Innovating Upon Infrastructure—Not Just Semantics

Kevin Monte de Ramos

INTRODUCTION

This is the first in a series of columns that will highlight and celebrate technological innovations addressing stakeholders within the energy and climate sectors. Not only do we see innovative end-use technologies being adopted by consumers and emergent grid components being deployed down utility distribution lines, but we are also witnessing energy companies and governmental agencies modernizing their internal information technologies to spawn new and optimize existing service offerings.

This first column begins with an evolution that is impacting not just the energy sector, but the whole ecosystem of companies supporting climate initiatives. We see governmental agencies and industry participants exploring the next generation of data warehousing initiatives. There exists a push to the proverbial "cloud," and many enterprises are amidst a digital transformation and actively exploring new data-aggregation platforms.

We begin our innovation series with a focus on an emergent data and analytics platform as they relate to the energy industry. Gartner, a global research and advisory firm providing information, advice, and tools for industry leaders in IT (**Figure 1**), has reported an evolution from data cubes to a centralized enterprise data warehouse with the introduction of data lakes, and landing on a platform that incorporates and complements all prior platform investments; this is known as logical data warehousing.

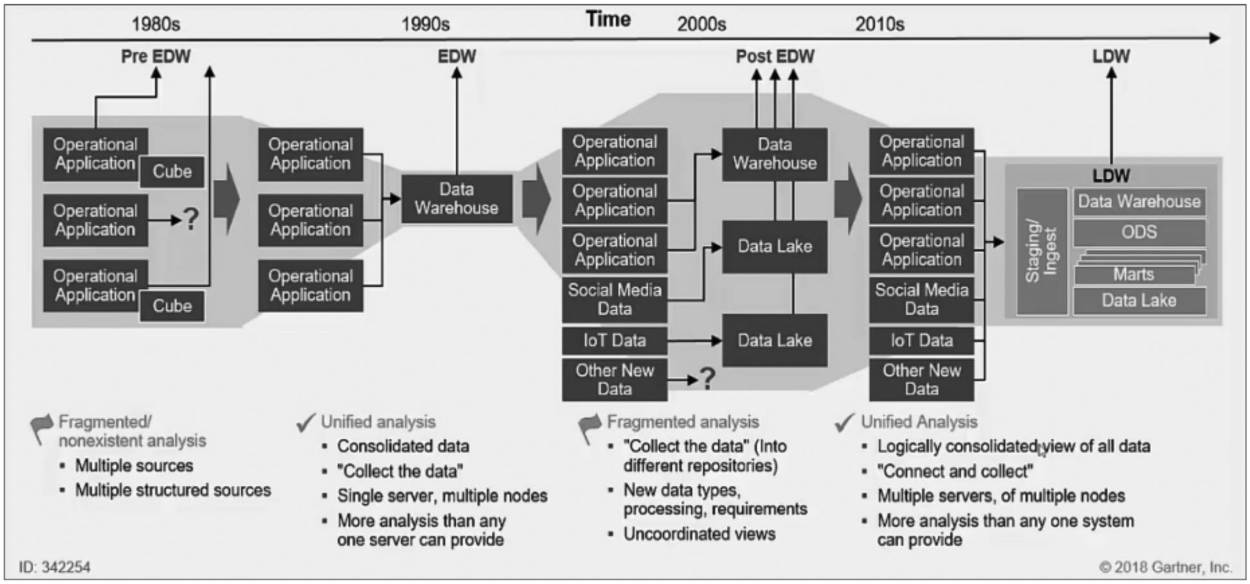
Prior to the 1990s, data were encapsulated within the operational systems performing each function. Where data were required to be shared between systems, extracts were created and loaded into custom databases for targeted activities. Then as those activities expanded, more efficient methods were required. This resulted in the birth of enterprise data warehousing in the 1990s, which automated the integration of data from disparate operational applications into centralized data stores. Internally, these data warehouses enabled the development of custom-built applications. A decade later, social media streams and the internet of things led to the addition of data lakes to the enterprise infrastructure landscape, whereby unstructured and structured data stores can coexist with pre-existing investments to empower the modern corporation.

Today, the logical data warehouse (LDW) introduces a semantic layer to the evolving

Kevin Monte de Ramos is a 30-year veteran of the utility industry. With a focus on service innovation, Monte de Ramos has leveraged data and analytics strategy to identify market opportunities and to qualify emergent revenue and business models. He currently serves as vice president of service line development for Climate Action Services Inc., where he advises on next-generation operations as a data and analytics strategist. He previously worked at Accenture as a management consulting senior principal advising on utility industry innovation, and also ran his own management consultancy serving the utility industry for more than 18 years. He can be reached at KMDR@ClimateActionServices.com.

