Cost Optimization Is Crucial for Modern Data Management Programs

Published 22 June 2020 - ID G00726143 - 33 min read

By Analysts Ankush Jain, Guido De Simoni, Adam Ronthal, Eric Thoo, Sally Parker, Donald Feinberg, Ehtisham Zaidi, Simon Walker, Malcolm Hawker, Melody Chien

Overview

Key Challenges

- Proven best practices for cost optimization exist for all five data management disciplines: data integration, data quality, master data management (MDM), enterprise metadata management (EMM) and database management. However, many organizations either are not aware of these practices or lack a structured plan for executing them to reduce costs and increase operational efficiency across their data management initiatives.

- Cost optimization initiatives are often evaluated only by their potential to reduce spending, even though many have shortcomings when it comes to long-term cost efficiency. Gartner sees organizations treating these initiatives as an afterthought.

- Cost optimization is frequently misunderstood as cost cutting. This can result in needless limiting of resources, failure to exploit digital business opportunities and weakening of competitive advantage.

Recommendations

Data and analytics leaders aiming to use cost optimization to strengthen their data management solutions should:

- Take control of their data integration tools’ licensing and rationalize capabilities by using converged data fabric techniques.

- Develop a cost-effective data quality program focused on a proactive business approach supported by augmented data quality.
Align MDM programs with prioritized business outcomes to keep these programs lean and focused.

Prioritize metadata management programs to support long-term data management cost optimization, reduce business risk and improve productivity.

Revamp their DBMS portfolio by migrating to the cloud, consolidating on-premises environments and instituting financial governance.

**Introduction**

Cost optimization is a business-focused discipline for reducing spending and, more generally, costs, while maximizing business value. This results in efficient allocation of spending to achieve desired outcomes. Forward-looking data and analytics leaders are investigating cost optimization as a discipline that systematically applies proven techniques across people, practices and technology assets in order to lower operating costs, drive value creation and equip organizations for a digital future. They are keen to control costs and maximize the business value of their data investments because these investments can run to several million dollars per year.

To do so, data and analytics leaders in charge of modernizing information infrastructure should apply cost optimization techniques in three broad categories (see Table 1):

- **People** (new and existing roles, personas, skills, training)
- **Practices** (team structures, modern architectures, team organization and collaboration)
- **Technology** (tool consolidation, new deployment options [such as the cloud], exploration of mature open-source solutions, and investigation of adaptive pricing strategies)

<table>
<thead>
<tr>
<th>People (Skills)</th>
<th>Practices (Processes and Architecture)</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1: Cost Optimization Across Data Management**
Typically, all three categories provide opportunities for cost optimization in relation to the five data management disciplines:

- Data integration
- Data quality
- MDM
- Enterprise metadata management (EMM)
- Database management

This report aims to help data and analytics leaders apply proven cost optimization techniques in all five disciplines.

Analysis

Take Control of Your Data Integration Tools’ Licensing and Rationalize Capabilities by Using Converged Data Fabric Techniques

Data and analytics leaders have an immediate opportunity to optimize their data integration program costs by effectively managing data integration tool licensing. They can begin by rationalizing their data integration approaches (by adopting bimodal, shared services approaches...
to data integration), and increasing the speed at which they integrate data silos through the use of data fabric techniques.

**People**

Many large organizations approach data integration in a somewhat fragmented manner — individual departments and business units may have implemented data integration tools in a project-specific fashion and therefore ended up with duplicated approaches and staffing. In addition, individual project teams often develop data integration skills in a disparate manner, particularly due to isolated procurement of tools and infrastructure. Skill and staff resourcing costs escalate when common data integration problems are addressed by wide-ranging architectural solutions that are not streamlined. Consequently, most large organizations lack rationalized techniques that use shared services and practices to optimize the productivity of staff running data integration processes. This can, however, be done by creating reference architecture and common capabilities for resourcing skills and staff.

Data and analytics leaders must address the skills gap between, on the one hand, development teams comprising architects, data engineers and extraction, transformation and loading (ETL) staff focused on creating data pipelines and, on the other, business teams that need integrated data for their data and analytics use cases. Additionally, less skilled roles, such as citizen integrators, analysts and citizen data scientists, must be empowered to integrate their own data in sandbox environments through the use of self-service data preparation tools that enable them to access internal/external, structured/multistructured and streaming data with minimal IT support. This will give IT teams, and particularly data engineering teams, more time to focus on strategic initiatives such as collaboration and data integration planning.

