EPISODE 746

[INTRODUCTION]

[00:00:00] JM: Chinese Internet companies operate at a massive scale. WeChat has over a billion users and is widely used as a primary means of payment by urban Chinese consumers. Alibaba ships 12 million packages per day, which is four times the amount of Amazon. JD.com, a Chinese e-commerce company has perhaps the largest production Kubernetes installation in the world. China's rapid adoption of internet services, combined with a large population and a growing middle-class has led to the creation of internet giants on par with the social networks, e-commerce sites and ridesharing startups of the United States.

Last November I attended the first KubeCon China and I saw firsthand how the Chinese internet companies are using open source software to scale their infrastructure. Despite the differences between the U.S. and China, the culture of technologists at KubeCon felt very familiar. In some ways it was just like any other Kubernetes conference that I've attended. There were a large number of engineers trying to find the cutting-edge of technology in learning how to solve the problems that they're facing back at the office.

There were presentations on scaling databases and service meshes and machine learning on Kubernetes and outside of each of these presentation halls there were tables where you could pick up a translation device so that Chinese only and English only presentations could be understood by the other nationality. So if you’re Chinese, you could put on a headset and listen to the English translation and vice versa.

Dan Kohn joins the show today. He is the executive director of the Cloud Native Computing Foundation and we talk about the Chinese internet companies and how they're adopting Kubernetes and the general culture of Chinese technology companies. We also talk about the Kubernetes landscape more broadly and the Cloud Native Computing Foundation, which is the organization within the Linux Foundation that organizes KubeCon. Before joining the CNCF, Dan has worked as an entrepreneur, and an engineer and an executive at several technology companies. He is a seasoned technologist and always great to talk to. We’ve done a couple of different previous episodes with Dan.

Before we get to the show I want to mention that we have a newsletter. You can go to softwareengineeringdaily.com/newsletter to sign up and get a weekly blast of news and information and commentary about what's going on in the world of software engineering. We also are conducting a listener survey. We would love to know what we're doing wrong and what we're doing right. Go to softwareengineeringdaily.com/survey and give us some feedback. If you want to, you can enter your email address for a chance to win some swag, like a Software Engineering Daily hoodie, or perhaps even a mug, or some socks, items from our store.

With that, let's get on with today's episode with Dan Kohn.

[SPONSOR MESSAGE]

[00:03:29] JM: Kubernetes can be difficult. Container networking, storage, disaster recovery, these are issues that you would rather not have to figure out alone. Mesosphere's Kubernetes-as-a-service provides single click Kubernetes deployment with simple management, security features and high availability to make your Kubernetes deployments easy. You can find out more about Mesosphere's Kubernetes-as-a-service by going to softwareengineeringdaily.com/mesosphere.

Mesosphere's Kubernetes-as-a-service heals itself when it detects a problem with the state of the cluster. So you don't have to worry about your cluster going down, and they make it easy to install monitoring and logging and other tooling alongside your Kubernetes cluster. With one click install, there's additional tooling like Prometheus, Linkerd, Jenkins and any of the services in the service catalog. Mesosphere is built to make multi-cloud, hybrid-cloud and edge computing easier.

To find out how Mesosphere's Kubernetes-as-a-service can help you easily deploy Kubernetes, you can check out softwareengineeringdaily.com/mesosphere, and it would support Software Engineering Daily as well.

One reason I am a big fan of Mesosphere is that one of the founders, Ben Hindman, is one of the first people I interviewed about software engineering back when I was a host on Software

Engineering Radio, and he was so good and so generous with his explanations of various distributed systems concepts, and this was back four or five years ago when some of the applied distributed systems material was a little more scant in the marketplace. It was harder to find information about distributed systems in production, and he was one of the people that was evangelizing it and talking about it and obviously building it in Apache Mesos. So I'm really happy to have Mesosphere as a sponsor, and if you want to check out Mesosphere and support Software Engineering Daily, go to softwareengineeringdaily.com/mesosphere.

[INTERVIEW]

[00:05:48] JM: Dan Kohn, you are the executive director of the CNCF. Welcome back to Software Engineering Daily.

[00:05:53] DK: Thanks, Jeff. Happy to be here.

