CHAOSSEARCH

THE BUSINESS CASE FOR

Switching from ELK to ChaosSearch

Harnessing log data for analytics has never been more important. IT Operations, DevOps, and SecOps teams need to gain the insights within their organizations' ever-growing volumes of log data to become more robust, resilient, and responsive.

Contents

INTRODUCTION

Harnessing log data for analytics has never been more important. IT Operations, DevOps, and SecOps teams need to gain the insights within their organizations' ever-growing volumes of log data to become more robust, resilient, and responsive.

However, the top analytics platforms in production today were introduced over 10 years ago and were not built for today's cloud scale environment. The amount of data that organizations must process and manage today far exceeds the intended design center of these legacy systems. Attempts to scale with legacy architecture can mean adding significant cost and complexity, with multiple disparate clusters and complicated database sharding. Indeed, these challenges are inherent in the most commonly deployed log management solution today: Elastic Stack, more commonly known as the ELK (Elasticsearch, Logstash, Kibana) stack.

As ELK stack environments scale up to incorporate more data sources and look to retain data beyond a few days, the overall deployment quickly becomes complex, which in turn, causes costs to rise rapidly. In order to move beyond a very small environment, the ELK stack's highly distributed architecture necessitates that data be partitioned and stored across numerous shards, and that separate servers be deployed in which each is responsible for its portion of the data. While it is easy to deploy and begin using the ELK stack with a low initial investment, most organizations quickly face ELK stack cluster sprawl—they're managing, and paying for, significant compute and storage resources.

The heightened importance of log data analytics, coupled with the inability to economically scale an ELK stack environment, is driving organizations to consider switching from ELK to a new platform for their log management and analysis needs.

Many organizations that value log analytics but are constrained by the costs and/or complexities of the ELK stack are migrating to the ChaosSearch Data Lake Platform. ChaosSearch overcomes the many inherent scalability constraints of the ELK stack, and delivers data analytics at scale while slashing the total cost of ownership (TCO).

This white paper will help you build the business case for moving from an ELK stack environment to ChaosSearch. The paper explains all of the cost drivers behind both the ELK stack and ChaosSearch, and demonstrates how to roll up the underlying costs into a TCO for each solution so you can develop your business case, using data from your own environment.

The calculations presented are based on ChaosSearch's TCO Tool, an interactive calculator which allows the user to input the key variables of a customer scenario and calculate the total cost of an ELK stack compared to ChaosSearch for the given scenario[1]. The tool uses real world data and published pricing information to calculate the detailed underlying costs of both solutions.

[1] For more information on the ChaosSearch TCO Tool, contact info@chaossearch.com.

IMPORTANCE OF LOG DATA ANALYTICS AT SCALE

Log data contains the insights an organization needs to run more effectively, and more securely. The sum of an organization's log data provides the details of the entire IT environment in real time, or at any point in time in history. This includes details on machine and network traffic, user access, changes to applications and services, and countless other pointers used to monitor the health and security status of the IT landscape.

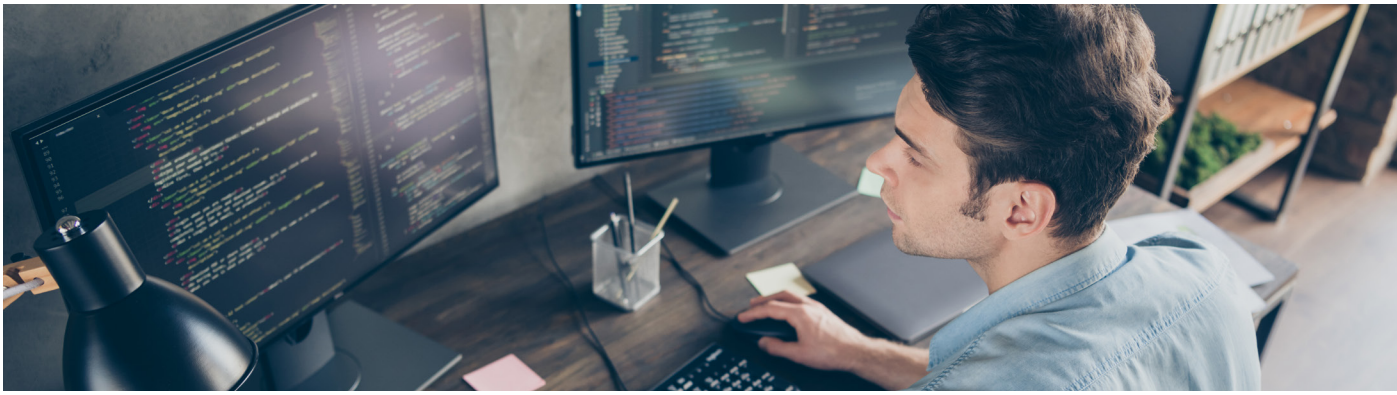
One of the first uses of machine data log analytics remains one of the most important—**IT monitoring**. Organizations that monitor and analyze logs on a continual basis can proactively ensure their systems and applications are performant during peak usage. Log analysis can help prevent disruptions, optimize operational performance, reduce required user support, help analysts better understand customer usage, and ultimately improve the bottom line.

Organizations that monitor and analyze logs on a continual basis can proactively ensure their systems and applications are performant during peak usage.

The most prevalent use of log analytics today is to support **cybersecurity**. Log analytics provide details on extreme traffic, unauthorized access, suspicious changes in activity, and many other pointers used to identify potential threats. These indicators sit at the heart of many core SecOps activities, including detection and alerting, forensic investigations, threat hunting, insider threat detection, and distributed denial of service (DDoS) attack prevention.

The most important factor to success using log analytics is having access to the right data. This includes collecting data from all relevant sources, allowing the ITOps and SecOps analysts to run sophisticated reports and queries that include correlations from a variety of data streams. It also includes accessing historical data, which is vital for anomaly detection, forensic investigations and other analyses that rely on comparing metrics over time.

