For better or worse, BI professionals live in a federated world. Despite a BI team’s best efforts, critical data and analytical logic is spread across multiple databases and applications in various parts of an organization. A federated architecture knits these disparate environments together virtually rather than physically. It is neither centralized nor decentralized, but a hybrid of the two, maximizing the benefits of both options, while minimizing their downsides. Moreover, a federated architecture is fluid, changing shape as an organization reinvents itself to better respond to new market realities. In this BeyeNET-WORK eBook, data warehousing expert Wayne Eckerson describes how to create a federated BI architecture that aligns with organizational imperatives.

READERS WILL:
- Learn the various options for architecting BI environments—centralized, decentralized, and federated—and their strengths and weaknesses using case studies.
- Discover the impact of organizational models on BI architectures and how the two evolve in lockstep over time.
- Learn how to balance corporate and departmental interests when creating a BI architecture.
- Understand the role and value of data virtualization tools in a federated architecture.
The Federated Option

“Business needs trump architectural purity every time.”
—DOUGLAS HACKNEY

For years, business intelligence (BI) experts have diagrammed data warehousing architectures in which a single set of enterprise data is extracted and transformed from various operational systems and loaded into a staging area or data warehouse that feeds downstream data marts. This classic architecture is designed to deliver a single version of high-quality enterprise data while tailoring information views to individual departments and workgroups. It also ensures an efficient and scalable method for managing the flow of operational data to support reporting and analysis applications.

That’s the theory, anyway. Too bad reality doesn’t work that way.

- Federated Reality. Most organizations have a hodgepodge of insight systems and non-aligned reporting and analysis applications that make achieving the veritable single version of truth a quest that only Don Quixote could appreciate. Let’s face it, our BI architectures are a mess. But this is largely because our organizations are in a constant state of flux. If an ambitious, tireless, and lucky BI team achieved architectural nirvana, their edifice wouldn’t last more than a few months. Mergers, acquisitions, reorganizations, new regulations, new competitors, and new executives with new visions and strategies all conspire to undermine a BI manager’s best laid architectural blueprints.

Thus, the BI reality in which most of us live is federated. We have lots of overlapping and competing sources of information and insight, most of which have valid reasons for existence. So rather than trying to roll a boulder up a hill, it’s best to accept a federated reality and learn how to manage it. And in the process, you’ll undoubtedly discover that federation—in most
incarnations—provides the best way to balance the need for enterprise standards and local control and deliver both efficient and effective BI solutions.

**UNDERSTANDING THE OPTIONS—EVOLUTION OF INTUIT’S BI ARCHITECTURE**

Federation takes the middle ground between centralized and decentralized BI approaches. Ideally, it provides the best of both worlds, while minimizing the downsides. TABLE 1 describes the pros and cons of centralized and decentralized approaches to BI. Not surprisingly, the strengths of one are the weaknesses of the other, and vice versa.

Founded in 1983, Intuit is a multi-billion dollar software maker of popular financial applications, including Quicken, TurboTax, and QuickBooks. Intuit’s BI evolution helps to illuminate the major characteristics of the two approaches and explains how they ended up with a federated BI architecture. (See Figure 1, page 4.)

**Decentralized Approach.** In the late 1990s, after a series of acquisitions, Intuit had a decentralized BI environment. Essentially, every division—most of which had been independent companies—handled its own BI requirements without corporate involvement. Divisions with BI experts built their own applications to meet local needs. Deployments were generally quick, 

<table>
<thead>
<tr>
<th>TABLE 1: Centralized versus Decentralized BI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>STRENGTHS</strong></td>
</tr>
<tr>
<td>Centralized BI</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Decentralized BI</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
The Federated Option

Centralized Approach. To obtain promised financial synergies from its acquisitions and create badly needed enterprise views of information, Intuit in 2000 implemented a classic, top-down, centralized data warehousing architecture. Here, a newly formed corporate BI team built and managed an enterprise data warehouse (EDW) to house all corporate and divisional data. Divisions rewrote their reports to run against the unified metrics and dimensions in the data warehouse. The goal was to reduce costs through economies of scale, standardize data definitions, and deliver enterprise views of data.