[00:05:54] JM: I had a great time at KubeCon in China and we spoke there at some length, and what was interesting about our conversation was some of the reflections on Chinese software companies and how they compare to those of the U.S. Give me your perspective on how Chinese software companies compare to those of the U.S.

[00:06:15] DK: Sure. I think the biggest thing that China and particularly the companies that are focused around software have going for them is that they don't have all the history that we've built up in Europe and in the U.S. So the biggest example of this is that they didn't spend 100 years investing into fixed line telephones. They jumped directly to mobile phones, and so that's meant that their usage has skyrocketed, but also the ubiquity has enabled things like mobile payments and such that in many ways are far ahead of the U.S. and Europe.

On the software company side, I think there's a newer phenomenon where a ton of companies just don't have the same 10 or 15-year investment into virtualization. They kind of skip right over the previous generation of technology, and so are really in a position to adopt cloud native much faster and are doing exactly that.

[00:07:10] JM: This is a leapfrog effect, I think you called it in China. What are some of the ways that the leapfrog effect impacts how these Chinese software companies operate?

[00:07:22] DK: Yeah. The other term we use is the second mover advantage, that coming later actually can give them a big heads-up oddly enough, and it really is just as extraordinary' thing. I actually made six trips to China in 2018 and a number of the year before and the year before that, and comparing it to my first visit there in 1997, it's just a completely different country. Everything about it has changed.

When you look at these modern technology companies, Baidu, which many people see is the Google of China and Tencent, which is kind of the Facebook and Twitter of China, and the Alibaba, which plays a similar role to Amazon both on the retail side and also with the cloud services that are just moving incredibly quickly and so eager to adopt technologies that have come out of the West. But one of the really exciting trends we're now seeing is that they're also very eager to contribute technologies back that really makes it a global open source community.

[00:08:29] JM: In the United States, we know the dynamics of the cloud providers. Most of the people listening do. Amazon is the biggest, and there's Azure, and there's Google, and then there's everything else. Tell me about the cloud provider market in China.

[00:08:44] DK: It's definitely different. So you have this great firewall of China that the government controls a huge amount of the content that's coming in and out of the country, and that means it's extremely unreliable to try and reach Chinese consumers from websites that are set up internationally. Now, you often can do it. I mean, CNCF.io is an example is one that they often are able to reach reliably. But one of the things that Kubernetes – The CCNF has been able to do for our projects is we just set up the content delivery network with Alibaba Cloud inside the great firewall so that copies all the content from Kubernetes.io to Kubernetes.cm, and that a Chinese site is just loads far, far faster inside China.

But from a high level, Alibaba has a similar role and that they are the dominant cloud company in China to what Amazon plays in the U.S., and then the other major companies like Huawei and Tencent and Baidu have cloud offerings and then there're a number of smaller ones as well, like

[inaudible 00:09:58], EasyStack and many others. But Alibaba is definitely the 800-pound gorilla there.

[00:10:04] JM: Possibly, the largest Kubernetes cluster in the world runs in China at JD.com. Have you learned anything unique about Kubernetes from seeing these large-scale Chinese deployments?

[00:10:19] DK: JD.com is the number two retailer in China after Alibaba and they're running Kubernetes on over 25,000 servers today, and their largest cluster has more than 5,000 servers. I think the biggest thing that I learned from it is just that the power of open source and this idea that this amazing technology is available and that anybody can download it and then learn about it and become an expert in it and administering and roll it out themselves without needing to pay a vendor, without needing to pay licensing fees or anything else is absolutely true. I mean, it is a truly extraordinary thing how JD.com just has seized on Kubernetes over the last four years has been using and upgrading each new version of it and is now beginning to contribute back and look at upstreaming some of their patches and such.

But it was just an amazing process to get to know them and to see that kind of adoption, which was not encouraged directly by the Kubernetes community or CNCF. Although we do have a team in China, but really just does speak to the underlying quality of the software and all the effort that's been put into it.

[00:11:34] JM: How has the open source ecosystem in China historically compared to that of the U.S. and how does it look today?

[00:11:41] DK: Well, I will give a quick shout out here to Professor Liu of [inaudible 00:11:43]. We gave him an award at KubeCon Shanghai last month where she was one of the real pioneers of bringing clinics to China 20 years ago and promoting open source and talking about how it could really help countries like China catch up and get ahead.

