

TIGHTENING THE DATA BELT

Creating value through better management of enterprise data logging

The era of big data has delivered real business impacts across industries. However, increased data gathering has a downside; ever-growing data storage volumes and compute requirements lead to budgetary and operational constraints. Automatically generated log files silently drive this rapid, enormous data growth. This Viewpoint identifies how companies can successfully improve holistic log file management to deliver benefits by controlling costs, ensuring compliance, and unlocking data value.

AUTHORS

Sri Rajagopal

Logan Jones

Mark Rowland

Michael Papadopoulos

Lokesh Dadhich

THE HIDDEN CHALLENGES OF BIG DATA

The age of big data is upon us. Organizations are unlocking its increasing power by growing the volume of data they collect, analyzing it with greater technical sophistication, and extracting insights that have real business impacts. Advanced tools such as artificial intelligence (AI) and the Internet of Things (IoT) are becoming increasingly prevalent, driving organizations to collect more and more data to enable better predictions, more insights, and more optimized operations. According to technology portal CDOTrends, approximately 64.1 zettabytes (ZB) of data were created globally in 2020, the equivalent of the total storage of over 250 billion average-sized personal computers. This volume is expected to grow rapidly, reaching 175 ZB by 2025 (a 23% CAGR).

The rise of the log file

Part of this data growth can be attributed to a silent, often overlooked source: log files, generated automatically by almost every connected machine, server, or application (see Figure 1).

Any organization tracking information in its IT environment or on a website, mobile app, or system is likely to be generating, storing, and analyzing logs. In the aggregate, logs represent a “watchtower” for organizations, as the data generated provides a basis for uses, including:

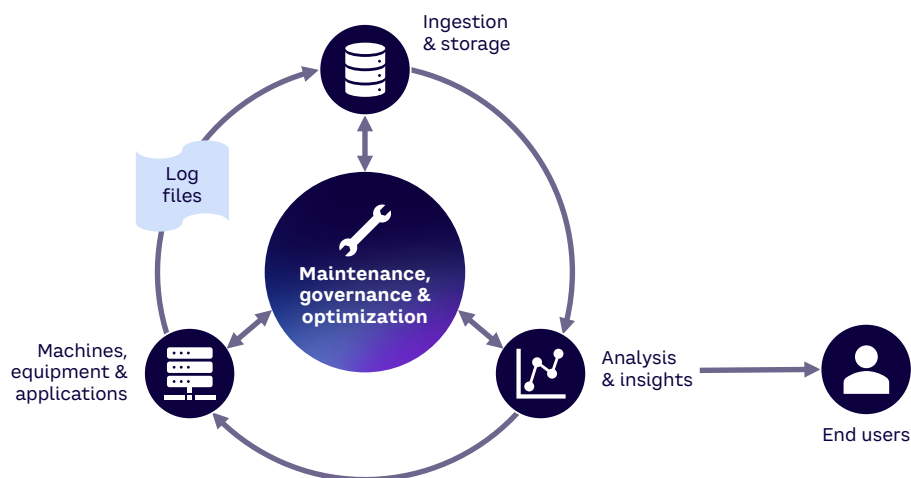
- Detecting errors and threats through firewall logs.
- Using analysis to enable optimized solutions.
- Troubleshooting and auditing systems.
- Establishing regulatory compliance.

What drives growth?

Three factors are generating growth in log file data volumes:

1. **Digitization.** Since log files are generated automatically, greater digitization increases their spread. 5G development, for example, requires that more network devices produce more logs. This provides telecom operators the opportunity to increase efficiency, reduce latency, and improve connectivity.

Figure 1. The role of log files in the data ecosystem



Source: Arthur D. Little

- 2. **Business insights.** Data logging brings significant strategic value to companies across sectors, as outlined in Figure 2. For example, analyzing log data on Web traffic helps companies better understand retail trends, uptime, and user behavior.
- 3. **Security and compliance.** Analyzing the data logs of security solutions can protect companies from malicious attacks and ensure regulatory compliance.