Although Intuit realized these gains, the corporate BI team quickly became overwhelmed with project requests, much to the chagrin of the divisions, which began to grumble about the backlog. When the divisions began to look elsewhere to meet their information requirements, the corporate BI team recognized it needed to reinvent itself and its BI architecture.

Federated Approach. To forestall a grassroots resurrection, Intuit adopted a federated architecture that blends the best of top-down and bottom-up approaches. To uncork its project backlog, the corporate BI team decided to open up its data warehousing environment to the divisions. It gave each division a dedicated partition in the

FIGURE 1: Intuit’s BI Architectural Evolution
EDW to develop their own data marts and reports. It also taught the divisions how to use the corporate ETL tool and allowed them blend local data with corporate data inside their EDW partitions. For divisions that had limited or no BI expertise, the corporate BI team continued to build custom data marts and reports as before. (See Figure 2.)

**Benefits.** The federated approach alleviated Intuit’s project backlog and gave divisions more control over their BI destiny. They could now develop custom tailored applications and reports without having to wait in line. For its part, the corporate BI team achieved economies of scale by maintaining a single platform to run all BI activity and maintained tight control over enterprise data elements. Today, the corporate BI team no longer tries to be the only source of corporate data. Rather, it focuses on capturing and managing data that is shared between two or more divisions and developing BI applications when asked.

**Challenges.** The main challenge with federation is keeping everyone who is
involved in the BI program, either formally or informally, on the same page. This requires a lot of communication and consensus building, or “social architecture” as Intuit calls it. The corporate BI team needs to relinquish some of its control back to the business units—which is a scary thing for a corporate group to do—while business units need to adhere to basic operating standards when using the enterprise platform. Ultimately, to succeed, federation requires a robust BI Center of Excellence with a well-defined charter, a clear roadmap, strong business-led governance, an active education program, and frequent in-house meetings and conferences. (See Table 2.)

Ultimately, managing a federated environment is like walking a tightrope. You have to balance yourself to keep from swinging too far to the left or right and falling off the rope. To succeed, you have to make continuous adjustments, ensuring that everyone on the extended BI team is moving in the same direction. Without sufficient communication and encouragement, a federated approach courts catastrophe.

### Table 2: Federated BI

<table>
<thead>
<tr>
<th>Federated BI</th>
<th>BENEFITS</th>
<th>CHALLENGES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Economies of scale</td>
<td>Communication</td>
</tr>
<tr>
<td></td>
<td>Quick deployments, flexibility</td>
<td>Building consensus</td>
</tr>
<tr>
<td></td>
<td>Consistency of shared data</td>
<td>Adherence to standards</td>
</tr>
<tr>
<td></td>
<td>Enterprise and local views</td>
<td>Business governance</td>
</tr>
<tr>
<td></td>
<td>Multiple career options</td>
<td>Education/training</td>
</tr>
</tbody>
</table>

Managing a federated environment is like walking a tightrope. You have to balance yourself to keep from swinging too far to the left or right and falling off.
Organizational Models That Impact BI

“Architecture must conform with corporate culture and evolve with the prevailing winds.” —WAYNE ECKERSON

In Chapter 1 of this eBook, we discussed the strengths and weaknesses of different architectural options, including federation. We used Intuit as an example of an organization that evolved through decentralized and centralized architectural models until it settled on federation as a means to marry the best of both worlds.

It’s important to recognize that the real problem that federation addresses is not architectural; it’s organizational. There is natural tension between local and central organizing forces within a company. Every organization needs to give local groups some degree of autonomy to serve customers effectively. Yet it also needs a central, guiding force to achieve economies of scale (efficiency), present a unified face to the customer (effectiveness), and deliver enterprise views of corporate performance. BI, as an internal service, needs to reflect a company’s organizational structure. But knowing where to draw the line architecturally is challenging, especially as a company’s business and organizational strategy evolves.

Organizational Models

At a high level, there are three basic organizational models that define the relationship between corporate and divisional groups: conglomerate, cooperative, and centralized. Although, this eBook treats each as a unique entity, the models represent points along a spectrum of potential organizational structures. The key to each model’s identity and defining characteristics, however, stems from a company’s product strategy. (See Table 3, page 8.)

