splunk> listen to your data™

Splunk at WhitePages

Real-time Insights Improve User Experience and Drive Revenue

white pages™

"I was able to get IT data into Splunk in a couple of hours, providing us with data warehouse-like capabilities without the data warehouse cost or complexity."

Hernan Alvarez
Senior Director of IT and Operations

OVERVIEW

INDUSTRY
- Online directory

SPLUNK USE CASES
- Application Management - troubleshooting and monitoring, real-time release management and log review
- Security - real-time monitoring and proactive incident detection

BUSINESS IMPACT
- Provided operational intelligence without investing in a data warehouse
- Optimized site content and banner ads support revenue creation
- Protected site content and brand value by monitoring for abusive site traffic
- Accelerated time-to-market by streamlining development effort

DATA SOURCES
- Server logs
- Router/switch/firewall syslog
- Web logs/middleware access logs
- Middleware/database exception logs
- DB long query logs
- Custom and packaged application events

The Business

As the largest on-line directory, WhitePages offers a suite of free mobile and Internet services that makes it easy to find and connect with people and businesses in the US. WhitePages offers access to more than 200 million adults and powers more than 2 billion searches annually on over 1,500 partner sites including MSN, Ask.com, and the United States Postal Service.

Challenges

WhitePages had always hosted its own people search function. However, its business search was hosted by a third party. In 2009, WhitePages made a strategic decision to bring this business search function in-house. Successfully making that transition required gathering and analyzing extensive operational data and user metrics. The challenge: how to manage all of this data without building and maintaining a data warehouse. WhitePages faced another challenge. Its rich database of information on people and businesses was under constant threat from attackers copying or "scraping" data and abusing the service. The site needed a faster, more efficient way to detect and mitigate scraping and abuse.

A third challenge was identifying bugs in their new middleware releases before they caused problems in the production environment by impacting the customer experience.

Enter Splunk

WhitePages deployed Splunk Enterprise to gain better visibility across their application infrastructure, to mitigate site abuse and to identify software bugs more quickly. Splunk collects a wide range of IT data including networking equipment syslog such as firewalls, switches, and routers, as well as front-end web access logs, middleware logs, query logs and database exception logs. Splunk easily handles these regardless if they are custom or packaged applications.

Splunk was instrumental in optimizing the new business search service. "I was able to get all our data into Splunk in just a couple of hours, providing data warehouse capabilities without the data warehouse," says Hernan Alvarez, Senior Director of IT and Operations. "Splunk helped us understand what questions we should be asking of our IT data to make informed business decisions and gave our teams access to the data through an easy-to-use dashboards."

Splunk allows WhitePages to deep-dive into their data looking for scripted behavior that identifies subnets or users who are abusing the system. WhitePages recently began monitoring in real time. In the near future, WhitePages will be transitioning the logic from Perl scripts directly into Splunk where Splunk will perform the logic and abuse detection. "The information we glean from Splunk enhances our automated abuse detection as well as blacklists," says Alvarez.

In addition, Splunk ingests all of WhitePages' exception logs in real time and looks for patterns that identify bugs in new releases of middleware that were missed in QA. "We can actually see exceptions displayed live on screens around the office."

Breakthroughs

Splunk enabled WhitePages to avoid the time, cost and effort of implementing and managing a full-blown data warehouse. It also allowed the Manager of the Business Search unit to gather the partner revenue information he needed without having to write his own scripts or distract engineers who were focused on other tasks.

Before Splunk, Business Development spent more than a day a month manually compiling spreadsheets. "With Splunk, in about an hour we ingested the information and wrote the reports," says Alvarez. "Now it's completely automated. Splunk sends a daily report and one at the end of the month."

The ability of Splunk to automate WhitePages' abuse detection was recently demonstrated. An exercise conducted by an engineer was able to eliminate 400,000 pages a day of abuse traffic to the site. "That's a significant number," says Alvarez. "And by reducing abuse traffic we increase our brand experience and can better manage site performance through reducing 400,000 visits a day."

With the help of Splunk, WhitePages has adopted the agile development. Their objective is to enable programmers to do daily releases without involvement from the operations team. By logging their releases in real-time, Whitepages can deploy a highly optimized site based on the latest user trends and activity. This real-time intelligence, delivered on a single dashboard, will help Whitepages increase conversions, ad click-throughs and ultimately drive new revenue.

"Having real-time logging aggregated into a single Splunk dashboard helps us achieve our agile programming goals."

Hernan Alvarez
Senior Director of IT and Operations

Free Download

Download Splunk for free. You'll get a Splunk Enterprise license for 60 days and you can index up to 500 megabytes of data per day. After 60 days, or anytime before then, you can convert to a perpetual Free license or purchase an Enterprise license by contacting sales@splunk.com.

250 Brannan St, San Francisco, CA, 94107   info@splunk.com | sales@splunk.com   866-438-7758 | 415-848-8400   www.splunkbase.com

splunk> listen to your data™

www.splunk.com