 DOI 10.1002/gas.22205 /

Figure 1. Gartner's Evolution of Data and Analytics Platforms



Ronthal, A., & Edjlali, R. (2018, March 9). *The practical logical data warehouse: A strategic plan for a modern data management solution for analytics.* Gartner Research. G00342308. Retrieved from https://www.gartner.com/en/documents/3867565/the-practical-logical-data-warehouse-a-strategic-plan-fo0.

enterprise IT reference architecture. By simply securely connecting data stores across organizations, greater transparency is assured while simultaneously securing data amidst stringent privacy laws. Not only do these logical data warehousing platforms enable service innovations, but LDWs justify their implementations through realized cost efficiencies and streamlined application development models.

By starting with these overarching global information technology trends, we will lay the foundation upon which we can drill down into more localized initiatives and targeted service innovations. At the heart of our evolving Smart Grid is a robust information technology infrastructure that allows datasets to be shared between private industry actors and governmental agencies.

COMMON ASPIRATIONS UNITE

Generally speaking, private- and public-sector entities are united by the need to organize their operational activities toward the production of goods and/or delivery of services. Many industries are extending capabilities through self-service kiosks, automation, and the deployment of connected devices. The utility grid will see all these elements implemented to monitor and speed the reaction of the electricity grid to consumer demand (most notably, electric vehicle charging systems). Vehicles are a battery storage device to both input energy to the grid when demand is high and draw energy from the grid when electricity demand troughs.

Critical to maximizing the achievements of these types of mandates is the cyclical integration of insights into the refinement of day-to-day activities and the rationale by which desired outcomes are to be realized. Achieving operational excellence requires novel, interesting, and understandable correlations to emerge from available data and an investigation into causality. Data science has emerged as an industry, which itself is held up by three pillars: (1) data aggregation, (2) applied sciences, and (3) domain expertise. We focus first on this first pillar, as domain expertise is required to apply the scientific method, and, of course, data are requisite before any analytics are facilitated.

In this column, we will explore the deployment of LDWs within a governmental agency and the use case for analytics within a utility distribution

Figure 2. Magic Quadrant for Data Integration Tools

Source: Gartner (August 2020)

company. This composite example is used to explore how emergent technologies are being used to overcome organizational constraints and meet evolving requirements.

A CALL FOR INNOVATION

Large enterprises and government agencies face similar technological challenges—identifying the best methods and practices to share information between business units to accomplish broad overarching objectives. The following pain points are often highlighted in this discussion:

1. The length of time to integrate disparate data within traditional relational enterprise data warehouses does not get implemented at the pace by which the market demands.
2. Communicating what data was available within and from other business units is often an exercise in futility between a small number of superusers.
3. Individuals seek to leverage their own preferred business reporting tools and visualization software without regard for organizational preferences to get the job done.
4. Information technology officers seek to unify support for data access, data analysis, and reporting under a single infrastructure platform.

To resolve these challenges, the city of Calgary published a Request for Proposals for a data virtualization solution in 2017. The stated objectives for this procurement were similar to those we hear from other organizations—namely, to achieve the following: greater agility in creating business intelligence solutions, capability to reduce data duplication, ability to deliver analytics and integration solutions more efficiently, and the integration of both structured and unstructured information sources.

There was an overarching call to implement a semantic layer between the data and the array of business intelligence tools used throughout the organization. This approach would allow ongoing enterprise data warehousing initiatives and the migration to a unified data analytics platform to continue uninterrupted. The city of Calgary ultimately selected an LDW solution known commercially as the Denodo Platform, a market-leading data virtualization platform.

To contextualize this selection, we present Gartner's vendor rankings around this space. The reader will note Denodo appears among other well-established enterprise software vendors, such as IBM, Informatica, SAP, and SAS,[1] that offer similar functional opportunities. We also note that other less prevalent firms, like Denodo, make the list—namely, the likes of SnapLogic, Information Builders, Qlik, Precisely, and Adeptia. This points to the continued need for innovation among our informational technology to build platforms that better integrate data, link with a breadth of analytics tools, and empower business unit employees who may lack experience with emergent data analytics platforms.