To meet the needs described above, today's IT environments depend on access to a scalable, centralized log data management system. Given how critical many of the use cases above are in the daily operations of many businesses, they simply cannot afford to limit log data access. But because of the rigidity of their architecture, legacy solutions like the ELK stack require organizations to either pay exorbitant costs to scale the environment, or make painful trade-offs by limiting the amount of data ingested daily and stored over time.



ELK STACK LIMITATIONS AND MOTIVATIONS TO SWITCH

ELK Stack Overview

The Elastic Stack, commonly referred to as the ELK stack, is the most widely deployed and well-known log analytics solution on the market today. There are many reasons behind the near-ubiquity of the ELK stack today. It was one of the first solutions in the market that comprehensively addressed organizations' requirements for log management and analysis. First introduced in 2012, the ELK stack came into existence when three previously independent open source projects joined forces within a single commercial solution, distributed by a single company: Elastic. As the name "ELK" implies, the solution is a combination of three complementary open source products:

a visualization tool for log search analytics

the underlying search and analytics database

a log ingestion and processing pipeline

A few years later, Elastic acquired another company and began distributing a fourth product, Beats, which is often deployed as part of the ELK stack environment. Beats is a set of agents that collect and send data to Logstash.

Adoption grew rapidly. The ELK stack was one of the first solutions to offer centralized log management functionality, and its open-sourced business model made deployment of a small environment relatively inexpensive. As an open-sourced project, the solution also benefited from a large community of contributors. This has allowed the ELK stack to remain feature rich, and offer powerful functionality that is on par with leading commercially distributed competitive solutions.

In many ways, the ELK stack played a key role in fostering digital transformation, which most companies began pursuing in the early 2010s, around the time ELK was launched. The ELK stack helped to unlock the value of data by making it more readily discoverable and shareable, and allowed companies to integrate data into their core operations.

Today, the ELK stack continues to be marketed as a free, easily accessible, and simple-to-install toolkit, and continues to see tens of thousands of downloads per month. According to Elastic, the company had 11,300 paying customers as of April 2021[2].

[2] Elastic-Annual Report Fiscal Year 2020

ELK Challenges

While the ELK stack has delivered considerable benefits to customers, it has also created some new risks. A byproduct of digital transformation is that companies have come to rely more heavily than ever before on accurate and timely data to drive most of their mission critical operations. This has created a heavy dependency on data accessibility—if the data is not accessible, or not complete, operations will be negatively impacted.

40%

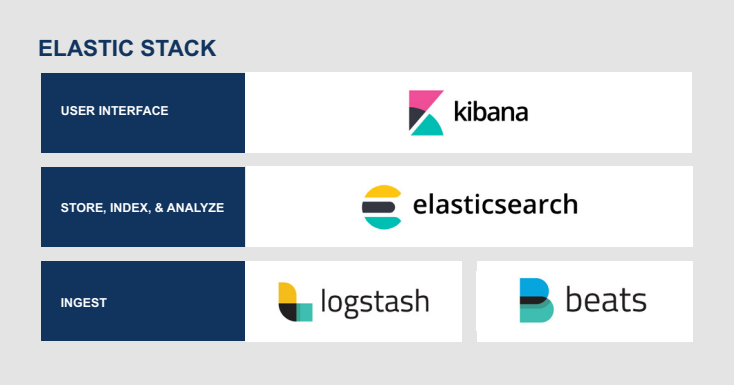
During the past 10 years, companies have seen an average annual data growth rate of 40%.

Today's ELK stack customers are too often realizing this challenge of data inaccessibility. The core problem is that exploding data growth rates have intersected with inherent limits of the ELK stack's ability to scale. During the past 10 years, companies have seen an average annual data growth rate of 40 percent. Every new business application and new IT solution that's introduced generates a new source of logs to capture and analyze. As organizations grow in size and scope, the log data volume growth accelerates.

When ingesting data, a typical ELK stack deployment requires a complex data pipeline consisting of parsing or schema changes which create a data stream bottleneck, limiting the ingest rate. Further, scalability is constrained by the allowable size of its underlying database. When pushed beyond its intended usage, a legacy ELK stack deployment can become unstable and experience outages. And, as discussed, these outages can have a severe impact on the business.

The only way to accommodate growth is complex and cumbersome—the IT Operations team must break up the database into many small shards and then set up and maintain multiple disparate clusters. To manage a large and growing ELK stack environment while mitigating instability challenges requires costly investments—to adding processing power, storage capacity, and more staff to manage it all. The result? A highly complex, expensive, and hard to manage environment.

New Alternatives Make the Case Compelling



The increasingly untenable cost of the ELK stack is driving organizations to seek alternatives. And with the availability of new scalable data platforms, like ChaosSearch, the case to switch away from the ELK stack has become compelling.

The ChaosSearch Data Lake Platform does what the ELK stack cannot—it allows customers to collect and analyze ALL of their data without tradeoffs, with virtually unlimited scalability, and with massive cost savings. ChaosSearch customers typically see **50 to 80 percent annual cost savings** compared to their legacy solution, as this paper exemplifies.

For example, by moving from ELK to ChaosSearch, a mid-size enterprise environment that processes 5 TB of log data per day reduces their 3-year TCO from $18.8 million to $6.9 million, avoiding close to **$12 million in costs over a 3-year horizon.**

How could this be possible? ELK is a "closed system" that ingests data into its own internal database, transforming the data on the way in, and maintaining custody of the data once ingested. In contrast, the ChaosSearch data lake platform simply connects to and indexes data within a customer's existing cloud data storage. With read-only access to the customer data, ChaosSearch builds a separate index without manipulating or taking custody of the underlying original data.

