TCO ANALYSIS

The Cost of Doing ELK on Your Own

By: Ariel Assaraf, CEO of Coralogix

Coralogix

Summary

Logs have become the bedrock of modern software observability. No deployment is confirmed without a healthy log output. This reliance on logs has created a challenge of scale.

A mid-size company's applications will output between 100-200 GB per day. That's 500,000-1 million log lines, every day. Some of those are duplicates that need to be filtered, and only a fraction of these logs are critical and need to be handled immediately. The importance of having an efficient log management tool cannot be understated.

Breaking Down Cost of Ownership

For any tool, open-source or commercial, there is more to consider than the initial price tag. While the purchase price is often visible and straightforward, there are still many other associated costs that need to be considered.

- Does the tool require additional hardware considerations?
- How long will it take to set up and onboard the team that will use the tool?
- Will the tool require ongoing maintenance and/or support staff?

Cost of Ownership can be broken down into *visible costs* and *hidden costs*. Visible costs are easily identified and quantifiable, while hidden costs are not immediately obvious nor easily quantifiable.

A tool may have a respectable price tag but have significant costs hiding beneath the surface. That's why it's important to consider all of the following cost areas when considering a new tool.

- Purchase cost
- Infrastructure
- Initial set up and customization
- Ongoing operational costs
- Intangible Costs

Estimating Cost of Ownership for the “Free” ELK Stack

Building your own log management solution means setting up, customizing and managing an ELK stack instance.

The associated costs will vary depending on several aspects such as: *how much log data is generated by your system? How much is your log volume likely to increase each year? How long do you need to retain that data? How accessible does your data need to be?*

We’ll break down the costs associated with building an ELK stack solution for a typical company based on the following assumptions:

- Produces approximately 100GB/day of log data for the first year (increases ~50% each year)
- Requires 14-day data retention
- Requires high data availability

Let’s look at how each of the 5 costs associated with ELK ownership play out over a 3-year period.

ELK Paid Subscription
(Gold, Platinum, Enterprise)

- Reporting
- Elastic Endpoint Security
- Stack monitoring
- Support services
- Machine Learning
- Cross-cluster replication
- etc.

PURCHASE COST

This is the tip of the iceberg, the most visible and “obvious” cost associated with any tool – its price tag.

For a free ELK Stack instance, the upfront cost is $0 (surprise!), easily setting it apart from commercial solutions. It’s important to remember that there are likely additional costs hiding beneath the surface. For example, Elasticsearch’s alerting capabilities are part of the X-Pack offering.

	Year 1	Year 2	Year 3
Purchase Cost	$0	$0	$0

In the following sections, we’ll take a look at what associated costs come with the setup, customization and operations of your homegrown ELK Stack.

INFRASTRUCTURE

For the company described, the requirements will include 3 master servers (since masters need an uneven number and you want them redundant) you can also use them for Kibana and Logstash, 2 data servers (for redundancy) and disk space for data storage (considering you don't want to cross 90% disk watermark with Elasticsearch). We'll make use of AWS pricing to provide a useful estimate.

One master server instance (c5.large.elasticsearch) costs $0.156, which comes out to about $1,367 each year.

For data instances, we'll use the r5.2xlarge.elasticsearch servers with an upper limit of 3TB per machine. For the first year, one machine covers our requirements plus one copy for data redundancy. In the second year, 2 instances are required plus redundancies, and in the 3rd year 4 instances will be needed (again, with redundancies).

In terms of disk space, that's $0.162 per GB of data per month plus one copy for redundancy, plus an additional consideration for system overhead.

	Year 1	**Year 2**	**Year 3**
Master Servers	$4,100	$4,100	$4,100
Data Servers	$14,472	$28,943	$57,886
Disk Storage (General Purpose SSD)	$7,893	$11,839	$17,758
TOTAL HOSTING COST	$26,464	$44,882	$79,744

SETUP COSTS

Unlike out-of-the-box commercial solutions, engineering costs for setting up an ELK stack are significant. These include the EC2 servers, mappings, Kibana and collectors. It is often disregarded as a sunk cost (i.e. engineering salaries are being paid regardless of how that time is used).

Although getting an ELK stack deployed can take a few minutes, ELK requires a significant investment of time and resources to make it production-ready.

