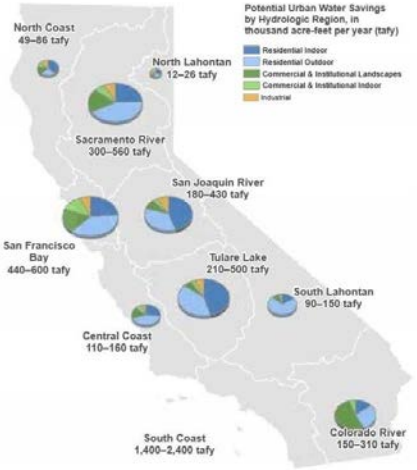
Functional requirements being tested

“Water Use” Reporter App Functional Categories:

1. Water flow/volume reporting (W. Rights, SGMA, DWMO, SWAMP, Citizen)
2. Streaming account and protocol declaration (W. Rights, SGMA, SWAMP, DWMO)
3. Water surface elevation reporting (W. Rights, SWAMP, SGMA)
4. Special projects and enforcement for public trust work (W. Rights, SWAMP)
5. Water quality reporting (NPDES)
6. Urban water supply use (Water Conservation)

Public App Functional Categories:

1. Geospatial overview & query
2. Summary statistics (performance) & report generation
3. Web services to report (we collect) and harvest (we report) machine readable data



Critical path and next steps

1. Form a team of partners to design (mostly done), build (not hard) and test (will be the real ask) a prototype of an electronic receiving system for water use data
2. Build it and test it.
3. Report back to the Delta consortium and other key regulatory stakeholders (eWRIMS, CEDEN, etc.) on business needs and data standards and structure.