This approach solves the two primary inhibitors of the ELK stack that prevent it from scaling efficiently. On ingest, ChaosSearch removes bottlenecks, as the data can stream directly into a customer's cloud storage in its native format. And because it avoids the burden of data custody, ChaosSearch has no internal database size constraint. ChaosSearch simply leverages the performance, scale and economics of the public cloud. This is the key that allows ChaosSearch to deliver unlimited scalability, industry-leading resiliency, and massive time and cost savings.

Importantly, once indexed, the ChaosSearch platform allows customers to conduct search and analytics on the data using existing tools in use today, including Kibana, leveraging the open APIs of these tools. In switching from ELK to ChaosSearch, customers can continue using Kibana seamlessly—running the same reports and queries that they do today. This allows customers to take advantage of the massive scalability and cost savings of ChaosSearch, without incurring significant switching costs.

5TB
Per Day

$11.9M
Savings

3YEARS
TCO

Low Switching Costs

The business case to switch to a new solution must include not only the cost advantages of the new solution over the legacy system, but also switching costs. As IT professionals know, any change to an environment—no matter how small—can be disruptive. Given that the most important role of IT is to keep existing systems online and data accessible, IT organizations are necessarily reticent to pursue disruptive projects. Often the burden of switching prevents IT from adopting new technologies, even when the advantages are clear and compelling.

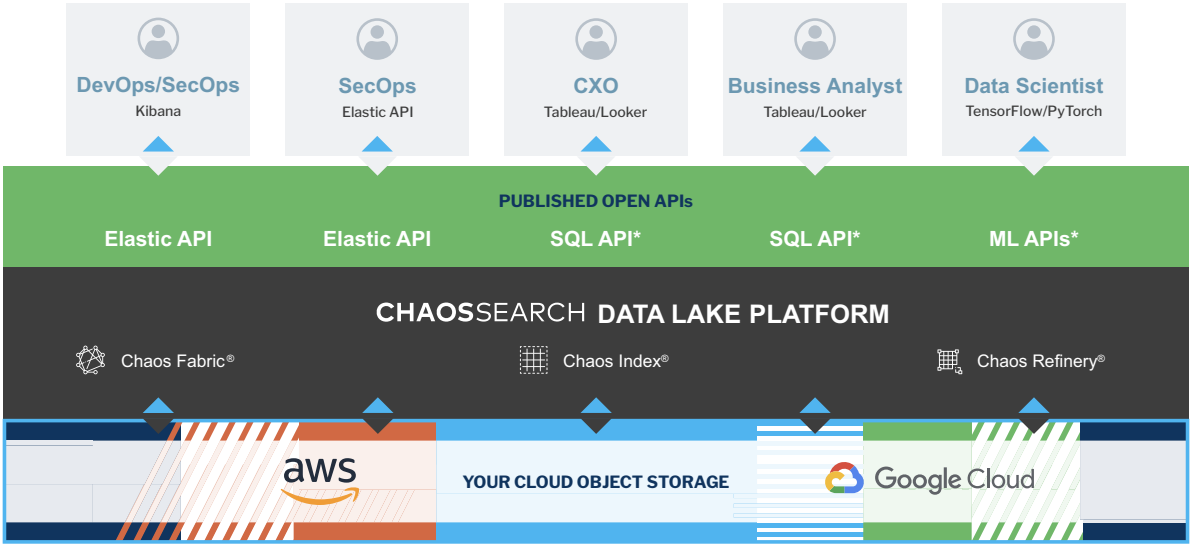
Thus, when considering a move to a new solution for log analytics—a mission critical system that powers many core IT and Security Operations—the actual project of migrating to the new solution must be well understood.

Fortunately, the move from ELK to ChaosSearch is seamless. Because ChaosSearch is tightly integrated with Kibana and supports the Elasticsearch API, customers can move their operations over to ChaosSearch quickly and easily. In doing so, customers maintain the same dashboards, visualizations and pre-staged queries that they use in their ELK environment, simply by exporting them from ELK and importing them into the ChaosSearch deployment.

Typical switching costs are avoided:

- **Very small learning curve to adopt and use**
- **Little to no training required for experienced Kibana / Elasticsearch users**
- **No manual conversions of existing queries to a new language, or recreations of existing reports and visualizations in a new system**

The result of the seamless transition is that customers can switch quickly, avoiding significant burden of staff time to manage a complex project, and avoiding a costly long-term transition project that entails maintaining (and paying for) two parallel systems at once.



ACCOUNTING FOR THE FULL COST OF THE ELK STACK

In this section, we'll provide details on all of the cost factors that go into a total cost of ownership calculation for an ELK stack deployment.

To do so, we'll assess two broad categories: **Infrastructure** and **Ongoing Operations**. The following analysis leverages Amazon's Elasticsearch Sizing guidelines[3].

Infrastructure

1. Primary Infrastructure Cost Drivers

The first cost to consider when calculating TCO is the infrastructure you need to run the ELK stack. Since the infrastructure varies based on the amount of data you expect to generate, estimating the daily log data volume accurately is very important. Over-provisioning means unnecessary costs. Conversely, under-provisioning leads to lost log data, causing the loss of critical insights that impact business performance.

Here are the key factors to consider when estimating capacity needs:

- **Baseline log data volume and retention**
- **Daily log volume generated by your applications, systems, and networks**
- **Anticipated log volume growth rate**
- **Data retention needs for both indexing and archiving**
- **Whether/when spikes in log volume typically occur**

In addition to planning for capacity, it's important understand additional factors that drive the configuration requirements:

- **The number of concurrent users and concurrent searches**
- **Fault-tolerance and redundancy requirements; a 1:1 ratio between server and a replicated backup is generally considered best practice**
- **The above factors drive the underlying infrastructure costs. The growth in users and underlying data will drive increases in the required infrastructure**

Another key consideration is that Elasticsearch and Logstash require significant capacity, availability, and redundancy. Because of this, Logstash and Elasticsearch should be run on different and multiple servers. Moreover, Kibana requires high availability so that users can reliably interact with and perform analysis on log data. Thus, to ensure that you don't slow performance for the user, it's also best to run Kibana on dedicated servers.