Getting Your ELK Stack Production-Ready:

1. Create your deployment and customize to best fit your use case
2. Configure index management
3. Configure Kibana, Elasticsearch clients, Beats, Logstash and APM Agents
4. Add additional plugins for desired functionality
5. Secure your deployment
6. Plan for scale and availability

Above all else, the system must be configured to ingest and parse logs coming from multiple systems and locations, with potentially different logging frameworks and formats. This data pipeline, feeding into Logstash, must also be optimized to ensure no data is lost. In some cases, this requires additional investment in hardware.

Completing the setup for our “typical” company would take the average engineer (who is familiar with the ELK Stack) about 5 full working days. This translates to $2,650 just for the initial setup, not including the investment in training engineers to use the tool.

If we assume that the training will take an additional 3 days, with 5 engineers being trained, that comes out to $7,950 for the first year. As the stack becomes more complex, this number will likely grow.

Due to the complexity of the ELK Stack architecture, companies will often dedicate at least one full-time employee to building and running it.

Considering that the average engineering salary is around $140K/year, this is no small expense.

	Year 1	Year 2	Year 3
Setup & Customization	$2,650	$5,300	$10,600
Training	$7,950	$15,000	$20,000
TOTAL SETUP COSTS	$10,600	$20,300	$30,600

ONGOING OPERATIONAL COSTS

Engineering costs don't end once your ELK Stack is set up. Additional considerations include monitoring performance of the ELK Stack and troubleshooting system errors and outages. As your system grows, logs tend to evolve and it's important to ensure that the logs are ingested and indexed properly.

The tooling itself is open-source but Elastic support comes with an additional cost. Elastic offers tier-based subscriptions that include varying levels of support. They also come with additional capabilities, including out-of-the-box integrations, alerting, endpoint security, some Machine Learning search and analysis capabilities, and much more.
Elastic uses a resource-based pricing model for their subscriptions so as your system scales, this cost will scale along with it.

Upgrades are also a serious consideration when it comes to cost, although Elastic made an effort in 2015 to coordinate the releases of the tools that make up the ELK stack.

Of course, there is still a fair amount of time and effort required to smoothly upgrade versions, including checking for deprecated features or any other breaking changes, backing up data, testing in pre-production environments and much more.

In the first year, one engineer dedicating a third of their time should be sufficient for necessary maintenance and scaling requirements, but as the stack grows and becomes more complex, more time must be invested.

Ongoing ELK Maintenance Tasks:

1. Performance monitoring and tuning of ELK clusters
2. Performing ELK upgrades when new versions are released
3. Maintaining proper ingestion and indexing of data
4. Creating and maintaining dashboards in Kibana
5. Creating and maintaining alerting
6. Handling change requests from teams and Executives
7. Troubleshooting system errors and/or outages

By the second year, that engineer will likely need to dedicate half of their time to maintaining the ELK Stack, and by the 3rd year, one full-time engineer should be allocated to managing the ELK stack.

	Year 1	Year 2	Year 3
Ongoing Operations (i.e. Engineering Time)	$46,200	$70,000	$140,000

INTANGIBLE COSTS

Intangible costs are those that we simply can't see coming. This could include things like cost of system downtime or critical issues, missed opportunities/value, etc.