- **Conglomerate Model.** A conglomerate, for example, has a highly diversified product portfolio that spans...
multiple industries. Here, each division sells radically different products to very different customers. For example, General Electric, is a conglomerate. It sells everything from financial services to household appliances to wind turbines. A conglomerate typically gives business unit executives complete autonomy to run their operations as they see fit using their own money and people. There is little shared data; and each business unit has its own BI team and technology and exchanges little, if any, data with other business units.

**Centralized Model.** The centralized model falls at the other end of the organizational spectrum. Here, an organization sells one type of product to anyone, everywhere. As a result, corporate plays a strong hand in defining product and marketing strategies while business units primarily execute sales in their regions. In a centralized model, IT and BI are shared services that run centrally at corporate. There is one BI team and enterprise data warehouse that supports all business units and users.

Dell, for example, sells personal computers and servers on a global basis. Although its regional business units had significant autonomy to meet customer needs during the company’s rapid expansion, Dell now has retrenched. Without dampening its

---

**TABLE 3. Organizational Models**

<table>
<thead>
<tr>
<th>Business Characteristics</th>
<th>CONGLOMERATE</th>
<th>COOPERATIVE</th>
<th>CENTRALIZED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Lines</td>
<td>Diversified</td>
<td>Related</td>
<td>Homogeneous</td>
</tr>
<tr>
<td>Business Units</td>
<td>Autonomous</td>
<td>Synergistic</td>
<td>Regional</td>
</tr>
<tr>
<td>Shared Services</td>
<td>None</td>
<td>Some</td>
<td>All</td>
</tr>
<tr>
<td>Customers</td>
<td>Minimal overlap</td>
<td>Moderate overlap</td>
<td>Complete overlap</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technical Characteristics</th>
<th>CONGLOMERATE</th>
<th>COOPERATIVE</th>
<th>CENTRALIZED</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT/BI Department</td>
<td>Decentralized</td>
<td>Distributed</td>
<td>Corporate</td>
</tr>
<tr>
<td>Shared Data</td>
<td>None</td>
<td>Some</td>
<td>Most</td>
</tr>
<tr>
<td>DW Architecture</td>
<td>Independent DWs</td>
<td>Federated DWs</td>
<td>Centralized DW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examples</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>General Electric</td>
<td></td>
<td>Intuit</td>
<td>Dell</td>
</tr>
</tbody>
</table>
entrepreneurial zeal, Dell has strengthened its centralized shared services, including BI, to streamline redundancies, reduce overhead costs, improve cross-selling opportunities, and deliver more timely, accurate, and comprehensive views of enterprise performance to top executives.

**Cooperative Model.** The cooperative model—which is the most pervasive of the three—falls somewhere between the conglomerate and centralized models on the organizational spectrum. Here, business units sell similar, but distinct, products to a broad set of overlapping customers. Consequently, there are significant cross-selling opportunities, which require business units to share data and work cooperatively for the greater good of the organization. For example, Intuit, with its various financial software products, can increase revenues by cross-selling its various consumer and business products across divisions. Most financial services firms, which offer banking, insurance, and brokerage products, also employ a cooperative model.

Although most divisions in a cooperative model maintain BI teams, corporate headquarters also fields a sizable corporate BI group. Given the distribution of BI teams and technology across divisions, the cooperative model by default requires a federated approach to BI. The challenge here is figuring out how to divide the BI stack between corporate and divisional units.

There are many places to draw the line. Often, the division of responsibility is not clear-cut and changes over time as a company evolves its business strategy to meet new challenges.

**Harley-Davidson.** For example, Harley-Davidson Motor Co., an iconic, U.S.-based manufacturer of motorcycles, decided to broaden its business strategy after the recent recession to focus more on global markets, e-commerce, and custom-made products. Consequently, the BI team created a new strategy to meet these emerging business requirements that emphasizes the need for greater agility and flexibility.