Furthermore, data engineering teams are spending a significant percentage of their time manually creating, maintaining and operationalizing data pipelines, which constrains their productivity. Fortunately, augmented data integration capabilities — which feature machine learning (ML) algorithms that utilize active metadata to inform and automate aspects of data integration design, delivery and orchestration — can help. When provisioned through data fabric architectures, they can immediately provide a much-needed productivity boost to integration and engineering teams by automating mundane and repeatable tasks (see “Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration”).

**Practices**

Data and analytics leaders must update their approaches to investing in data integration tools and gain the extra agility required to optimize costs. To do this, they should:

**Rationalize Data Integration Tools and Technology-Licensing Techniques**

Many organizations pay much more than they need for data integration tools, because they do not pay enough attention to evolving pricing models and licensing terms, or to negotiation best practices, when dealing with tool vendors (see “Focus on These Key Contract Components When...”)
Negotiating a Data Integration Tool Purchase and “Adopt 10 Best Practices to Become a Contract Negotiation Expert for Data Integration Tools”).

Data and analytics leaders should:

- Assess and use the data integration tools included, at no additional cost, in products that have already been deployed, such as DBMS tools and data warehouse tools (see “Magic Quadrant for Data Management Solutions for Analytics”).
- Assess alternative integration tool models, such as open-source, freeware and freemium options, all of which offer low or no upfront costs. For some organizations, tools that deliver “good enough” capabilities for specific project goals will be acceptable.
- Evaluate the benefits of a subscription-based, mature integration platform as a service ( iPaaS) solution to lower infrastructure and support costs and enable business users and citizen integrators to perform integration without depending on IT staff (see “Magic Quadrant for Enterprise Integration Platform as a Service”).
- Negotiate aggressively for nonproduction tool licenses, which can cost 50% less than production licenses and gain an even higher percentage reduction in big transactions.
- Protect future discounts for incremental purchases by negotiating effective price protection over a period of three to five years.

Create and Empower a Bimodal Data Integration Practice to Bring Agility, Reusability and Faster Time-to-Insight to Slow and Costly Data Integration Programs

Data integration has often been described as slow and costly due to its IT-centric focus and tools that do not focus on business users. Today, data integration is no longer the preserve of IT teams. Business users are demanding tools and platforms that enable them to integrate a variety of data sources (both on-premises and cloud) with minimal coding knowledge or IT support, in order to reduce the latency and costs associated with IT-centric data integration.

Data and analytics leaders need to create a bimodal data integration strategy with equal representation from IT and business functions. They must do so to empower business users (such as citizen integrators) to undertake new projects that require faster delivery of integrated data without much IT support or knowledge of coding. These projects need tools that are low-cost and easy to implement and support, because most could be one-off, “fail fast and move on,” Mode 2 projects. For these, organizations do not need the kind of enterprisewide, comprehensive and expensive data integration tools used for mission-critical, Mode 1 projects that demand IT support. Instead, they are often ideally suited to modern lightweight and/or self-service integration tools, which can also reduce costs substantially. There is an opportunity here for data and analytics leaders to invest in these tools for Mode 2 projects in order to rationalize costs and empower the business. These tools include:

Assess and use the data integration tools included, at no additional cost, in products that have already been deployed, such as DBMS tools and data warehouse tools (see “Magic Quadrant for Data Management Solutions for Analytics”).

Assess alternative integration tool models, such as open-source, freeware and freemium options, all of which offer low or no upfront costs. For some organizations, tools that deliver “good enough” capabilities for specific project goals will be acceptable.

Evaluate the benefits of a subscription-based, mature integration platform as a service ( iPaaS) solution to lower infrastructure and support costs and enable business users and citizen integrators to perform integration without depending on IT staff (see “Magic Quadrant for Enterprise Integration Platform as a Service”).

Negotiate aggressively for nonproduction tool licenses, which can cost 50% less than production licenses and gain an even higher percentage reduction in big transactions.

