Wavefront Enables Algolia to Optimize Operations at Scale

The Challenge
Algolia provides a hosted search service for a wide variety of customers. Because they replicate their engines geographically, their customers experience search results instantaneously. Given that both the customer base and the service usage are growing, Algolia must not only watch current performance to detect unexpected anomalies, but they must also trend data to ensure that, as they scale, their performance doesn’t bog down. The infrastructure components that must be monitored include:

- CPU usage
- RAM usage
- Disk usage
- Read and write I/Os
- SSD state
- Indexing speed
- Queue size
- Network latency
- Inter-provider latency
- Number of crashes

Algolia also exposes an API both for setting and getting data, so another growing performance parameter is the number of hits the API gets over time.

They tried various monitoring solutions, but were stymied by several issues:

- While these tools performed at smaller scales, they bogged down as activity grew.
- The tools tended to allow alerting only on single metrics, based on a threshold.
- These tools required up-front metrics definition. If an ad-hoc metric was needed to solve an immediate problem, there was no historical record of that metric available for trending or comparison.
Additionally, while they saw ways to automate the resolution of certain issues, they couldn’t execute on that because the amount of “alert noise” – that is, false alarms – was too high.

Finally, they wanted a way to give their own customers access to real-time performance metrics.

The Solution
Other large-scale SaaS providers recommended Wavefront to Algolia. The Wavefront SaaS based metrics monitoring and analytics platform lets them define complex metrics and alerts based on multiple streams of data and complex functions like moving averages and derivatives.

Because the metrics and alerts are derived from complete historical data streams, they can define dynamic metrics for use in debugging an issue, and still view the history of that newly defined metric. In fact, because that historical data is fully stored in raw form, it also serves as a full performance archive.

Algolia has additionally been able to tap into Wavefront metrics for giving their customers access to performance data. They have wrapped the Wavefront queries in an interface that their customers can use.

The Results
Algolia immediately got visibility into performance information – current and historical – that simply wasn’t available from other tools. This meant that they were able to optimize their operations at any point based on arbitrary metrics that could be defined on the spot, leveraging the data archive.

This made it much easier for them to debug obscure issues as they arose. For example, they were able to isolate a DNS response issue caused by one provider’s subtle bug and compounded by an exotic IPv6-related problem on some DNS servers.

Their customers now have access to real-time metrics, which are really Wavefront metrics passed through Algolia’s metrics interface.

And, going forward because the alert noise has practically vanished, they will be able to automate the resolution of certain issues, freeing up critical developer and operations time for solving the really hard problems.