The data dilemma

Log volume growth poses a dilemma: more data enables new insights, better-informed business decisions, and new opportunities for innovation. However, this growth simultaneously leads to:

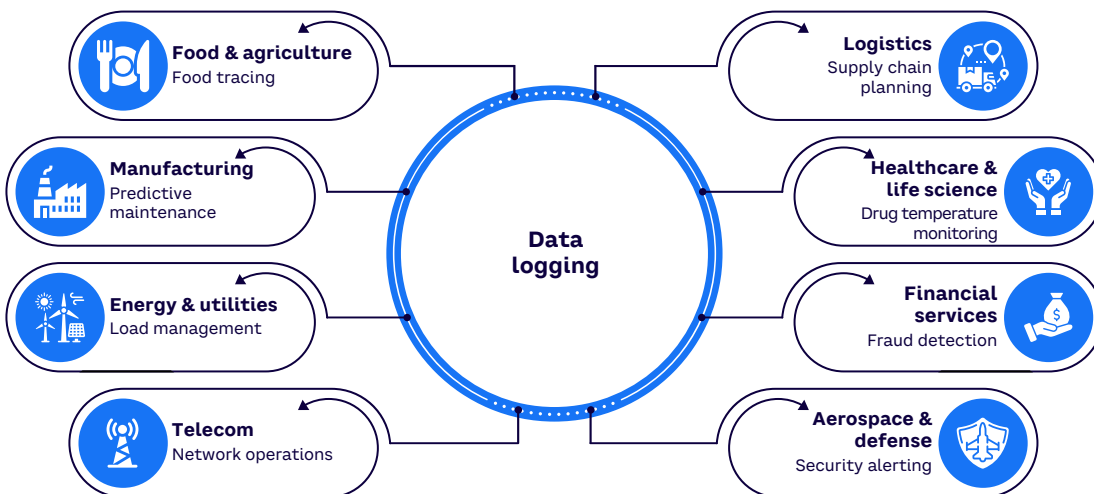
- Ever-rising storage and compute costs (e.g., a 10% increase in log volumes translates directly into 10% higher storage costs).
- Operational growing pains when legacy management systems reach their budgetary and operational limits.
- Strains on existing data governance systems, which potentially impact adherence to compliance requirements.

CONSIDERATIONS FOR HOLISTIC LOG MANAGEMENT

CTOs and CIOs are increasingly facing the inevitable data management side effects of a “log everything” approach. How can they transform management practices to increase efficiency while expanding the value gained from log data? Arthur D. Little’s (ADL) experience has revealed that CTOs and CIOs need to focus on four key considerations.

MOST ORGANIZATIONS HAVE ADOPTED A “LOG EVERYTHING” APPROACH, WITH ALL LOGS INGESTED AND STORED, REGARDLESS OF THEIR POTENTIAL FUTURE NEED OR VALUE

Figure 2. Business use cases of data logs



Source: Arthur D. Little

1. Governance is your greatest tool

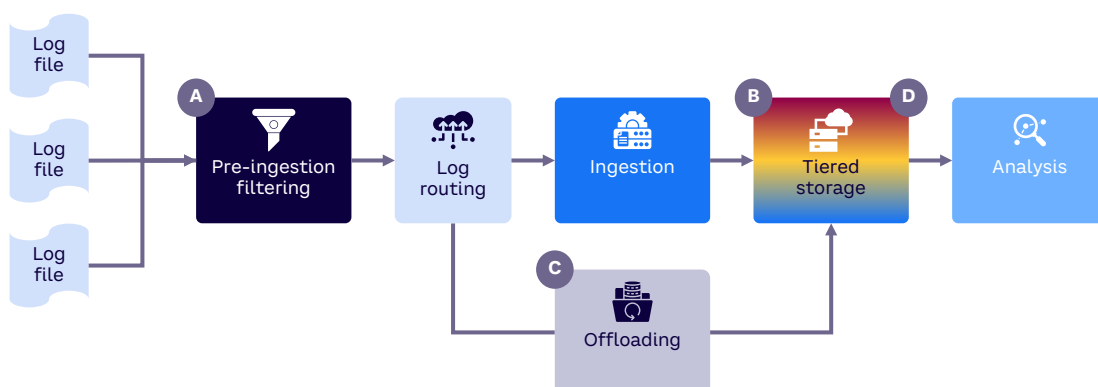
According to the big data era, collecting more data means gaining more value. By default, this means most organizations have adopted a “log everything” approach, with all logs ingested and stored, regardless of their potential future need or value.