Protect future discounts for incremental purchases by negotiating effective price protection over a period of three to five years.
The choice of which tools to add or eliminate as part of a rationalization effort is greatly influenced by the pressure to reduce costs and simultaneously provide much-needed flexibility for sharing data, enable less complex implementations and achieve faster time to value. Data and analytics leaders should:

**Use Data Virtualization to Consolidate Data Silos, Lower Costs and Boost Business Agility**

Data virtualization has become a mature and common style of data delivery used to reduce the physical consolidation of data in silos. Most data integration vendors now provide mature data virtualization offerings that enable organizations to move away from rigid and expensive enterprise data warehouse architectures to more flexible logical data warehouse (LDW) architectures. With an LDW architecture, data virtualization enables the federation of data from multiple silos (data marts, for example) into an integrated, shared services layer and eventually into one integrated data model that can be used by many different business functions (see "Market Guide for Data Virtualization").

Gartner finds that, in some cases, a significant proportion of the overall cost incurred in supporting individual data marts is redundant, because each mart requires its own attendant infrastructure, storage and database administration resources. Much, if not all, of that redundant cost can be eliminated by connecting those marts to application-neutral virtual views using data virtualization. Moreover, data virtualization can enable organizations to connect to new and upcoming data sources (such as Hadoop, NoSQL, the cloud and the Internet of Things [IoT]) without having to physically move data into repositories. This, in turn, enables them to quickly produce the integrated views of data needed by the business in order to achieve faster time to analysis — which can lead to huge cost savings.

**Apply a Data Fabric Design When Adopting Data Integration Technologies**

Most organizations are struggling to maintain numerous point-to-point integration flows and multiple redundant tools for data integration. Many organizations are starting to require the capabilities to combine data delivery styles (batch with streaming, for example), but are struggling to provide a comprehensive architecture that supports this requirement.

The cost of gaining the required flexibility and agility can be optimized with the help of a data fabric design. Data fabric is an emerging concept for the optimal combination of integration and data management technologies, related architectural designs and services delivered via orchestrated platforms and practices (see "Data Fabrics Add Augmented Intelligence to Modernize

Gartner, Inc. | 726143
Data and analytics leaders should focus on increasing the collaborative deployment of integration patterns augmented by meaningful metadata exchange and graph-based semantic enrichment on required data service compositions. These are needed for fluid sharing of data and applications to enhance their integrated use as a solution for supporting data management. This approach avoids any haphazard reactive measures such as point-to-point data integration delivery where costs escalate uncontrollably.

Further Recommendations

- Ensure you are making the most of your current investments in data integration tools and investigate whether your current vendor can meet your upcoming data integration requirements (for data virtualization, for example) before investing in new integration products.

- Create and empower a bimodal data integration practice, with separate tools for business users and citizen integrators, in order to provide the agility that has long been missing from data integration teams. This will help to keep IT costs down.

- Investigate and adopt low-cost iPaaS solutions and open-source data integration tools for new projects that require agility, flexibility and a lower time to solution. Use a data fabric design as a guide.

Develop a Cost-Effective Data Quality Program Focused on a Proactive Business Approach Supported by Augmented Data Quality

Data quality is the foundation of everything that is built on an organization's data assets. Poor data quality destroys business value. A recent Gartner survey of reference customers for the forthcoming 2020 edition of “Magic Quadrant for Data Quality Solutions” found that organizations estimate the average cost of poor data quality at $12.8 million per year. This number is likely to rise as business environments become increasingly digitized and complex. Also, another recent survey found that the need to strive for data quality across data sources and landscapes is the joint-biggest challenge to data management practice (see Figure 1 and “Survey Analysis: Data Management Struggles to Balance Innovation and Control”).

Figure 1: Key Challenges to Data Management Practice
Data quality is a critical aspect of data management. To optimize costs effectively for data quality programs, Gartner recommends actions in three key areas, as described below.

**People: Position Key Data Quality Roles Efficiently to Have the Biggest Impact on the Business and Optimization Efforts**

Data and analytics leaders should optimize data quality efforts by positioning staff with data quality skills in the business units where poor data quality is costing the most money and where those staff can have the greatest impact on the business. Responsibility for defining and managing data quality skills, data quality responsibilities and placements of data quality staff within organizations is shifting from the IT department to the business or an IT-business hybrid.

To select and position data quality roles and responsibilities cost-effectively, data and analytics leaders should take three steps:

1. **Create an adaptive data governance strategy.** When talking about data governance, it is important to understand that not all data is equal. Fortunately, adaptive data and analytics
governance can help to create a flexible framework for handling data of whatever nature (see “Data and Analytics Leaders Must Use Adaptive Governance to Succeed in Digital Business”). The data quality requirement is not the same for all data. Some data, such as master data, needs to be highly controlled or, as in the case of personally identifiable information (PII), regulated in accordance with clear compliance requirements. For other data, such as that associated with many business unit projects, a more empirical governance method involving minimal processes and controls is often appropriate. Used together, the two approaches form agile information governance, which defines and communicates a common decision framework for data quality policies, standards and processes for both Mode 1 and Mode 2 projects, as well as for adaptive sourcing strategy.