[3] Sizing Amazon ES Domains

2. Infrastructure Configuration Details

The above factors establish the high-level sizing requirements for a given deployment. In addition, there are a number of detailed infrastructure considerations that go into deploying an ELK stack into production. Production-ready deployments require the team to:

- **Configure the stack to ingest and parse logs from all logging components.** This includes maintaining all of the configurations needed to accommodate the variety of logging frameworks, data formats, and log sources. Many organizations must account for hundreds of different data configurations.
- **Build a resilient data pipeline ensuring there is no log data loss if the system generates events faster than Elasticsearch can index them.** This typically requires placing a buffer in front of Logstash that acts as the entry point for log events. Doing so will enable you to accumulate the data until it can be pushed to Elasticsearch. Some organizations use Apache Kafka, Redis, or RabbitMQ for buffering logs. However, this requires you to host and maintain yet another piece of software.
- **Handle mapping exceptions.** To ensure that Elasticsearch indexes documents instead of returning failure messages and dropping logs that don't fit into the automatically generated mapping, you have to keep log formats consistent and continually monitor Elasticsearch exceptions.
- **Ensure log data consistency.** Applying relevant parsing abilities to Logstash is critical to ensuring you have correct fields for Elasticsearch and Kibana. But it is also challenging. It's easy to make mistakes using Logstash, so you need to devote time to testing all log configurations before your ELK stack goes into production.
- **Implement monitoring and alerting capabilities that notify you of performance and potential security issues.** You must research and evaluate the many open source and commercial solutions on the market, and then dedicate resources to implementing and integrating them into your ELK stack.

Ongoing Operations

Operating Cost Drivers

Your organization's work on the ELK stack doesn't end after it's rolled out in production. As log data volumes increase, more resources are consumed, and new complexities and issues arise. The distributed shard-based architecture can create issues of consistency and durability due to the complex dependencies and failure modes across shards. Additionally, it's not uncommon to run into problems with Logstash not running or not shipping data, and Kibana not fetching mappings or not connecting with Elasticsearch. The myriad of potential problems requires ELK stack experts to be on hand to respond to these issues and perform significant day-to-day maintenance just to ensure the system is working properly. The number of people required to handle it will grow as your stack expands, and potentially cause you to commandeer engineers from other priorities.

The above factors demonstrate the need for active daily management of the ELK stack environment. The human resources required are directly correlated with the size of the complexity of the environment.

ELK stack management activities include:

- **Maintaining the infrastructure and planning capacity increases.** Staff must continually monitor performance and capacity, and plan ahead for increases in order to avoid failures. Adding or removing servers from an ELK cluster, for instance when data volumes grow, is a non-trivial task and may require the process of rebalancing shard allocations.
- **Reindexing outdated indices so that you stave off potential failures and log data losses.** Logs are dynamic. Their formats change over time and require configuration adjustments. In addition, the number of indices handled by Elasticsearch impacts performance, so staff needs to continually monitor them, removing or freezing old and unused indices.
- **Monitoring cluster health and responding to failures.** Staff needs to track information about the status of the cluster, the number of nodes, and the counts of active shards. The team must also monitor counts for relocating shards, initializing shards, and unassigned shards. All of this is needed to tune the cluster for better performance.
- **Handling software upgrades.** Upgrading an ELK stack can be a large undertaking, depending on the size and complexity of its deployment. Even though new versions are released on a regular basis, the team must thoroughly research what changes a given upgrade contains for each ELK component before deciding whether to expend the effort required to implement them. To ensure that no data is lost during upgrades, the team must run tests in a non-production environment first, and validate all changes run smoothly without negative impacts. When upgrading Logstash, compatibility between it and the Elasticsearch version running must be verified. When upgrading Kibana, it's important to note that plugins often break, and visualizations sometimes require total rewrites. Backup your objects first and test the Kibana upgrade process before rolling it out, making sure to plan for any reconfiguration or rewrites that might be required.

Operational Costs in the TCO Calculation

Staffing

Staffing needs can vary significantly from one environment to another, depending on the factors above. A good rule of thumb for an average environment that rolls up all of the underlying operating costs into a single number is to assign one full-time employee (FTE) for every 5 TB of log data ingested daily. The FTE must be an expert in ELK stack and know how to manage the performance and indices under various production workloads. In the following analysis we assume an annual cost per FTE of $130,000.

Support

Elasticsearch may be free, but support is not. Many enterprises will also see the value in a support plan from Elastic. Elastic offers Basic, Gold, Platinum, and Enterprise subscription plans, which include SLA-based support and dedicated support contacts. If you want 24/7 coverage, then you must factor in this cost as well. Remember, pricing varies based on the number of nodes you have. So if your system is continuously growing, then your cost of support will grow too.

Training

Elastic has produced many instructional videos on how to use and manage ELK stack. The company also offers in-person and virtual classroom training for both administrators and end users. Such offerings may cost you approximately $5,000 per participant, per training. You can choose to save money by developing and maintaining your own training materials in-house and conducting internal training sessions for the rest of the team, but that's another task that distracts your ELK experts from managing your production-grade system.