So I don't want to by any means make this out as a new phenomenon necessarily, but I do think that it's the case that China' tends to be a very young country, the vast majority of software developers that we run into are younger or have just getting started in the field or in their 20s or

30s, and open source is seen as an extremely standard way of doing things. So they're very used to the idea of going out, I'd say it on the web and looking at different options and finding the best ones.

I think the change in just the last three or four years has been major initiatives by these big Chinese companies that are really interested in having it be a two-way street in beginning to have some of the code that they've developed internally to solve their problems get open sourced and then promoting it to be more widely adopted around the world.

[00:12:59] JM: Do Chinese and American engineers collaborate over GitHub or is there some other social coding tool?

[00:13:07] DK: It's definitely over GitHub. GitHub is not totally reliable in China, but seems reliable enough. Also a lot of the Chinese developers we talked to do have VPMs in order to be able to get access to it if they do have trouble. My perception is that the technical English skills of most Chinese developers are relatively strong, that kind of I want to look at a GitHub readme and understand the gist of it is something that they're very comfortable with, or looking at comments in code or variable names or method names or such. I think they're also often assisted by the pretty decent online translation tools that are out there both Baidu, and amazingly the only Google service that is not blocked by the great firewall is Google translate.

So I think the main answer to the question is GitHub. Within Kubernetes, in addition to GitHub, you have the mailing list like Kubernetes dev. Interestingly, the mail isn't blocked by the great firewall. That still gets through. The only issue is connecting to the website in order to be able to subscribe or unsubscribe or search the archives, and that remains an issue.

[00:14:26] JM: When you and I spoke in China, you mentioned that there is not this Stark distinction between Chinese engineers and American engineers, because often times the Chinese engineers have spent some time in the United States. They work at a major tech company or they go to university and then they might come back to China, they might start a company, they might join a company. What has been the result of that cross-pollination between U.S. and Chinese technology?

[00:14:53] DK: Oh! I think it's just extraordinarily positive, and I really can't say enough about it. My wife was recently a midcareer Fulbright scholar, and if you look at the history of that program, it was all about enabling academics and other kinds of people to travel around the world and understand each other and increase communication, and I think the tech world is essentially making that happen in a very similar way.

So I'll just call out two specific examples in each direction. One is startup in Hungzhou. They were keynote speakers at KubeCon China, which is Xin and Julia are the cofounders of [inaudible 00:15:37] Cloud, and they both worked at Google for a while and then went back to Hungzhou and have a startup that's focused around commercializing Kubernetes there.

But the other direction I have gotten to know little bit, Evan Yu, who is the creator of Vue.js, and he grew up outside of Beijing and speaks spectacularly good English. He says his father made him take English lessons as a kid, which he's now grateful for, but didn't appreciate it at the time. He now lives in New Jersey, near New York, and Vue.js is just an incredible open source project as a competitor to React and Angular, but where both of those have huge companies, Google and Facebook backing them and are two of the 30 highest philosophy projects in open source. Vue.js is really just him and contributors who contribute code that he reviews. So I really just think incredibly highly of having this two-way communication and travel and in interaction back that I think has been so positive for both countries and really for the rest of the world.

[SPONSOR MESSAGE]

[00:16:58] JM: Managed cloud services save developers time and effort. Why would you build your own logging platform, or CMS, or authentication service yourself when a managed tool or API can solve the problem for you? But how do you find the right services to integrate? How do you learn to stich them together? How do you manage credentials within your teams or your products?

Manifold makes your life easier by providing a single workflow to organize your services, connect your integrations and share them with your team. You can discover the best services for your projects in the manifold marketplace or bring your own and manage them all in one dashboard. With services covering authentication, messaging, monitoring, CMS and more,

Manifold will keep you on the cutting-edge so you can focus on building your project rather than focusing on problems that have already been solved. I'm a fan of Manifold because it pushes the developer to a higher level of abstraction, which I think can be really productive for allowing you to build and leverage your creativity faster.

Once you have the services that you need, you can delivery your configuration to any environment, you can deploy on any cloud, and Manifold is completely free to use. If you head over to manifold.co/sedaily, you will get a coupon code for $10, which you can use to try out any service on the Manifold marketplace.