2. **Evolve toward a business-driven organizational structure.** Such a structure clearly identifies, defines, positions and quantifies data quality tasks and the roles responsible for them (for example, data steward, data owner and developer). With such a structure, data and analytics leaders can enable the application and enforcement of data quality policies and standards. They can also enable more flexible reallocation of data quality staff, thus optimizing personnel costs.

3. **Adopt the right operating model.** The model used should establish the optimal workflow between critical roles and directly measure the performance of those roles against business outcomes — through quantifiable data quality improvement (see “Build a Data Quality Operating Model to Drive Data Quality Assurance”).

**Practices: Embed Data Quality Tasks in Business Processes and Monitor Progress Over Time**

The greatest impact of cost optimization comes from treating data quality assurance as a proactive business process (see “5 Steps to Build a Business Case for Continuous Data Quality Assurance”). Most organizations merely react to data quality problems as they arise. This is a time-consuming approach that fails to find and fix the root causes. Instead, data and analytics leaders should recast data quality as a business process by designing and executing a series of tasks to be performed throughout the life cycle of data. These tasks should include steps for continuously monitoring, identifying and resolving data quality issues as early as possible. Such a proactive, business-process-oriented approach should focus on three activities:

- **Identify and fix data quality problems as early as possible.** This will avoid the much higher cost (in time and money) of fixing them later and minimize the follow-on costs when poor data quality affects other processes and systems.

- **Proactively monitor the data quality process.** This should improve the productivity of the entire organization by reducing uncertainty, confusion and arguments about what constitutes data quality, whether quality has been compromised, how problems should be resolved and who should be blamed for them.
Establish and enforce data quality standards. This will ensure quality throughout the data life cycle, thereby reducing the downstream cost of controlling and managing data assets.

Technology: Prioritize Adoption of Modern Data Quality Solutions in a Cost-Effective Manner

Gartner’s survey of reference customers for the forthcoming “Magic Quadrant for Data Quality Solutions” found that the average annual spending on data quality tools is nearly $250,000 per organization. To reduce these costs, leading organizations embrace an adaptive sourcing approach that improves sourcing flexibility, drives business value and maintains fit-for-purpose governance. This approach involves an array of proven tactics:

- **Establish an enterprise standard** for data quality tools for Mode 1 projects, with a focus on performance, reliability and reuse. This lets enterprise data quality tools become part of shared services for both Mode 1 and Mode 2 projects, which reduces tool and training costs.

- **Prioritize adoption of data quality SaaS or PaaS products**, if the organization lacks enterprise tools, especially for use cases involving the IoT, cloud or mobile data.

- **Evaluate alternative deployment models** in collaboration with an organization’s existing data quality vendor(s), and renegotiate pricing for hybrid deployments.

- **Exploit new licensing models**, such as subscription-based, pay-as-you-go and consumption-based pricing.

- Participate in technology trends that include augmented (machine learning [ML]-enabled), unified and synchronized data quality solutions. Modern data quality tools deliver an array of intelligent capabilities by using ML and natural language processing (NLP). ML uses algorithms to learn from experience and to provide new insights without being explicitly programmed to do so. The productivity of these solutions will be greater, so fewer licenses will be needed.

- **Use the convergence of data management functions as leverage** to get greater discounts on data quality solutions and achieve better integration. This approach can also result in tighter workflow integration, better metadata management, and easier tool integration and vendor management.

Further Recommendations

- Define an adaptive data governance strategy, evolve toward a business-driven organizational structure and establish an operational model for data quality initiatives by using the best practices identified above.

- Manage data quality as a business process by designing and executing a series of steps that constantly watch for, and identify and resolve, data quality issues.
Align Master Data Management Programs With Prioritized Business Outcomes to Keep These Programs Lean and Focused

The discipline of MDM focuses on the consistency and quality of data that describes the core entities of an organization — its customers, citizens, patients, assets and products, for example. At its core, MDM is about optimization — through the breaking down of siloed data management and decision making — which leads to lower costs and higher business value.

