

Q: Pipeline my S3 data into Elastic for search?

A: Not anymore...

Conventional wisdom says use Elasticsearch

The answer to “Where should we move our S3 data for search?” is almost always “Elasticsearch.” Or maybe a cloud alternative like Sumo Logic.

Maybe a better question is “Why move our S3 data anywhere?”



Amazon S3 does the hard stuff better and cheaper than Elastic

Elasticsearch provides a popular search API and tooling (Kibana). As your data grows, the following become Elasticsearch operational challenges, which will force you to add staff, grow expensive hosting infrastructure, or limit your data retention:

- Scalability
- High availability
- Durability
- Security

Amazon S3 securely scales storage without limits, with high availability, and at very low cost. If there was a search interface for S3, you would not face painful Elastic operational costs.

Search S3 data in place with CHAOSSEARCH

CHAOSSEARCH has developed a search service for S3 that decouples compute from storage so you can scale up and down as needed, without delay or disruption. It features:

- One-click schema detection, normalization, and indexing of S3 data
- Elasticsearch-compatible API with fully integrated Kibana visualization
- Dynamic scalability

ELB (1.35 billion)	 CHAOSSEARCH	 elasticsearch
Index - Time	1.6 hours	4.2 days
Index - Size	150 GB	2.1 TB
Index - Cost	c4.2xlarge (< \$1.00)	c4.2xlarge (\$38.00)

Better performance, economics, and insights

Searching S3 data in place with **CHAOSSEARCH** eliminates Elasticsearch operational costs — licensing, hosting infrastructure, and people costs for mapping and ETL, scaling, performance tuning, etc.

Save \$200K per year on a modestly-sized search project (150 GB new data per day)

Running the Elastic Stack to process a modest 150 GB of daily new data with 30-day retention costs over \$250K (USD) per year, which is 6 times more costly than simply using **CHAOSSEARCH** to analyze data in place on Amazon S3.

Expense	Elastic Stack	CHAOSSEARCH	Savings	Notes
Hosting	\$157,365 / year (18 i3.2xlarge.elasticsearch instances)	\$39,042 / year (4500 GB = \$1,242 for S3 and \$37,800 for CHAOSSEARCH)	\$118,323	Elastic requires a separate software and hardware stack from S3. Build your own or use a service like Amazon Elasticsearch like we've done here. CHAOSSEARCH searches S3 data in place, and therefore requires no separate datastore stack.
Data prep	\$52,000 / year (40% technical FTE @ \$130K/year)	\$5,000 / year (5% analyst FTE @ \$100K/year)	\$47,000	Elastic requires ETL development, parsing, conflict resolution, schema mapping. CHAOSSEARCH analyzes S3 data in place, so it requires no ETL development — just a little time to initiate and review automatic metadata and schema discovery.
Operations	\$52,000 / year (40% technical FTE @ \$130K/year)	\$0 / year	\$52,000	Elastic requires index management, sharding, cluster sizing, disaster recovery setup, archiving. CHAOSSEARCH searches S3 data in place, and therefore requires no datastore admin.
Support	\$22,100 / year	\$0 / year	\$22,100	Cost for annual technical support from Elastic. CHAOSSEARCH has no extra fee for support; it is included with the storage-based pricing.
Total	\$283,465 / year	\$44,042 / year	\$239,423	

With the money you save...

As you increase the amount of data you store, you can expect your Elastic Stack costs to grow linearly, doubling in cost as your data doubles in volume.

The money you save by analyzing S3 data in place with **CHAOSSEARCH** enables you to:

- **Retain and analyze more data for less.** You could retain years of data in Amazon S3 — or integrate additional data from many more sources, and not come close to the Elastic Stack costs above.
- **Use IT people more productively.** Invest savings in new development instead of systems administration and data wrangling. You could hire two full-time software engineers with the savings in the example above.

The **CHAOSSEARCH** “no-migration on-boarding” experience

If your data is in S3 buckets, you activate **CHAOSSEARCH** against it as-is. There is no new data store to install. No data to migrate. No ETL to set up. Just index and search.

To begin your free trial of **CHAOSSEARCH**, visit chaossearch.io today.



© 2019, CHAOSSEARCH™, Inc.

Elasticsearch, Logstash, and Kibana are trademarks of Elasticsearch B.V., registered in the U.S. and in other countries. Elasticsearch B.V. and CHAOSSEARCH™, Inc., are not affiliated.

CS-OP-02