REPORT REPRINT

Why Box ditched its homegrown monitoring tool for Wavefront

NANCY GOHRING

9 MAY 2017

Box initially built its own tool for monitoring and troubleshooting performance of its service, but as it grew, that internally built tool struggled to keep up. The company turned to Wavefront, which collects roughly 400,000 data points per second for Box.

451 Research®

Using some open source software, Box initially built its own tool for monitoring and troubleshooting performance of its service, but as the company began to grow, that internally built tool struggled to keep up. Rather than continuing to spend the time maintaining it – time that could better be spent on more important tasks such as new feature development – Box turned to Wavefront, one of a new breed of monitoring tool vendors, one that is in the process of being acquired by VMware.

THE 451 TAKE

Box faced a classic 'build vs. buy' decision for monitoring. We commonly hear of scenarios such as Box's where a company outgrows its DIY solution, although some of the pain points that Box faced, such as that it lacked documentation, were avoidable. However, homegrown monitoring tools such as Box's often stumble under the volume of data that modern application environments produce. One downside to choosing a vendor over an internally built tool is that the user is subject to the vendor's priority list for adding new features and functionality. In Box's case, the benefits of using Wavefront – particularly around valuable business use cases – seem to outweigh that downside.

EARLY ADOPTER SNAPSHOT

Box has 52 million registered users. While the bulk of them don't pay, the company supports more than 71,000 paying organizations, some of which rely on Box as their primary source of document and other content storage. As such, the service must be available and performant. Like many large organizations, Box's site reliability engineering (SRE) team uses multiple monitoring tools to discover what problems are occurring and to help identify their root cause. Tools include log management from Splunk, as well as the open source Elastic Stack, plus Amazon Cloudwatch, the metrics service for monitoring AWS performance. Box also uses open source software including MRTG and Cacti for network monitoring.

BUSINESS CHALLENGES

About three years ago, Box's SRE group, the team responsible for ensuring quality of service, was relying on a homegrown monitoring system that was difficult to use, limited in terms of the visibility it offered and accessible only by engineers. The tool, built on the open source time-series database OpenTSDB, was particularly difficult for new team members because there was no documentation about how to use it. That meant new team members faced a slow learning curve, relying on coworkers to learn what metrics were available to search for when problems occurred.

About two years ago, the system began to struggle with the volume of data Box was sending to it. It had both capacity and throughput problems, and the SRE team found itself spending too much time shoring up the tool, at times postponing releases and maintenance of Box products in order to work on the tool itself.

DEPLOYMENT SUMMARY

At the same time that the OpenTSDB deployment began to buckle under the weight of Box's scale, senior executives at the company were introduced to Wavefront and arranged for some members of the SRE team to see a demo. The engineer we spoke with was immediately interested, noting that Wavefront's visualizations were compelling, as was the speed with which the user could query and view results.

Although some at Box had initial concerns about whether Wavefront, itself a young startup at the time, could handle the volume of data Box was likely to want to process, it moved forward with a proof-of-concept installation, followed by a full rollout. Box is currently sending roughly 390,000 data points per second to Wavefront, data that Box can query as far back as the day it started using Wavefront. Box reports that it has not had any capacity issues. It retained parts of the OpenTSDB-based system, namely the collectors that ship metrics to Wavefront agents.

Wavefront has many more users (and use cases) than Box's former homegrown tool. In addition to troubleshooting use cases – Box has monitors hanging throughout its network operations center displaying important metrics in Wavefront visualizations – Box has found the service valuable in business and planning use cases. For instance, engineers use Wavefront to establish a baseline of performance to compare to after making a change to the service to discover whether the change had the anticipated impact. The engineering team has also built dashboards for sharing with business executives as a way to demonstrate cost savings – comparing the cost and performance of one service provider against another, for instance. Because Box enables access to Wavefront via Okta's single sign-on product, it's easily accessed by employees throughout the company.

About 400 people at Box access Wavefront each day, mostly engineers and some business users. Such broad usage of a monitoring tool is unusual among older-generation products, but we hear of large user bases with some newer-breed tools such as Wavefront that are designed to ingest an array of data and are flexible enough to support many use cases.

INNOVATION AND ROADMAP

One feature request that Box has sent to Wavefront would allow users to see when someone has viewed a graph that's been shared. In addition, while Box can itself set up data ingestion from most sources, it has requested some technology integrations that it hopes Wavefront will build so that it doesn't have to. Box currently feeds metrics from CloudWatch, AWS's monitoring service, into Wavefront and expects to continue to consolidate metrics from other sources into Wavefront.

We think Box would benefit from some additional administrator controls in Wavefront. The company has had a couple of instances where it began to flood Wavefront after someone inadvertently began sending a huge volume of metrics. While Wavefront didn't technically have issues handling the data, Box began exceeding its agreement with Wavefront, which is based on volume of data points sent per second. Box reports that the situations were easily resolved.

The Box site reliability engineer we spoke with said that any time a vendor gets acquired – Wavefront is in the process of being acquired by VMware – there's some uncertainty for users about product development. When we spoke with VMware at the time the acquisition was announced, it indicated that it expected to continue selling Wavefront as a stand-alone product and that part of the motivation for the deal was to help VMware better attract DevOps pros and developers, a good indication that it may continue along the path Wavefront has set.

COMPANY NAME
Box

ACTIVITIES
Content management

HEAD OFFICE
Redwood City, California

NUMBER OF EMPLOYEES
1,500

LY REVENUE
$399m for FY17

LY NET INCOME
($151.8) for FY17

KEY SUPPLIERS
Wavefront