Organizations that have successfully implemented MDM programs benefit from greater agility with which to respond to unexpected and unprecedented events. They have, for example, more ability to predict and respond to changing customer buying patterns in relation to the COVID-19 pandemic.

MDM is key to digital business success. However, organizations keen to achieve MDM-related cost optimization are at different stages in their MDM journeys:

- Some organizations are seeking to use MDM for the first time, and to do so cost-effectively.
- Other organizations are seeking to optimize existing MDM programs.

MDM is not an “out of the box” technological fix. It is a technology-enabled business discipline that requires due consideration to be given to all three of the broad categories described in this document: people, practices and technology.

To effectively optimize costs for programs, Gartner recommends the following actions to data and analytics leaders for the two scenarios:

**Scenario 1: Organization Exploring MDM for the First Time**

**People:**

- MDM supports business goals. Therefore, the rightful custodians of data stewardship activity are business stakeholders. Engage these stakeholders early to ensure they understand and accept their role. Use tools to automate and simplify stewardship tasks (see “Toolkit: Data and Analytics Governance Role Descriptions”).

- Engaging third parties with a track record of expertise in your industry and use case is a proven way to shorten the time to value. Use their frameworks and toolkits, and incorporate skills transfer into your contract (see “Market Guide for MDM External Service Providers”).

**Practices:**
Evaluate whether simply fixing a process can resolve a master data pain point, without a need for any new technology, by reviewing existing processes.

Align all MDM activity behind an agreed business outcome. Wherever possible, articulate and seek agreement on goals in business terms. Take an iterative approach and prioritize a short-term focus based on the agreed goals (see “Data & Analytics Operational, Data Quality, and Data Management Metrics”).

Technology:

Assess the viability of using simpler, discrete solutions, such as application data management (ADM) tools, customer data platforms (CDPs), data quality tools and stewardship tools, as a response to specific short-term needs, and even their adequacy for business needs in general. Have a long-term strategy to protect short-term investments (see “Think Big, Start Small, Be Prepared — Master Data Management”).

Where MDM is deemed necessary, explore solutions that present a lower barrier to entry — look for subscription pricing, cloud deployment options and solutions that require configuration rather than coding. Many MDM software vendors offer starter MDM solutions with an option for future expansion with more advanced functionality (see “Magic Quadrant for Master Data Management Solutions”).

Take stock of what you already have. Are you using all the features of software you have already purchased? Is there potential to extend existing relationships to include MDM as part of a preferential agreement?

Scenario 2: Organization Seeking to Optimize an Existing MDM Program

People:

Look for ways to drive down data stewardship costs by automating more of your existing governance processes, or more of the master record creation process itself.

Carry out a “health check” with a view to making resource savings by consolidating or even ending certain functions (see “Toolkit: Data and Analytics Governance Role Descriptions”).

Practices:

Carry out a health check with a view to making further gains in terms of processes. What can you refine, automate or stop?

Work with business colleagues to understand where their use cases may not require a “one size fits all” approach to governance, and where AI/ML capabilities could provide “good enough” data stewardship and maintenance.
Partner with third-party data providers or data consortiums that may offer cost-effective solutions for reference data as an alternative to in-house stewardship.

Technology:

- Investigate whether your current data management tool vendor has the capabilities to meet your upcoming MDM requirements, before investing in new products (see “Magic Quadrant for Master Data Management Solutions”).

- Compare your assumptions about the volume of records you will have to manage in your hub against post-COVID-19 forecasts. If your contract is due for renewal shortly, negotiate based on the lower numbers. If you are in the middle of a contract, request that the vendor lower its in-year costs (based on a forecast of reduced volume) in exchange for an extension of the contract.

- Evaluate lower-cost technologies, such as CDPs, which may meet many of your requirements at a fraction of the current cost. This is especially the case for analytical forms of MDM, where MDM is used primarily to generate “360-degree” customer views. (see “Think Big, Start Small, Be Prepared — Master Data Management” and “Choose Between Customer Data Platforms and MDM Solutions for 360-Degree Customer Insights”).

Further Recommendations

- Secure agreement with stakeholders about prioritization of the business outcomes that the MDM program will support. MDM program goals must align with the business strategy — they should justify the investment in the MDM program.

- Think big, but start small. Only if process improvements alone do not fix your problems should you invest in new tools to continue your MDM journey.

- Agree on what constitutes master data — stay “lean and clean” (see "MDM Primer: How to Define Master Data and Related Data in Your Organization").