Thanks to Manifold for being a sponsor of Software Engineering Daily, and check out manifold.co/sedaily. Get your $10 credit, shop around, look for cool services that you can use in your next product, or project. There is a lot of stuff there, and $10 can take you a long way to trying a lot of different services. Go to manifold.co/sedaily and shop around for tools to be creative.

Thanks again to Manifold.

[INTERVIEW CONTINUED]

[00:19:13] JM: In the United States software market there is this distinction between "enterprises" that have been around for a long time and startups, and obviously there's a blurry line between those two things. But in the U.S., the enterprises are often harder to sell to. They might have regulatory requirements, startups are a little bit more willing to take risks, and when I go to the KubeCons, I've been to a number of them at this point, I think three. I really like to walk around the vendor booths and just talk to people about their sales process and what has been hard. What kinds of businesses have been easier to sell to? How willing are people to adopt Kubernetes. There has been I think around KubeCon Copenhagen time, I really feel there is an increase in willingness of enterprises to buy Kubernetes. Everybody was looking for a Kubernetes vendor to help them.

I think in China, something similar is happening. Although the re-factoring of older enterprise software is somewhat different, and I think the relationship between the – I think it's called ISV,

like the vendor that will help the enterprise adopt Kubernetes. That relationship is a little bit different than perhaps an Accenture and an enterprise in the United States. Can you tell me about the relationship between enterprises and software vendors in China?

[00:20:37] DK: I am not sure that it actually winds up being all that different. When we deal with more traditional industries in China, like airlines and banks and kind of big business, they have many of the same slow enterprise buying patterns that you see in the U.S. I mean I do think in general, the typical midsize bank just seems to move much faster in China that there is just a sort of – I don't know, almost cultural difference of things expecting to move faster and decisions getting made at the last minute and then getting implemented extremely quickly.

Maybe my favorite example to this is at an end user event I went to I think in 2016. I met a vendor that – It was a hotel chain not targeting to Western tourists or business people. So I wouldn't stay there more like a three or four star kind of hotel and in second and third tier cities, but they had a million hotel rooms across China and their entire backend platform was all running on Kubernetes. At the time, Kubernetes was quite new. It was really just two-years-old. So that was a pretty bold thing for them to do.

Then you absolutely have tons of startups just like you do in the U.S. that are doing that kind of hill climbing process of trying to search out different business models and figure out works. So there're people that are hosting a managed hosting in the cloud and managed distributions, different combinations of services and support and training them. Then you also have these tech giants, which are just really crazy how fast they've grown, the Tencents and Alibaba, Baidu and much. When you visit those campuses, it feels very much like being at a Google or a Facebook or something even down to the metal slides. So maybe those move a little bit faster than their Western equivalents, but those are probably the ones that seem most just transplanted over.

[00:22:49] JM: It's been a little more than a month since KubeCon China, and I know your amount of work to organizing that conference was tremendous. Now that you've had some time to reflect on the conference, what takeaways do you have from it?

[00:23:05] DK: If you wouldn't mind, I might actually just ask to turn it back to you, especially because I think you were in Austin and Copenhagen and Shanghai. I'd love to hear your

thoughts on some of the differences between them. From my perspective, I was thrilled. We had a lot of aspirations going into it of having a real KubeCon and not a sort of slimmed down or dumbed down version, and that we have this aspiration of bringing a lot of Western developers into China and then having a lot of local content as well and really having that two-way communication. I think we really hit all of our expectations along those lines. Then the fact that we did so out with 2,500 tickets was also just incredibly satisfying and just a really positive sign for us.

Now there's an example where on a last minute thing, I believe most it was something like 900 of those tickets were sold the final week, which makes it extremely challenging as a conference organizer to know what to expect, but just does seem to be the way that a lot of decisions get made in China right now.

[00:24:16] JM: My experience was overwhelmingly positive. So from what you said about things just moving culturally faster there, what stood out to me even just coming into the country was the airports and how fast things moved through security. I was just like really pleasantly surprised. There was all of these militaristic speed to which you are ushered through security, which I'm very happy. The faster I can get to my terminal and sit down and read a book, that's better for me.

[00:24:45] DK: I've consistently commented that I feel – I mean, I'm not an expert in it, but that just the Chinese airport security is better than anywhere else in the world. I love the fact that every single person who goes through the metal detector also gets patted down so you don't have this issue of, "Oh! Are they patting you down because you're Muslim or because you look suspicious or something else?" It's just a universal thing, but they do it quickly and efficiently, and they have the staffing and the equipment there that they can do it for a hundred percent of people without really slowing things down much at all.