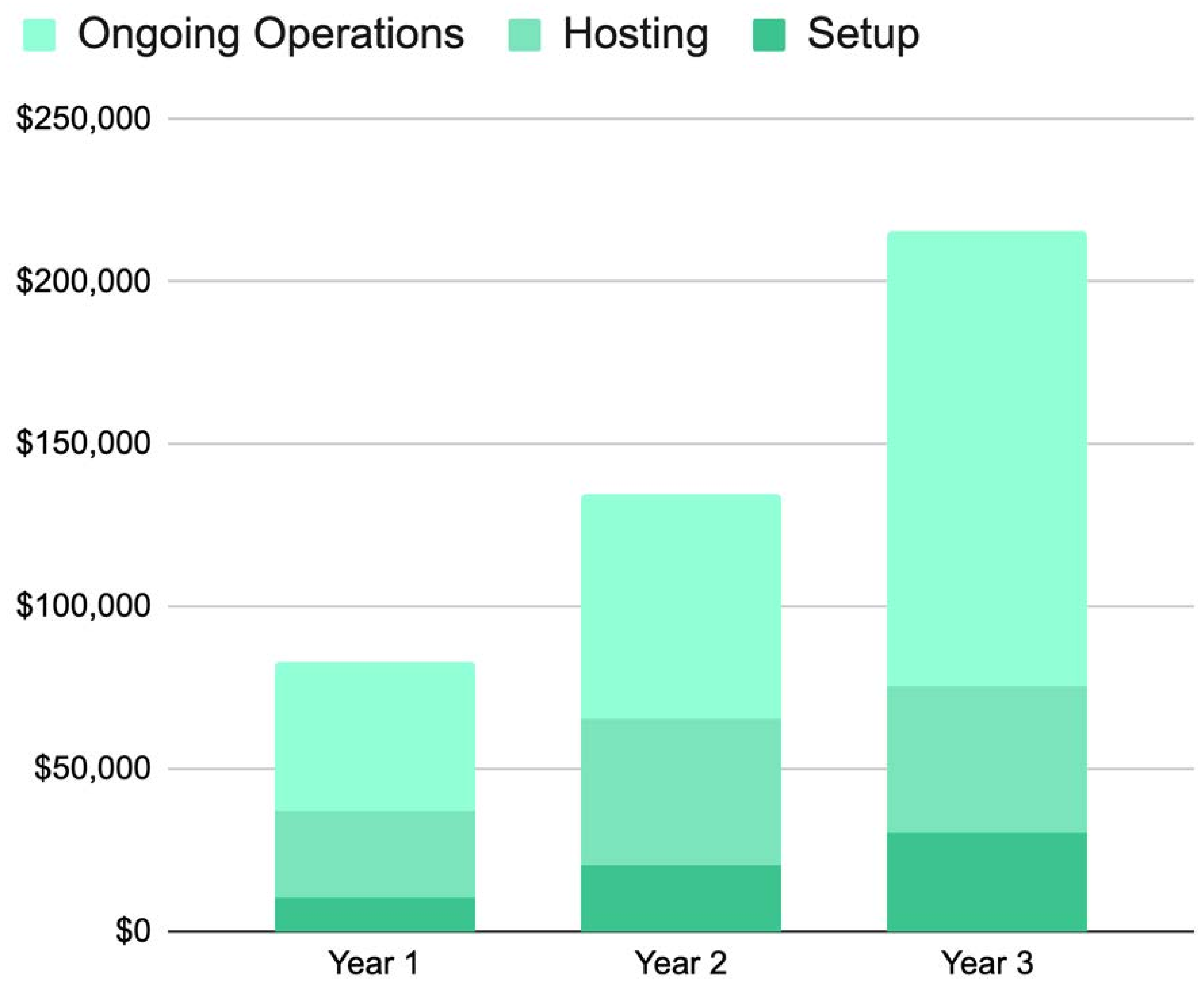
For example, the difference between a solution with Machine Learning capability compared with a stack equipped with static alerting thresholds. An anomaly in error rates may be flagged in the ML solution, but if the actual volume of errors doesn't pass a certain threshold it would be missed by static alerting.

If that piece of information is the difference between 1 hour of downtime and 1 day of downtime, the cost difference can be huge in terms of customer satisfaction and retention, SLA violations, etc.

These are things that ultimately cost companies money, though they may not be easily quantified or attributed to tooling investments.

Although not easily measured, intangible costs cannot be ignored when considering cost of ownership and return on investments. Ultimately, any trade-off that impacts the return you see on your investment in a tool IS part of the cost of ownership.

ELK Stack Total Cost of Ownership Summary



At first glance, *free log management* is hard to ignore but remember that with open-source tooling the price will likely come in the form of engineering hours. There are many commercial log management solutions on the market, each with their own advantages and disadvantages. Aside from looking for a tool that matches your team's functional requirements, cost is still a major factor to be considered.

	Year 1	Year 2	Year 3
Total Cost of Ownership for "Free" ELK Stack	$83,264+	$135,182+	$250,344+

Gain Observability, Without Paying For The Noise

Most log data is never queried or analyzed and only contributes to high data storage costs (as seen above), but at the same time, logs do contain important information. It just doesn't make sense for you to pay the same price for critical log data that you pay for irrelevant log data.

Our team has spent the past year focused on creating a new model for logging, allowing you to prioritize your log data and define how the data is routed and stored according to function and importance.

Disguised as a simple triaging capability, Coralogix's new Total Cost of Ownership (TCO) Optimizer drastically reduces logging costs while also improving your ability to query, monitor, and manage your data.

CENTRALIZE

Collect log data from all on-perm or cloud services and format it in any fashion for powerful analytics

REDUCE

Pay by data prioritization rather than volume by triaging logs based on business value

SOLVE

Blazing fast powerful search and automatic version bench-marks to solve problem faster

MONITOR

Loggregation™, Smart alerts, Anomalies, Kibana dashboards and Grafana

FREE TRIAL

SCHEDULE A DEMO