- Make master data governance part of your usual business practice, and ensure there is accountability for the provision of sustainable governance.

Prioritize Metadata Management Programs to Support Long-Term Data Management Cost Optimization, Reduce Business Risk and Improve Productivity

Metadata has a key role to play in optimizing data management costs. It focuses on the data associated with information assets. Metadata management ensures that information can be integrated, linked, maintained, accessed, shared and analyzed to best effect across an organization. The goal is to improve information assets’ usability for specific data management and analytics programs, such as data integration, data quality, MDM and analytics governance.
Sharing key metadata across these programs, through EMM, enables an organization to derive business value from information assets. EMM encompasses roles, skills, responsibilities, processes, organizational units and technologies, and for each offers opportunities for cost optimization (see “Implement Enterprise Metadata Management to Drive Effective Enterprise Information Management”).

EMM reduces business risk and increases productivity by enabling more effective information linkage, integration and exchange. This, in turn, leads to lower costs and higher business value. Metadata management maintains consistent definitions of data elements and their end-to-end lineage, leading to greater data sharing across business processes and business units. Greater sharing drives greater collaboration, which in turn drives the optimization of business processes.

In addition, augmented data catalogs assist with ML-enabled automation of four broad categories of metadata management tasks that are essential to reduce the time to insight for data and analytics use cases: (1) Discover, (2) Understand, Enrich and Trust, (3) Contribute and Govern, (4) Consume (see “Augmented Data Catalogs: Now an Enterprise Must-Have for Data and Analytics Leaders”).

Cost-Effectiveness and Reduced Business Risk

EMM consistently and clearly defines data assets, their end-to-end data lineage and relevant transformation rules. This enables an organization to trace and audit data assets and create a robust framework for compliance reporting. Among EMM’s other benefits, two are especially important:

- EMM supports the information catalogs that hold common descriptions of enterprise data assets and their semantics.
- EMM saves business analysts and end users time by enabling them quickly and reliably to find the right data and the right report filters to gain more accurate insights.

Cost Optimization Through Improved Productivity

A unified and consistent view of metadata saves a considerable amount of analysis and data collation time, thereby enabling greater productivity. Metadata management can support significant productivity improvements in several ways:

- Reducing data integration process costs by using data lineage to identify process redundancies, such as the use of different processes to update the same attributes or tables. Data lineage can also be used to support impact analysis, thus reducing development time.
- Exchanging data easily across business units leads to better and faster decision making, thus improving the effectiveness of a wide range of business processes, including customer service and operational processes.
Creating transparency concerning the data assets in an inventory and driving insights about them in order to show how they can be used for business decision making.

Improving business and technical queries through effective rules management that significantly reduces the amount of manual work required to analyze critical data (see “Magic Quadrant for Metadata Management Solutions”).

Figure 2 shows an example of cost optimization achieved with metadata management.

Further Recommendations

Gartner recommends the following actions in the three key areas to support cost optimization through the use of metadata management:

- **People**: Reach a consensus internally about what constitutes metadata, based on the information assets it describes and how your organization uses those assets to create value (see “Create a Business Case for Metadata Management to Best Fulfill Your Data and Analytics Initiatives”).

- **Practices**: Develop a metadata management strategy that incrementally improves the standardization and sharing of metadata and therefore helps exploit the value of your organization’s information assets, reduces business risks and increases productivity (see “Implement Enterprise Metadata Management to Drive Effective Enterprise Information Management”).
Revamp Your DBMS Portfolio by Migrating at Least Some Elements to the Cloud, Consolidating On-Premises Environments and Instituting Financial Governance

The DBMS market continues to move to the cloud, with many new offerings and enhancements to existing offerings reflecting a cloud-first, or even a cloud-only, mindset. Any cost optimization initiative must be undertaken in light of this broad move to the cloud.

There are five key ways to optimize DBMS costs:

1. **Identify and seize opportunities to move DBMSs and the applications that run on them to the cloud.** Applications that draw on database platform as a service (dbPaaS) offerings will benefit most from a cloud migration initiative.

2. **Rationalize and consolidate the overall DBMS portfolio.** Reducing the number of supported platforms for data management can simplify operations, make the most of skills, reduce spending, and streamline data integration and governance overhead.

3. **Align workloads with appropriate pricing models for cloud DBMS services.** Exploratory or experimental workloads will benefit from consumption-based pay-as-you-go pricing, whereas established production workloads will favor more predictable pricing.