Data virtualization platforms serve as a proof of concept within a rich data warehousing landscape. The Denodo Platform was found to be highly effective at linking disparate data stores and documenting data elements stored across the whole enterprise were less effective at aggregating high-volume data requiring highly complex data transformations. The virtual nature of the platform puts a premium on CPU cycles and competes for available memory. As such, other tools and platforms are deployed and maintained to accomplish these functions.

Still, Denodo maintains other advantages—namely, the aggregation of disperse transactional data without complicated planning cycles often associated with other data warehousing tools. Denodo also makes available a data catalogue for its users. This transparency alleviates exploratory requests for user credentialization into other data warehousing platforms. This embedded data catalogue is dynamically maintained, allowing users to explore additions and definitions in real time as data were connected to the data virtualization platform.

Once the value of the data elements is exposed through the data catalogue, business units could formalize their requests and be granted access to credentialed data views. This is important, as Denodo creates a logical data warehouse that spans the entire organization. This feature of the Denodo Platform allows a diverse set of business intelligence software to connect seamlessly to shared data views by which a wide range of rich data visualizations can be activated.

Figure 3 shows how data virtualization offers a generic semantic layer that connects all business units to an organization's vast array of the data. The reader should note that other systems and tools remain in place. Experienced implementers understand and appreciate that each platform offers its own strengths and weaknesses, providing a niche for market software within the technical reference architecture illustrated by a representative from Denodo.

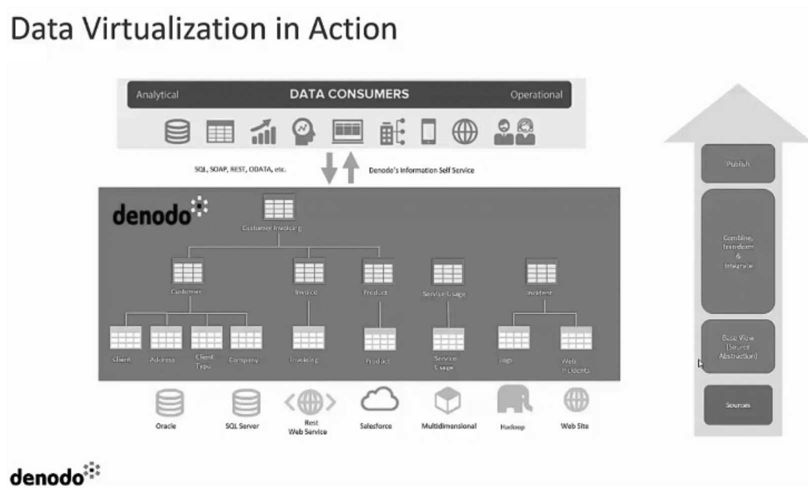
Denodo represents business units generically referred to as data consumers. In darker shade is Denodo's data virtualization platform consisting of the shared data views in the center and the data catalogue highlighted amidst other administrative tools on the right. Below the data virtualization layer are the actual data stores of the organization itself, which include databases, analytics platforms, web services, and websites.

Even before organizations can assess and quantify the value generated for themselves, representatives from the impacted business units frequently sponsor additional implementations and express enthusiasm for the results achieved early in the implementation process. When asked about the lessons learned regarding data virtualization, client references for Denodo shared the following insights.

First, data virtualization provides a single point of access for data auditing and aids in determining what business units will be impacted by application maintenance being performed. Second, data virtualization is like any enterprise software; its adoption must be planned, implemented, and tuned, and resources calibrated, to optimize the solution for organizational requirements. While time and resources are required, data virtualization involved fewer commitments than other options considered to achieve similar results. Third, the implementation of data virtualization requires experience with the software chosen. Consultants experienced with the Denodo Platform were initially hired to implement and coach the

[1] Zaidi, E., Thoo, E., Heudecker, N., Menon, S., & Thanaraj, R. (2020, August 18). *Magic quadrant for data integration tools.* Gartner Research. G00450251. Retrieved from https://www.gartner.com/en/documents/3989223/magic-quadrant-for-data-integration-tools.

Figure 3. Data Virtualization Platform Integration



Denodo. (2020). *Introduction to data virtualization for architects.* DEN70DUA3001VID01.

internal teams through error handling, naming conventions, design patterns, best practices, etc.

In the end, the clients of Denodo procure and implement data virtualization solutions that empowered data-consuming business units to create their own data views, author robust visualizations using their preferred business intelligence tools, and compose reports that enable service delivery through a semantic layer that promotes information sharing necessary to originate value for the sponsoring business units.