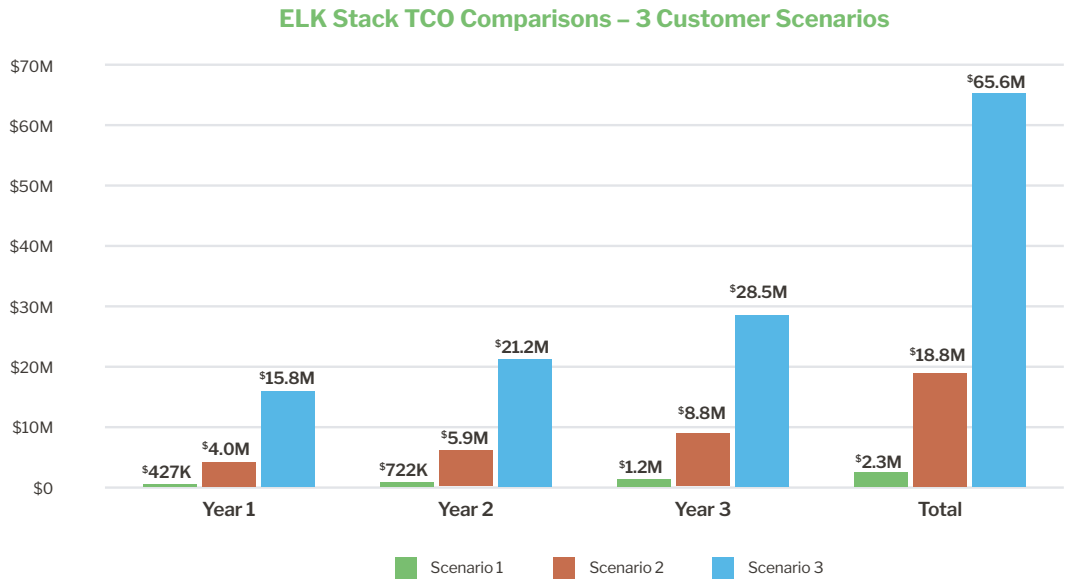
CALCULATING THE DETAILED ELK STACK TCO

To demonstrate how the costs described in the previous section form the ELK stack total cost of ownership, we'll quantify the full three-year TCO of three different representative customer scenarios.

The table below summarizes the three customer scenarios.

	SCENARIO 1 SMALL	SCENARIO 2 MEDIUM	SCENARIO 3 LARGE
Year 1 Daily Ingest Volume	500GB	5TB	20TB
Annual Data Growth Rate	75%	50%	35%
Active Data Retention	60 Days	60 Days	60 Days
FTE per TB daily ingest	.2	.2	.2
Replication Factor	1	1	1
Availability Zones	2	2	2
Elastic Support Level	Gold	Gold	Gold

The chart below shows the annual costs and the 3-year TCO for each scenario.



The summary table below provides the details behind the overall TCO, showing the breakdown by cost category for each of the three scenarios. This helps demonstrate the high cost of infrastructure compared to the other categories, and shows how costs dramatically rise based on the amount of data ingested daily. Importantly, the table demonstrates that even a small environment that starts with 500 GB of data per day reaches a 3-year TCO of $2.4 million.

ELK Stack 3-Year TCO for 3 Customer Scenarios

	SCENARIO 1	SCENARIO 2	SCENARIO 3
AWS Compute and Storage	$2,052,894	$16,676,895	$58,178,381
Operations Staffing	$94,313	$636,250	$2,188,450
Elastic Software Support	$220,000	$1,528,000	$5,256,000
Total 3-year TCO	$2,367,206	$18,841,145	$65,622,831

To allow you to begin putting together your own TCO calculations, the tables below provide the detailed year-by-year cost breakdown for each customer scenario. These details highlight not only the high cost of the infrastructure overall, but also how dramatically infrastructure costs rise year-over-year based on the data growth rate.

Scenario 1

	YEAR 1	YEAR 2	YEAR 3	TOTAL 3-YEAR TCO
AWS Compute and Storage	$361,404	$627,095	$1,064,394	$2,052,894
Operations Staffing	$21,750	$27,750	$44,813	$94,313
Elastic Software Support	$44,000	$68,000	$108,000	$220,000
TOTAL	$427,154	$722,845	$1,217,207	$2,367,206

Scenario 2

	YEAR 1	YEAR 2	YEAR 3	TOTAL 3-YEAR TCO
AWS Compute and Storage	$3,576,542	$5,282,009	$7,818,344	$16,676,895
Operations Staffing	$138,750	$200,000	$297,500	$636,250
Elastic Software Support	$328,000	$484,000	$716,000	$1,528,000
TOTAL	$4,043,292	$5,966,009	$8,831,844	$18,841,145

Scenario 3

	YEAR 1	YEAR 2	YEAR 3	TOTAL 3-YEAR TCO
AWS Compute and Storage	$14,057,743	$18,824,305	$25,296,333	$58,178,381
Operations Staffing	$528,750	$707,000	$952,700	$2,188,450
Elastic Software Support	$1,264,000	$1,700,000	$2,292,000	$5,256,000
TOTAL	$15,850,493	$21,231,305	$28,541,033	$65,622,831

CHAOSSEARCH COST ADVANTAGES

This section demonstrates the ChaosSearch TCO for each of the three customer scenarios, keeping all variables the same in order to allow for an apples-to-apples comparison.

Cost Savings Explained

ChaosSearch delivers massive TCO savings in two primary ways. First, the simplified architectural approach that ChaosSearch takes results in dramatically reduced infrastructure requirements. With ChaosSearch, customers need only pay for their cloud object storage environment, and can eliminate all spending on compute and block storage infrastructure associated with the ELK stack environment. Secondly, ChaosSearch is delivered to customers as a managed service, with a single monthly fee based on the daily ingest rate. This SaaS approach reduces the amount of customer operations personnel required to operate the environment down to a fraction of one FTE.

