InfoWorld

AHEAD OF THE CURVE

By Eric Knorr, Editor in Chief, InfoWorld
MAR 7, 2016

What eBay looks like under the hood

eBay CTO Steve Fisher offers a guided tour of the architecture, technologies, and best practices of one of the world's largest e-commerce operations

Want more InfoWorld?

The Internet giants, from Google to Facebook to Amazon to Twitter to eBay, lead the way in enterprise technology. Their enormous scale presents them with problems that require new solutions and compel them to squeeze maximum utilization from their infrastructure.

That's one reason I jump at the chance to see what these operations look like under the hood. What can enterprises learn from their advances? At the same time, it's important to remember that the Internet giants tend to focus their technology development and maintenance efforts on a single, hyperscale application.

[See what hardware, software, development tools, and cloud services came out on top in the InfoWorld 2016 Technology of the Year Awards. | Cut to the key news in tech with the InfoWorld Daily newsletter, our summary of the top tech happenings.]

As eBay CTO Steve Fisher put it in a recent interview with InfoWorld, in contrast to enterprise IT, "We're not supporting the business -- we *are* the business." The whole company is focused on this gargantuan e-commerce site's road map and how it should be implemented.

How the monster app lives

Like most mammoth e-commerce operations, eBay uses microservices architecture rather than a monolithic design. Fisher says eBay runs more than 1,000 services, with "front-end experiences that call the APIs for those services. ...There's a back-end

service for shipping, a back-end service for every job." The "experiences" he refers to are five Android apps, all of which call intermediate orchestration services that then talk to back-end services.

Each development team is responsible for its own set of services. When a team wants to spin up a new service, it uses an internal cloud portal to provision dev/test, staging, and production servers.

"We'll spin up a continuous integration environment for them and, basically, you tell us what you want. You push a button, and within our internal cloud across the various security zones that we have, we will give you the infrastructure that you need," says Fisher. "Then basically the team manages that infrastructure."

This service orientation removes unnecessary dependencies and makes it easier to divide the work to make improvements and add new functionality. At an operation like eBay, that sort of activity never stops.

[Learn Java from beginning concepts to advanced design patterns in this comprehensive 12-part course!]

Taking search to the next level

According to Fisher, eBay is in the middle of a "transformation on the technology side."

The most important current initiative goes to the heart of the eBay value proposition from the beginning: the concept of a listing. With 800 million items for sale at any given time, "we've given the people listing those things the flexibility to describe them as they deem appropriate." But that flexibility has always had some drawbacks for buyers. Fisher provides an example:

> *We may have 20,000 iPhones 5s -- 64GB, slate gray, AT&T -- on eBay for sale but we haven't really been able to, in the past, connect them all together so that we knew that these were all actually the same thing ... We have millions of great deals ... but it's difficult to identify a great deal when we don't know that these 20,000 iPhones are actually the same thing.*

The answer to the problem, says Fisher, is to layer structured data on top of free-form data, to be able to determine that all those iPhones are the same. "That allows us to understand pricing and supply and demand, and identify deals and give better recommendations and better search results and make onboarding inventories much easier."

Doing that on top of the diversity of eBay's inventory has been an "interesting technology challenge," says Fisher, to which eBay has applied machine learning and deep learning. How important is this initiative to the business? Important enough that the current generation of search technology required a "multi-hundred-million-dollar investment" to deliver the scalability and reliability that eBay needed. No existing search solution, open source or commercial, came close to filling the bill.

Open source engagement

Not surprisingly, eBay has no inclination to open-source this huge proprietary search effort -- and besides, it's specific to the way eBay works. But like many Internet giants, eBay regularly contributes open source projects to the community.

One recent, powerful example is Apache Kylin, a distributed analytics engine that provides a SQL interface and OLAP (online analytical processing) on top of Hadoop. "We have a ton of data, we do a ton of analytics. We're an extremely data-driven company, and we've been migrating from more traditional data warehousing technologies over to Hadoop -- but we still wanted to be able to leverage existing BI tools," Fisher says. eBay created Kylin for that purpose and ultimately handed the code to the Apache Foundation.