Currently, Harley-Davidson has well-defined, centralized processes for sourcing, validating and delivering data. Any application or project that requires data must be approved by the Integration Competency Center (ICC), which generally also performs the data
integration work. The ICC’s job is to ensure clean, consistent, integrated data throughout the company.

Currently, the ICC handles all business requests the same way, regardless of the nature or urgency of the request. In the future, to go as fast as the business wants, the ICC will break free from its “one size fits all” methodology and take a more customized approach to business requests. “We need to make sure our processes don’t get in the way of the business,” said Jim Keene, systems manager of global information services.

Dell. While Harley-Davidson is loosening up, Dell is tightening up. After experiencing rapid growth during the 1990s and 2000s through an entrepreneurial business model that gave regional business units considerable autonomy, Dell now wants to achieve both fast growth and economies of scale by strengthening its central services, including BI. In the mid-2000s, Dell had hundreds of departmental systems, thousands of BI developers, and tens of thousands of reports that made it difficult to gain a global view of customers and sales.

As a result, Dell kicked off an Enterprise BI 2.0 initiative chartered to consolidate and integrate independent data marts and reports into a centralized data warehouse and data management framework. It also created a BI Competency Center that oversees centralized BI reporting teams that handle report design and access, metric definitions, and governance processes for different business units. By eliminating redundancies, the EBI program has delivered a 961% return on investment, improved data quality and consistency, and enhanced self-service BI capabilities.

Although Harley-Davidson and Dell represent two ends of the spectrum—a centralized BI team that is rethinking its centralized processes and a decentralized BI environment that is consolidating BI activity and staff—most organizations are somewhere in the middle. Chapter 3 of this eBook will map our three organizational models to a BI stack and show the various ways that a BI team can allocate BI architectural elements between corporate and divisional entities.
The Dividing Line

“If it can’t be my design, tell me where do we draw the line.”
—POETS OF THE FALL

In Chapter 2 of this eBook, we discussed the impact of organizational models on BI development. In particular, we described conglomerate, co-operative, and centralized models and showed how two companies (Harley-Davidson and Dell) migrated across this spectrum and its impact on their BI architectures and organizations.

As any veteran BI professional knows, BI architectures come in all shapes and sizes. There is no one way to build a data warehouse and BI environment. Many organizations try various approaches until they find one that works, and then they evolve that architecture to meet new business demands. Along the way, each BI team needs to allocate responsibility for various parts of a BI architecture between corporate and business unit teams. Finding the right place to draw the proverbial architectural line is challenging.

The BI Stack. Figure 3 (see page 12) shows a typical BI stack with master data flowing into a data warehouse along with source data via an ETL tool. Data architects create business rules that are manifested in a logical model for departmental marts and business objects within a BI tool. BI developers who write code and assemblers who stitch together predefined information objects, create reports and dashboards for business users. Typically, most databases and servers that power operational and analytical systems run in a corporate data center.
Using this high-level architectural model, we can study the impact of the three organizational models described in Chapter 2 of this eBook on BI architectures.

**Conglomerate—Shared Data Center.** In a conglomerate model, business units have almost complete autonomy to design and manage their own operations. Consequently, business units also typically own the entire BI stack, including the data sources, which are operational systems unique to the business unit. Business units populate their own data warehouses and marts using their own ETL tools and business rules. They purchase their own BI tools, hire their own BI developers, and develop their own reports. The only thing that corporate manages is a data center that houses business unit machines and delivers economies of scale in data processing. (See **Figure 4**, page 13.)

**FIGURE 3: Typical BI Stack**

In a typical BI stack, end users view a report created by a BI developer using a BI tool that queries a data mart or data warehouse that contains business rules built into the data model, BI semantic layer, or ETL tools. ETL tools load data from source systems into a data warehouse, which increasingly acts as a hub feeding downstream systems. Master data management systems manage reference data (e.g., customer, product, geography data as well as business hierarchies), which it feeds to the DW.
Cooperative BI – Virtual Enterprise. In a cooperative model where business units sell similar, but distinct, products, business units must work synergistically to optimize sales across an overlapping customer base. Here, there is a range of potential BI architectures based on an organization’s starting point. In the Virtual Enterprise model, an organization starts to move from a conglomerate business model to a more centralized model to develop an integrated view of customers for cross-selling and upselling purposes and maintain a single face to customers who purchase products from multiple business units.