INNOVATION IS NOT JUST TECHNOLOGY

The technical architecture discussed above represents an innovative advanced analytics platform, which empowers its users with the insights necessary to transform day-to-day operations. Aviroop Dutt-Mazumder, who serves as a data scientist with Consumers Energy, shared with us the analytics used to prioritize utility investments and the modeling leveraged to direct business activities.

Electric utilities must match electricity demand with available generation. To manage electricity demand down feeder lines, data scientists are tasked with generating multiple demand forecast scenarios to target maintenance of key assets and develop propensity models to enroll customers into demand management initiatives.

Maintenance of key assets requires utilities to deploy capital to renew and upgrade the utility distribution networks. The capital available for the renewal of the electricity grid is necessarily finite and requires discipline to direct its allocation. This is where "nonwire options" are necessary to mitigate electricity demand that would exceed the substation's ability to deliver reliable electricity service.

Consumers Energy explores these nonwire alternatives through a series of statistical models. The goal is to identify opportunities to shift demand and minimize peaks along the feeder lines. If this can be achieved, then capital investments can be deferred and deployed elsewhere. The methodology to accomplish a shift of electricity demand is through the introduction of demand-side management initiatives and incentivizing consumer investments in renewable generation.

The activity starts by prioritizing thousands of asset clusters into just a few handfuls of targeted clusters. Dutt-Mazumder and his team begin to assess substation efficiencies, iterating through line-loss calculations and modeling the introduction of the battery storage, solar gardens, and renewable generation options. The result identifies the substations requiring attention and the types of interventions to be deployed within the targeted communities.

 DOI 10.1002/gas /

Utilities can directly invest in grid assets, install large-scale electricity storage, and purchase renewable generation to extend substation capacity to support growing demand for electricity service. And utilities can and do develop market initiatives that slow electricity demand within targeted communities to ensure electricity remains reliable and affordable. These initiatives take the form of energy efficiency, demand response, and distributed generation programs targeting all classes of customers. The challenge then turns to how best to encourage participation in these programs and compliance with the policies being promoted.

Consumers Energy develops propensity models to support the marketing of these types of initiatives. Propensity models identify key characteristics of past participants. These characteristics are then mobilized to target advertising through both broadcast and print media, as well as to digitally engage customers online. Finally, the team will evaluate click-throughs and customer enrollments for additional insight. Where these propensity models accurately predict program enrollment and participant success within those programs, then the general ratepayer benefits from these demand-side management initiatives and the efficient allocation of capital to grid maintenance.

CONCLUSION

Known but often unspoken is that energy and climate policy require an ecosystem of actors to work collaboratively toward a common end. The objective is to effectively manage finite energy resources, while enabling a quality of life that provides individuals with security and opportunities to fulfill individual ambitions. If we are to achieve these outcomes, many industry stakeholders must act synergistically. This requires an unprecedented level of information sharing over a timeframe that traditional technology transformational roadmaps do not allow.

The examples presented earlier highlight the fact that innovation in infrastructure can expose data from one part of the organization to another within the same organization. This innovation in information technology is necessary for organizations seeking to share operational data between organizations, with external industry stakeholders, and openly among the general populace. What is needed is a coordinated data and analytics strategy that evolves our information technology to meet the demands of the ecosystem supporting the energy and climate industry.

Known but often unspoken is that energy and climate policy require an ecosystem of actors to work collaboratively toward a common end.

New methods and analytical practices place the largest demands on our information infrastructure. Data scientists are frequently the authors of these techniques as they work cooperatively with corporate leaders to define necessary data requirements that help prioritize investments in information technologies. It is these superusers who drive the need for infrastructure innovations, which trickle down through the organization as proofs of concept and pilot projects for service innovations.

We need only explore the number of service innovations that rely upon technology innovations that form a new delivery channel: Airbnb, Lyft, Uber, Amazon, McDonald's self-service kiosks, and advanced charging stations are just a few examples of service organizations built upon a technology innovation. It is more than semantics that business must innovate to survive. As we heard from client referrals, the deployment of Denodo's Data Virtualization Platform serves as a semantics layer to integrate data between business units. And we saw how superusers at Consumers Energy leverage the aggregation of data to empower and direct day-to-day operational decisions.

It is this synergy between business units and the service organizations that creates the tangible value from the firm's infrastructure investments. Such is the cycle of continued improvements in constant pursuit of operational excellence. Where energy industry participants support government commitments to climate policy, underlying any impactful outcome will be innovations upon the existing information technology infrastructure to scale service innovations demanded by the consumer.