Cost Savings Quantified

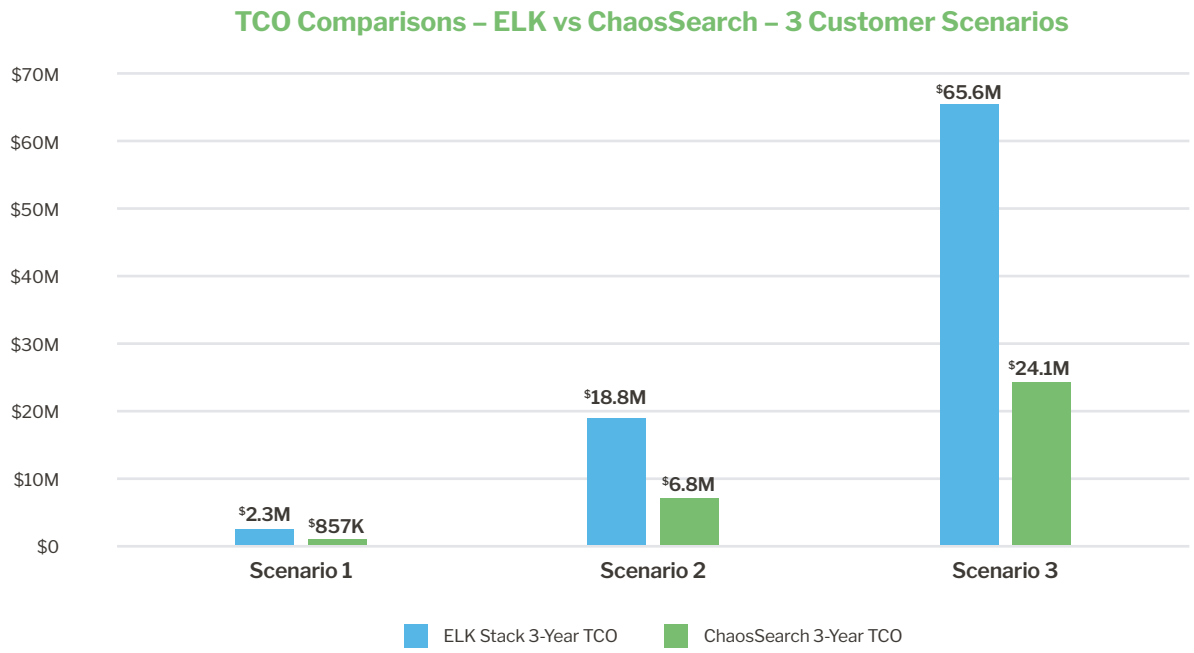
Quantifying the 3-year cost of ChaosSearch for each of the three customer scenarios demonstrates how ChaosSearch delivers significant savings compared to the ELK stack across all three. The summary table below shows the breakdown by cost category when using ChaosSearch in each of the three scenarios. As the table demonstrates, most of the cost of ChaosSearch is bundled into a single fee for the managed service, which bundles the cost of the software, infrastructure and management. Beyond that, customers need only pay a small amount for their cloud object storage environment, and for some operations staff time, which equates to a little over 1/10th of an FTE or about 4 hours per week.

ChaosSearch 3-Year TCO for 3 Customer Scenarios

	SCENARIO 1	SCENARIO 2	SCENARIO 3
Managed Service	$837,000	$6,840,000	$24,033,600
AWS Compute and Storage	$4,813	$39,330	$138,193
Operations Staffing	$15,500	$15,500	$15,500
Elastic Software Support	Included	Included	Included
TOTAL 3-YEAR TCO	$857,313	$6,894,830	$24,187,293

TCO COMPARISON OF 3 CUSTOMER SCENARIOS

Comparing the ChaosSearch TCO directly to the ELK stack TCO for each of the three customer scenarios demonstrates the tremendous cost savings in each.



The summary table below shows the 3-year savings that ChaosSearch delivers over the ELK stack for each scenario.

	SCENARIO 1	SCENARIO 2	SCENARIO 3
ELK Stack 3-Year TCO	$2,367,206	$18,841,145	$65,622,831
ChaosSearch 3-Year TCO	$857,313	$6,894,830	$24,187,293
ChaosSearch Cost Savings vs ELK	$1,509,894	$11,946,315	$41,435,537
% SAVED VS ELK	64%	63%	63%

GOING BEYOND APPLES-TO-APPLES SAVINGS

The section above provides apples-to-apples comparisons of ChaosSearch to ELK in which all variables of a given customer deployment are kept exactly the same. This analysis is useful because it allows us to isolate the different TCO outcomes for ChaosSearch vs. ELK when all else is held constant. However, in the real world, customers often intentionally change variables when moving to ChaosSearch. For example, a significant driver of customer adoption of ChaosSearch is to increase the data retention period far beyond what they have in their existing ELK stack environment.

ChaosSearch's architecture not only enables massive scalability, but it also eliminates the high cost of data retention that the ELK stack imposes. Whereas increasing retention while staying within an ELK environment drives the overall cost up dramatically, with ChaosSearch increasing retention has a very minor impact on the total cost.

To exemplify this, let's take Scenario 2 which has a starting daily ingest volume of 5 TB and a retention rate of 60 days. Keeping all parameters the same shows that ChaosSearch delivers almost $12 million in savings over three years.

Scenario 2 – 5 TB/Day, 60-Day Retention

	SCENARIO 2
ELK Stack 3-Year TCO, 60-Day Retention	$18,841,145
ChaosSearch 3-Year TCO, 60-Day Retention	$6,894,830
ChaosSearch Cost Savings vs ELK	$11,946,315
% SAVED VS. ELK	63%

What if the customer wants to change retention to 180 days? As the table below shows, with ChaosSearch, this only adds a total of $78,660 to the 3-year TCO, increasing it from $6,894,830 to $6,973,490, a mere 1.1% increase. However, increasing retention to 180 days within the ELK stack environment increases the TCO from $18.8 million to $54.7 million—nearly 3X the original TCO!