In a Virtual Enterprise, business units still control their own operational systems, data warehouses, data marts, and BI tools and employ their own BI staff. Corporate hasn’t yet delivered enterprise ERP applications but is thinking about it. Its first step towards centralization is to identify and match mutual customers shared by its business units. To do this, the nascent corporate BI team develops a master data
management (MDM) system, which generates a standard record or ID for each customer that business units can use in sales and service applications. (See Figure 5.)

The corporate group also creates a fledgling enterprise data warehouse to deliver an enterprise view of customers, products, and processes common across all business units. This enterprise data warehouse is really a data mart of distributed data warehouses. That is, it sources data from the business unit data warehouses, not directly from operational systems. This can be a persistent data store populated with an ETL tool or virtual views populated on the fly using data virtualization software. (See Chapter 4: “Data Virtualization: The Time Has Come.”) A persistent store is ideal for non-volatile data (i.e., dimensions that don’t change much) or when enterprise views require complex aggregations, transformations, multi-table joins, or large volumes of data. A virtual data store is ideal for delivering enterprise views quickly at low cost and building prototypes and short-lived applications.

Cooperative BI—Shared BI Platform.

The next step along the spectrum is a Shared BI Platform. Here, corporate expands its appetite for data processing. It replaces business unit operational applications with enterprise resource

**FIGURE 5: Cooperative BI—Virtual Enterprise**

*In a cooperative model, the corporate group and business units share responsibility for information and systems. Early in the evolution, a corporate group may only manage a data center and common reference data about customers. For its own benefit, it may create a mart of marts to get an enterprise view of business activity across the business units.*
planning (ERP) applications (e.g., finance, human resources, sales, service, marketing, manufacturing, etc.) to create a more uniform operating environment. It also fleshes out its BI environment, creating a bonafide enterprise data warehouse that pulls data directly from various source systems, including the new ERP applications, instead of departmental data warehouses. This reduces redundant extracts and ensures greater information consistency. (See Figure 6.)

Meanwhile, business units still generate localized data and require custom views of information. While they may still have budget and license to run their own BI environments, they increasingly recognize that they can save time and money by leveraging the corporate BI platform. To meet them halfway, the corporate BI team forms a BI Center of Excellence and teaches the business units how to use the corporate ETL tool to create virtual data marts inside the enterprise data warehouse. The business units can upload local data to these virtual marts, giving them both enterprise and local views of data without having to design, build, and maintain their own data management systems.

In addition, the corporate BI team builds a universal semantic layer of

---

**FIGURE 6: Cooperative BI—Shared Platform**

In a more advanced cooperative model, the corporate group builds an enterprise data warehouse to house all shared data and makes this data management platform, along with ETL and BI tools, available to the business units, which they can use to build their own unique applications, or not, depending on whether the corporate BI platform enables them to work more efficiently and cost-effectively.
shared data objects (i.e., a semantic layer) that it makes available within the corporate BI tool. Although business units may still use their own BI tools, they increasingly recognize the value of building new reports with the corporate BI tools because they provide access to the standard business objects and definitions, which they are required or highly encouraged to use. The business units still hire and manage their own BI developers and assemblers, but they now have dotted line responsibility to the corporate BI team and are part of the BI Center of Excellence. This is the architectural approach used by Intuit (see Chapter 1.)

- Centralized BI—Shared Service.
Some organizations, like Dell, don’t stop with a Shared Platform model; they continue to centralize BI operations to improve information integrity and consistency and squeeze all redundancies and costs out of the BI pipeline. Here, the corporate BI team manages the entire BI stack and creates tailored reports for each business unit based on requirements. For example, Dell’s EBI 2.0 program (see Chapter 2) reduced the number of report developers in half and reassigned them to centralized reporting teams under the direction of a BI Competency Center where they develop custom reports for specific business units. The challenge here, as Harley-Davidson discovered, is to keep the corporate BI bureaucracy from getting too large and lumbering and ensure it remains responsive to business unit requirements. This is a tall task, especially in a fast-moving company whose business model, products, and customers change rapidly.
To this point, we’ve examined how to design an information architecture to support federation, and we’ve largely avoided discussion of tools. However, an emerging class of software, called data virtualization tools, can play a critical role in managing the complexity of federated data environments.

