



SnapLogic CEO Talks



Chang

SECTIONS ▼



- [Home](#)
- [News & Commentary](#)
- [Authors](#)
- [Slideshows](#)
- [Video](#)
- [Reports](#)
- [White Papers](#)
 - [Events](#)
 - [Interop](#)
 - [University](#)
- [Strategic CIO](#)
 - [Software](#)
 - [Security](#)
 - [Cloud](#)
 - [Mobile](#)
 - [Big Data](#)
- [Infrastructure](#)
- [Government](#)
- [Healthcare](#)
- [IT Life](#)



- [Login to your account](#)
 - [Register](#)
 - [About Us](#)
 - [Contact Us](#)
- [Digital Subscription](#)
- [Advertise with Us](#)



Search InformationWeek



- [Facebook](#)
- [Twitter](#)
- [LinkedIn](#)
- [Google+](#)
- [RSS](#)



[Strategic CIO](#)[Software](#)[Security](#)[Cloud](#)[Mobile](#)[Big Data](#)[Infrastructure](#)[Government](#)[Healthcare](#)[IT Life](#)

Big Data // **Big Data Analytics**

NEWS

10/24/2012

04:40 PM



Doug Henschen
News

Connect Directly



4

comments

[Comment Now](#)[Login](#)

50% 50%

Like

1

Tweet

108

Share

284

g+1

27

Why Sears Is Going All-In On Hadoop

Sears pushes the cutting edge with some big data techniques, while trying to sell its big data services. Can emerging tech drive change in old-school companies?

Like many retailers, Sears Holdings, the parent of Sears and Kmart, is trying to get closer to its customers. At Sears' scale, that requires big-time data analysis capabilities, but three years ago, Sears' IT wasn't really up to the task.

"We wanted to personalize marketing campaigns, coupons, and offers down to the individual customer, but our legacy systems were incapable of supporting that," says Phil Shelley, Sears' executive VP and CTO, in a meeting with InformationWeek editors and his team at company headquarters in suburban Chicago.

Improving customer loyalty, and with it sales and profitability, is desperately important to Sears as it faces fierce competition from Wal-Mart and Target, as well as online retailers such as Amazon.com. While revenue at Sears has declined, from \$50 billion in 2008 to \$42 billion in 2011, big-box rivals Wal-Mart and Target have grown steadily, and they're far more profitable. Meantime, Amazon has gone from \$19 billion in revenue in 2008 to \$48 billion last year, passing Sears for the first time.

A Shop Your Way Rewards membership program started by Sears in 2011 is part of a five-part strategy to get the company back on track. Behind the scenes is a cutting-edge implementation of Apache Hadoop, the high-scale, open source data processing platform driving the big data trend. Despite Sears' less-than-cutting-edge reputation as a retailer, the company has been an innovator in using big data. In fact, Shelley is leading a Sears subsidiary, MetaScale, that's pitching services to help companies outside retail use Hadoop.

But will companies be interested in buying big data cloud and consulting services from Sears? And can Sears' own big data efforts help the company regain its footing in the retail industry?

Fast And Agile

Sears' process for analyzing marketing campaigns for loyalty club members used to take six weeks on mainframe, Teradata, and SAS servers. The new process running on Hadoop can be completed weekly, Shelley says. For certain online and mobile commerce scenarios, Sears can now perform daily analyses. What's more, targeting is more granular, in some cases down to the individual customer. Whereas the old models made use of 10% of available data, the new models run on 100%.

"The Holy Grail in data warehousing has always been to have all your data in one place so you can do big models on large data sets, but that hasn't been feasible either economically or in terms of technical capabilities," Shelley says, noting that Sears previously kept data anywhere from 90 days to two years. "With Hadoop we can keep everything, which is crucial because we don't want to archive or delete meaningful data."

Sears is still the largest appliance retailer and appliance service provider in the U.S., for example, so it's in a strong position to understand customer needs, service trends, warranty problems, and more. But Sears has

only been scratching the surface of using available data.

Enter Hadoop, an open source data processing platform gaining adoption on the strength of two promises: ultra-high scalability and low cost compared with conventional relational databases. Hadoop systems at 200 terabytes cost about one-third of 200-TB relational platforms, and the differential grows as scale increases into the petabytes, according to Sears. With Hadoop's massively parallel processing power, Sears sees little more than one minute's difference between processing 100 million records and 2 billion records.

The downside of Hadoop is that it's an immature platform, perplexing to many IT shops, and Hadoop talent is scarce. Sears learned Hadoop the hard way, by trial and error. It had few outside experts available to guide its work when it embraced the platform in early 2010.

The company is now in the enviable position of having big data experience among its employees in the U.S. and India. MetaScale will leverage Sears' data center capacity in Chicago and Detroit, just as Amazon Web Services takes advantage of Amazon's massive e-commerce compute capacity.



Open Source Moves In

CTO Shelley: big data zealot

Sears' embrace of an open source stack began at the operating system level, with Linux. Sears routinely replaces legacy Unix systems with Linux rather than upgrade them, Shelley says, and it has retired most of its Sun and HP-UX servers. Microsoft server and development technologies are also on the way out.

Moving up the stack, Sears is consolidating its databases to MySQL, InfoBright, and Teradata--EMC Greenplum, Microsoft SQL Server, and Oracle (including four Exadata boxes) are on their way out, Shelley says.

Hadoop's power comes from dividing workloads across many commodity Intel x86 servers, each with multiple CPUs and each CPU with multiple processor cores. Since early 2010, Sears has been moving batch data processing off its mainframes and into Hadoop. Cost is the big motivator, as mainframe MIPS cost anywhere from \$3,000 to \$7,000 per year, Shelley says, while Hadoop costs are a small fraction of that.

