White Ops Uses **Snowflake** to Make Real Time Decisions Against Digital Ad Fraud

whiteops

CUSTOMER White Ops

PARTNER Snowflake

snowflake

CASE STUDY

White Ops is a leading provider of cyber security services for the detection and prevention of advanced bot and malware fraud. Unlike traditional approaches that employ statistical analysis, simple blacklisting or static signatures, White Ops effectively combats criminal activity by actually differentiating between robotic and human interaction within online advertising and publishing, enterprise business networks, e-commerce transactions, financial systems and more, allowing organizations to remove and prevent fraudulent traffic and activity.

THE CHALLENGE

Analyzing Lots of Data, Fast, to Separate Friend from Foe

Ad fraud is a dynamic target, with new threats and methods of fraud appearing continually on the horizon. White Ops' solutions must analyze a torrent of real-time data—intercepting the data; interpreting whether the data indicates fraud; and aggregating statistical data for confidence, reliability, and impact reporting.

Because threats are morphing all the time, continuous monitoring of ad traffic is required to stay on top of both new and changing threats. White Ops is constantly working to uncover and characterize new fraud patterns, which requires storing and processing massive amounts of data. To differentiate itself from its competitors, White Ops relies on continuous generation of new detection algorithms.

"In very real ways, Snowflake helps us improve our competitive advantage and better focus on our core competency: the fast delivery of a broad spectrum of ad-fraud detection algorithms for our customers."

Tamer Hassan Co-founder and CTO

White Ops had previously relied on no SQL systems including Hadoop and MongoDB to store and process that data. However, that approach created challenges for their researchers.

"Our researchers and security engineers are frequently looking at new data and working to find new patterns in the data," says Tamer Hassan, co-founder and CTO of White Ops. "Their work generates 'big data' questions that used to always require writing new map-reduce jobs. As a result, asking a question of big data had historically been a difficult problem. In fact, until recently it would have been a dead request."

When a White Ops security engineer faced a big data question, he or she had to send a request to the Hadoop team and wait until a developer could build a custom map-reduce job.

"The latency for results was at least 24 hours, depending on workload at that time. The more requests, the greater the delays," says Hassan.

The bottleneck was the result of needing to write custom map-reduce code for each request, which required a level of programming skills that only a small number of engineers possessed. This bottleneck meant that many times, requests for data and new analytics were not made in the first place or didn't complete fast enough.

White Ops needed a way to simplify and standardize the process supporting this in order to increase the productivity of their research and development teams, which in turn would help White Ops deliver the responsiveness its customers depend on.

THE SOLUTION

Snowflake Elastic Data Warehouse

To get there, White Ops implemented the Snowflake Elastic Data Warehouse. Because Snowflake uses SQL as its core language and is delivered as a data warehousing cloud service, it is now possible for not only data scientists but also any analyst to access data directly.

Using the Snowflake data warehouse, White Ops can:

- Have all result data in one place
- Scale elastically
- Query diverse data with standard SQL

Consolidating data in one place

"Snowflake absorbs just about all of our data in various levels," says Hassan. "Last year, we did one of the largest studies ever on ad fraud, and we leveraged Snowflake to do the research. We put all our data in there, we gave all of our analysts and researchers access to it, and we were able to build insights much faster than before with everybody queuing up queries on the Hadoop cluster."

Key to making that possible is the Snowflake architecture, which made it easy for White Ops to bring its data together in one place.

"One thing that we really use is Snowflake's separation of compute and storage," says Hassan. "Our detection team now can compare trends in ad fraud in one place and perform statistical studies directly with that data, both historical and current. They're able to look at large- scale data sets in Snowflake and query all the data. That is a very powerful capability and was one of the key capabilities that drew us to Snowflake."

Scaling where and when needed

White Ops also takes advantage of Snowflake's separation of compute and storage to scale up and down on an on-demand basis. For instance, White Ops can adjust the computing power for its users based on their needs.

"That approach works great," says Hassan. "We turn on our standard amount of data warehouse compute capacity

for most of the day across east and west coast U.S. time. But any time we have to do heavier research or some bigger jobs, we will simply spin up a larger virtual warehouse in Snowflake."

Querying large, diverse data with SQL

White Ops appreciates that the Snowflake Elastic Data Warehouse lets it use standard SQL so that data access is possible for users across the organization. Additionally, Snowflake's native support for optimized storage and processing of both structured and semi- structured data allows White Ops to add data from multiple sources into a single place for analysis .

"Being able to query semi-structured data and large data sets with SQL is new," says Hassan. "Most people don't have the spare time to learn a completely new query language. Using SQL gives broad accessibility to a lot of people. Most people in my organization can write basic SQL. Adding that capability on top of semi-structured data, as Snowflake does, is even more powerful.

THE RESULTS

Better Fraud Detection and Expanded Data Access

Implementing the Snowflake Elastic Data Warehouse helped White Ops accelerate the pace of iteration and evolution of its fraud prevention offerings.

"For us, the value of Snowflake to our company is clear," says Hassan. "We have seen three big areas of improvement. The primary one is accelerating our output of algorithms that detect ad fraud--we can write more algorithms in less time. Second, we have improved both system health and the quality of results via our QA process. Third, the resulting data is deeper, more accurate, and much more accessible and easier to monitor using SQL everywhere across different user groups."

Accelerating productivity

To keep ahead of rapidly evolving fraud, Snowflake has helped White Ops develop and release new algorithms in a fraction of the time previously required.

WHITEOPS RESULTS

- Accelerated ad fraud detection algorithms, from 24+ hours to less than 2 hours
- Improved system health and quality of QA
- Deeper, more accurate and more accessible data results

That improvement is due in no small part to the fact that Snowflake enables diverse analytics using SQL without requiring specialized programming. "Our development team uses Snowflake as a tool in developing the algorithms they need to answer a question," says Hassan. "With Snowflake they can perform this without having to go through an entire team process to write and submit a Hadoop job. The time to put new algorithms into production has gone from 24 hours latency to two hours, dramatically reducing cycle time for our product releases."

Improving QA

"Snowflake is helping us simplify testing and QA," says Hassan. "The performance we get with Snowflake allows us to check the results of each change in code, rather than only doing that for a batch of changes."

Enabling richer analytics across the company

By decentralizing access to data and extending data exploration capabilities across the entire company, Snowflake allows results and reports to be provided to customer support, sales, and anyone else at the company who can use the data.

"Democratizing access to data increases involvement and alignment within White Ops," says Hassan. "This breadth of access is important for the complex challenges we encounter every day, because we each have different

insights that can lead to powerful results. In very real ways, Snowflake helps us improve our competitive advantage and better focus on our core competency: the fast delivery of a broad spectrum of ad-fraud detection algorithms for our customers."

Using Snowflake, White Ops also can experiment easily and safely in a manner that was not practical or cost-effective before. "What we have with Snowflake is the flexibility to build a sandbox, try something out with it, then blow it away," says Hassan. "There's too much overhead to make that process easy and convenient in a legacy system—you'd have to size hardware, buy extra software licenses, load data, partition data, and build indexes just to get it going, and after all that work you're unlikely to want to blow it all away. Even if you do delete that data, you now have this purpose-built installation in need of a new purpose—it's idle capacity otherwise."

CONCLUSION

By implementing the Snowflake Elastic Data Warehouse, White Ops has been able to work more effectively as an organization and sharpen its competitive edge. Using Snowflake to consolidate and scale its massive amounts of data, enable access to the data without waiting for specialists with deep programming skills, and work both faster and smarter has made White Ops better positioned than ever to help its customers avoid the potentially devastating effects of online advertising fraud.