4. **Align workloads with appropriate service offerings.** The key metric to consider as part of any migration is price/performance. Different offerings may provide better results for workloads with different characteristics.

5. **Assess open-source alternatives or subscription pricing options that may provide cloud-like benefits.** This applies to those workloads that remain on-premises.

Each of these opportunities aligns with one or more of the three strategic focal points — people, practices and technology — for overall cost optimization of information infrastructure.

**People**
Use of the cloud will have a positive impact on infrastructure and DBMS operations teams. The cloud deployment model reduces the need for infrastructure support, system hardware and software support staff, and database administrators (DBAs) who focus on enterprise platform integration (backups, servers and provisioning). As a result, data and analytics leaders can reallocate resources to the core business-focused activities associated with the DBAs who focus on the application development tasks associated with data management.

The replacement of older DBMS technology with cloud-based solutions also helps organizations align with the evolving skills in the market. Skills to support legacy or aging technology are difficult to find and, when found, command a cost premium.

Finally, rationalization of the DBMS portfolio will have an impact on the skills required to support the database management environment. The use of fewer technologies may lead to a more streamlined support organization (see “Overcoming Resistance to Data Infrastructure Modernization”).

**Practices**

The process of cost optimization is iterative. Data and analytics leaders need to decide whether it is worth consolidating all workloads on a single platform — even if it does not provide the most optimal price/performance for all of them — or splitting workloads across multiple service offerings. In the latter case, each workload may run on the optimal platform for price/performance, but additional overhead is incurred because of the more complex integration, governance and operational management.

The use of dbPaaS as a deployment model gives rise to cost optimization opportunities. These include deployment of transient environments (such as development and testing environments), rightsizing of production environments through cloud elasticity, and reduction of high-availability/disaster recovery (HA/DR) investments by using the cloud for contingency planning and business continuity.

In the cloud, appropriate alignment of pricing models with workloads can be a powerful means of cost optimization. Aligning consumption-based, serverless, metered pricing models with exploratory, experimental or transient workloads will ensure that resources are used efficiently and that idle resources are not adding to pricing overhead. Production workloads that have specific SLAs and known characteristics will favor pricing predictability and align better with node-based or serverless-metric pricing models. Look for cloud offerings that blend pricing models or provide the ability to switch between models as workloads change or shift (see “New Pricing Models for Cloud DBMSs Provide Cost Optimization Opportunities”).

Although most important in the cloud, price/performance is the most overlooked metric. Informed data and analytics leaders should not select offerings solely on the basis of list prices or discounted prices, or the presence of specific features, although these should influence the buying decision. Rather, they should focus on assessing how much it costs to run a given workload at the
desired level of performance. They should also bear in mind that, in the cloud, metrics will likely change over time as cloud service providers and independent software vendors optimize their environments and add new performance-oriented features on a regular basis.

The form of DBMS consolidation depends on the deployment environment. In the cloud, there is a choice between a multimodal and a best-fit approach. If a single multimodel DBMS can efficiently support multiple workloads (relational, document store and graph, for example), it may greatly simplify the data management landscape. This may offer a cost advantage over a best-fit approach, for which integration of services may entail higher costs. It is important to check, however, that price/performance compromises do not outweigh the multimodel approach’s benefits.

On-premises consolidation efforts focus on reducing hardware infrastructure and licensed cores. They can yield significant cost savings. Consolidation can be achieved by using virtualization or container approaches, or using multitenant capabilities native to the DBMS itself. Best practices for consolidation of DBMS instances are based on isolation levels, performance requirements and complementary workloads that can share physical hardware. They can reduce costs across the stack — from data center footprint to administrative overhead. Additional opportunities present themselves in DBMS migration efforts. The promise of lower-cost subscription model pricing (associated with open-source offerings, but also emerging in commercial offerings) is appealing. In these scenarios, extra care should be taken to ensure that the potential cost reductions outweigh the migration costs (see “What to Consider When Migrating to a New DBMS”).

Technology
The rise of multimodel capabilities enables a single modern DBMS to accommodate both relational models and nonrelational models (such as document, graph and key-value models). They are often sufficient for modest workload requirements. Data and analytics leaders may therefore consolidate DBMSs for different workloads, leading to significant potential cost savings in relation to application integration and skills. But it is important to remember that a best-fit approach may deliver better performance.