Sears says it has surpassed its initial target to reduce mainframe costs by \$500,000 per year, while also delivering "at least 20, sometimes 50, up to 100 times better performance on batch times," Shelley says. Eliminating all of the mainframes in use would enable it to save "tens of millions" of dollars, he says.

More Insights

Webcasts

[How to Build Search into Your Database](#)

[NoSQL and Big Data in the Government: What You Need to Know](#)

More Webcasts

White Papers

[How to Avoid the Need for SLA Upgrades](#)

[Every Business Is a Mobile Business](#)

More White Papers

Reports

[Ponemon Institute Research: How IT security is addressing threats to structured and unstructured data](#)
[2015 App Dev Priorities Survey](#)

More Reports

Comments

[Newest First](#) | [Oldest First](#) | [Threaded View](#)

Soozy G. Miller,
User Rank: Apprentice
12/5/2012 | 3:34:31 PM

[Login](#)

50%50%

re: Why Sears Is Going All-In On Hadoop

"But will companies be interested in buying big data cloud and consulting services from Sears?"

So, what, is Sears planning on entering a whole new market? Then again, The Gap started by selling music and jeans.

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)



Soozy G. Miller,
User Rank: Apprentice
12/5/2012 | 3:31:54 PM

[Login](#)

50%50%

re: Why Sears Is Going All-In On Hadoop

@srentner good point. It's great that Sears has a great big data solution; it's quite another to get the analytics to put that big data to use.

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)



srentner,
User Rank: Apprentice
10/31/2012 | 7:59:28 PM

[Login](#)

50%50%

re: Why Sears Is Going All-In On Hadoop

The final questions are extremely apropos and often cause the most confusion: "Could quick analytical access to an entire decade of medical record data change how doctors diagnose and treat patients? Could faster processing spot financial services fraud more effectively?" This is not what Hadoop does. It is not an analytics technology, as pointed out in page 1. Extracting this type of valuable insight from the data requires a new class of analytics technologies, and the more powerful the mathematical algorithms, the faster and more accurate the insight.

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)



Ellis Booker,
User Rank: Moderator
10/31/2012 | 3:07:19 PM

[Login](#)

50%50%

re: Why Sears Is Going All-In On Hadoop

This is a big story on a number of fronts. First, it clearly expresses the value of big data analysis for retailers. As one of the Sears executives puts it, "With Hadoop we can keep everything, which is crucial because we don't want to archive or delete meaningful data." Second, it addresses the oft-heard complaint that big data solutions are prohibitively expensive--in fact, Sears says it reduced mainframe costs by more than \$500,000 per year. Finally, the installation moves the retailer closer to real-time analysis: "Sears can develop in three days interactive reports that used to take IT six to 12 weeks."

--Ellis Booker, InformationWeek Community Editor

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)

[Reports](#)[Infographics](#)

6 Tools to Protect Big Data

Most IT teams have their conventional databases covered in terms of security and business continuity. But as we enter the era of big data, Hadoop, and NoSQL, protection schemes need to evolve. In fact, big data could drive the next big security strategy shift.

[Download Now!](#)

[More Reports](#)



[Subscribe to Newsletters](#)

[Live Events](#)[Webinars](#)

**Interop Las Vegas Full & Half-Day
Workshops**

Interop Las Vegas Conference & Expo**Aligning emerging technology and
ITSM**More UBM Tech
Live Events

White Papers

Single Sign-On: with Passwords, Less is More**How to Avoid the Need for SLA Upgrades****Every Business Is a Mobile Business****Find 10% in Excess Capacity in Your Operations****Secure Cloud Networking**

More White Papers

Current Issue

**InformationWeek Tech Digest, Nov. 10, 2014**

Just 30% of respondents to our new survey say their companies are very or extremely effective at identifying critical data and analyzing it to make decisions, down from 42% in 2013. What gives?

[Download This Issue!](#)[Subscribe Now!](#)[Back Issues](#) | [Must Reads](#)[Video](#)

[All Videos](#)[Slideshows](#)

10 Windows Tablets, Laptops Under \$200: Holiday Steals

[22 comments](#) | [Read](#) | [Post a Comment](#)[8 Tech Turkeys To Avoid As Gifts](#)[28](#)[Fitness Trackers: Holiday Gift Guide](#)[16](#)[More Slideshows](#)[Twitter Feed](#)[Tweets about "from:InformationWeek OR #InformationWeek OR @InformationWeek"](#)[InformationWeek Radio](#)[Archived InformationWeek Radio](#)

This InformationWeek

Join us for a roundup of the top stories on InformationWeek.com for the week of November 16, 2014.

UPCOMING!**Friday, December 5, 3pm EST****InformationWeek Live For the Week of November 30, 2014**[Sponsored Live Streaming Video](#)**UPCOMING!****Friday, December 12, 3pm EST****InformationWeek Live For the Week of December 7, 2014****UPCOMING!****Friday, December 19, 3pm EST****InformationWeek Live For the Week of December 14, 2014****InformationWeek**[FULL SCHEDULE](#) | [ARCHIVED SHOWS](#)

Everything You've Been Told About Mobility Is Wrong

Attend this video symposium with Sean Wisdom, Global Director of Mobility Solutions, and learn about how you can harness powerful new products to mobilize your business potential.

 **465 comments** | [Read](#) | [Post a Comment](#)

[Full schedule](#) | [Archived Shows](#)

[About Us](#)
[Contact Us](#)
[Customer Support](#)
[Reprints](#)
[Editorial Calendar](#)

[Twitter](#)
[Facebook](#)
[LinkedIn](#)
[Google+](#)
[RSS](#)



All rights reserved