Scenario 2 – 5 TB/Day, 180-Day Retention

	SCENARIO 2
ELK Stack 3-Year TCO, 180-Day Retention	$54,753,546
ChaosSearch 3-Year TCO, 180-Day Retention	$6,973,490
ChaosSearch Cost Savings vs ELK	$47,780,056
% SAVED VS. ELK	87%

Whereas the cost of increasing the retention period to six months is prohibitive in the ELK stack environment, with ChaosSearch, the incremental cost is negligible. This exemplifies how customers can not only achieve dramatic cost savings, but can also significantly improve their IT and Security Operations by significantly increasing data access and retention periods.

Thus, the business case to switch from ELK to ChaosSearch should include not only a cost comparison, but also an assessment of the ChaosSearch benefits that the customer can take advantage of, and the impact that ChaosSearch will have on the operations that rely on log data analytics.

> “Before we partnered with ChaosSearch in late 2020, our SRE teams used to struggle with managing the vast amount of logs it takes to support millions of users in real time in a consistent manner across all our product lines. With ChaosSearch, we are able to use a singular solution for our various logs without the hassle of managing the logging tools as well.”

Joel Snook
Director of DevOps Engineering at Blackboard

ASSESSING THE SWITCHING COSTS

The business case to switch to ChaosSearch should include the cost of the migration itself.

As described in section 3 of this paper, the migration from ELK to ChaosSearch is seamless, as customers can continue to run the same Elasticsearch queries and maintain the same Kibana visualizations that they have today. This ease-of-adoption allows customers to safely move their operations to ChaosSearch during a brief transition period, thereby limiting the costs of the effort to make the move.

Most customers pursue a transition period during which ChaosSearch is deployed in parallel to the ELK stack environment, allowing an orderly process of testing and moving workloads to ChaosSearch over time. The duration of the transition is based on the number and variety of the workloads to be migrated. While some customers pursue rapid transitions, moving all workloads to ChaosSearch within a two-week period, the average transition time is 30 to 60 days.

Once the length of the transition period is estimated, quantifying the switching costs is straightforward. The cost can be calculated based on the daily cost of operations for the ELK environment, which can be pulled directly out of the TCO Tool.

Let's look again at Scenario 2.

	YEAR 1	YEAR 2	YEAR 3	TOTAL
ELK Stack	$4,043,292	$5,966,009	$8,831,844	$18,841,145
ChaosSearch	$1,453,780	$2,177,420	$3,263,630	$6,894,830

In Year 1, the customer is processing 5 TB of data per day. In the ELK stack environment, the total cost for Year 1 is $4,043,292 per year, or $11,077 per day.

If the customer plans for a 45-day transition period, the cost to continue running the ELK environment in parallel to the ChaosSearch one will be: 45 X $11,077, or $498,465. In the context of a business case, this $498,465 transition cost can be considered the initial investment required that allows you to realize the overall TCO savings of the three-year period.

Note that this calculation is a simplified and conservative example. In many cases, the actual transition cost will be much lower. Costs typically decline each day during the transition period as workloads shift and the infrastructure costs decline accordingly. Moreover, customers often take advantage of the terms of their contracts with their cloud provider, which typically include consumption credits, incentives and discounts. Thus, the actual transition cost is typically a fraction of the simple calculation of cost per day multiplied by the number of anticipated days of transition.

PUTTING IT ALL TOGETHER—BUILDING THE BUSINESS CASE FOR SWITCHING FROM ELK TO CHAOSSEARCH

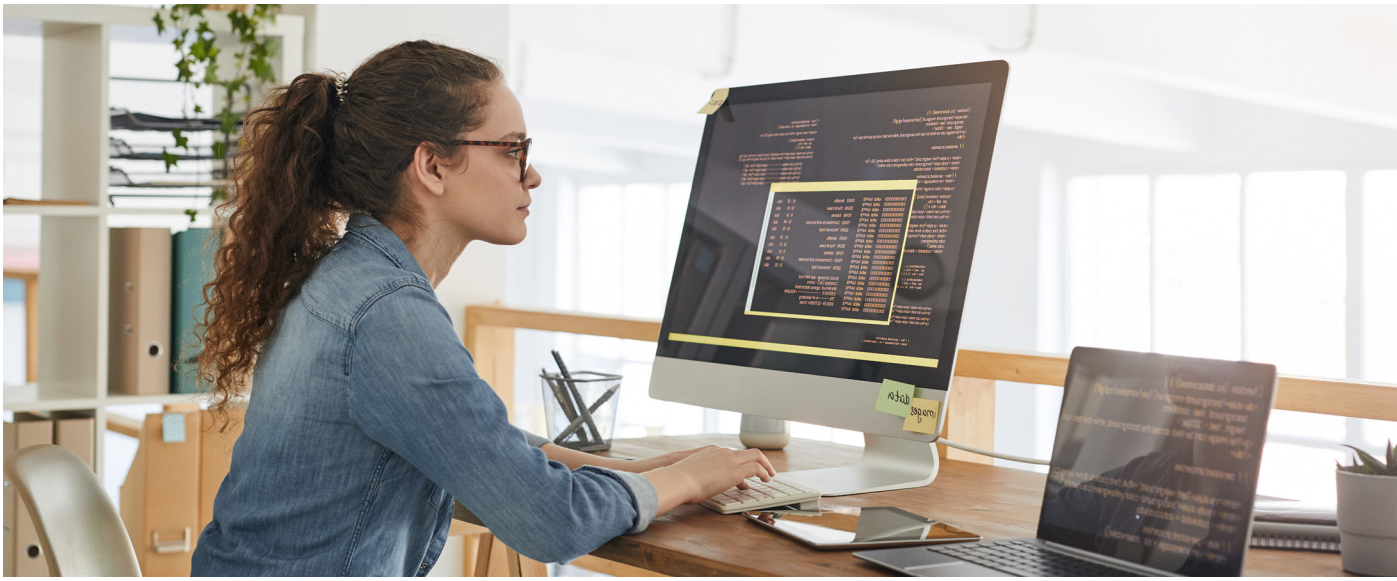
Unlike many technology purchases in which projected benefits are difficult to quantify and often rely on shaky assumptions, the business case for switching from ELK to ChaosSearch is easy to calculate, and rests on real world, verifiable data.