Therefore, log volumes have grown, impacting resources and budgets (see Figure 3). While switching to a different logging solution can lower costs, the most powerful (and often overlooked) cost reduction tool available to organizations is improved governance, with a specific focus on cost discipline. Four key cost-reduction levers can be utilized within any logging governance framework to reduce total cost of ownership (TCO) and improve operational efficiency without reducing data value:

A. Pre-ingestion filtering. Removing superfluous fields and data points automatically reduces costs. Application teams and users should review their needs and decide which data is superfluous. Enforcement of this improved governance framework must include the requirement for regular reviews of log utilization to identify opportunities for pre-ingestion filtering.

- B. Tiered storage.** Not all log data needs to be easily accessible. While some data sets must be readily available for regular analysis, other data may only be required sporadically. Adopting tiered storage reduces storage costs by transferring underutilized data to cheaper “cold” storage locations. As part of data governance process, teams are again required to isolate data that is less likely to be queried.
- C. Pre-ingestion offloading.** Data identified as less critical or is seldom accessed can be placed immediately into colder storage before being ingested, avoiding the need for initial placement in more expensive hot storage.
- D. Retention policy optimization.** Retention policies govern how long post-ingestion data must be stored. Although some teams match retention periods to minimum compliance requirements (a “lean” state), many simply adopt the default settings of their logging solution provider. Data is stored longer than necessary, accruing unjustifiable costs. Optimizing retention periods must be a thoughtful process to avoid deleting vital data, but it can significantly reduce costs when implemented in a well-governed manner.

Figure 3. Data flow in log management



Source: Arthur D. Little

These levers require organizations to establish a larger governance framework, with all teams taking concrete steps to identify opportunities for filtering, offloading, or tiering, specific to each log type, source, and destination. All application teams should use clear guidelines to define which data matters and have associated KPIs in place to incentivize reduced ingestion volume and decrease the use of expensive “hot” storage.

2. Vendor selection: Strategic & inform governance

The logging solution landscape is enormous and wide-ranging; every vendor aims to distinguish themselves and unique abilities, such as wide capability sets, consistent user support, or cost-reducing creative storage methods. As the use of logging solutions has grown, many organizations have delegated vendor selection to individual teams to choose the best solution for their specific needs. However, over time this approach leads many organizations to buy multiple, disconnected logging solutions. Minimal enterprise-wide coordination of license management or vendor negotiation represents missed opportunities to reduce costs and complexity. While a multi-vendor strategy is not inherently problematic, an unplanned approach will fail to evaluate feature needs, volume requirements, and use cases for all teams.

As a consequence, opportunities to consolidate licenses or leverage purchasing power during vendor negotiations are overlooked. Additionally, an organization characterized by disconnected teams typically engages in minimal knowledge sharing. Teams develop their own logging solutions and best practices but do not share their knowledge widely. Reviewing how every team operates will uncover potential examples of excellence, to be shared among teams with similar use cases or solution needs, and improving overall efficiency and effectiveness.

3. Explore all available solution features

Logging solution providers are rapidly adding new features as they seek to differentiate themselves and provide additional value to users. However, these innovations remain underutilized or simply untapped by many organizations, as they are wary of changing how they operate and use their technology.

This lack of feature experimentation represents another missed opportunity, as teams may have access to new features that could provide significant cost savings or deeper insights with the proper training and implementation. Feature experimentation delivers value in the following areas:

- **Teams retaining their existing solution** should make deliberate efforts to test features not currently utilized to identify new use cases or efficiencies. Working with vendor support teams can discover opportunities for potentially relevant features and support teams can help implement these features effectively.
- **Teams exploring new solutions** should look beyond current needs and use cases and instead examine how requirements will change as data volumes grow and applications evolve. This planning will provide major dividends to teams looking to identify new features proactively, rather than playing catch-up.