For example, the Virtual Enterprise discussed earlier in this series often uses a data virtualization tool to create an enterprise view of customers, products, and processes common across all business units. This enterprise view is really a data mart of distributed data warehouses pulled together dynamically using a data virtualization tool. This virtual data warehouse is ideal for delivering enterprise views quickly at low cost.

Besides creating virtual data warehouses, organizations use data virtualization tools for niche applications, such as creating a quick and dirty prototype or augmenting the data warehouse with real-time data or accessing data outside the corporate firewall. But today, some forward-thinking data architects are using data virtualization tools to create agile, cost-effective data management environments that simplify data access and speed project delivery. In an ideal world, data architects should design and deploy a data virtualization layer prior to building any data management or delivery artifacts.

What is Data Virtualization? Data virtualization software makes data spread across physically distinct systems appear as a set of tables in a local database. Business users, developers, and applications query this virtualized view; and the software automatically generates an optimized set of queries that fetch data from remote systems, merge the disparate data on the fly,
and deliver the result to users. (See **Figure 7**.) Data virtualization software consumes virtually any type of data, including SQL, MDX, XML, Web services, and flat files and publishes the data as SQL tables or Web services. Essentially, data virtualization software turns data into a service, hiding the complexity of back-end data structures behind a standardized information interface.

With data virtualization, organizations can integrate data without physically consolidating it. In other words, they don’t have to build a data warehouse or data mart to deliver an integrated view of data, which saves considerable time and money. In addition, data virtualization lets administrators swap out or redesign back-end databases and systems without affecting downstream applications. The upshot is that IT project teams can significantly reduce the time they spend sourcing, accessing, and integrating data. In other words, data virtualization...
tion speeds project delivery, increases business agility, reduces costs, and improves customer satisfaction. What’s not to like?

**LONG TIME COMING**

Data virtualization has had a long history. In the early days of data warehousing (1995), it was called virtual data warehousing (VDW) and advocates positioned it as a legitimate alternative to building expensive data warehouses. However, data warehousing purists labeled VDW as “voodoo and witchcraft” and chased it from the scene. During the next 10 years, data virtualization periodically resurfaced, each time with a different moniker, including enterprise information integration or EII and data federation, but the technology never got much traction, while vendor providers came and went.

- **Drawbacks.** The traditional drawbacks of data virtualization are performance, scalability, and complexity. That’s because the engineering required to query two or more databases is complex. As you increase data volumes and query complexity, the challenge increases exponentially. As a result, data virtualization tools historically have been confined to niche applications involving small volumes of clean, consistent data that require little to no transformation or complex joins.

  The other major drawback is political. Source systems owners don’t want BI tools submitting ad hoc queries against their operational databases. And these administrators have the clout to lock out applications they think will bog down system performance. This is still a major reason that organizations build data warehouses and marts.

**But does this mean data virtualization is ready to take center stage in your data management architecture? I think yes.**

**THE MISSING LAYER**

Today, however, data virtualization is making a resurgence. Thanks to the popularity of data center virtualization, many organizations are exploring the possibility of virtualizing their data as well. More importantly, advances in network speeds, CPU performance, and available memory have significantly increased the performance and scalability of data virtualization tools, expanding the range of applications they can support. Moreover, data virtualization vendors continue to enhance their query optimizers to handle more complex queries and larger data volumes.

But does this mean data virtualiza-
Federated architectures for BI professionals

THE FEDERATED OPTION

ORGANIZATIONAL MODELS THAT IMPACT BI

THE DIVIDING LINE

DATA VIRTUALIZATION: THE TIME HAS COME

SUMMARY

The dividing line
data virtualization: The Time has Come

Data Federation is ready to take center stage in your data management architecture? I think yes. Recently, I listened to data architects from Qualcomm, BP, Comcast, and Bank of America discuss their use of data virtualization tools at “Data Virtualization Day,” a one-day event hosted by Composite Software, a leading data virtualization vendor. After hearing their stories, I am convinced that data virtualization is the missing layer in our data architectures.