Projecting the financial benefits of migrating from the ELK stack to ChaosSearch forms the basis of the business case. The prior sections of the paper have demonstrated how to calculate the TCO of both solutions, and how to account for the transition period.

SAMPLE 3-YEAR TCO COMPARISON	
3-Year ELK Stack TCO	$18,800,000
3-Year ChaosSearch TCO	$6,900,000
45-Day Transition Cost	$498,000
3-Year ROI	$11,402,000
3-Year Rate of Return	165%

Here's how to put together your own business case:

1. **Start with a detailed three-year TCO of the "status quo"** -- that is, the total projected costs of continuing to operate your existing ELK stack environment. This will include the total projected spending on personnel, infrastructure, and software support as shown earlier in the paper.
2. **Next, calculate the projected TCO for ChaosSearch in the same environment.**
3. **Then, calculate the cost of maintaining the ELK stack environment during the transition period.**
4. **Calculate the delta between the cost of the ELK stack and the cost of ChaosSearch, including the transition cost, to show the total return on investment (ROI).** The rate of return is simply the net gains delivered by ChaosSearch, divided by the cost of deploying and running ChaosSearch. Using our Scenario 2 customer example demonstrates this, and shows a 165% Rate of Return in making the move from ELK to ChaosSearch.



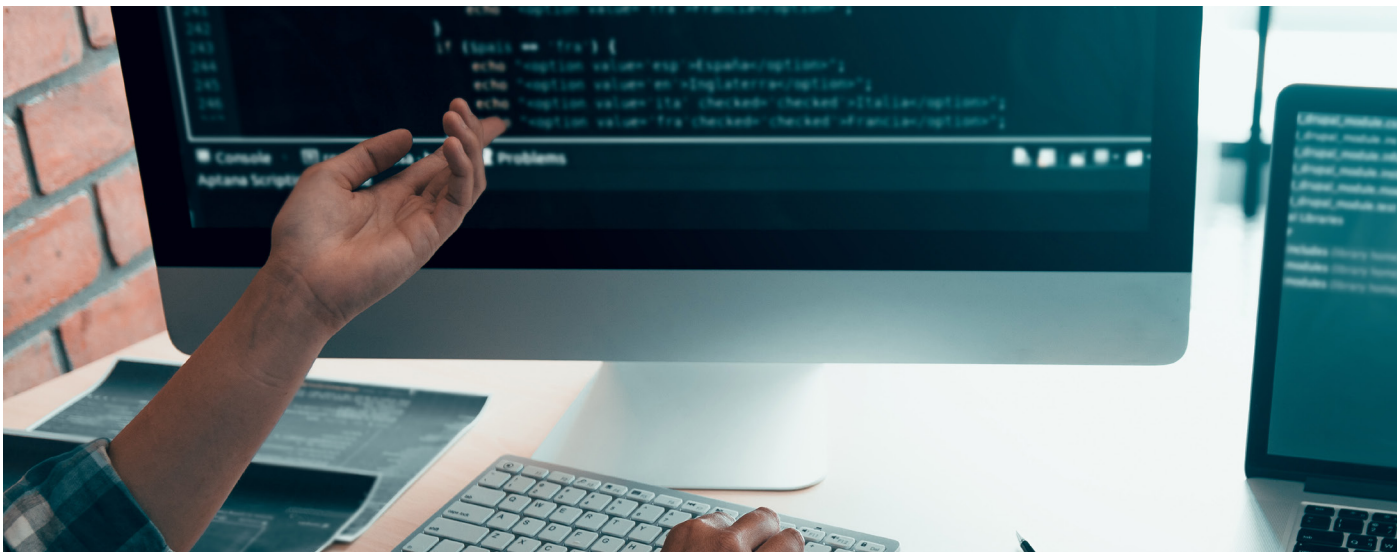
CONCLUSIONS

Motivations to move off of the ELK stack run very high as customers struggle with its underlying scalability limits and the overall costs that are required to maintain their environments.

ChaosSearch provides the ideal replacement for the ELK stack: It delivers massive reductions in cost and complexity, solves the scalability problem, and enables a seamless transition given that it enables Kibana and the Elasticsearch query language.

The potential benefits of migrating from an ELK stack environment to ChaosSearch are extensive. The examples in this guide demonstrate that the business case for moving to ChaosSearch can be built on a solid, quantitative analysis that demonstrates the cost savings attainable over a three-year period, and results in a very high ROI. Furthermore, the business case should include the significant—albeit harder to quantify—benefits of running ChaosSearch, including providing access to a much larger amount of data for analysis, enabling much longer data retention periods, and delivering a much higher level of resiliency than customers typically experience in their ELK stack environments.

The details in this guide on how to calculate the TCO of your ELK stack environment, and how to assess the potential cost savings of ChaosSearch, will help you assess making the switch in your organization. The ChaosSearch team can assist you in using the TCO Tool to develop a customized business case for your environment. We'd like to hear from you about your plans and any unique challenges you are facing. For any questions or requests, or to simply learn more, visit us online or send us an email chaossearch.io/contact.



ABOUT CHAOSSEARCH

ChaosSearch empowers data-driven businesses like Blackboard, Equifax, and Klarna to Know Better™, delivering data insights at scale while fulfilling the true promise of data lake economics. The ChaosSearch Data Lake Platform indexes a customer's cloud data, rendering it fully searchable and enabling data analytics at scale with massive reductions of time, cost and complexity. The Boston-based company raised $40M Series B in December 2020 and is hiring to support its hyper growth.

For more information, visit ChaosSearch.io or follow us on Twitter @ChaosSearch and LinkedIn.

info@chaossearch.com | www.chaossearch.io