Myths persist about the cost of cloud services. The initial investment in the cloud infrastructure required for a DBMS is minimal, especially compared with the required investment in infrastructure, support and licenses for a comparable on-premises DBMS. This alone will deliver a short-term cost saving. But the real benefit of a cloud dbPaaS comes from its operational flexibility, which:

- In response to changing workload requirements allows for optimal provisioning of compute and storage tiers, enabled by the separation of resources that is common to cloud architectures.

- Enables an organization to dispense with long-term infrastructure investment (see “Toolkit: Comparative Total Cost of Ownership Calculator for Cloud and On-Premises DBMS Deployments”).
Achieving long-term cost savings in the cloud requires application systems to use or migrate to native cloud dbPaaS systems, as opposed to hosting DBMS systems on cloud IaaS. Data and analytics leaders may find that reengineering on-premises systems to take full advantage of cloud capabilities is more cost-effective than migrating to cloud dbPaaS systems.

Finally, open-source DBMSs have matured significantly and are now suitable for a growing percentage of workloads. They often offer significant license savings in comparison with commercial, proprietary DBMS options (see “State of the Open-Source DBMS Market, 2019”).

**Further Recommendations**

- Align workloads with appropriate pricing models to ensure optimal use of cloud infrastructure.
- Institute strong financial governance models when using cloud DBMS platforms to ensure predictable costs without major surprises.
- Weigh the trade-off between multimodal and best-fit approaches in terms of price/performance for a given workload. Recognize that the additional overhead associated with full alignment of a workload with a service offering may not be cost-effective.
- Move DBMS applications to cloud dbPaaS offerings that have lower initial risk profiles. This applies to data management solutions for analytics use cases, development and test environments, nonsensitive data, new initiatives and technology exploration.
- Consolidate current DBMS instances to reduce license costs (and required resources and risks), if you have the ability to retire DBMS licenses and if virtualization works in your favor.
- Add open-source DBMSs to your organization’s DBMS standards for new uses, and replace commercial relational DBMSs where these are not specifically required.

**Evidence**

Insights in this document derive from sources including:

- **Gartner’s client inquiry service:** Each year, Gartner’s data and analytics team receives thousands of inquiries from Gartner clients. During the past 24 months, inquiries specifically about data management pricing, licensing and cost optimization totaled more than 2,000.
- **Survey of reference customers for the 2020 “Magic Quadrant for Data Quality Solutions”:** The survey of vendors’ reference customers was conducted in the first quarter of 2020 and formed part of a data-gathering effort to help Gartner build on its knowledge of vendors in the market for data quality solutions. At the start of the Magic Quadrant research process, all contacted vendors were asked to identify reference customers that generally represented the inclusion criteria. Vendors provided contact information that was used to invite their reference customers to complete a 35- to 40-minute online survey. A total of 154 reference customers from 20
vendors completed the survey. Note: vendors’ reference customer data is different from primary research and is not a representative knowledge base for the data quality tool market.

- Gartner’s Data Management Strategy Survey was conducted online from 19 August through 4 September 2019 with 129 Gartner Research Circle members — a Gartner-managed panel. The survey was developed by a team of Gartner analysts. It was reviewed, tested and administered by Gartner’s Research Data and Analytics team.

**Recommended by the Authors**

- Cost Optimization in 2020 Resource Center Primer for 2020
- Modern Data and Analytics Requirements Demand a Convergence of Data Management Capabilities
- Apply These Techniques to Maximize Your Return on Data Quality Technology Investments
- 7 Steps to Build a Successful Business Case for MDM Programs
- Metadata Is the Fish Finder in Data Lakes
- 5 Ways to Use Metadata Management to Deliver Business Value From Data
- Understanding and Planning for Database Management Transformation to the Cloud
- How to Succeed at Database and DBMS Migration
- Emerging Technology Trends Create Opportunities for DBMS Cost Optimization
- The Future of the DBMS Market Is Cloud

**Recommended For You**

- Reduce IAM Program Risk and Redundancy by Formalizing an IAM Program Architecture
- Choosing Event Brokers: The Foundation of Your Event-Driven Architecture
- Summary Translation: Case Study: Product Management to Accelerate Business Transformation (PSEG)
- Summary Translation: 5 Fatal Pitfalls for New CIOs
- Summary Translation: How to Cut WAN and Internet Costs During a Crisis Such as COVID-19
believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by Gartner’s Usage Policy. Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "Guiding Principles on Independence and Objectivity."