Fisher notes that eBay has a very large Hadoop infrastructure. "In the consumer Internet world, little tiny changes can actually make a huge difference, and we do a ton of A/B testing" using Hadoop analytics to interpret the results, he says. The company has been making big investments to move from large batch jobs to near-real-time to help "make sense of this enormous amount of extremely interesting data." eBay has also plunged into the Hadoop ecosystem, leveraging Storm, Kafka, Spark, and more.

But eBay's open source adventures don't begin and end with analytics. For example, eBay donated a popular suite of JavaScript tools, RaptorJS, to Apache (as the Register). "We take advantage of open source, and we think it makes sense to also contribute to open source to be good members of the technical community," says Fisher. "It helps build our reputation as a technology community, and people all over the world, even our competitors, are using technology that came from us. "

Wrestling with OpenStack

eBay's use of one open source technology in particular is the stuff of legend: OpenStack. Nearly four years ago, InfoWorld broke the story that eBay was using OpenStack to manage a high-volume dev, test, and experimentation environment. Today, eBay is one of the largest OpenStack users in the world.

Across the industry, initial excitement over OpenStack's ambitious "cloud operating system" has given way to disappointing adoption rates and complaints about the project's sprawling complexity. But according to Fisher, eBay has made OpenStack work:

> *I think the secret was that we were just committed to it and we put a large team on making it work. We built a lot of infrastructure where, if it wasn't doing what we needed, we built what we needed and then contributed back … In a lot of areas it was more just like building a platform as a service around it. That's not OpenStack. But OpenStack is a lot less productive if you don't make it really easy for your users.*

I asked Fisher about the problem of "upgrades in place" with OpenStack. New OpenStack versions arrive twice a year and you want to take advantage of improvements and new features without bringing down the entire infrastructure to do so. He offered a frank response:

> *It's not the easiest thing in the world to do, I'll be honest. It's actually a challenge we also just faced somewhat for our own services as we upgrade them and it's something that we're working to get better at. It's not easy. I wouldn't say we've got it down so that major changes coming in just kind of magically show up and everything works fine.*

Going forward, Fisher is excited about the flexibility that Docker containers bring to the entire development lifecycle -- he's a big fan of clustering with Kubernetes, too. "We're rebuilding our broader continuing integration infrastructure to be fully

container-dependent," he says. "It's one of the biggest things that we're doing, ... e this year and probably into next year."

Where does that leave OpenStack? Fisher says that OpenStack will remain the central coordinating system for managing eBay's infrastructure, although "it just will be moving and provisioning containers," he says. But what about the dramatic utilization benefits of containers on bare metal that, say, Google currently enjoys?

> *I think that ultimately we'll be able to take advantage of that as well, but probably not this year. This year we'll be more using containers [for whatever framework or whatever it is that you need to deliver your service on a server ... Someday I think virtualization will no longer be needed for us.*

Want more InfoWorld?

As Fisher observes, virtualization was really created to slice up a single PC, and that's obviously not the world we're in.

Managing scale

How much of what eBay does can be applied to enterprise IT? The most obvious answer is that many enterprises are growing their own large-scale Internet operations, in which case, obviously, they can learn from the Internet giants like eBay.

Inside Netflix, Amazon, eBay, and others, we see the benefits of microservices architecture, not to mention immersion in open source, again and again. At hyperscale, large investments in automation for both developers and operations almost always pay off. The same goes for big data analysis of Web clickstreams to optimize the user experience.

At eBay and others, perhaps the most important lesson is that no matter how "modern" you are, you always need to look ahead, experiment, and prepare for the next technology shift. That cutting-edge technology you adopted early will be legacy in no time, and in hypercompetitive areas like the business-to-consumer space, you can't afford to fall behind.

Eric Knorr is the editor in chief of IDG Enterprise. Eric has received the Neal and Computer Press Awards for journalistic excellence.

Follow