4. Comprehensive view of TCO is critical

More organizations face cost constraints from increasing data volumes; detailing current TCO is a natural next step to provide a basis for future cost-reduction initiatives. These efforts to calculate TCO are valuable, but organizations often view total cost simplistically, focusing only on license, compute, and storage costs while often overlooking other equally important costs.

For example, maintenance and overall infrastructure-related costs comprise a non-negligible share of an organization's TCO and should be tracked in detail across teams to recognize high performers to share their lessons accordingly. When considering migration to a new solution, build an informed business case to justify the move that factors in detailed people-related costs. Overall, calculating and tracking TCO requires a diligent, granular, and multidimensional approach to ensure all investments or cost-reduction initiatives are backed by robust, plausible figures.

BELT TIGHTENING DOES NOT MEAN SACRIFICING DATA VALUE

Transitioning to a better-governed, more cost-aware enterprise-wide log management system does not necessarily imply sacrifice. While implementing the steps above will require an initial investment of time and effort from teams, these actions will give them notable longer-term financial and operational improvements when managing their logging environments.

More importantly, cutting costs and using well-governed systems does not undermine the need to extract deeper insights or ensure compliance — in fact, improving governance, reducing costs, and embracing new features will deliver major benefits in these areas.

A holistic, enterprise-wide strategy can help organizations increase long-term value by improving cybersecurity, supporting product and service innovation, promoting operational efficiency, and providing insights into customer needs and buying behavior. Smart, centralized log management enables companies to boost their cyber defenses by tackling collection, processing, and correlation challenges. Log analysis provides access to important metrics to help mitigate risks and initiate automated responses when needed. Insights about online user behavior enables greater customer-centricity; identifying and sharing existing knowledge accelerates innovation and overall competitiveness.

Case study — Delivering a holistic logging strategy to a major organization

A *Fortune* 100 company engaged ADL to help evaluate its data logging usage and solution architecture. The costs of the company's data logging solutions had been increasing drastically and becoming prohibitive, especially given the anticipated additions to logging requirements. ADL analyzed the organization's current usage of data logging solutions. It found that the company utilized a mix of software as a solution, private cloud, and on-premises solutions, making its logging ecosystem extremely complex, siloed, and lacking observability.

With ADL's help, the organization established a governance framework for log management that focused on quantifiable ROI. Establishing clear rules allowed logs to be pre-filtered while routing noncritical logs to cheaper storage. Concurrently, ADL reviewed the organization's current and future requirements and conducted a request for proposal to evaluate logging solutions and developed a comprehensive TCO model and business case which identified one solution for the entire organization. We used our established logging governance framework to help establish an internal center of excellence for logging and data platforms and continued assisting the client as it worked to achieve its medium- and long-term strategic goals.

CONCLUSION

PRACTICAL STEPS TO DISCIPLINED LOG MANAGEMENT

ORGANIZATIONS CAN NO LONGER DELAY LOG MANAGEMENT IMPROVEMENTS

Data volumes are growing rapidly, and organizations can no longer delay log management improvements. Take the following four practical actions to make meaningful, rapid improvements:

- 1 Conduct a detailed audit** of how teams utilize existing data. Develop maps of associated original log files, utilization, and retention paths to uncover cost-reduction opportunities.
- 2 Review existing logging solutions** and document current use cases and future needs. Identify opportunities for license consolidation.
- 3 Define criteria for outstanding logging centers** and schedule regular knowledge-sharing sessions and similar learning opportunities to make excellence contagious.
- 4 Model and track TCO** by cataloging all existing, relevant costs across teams. Identify potential savings from new solution vendors under consideration by incorporating their cost structure into existing ingestion and storage models.



Arthur D. Little has been at the forefront of innovation since 1886. We are an acknowledged thought leader in linking strategy, innovation and transformation in technology-intensive and converging industries. We navigate our clients through changing business ecosystems to uncover new growth opportunities. We enable our clients to build innovation capabilities and transform their organizations.

Our consultants have strong practical industry experience combined with excellent knowledge of key trends and dynamics. ADL is present in the most important business centers around the world. We are proud to serve most of the Fortune 1000 companies, in addition to other leading firms and public sector organizations.

For further information, please visit www.adlittle.com.

Copyright © Arthur D. Little - 2023. All rights reserved.