Flexible Caching. These architects reported no performance or scalability issues with data virtualization. If they encounter a slow query, they simply persist or cache the target data in a traditional database and reconfigure the semantic layer accordingly. In other words, they virtualize everything they can and persist when they must. This approach overcomes all physical and political obstacles to data virtualization, while improving performance and agility. And some hard-core “data virtualizers” do away with a data warehouse altogether, preferring to persist snapshots of data that requires an historical or time-series view.

Today, the biggest obstacles to the growth of data virtualization are perceptions and time. Given the innate bias among data warehousing professionals to persist everything, most doubt that data virtualization tools offer adequate performance for the majority of query workloads. In addition, it takes time to introduce data virtualization tools into an existing data warehousing architecture. The tools must prove their worth in an initial application and build on their success. Since enterprise-caliber data virtualization tools cost several hundred thousand dollars, they need a well-respected visionary to advocate for their usage. In most organizations, it’s easier to go with the flow than buck the trend.

Nonetheless, the future looks bright for data virtualization. Since most of our data environments are heterogeneous (and always will be), it just makes sense to implement a virtualization layer that presents users and applications with a unified interface to access any back-end data no matter where it’s located or how it’s structured. This layer of abstraction, which balances federation and persistence, can do two things that every IT department must do today: reduce costs and increase agility.
Summary

Federation is the most pervasive BI architectural model, largely because most organizations cycle between centralized and decentralized organizational models. A BI architecture, by default, needs to mirror organizational structures to work effectively. Contrary to popular opinion, a BI architecture is a dynamic environment, not a blueprint written in stone. BI managers must define an architecture based on prevailing corporate strategies and then be ready to deviate from the plan when the business changes due to an unanticipated circumstance, such as a merger, acquisition, or new CEO.

Federation also does the best job of balancing the twin needs for enterprise standards and local control. It provides enough uniform data and systems to keep the BI environment from splintering into a thousand pieces, preserving an enterprise view critical to top executives. But it also gives business units enough autonomy to deploy applications they need without delay or IT intervention. Along the way, it minimizes BI overhead and redundancy, saving costs through economies of scale. Data virtualization software complements federated information architectures by providing a layer of abstraction between data sources and applications, simplifying data access and management.
About the Author

WAYNE ECKERSON has been a thought leader and consultant in the business intelligence (BI) field since 1995. He has conducted numerous in-depth research studies and is a noted speaker and blogger. He is the author of the best-selling book *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*. For many years, he served as director of education and research at The Data Warehousing Institute (TDWI), where he chaired its BI Executive Summit and created a popular BI Maturity Model and Assessment. Wayne is currently director of research at TechTarget and founder of the BI Leadership Forum, a network of BI Directors that exchange ideas and educate the larger BI community. He can be reached at weckerson@techtarget.com. This content originally appeared on Wayne’s World, a BeyeNetwork blog that illuminates the latest thinking about how to deliver insights from business data and celebrates out-of-the-box thinkers and doers in the business intelligence, performance management and data warehousing fields.
COMPOSITE
SOFTWARE

• Introducing the first book ever written on Data Virtualization
• Follow the Latest Analyst Research on Data Virtualization
• Learn more about how enterprise IT is using data virtualization

About Composite Software, Inc.:
Composite Software, Inc. is the data virtualization performance leader. Backed by a decade of pioneering R&D, Composite Software is the data virtualization gold standard at ten of the top 20 banks, six of the top 10 pharmaceutical companies, four of the top 5 energy firms, major communications providers and the world’s largest IT organization, the US Army. These and hundreds of other global organizations rely on Composite to fulfill their ever-changing information requirements with greater agility and lower costs. To contact Composite, please call (650) 227-8200, visit us on the Web at http://www.compositesw.com, or follow us on twitter http://twitter.com/compositesw and Facebook http://www.facebook.com/pages/Composite-Software/284366864940945.