[00:25:17] JM: I think that a metal detector is less risky than like the irradiation booth that they put you through in the United States. I guess I'm not sure about that, but –

[00:25:27] DK: Yeah, that's my understanding as well.

[00:25:30] JM: Anyway, the conference though was awesome, because – So here's the thing. It's like when I first got to the conference, I felt a little bit – I was like – Okay. So I start walking up to people and started making conversation and very quickly I realized that the fact that I did not speak any Chinese was going to be an issue, because it was just a real big barrier there.

So I end up talking to some – There would be little groups of people where there were people who were not at least on the outside I could tell were not Chinese, and I had my conversations with them and then sometimes there were people who spoke English and Chinese in the conversation. What was really kind of an inspiring moment for me was when I went to the first session where I used a translator. So you had these little translator things. So you walk into the conference session and they give you a translation device. So you put in your ear and so you can hear the live translation that you had translators who are actually translating what the people were presenting.

I put it in and it goes from what to me is completely unintelligible Chinese to Kubernetes in the cloud and the same stuff that I hear at KubeCon North America.

[00:26:48] DK: Just to be clear. We did have the simultaneous interpretation in both directions. So the Chinese attendees were also able to hear the English talks and have that translated into Chinese.

[00:26:58] JM: Yeah. So this is – It made me feel there were some other smaller things, but that was the kind of the biggest thing that made me feel that there is a global hunger for technology and a global hunger for building stuff, and there was enough such an absence of whatever Chinese-U.S. exists in the business news and the people searching the news for conflict. What I saw was a bunch of people who are just builders and that is an international human sensation. It made me optimistic, frankly.

[00:27:41] JM: That's great. Yeah, I definitely feel the same way in the time that I spend there. Now I've often had an additional advantage, which is to have a simultaneous interpreter who walks with me and sometimes the people – I don't speak Chinese unfortunately. So sometimes the people I speak with speak shockingly good English, especially when they've never lived in the U.S. But when they're not, and I can just have that person come over and begin translating,

it really is amazing how much we do have in common and how interested they are in what the CNCF and these cloud native communities offering and then of course vice versa.

I will say just based on your comment, one thing that occurs to me is that – Because for Seattle coming up next week, one of the things that we really do emphasize and talk about is the value of the hallway track. So it occurs to me that we could try and formalize that a little bit more when we go back to Shanghai next year in June 2019 and just have some live simultaneous interpreters who are stationed at different spots along the hallway. No guarantee that they'd be available, but just that you could grab some people and walk over and would allow you to have more of that hallway conversation than you might be able to otherwise.

[00:29:00] JM: Another thing I regret not doing is trying to schedule, because you had a simultaneous interpreter that was made available to me, but I had to schedule it and I wasn't exactly sure who to schedule it with or how to schedule it, but in retrospect, I wish I would've done that a little bit more aggressively, because I was unable to do my walk around the Expo Hall and converse with people because it was very rare that there was both a person that was an engineer that spoke English at those booths. I saw their diagrams. I saw some of their presentations that they had at the Expo Hall, but I was not able to really get the same kind of sense that I would walking around KubeCon North America and talking to the Kubernetes businesses there.

[00:29:50] DK: Yeah. So we would love to try and formalize that with you will bit. I think just on your second visit, you would find it easier to ask for those things and just, “Hey, I’m just going to reserve the person for half an hour, an hour.”

It definitely makes a little bit more challenging to have that kind of serendipity where I do feel like it's often just this hallway conversations or running into someone that could be the highlight of the event. Of course for most of the people there speak Chinese. So they're all able to interact perfectly well.

[00:30:21] JM: So let's talk some about Kubernetes more broadly and the CNCF. The CNCF exists to make cloud native ubiquitous. You want to be a neutral home for cloud native projects

to build cloud native solutions. How has the CNCF adapted as an organization over the last year?

[00:30:42] DK: We have been trying to deal with hyper growth. So CNCF is the biggest and fastest growing open source organization ever. When we got started three years ago, and that's literally three years ago this week, we had 28 members in the organization and zero projects, and then soon after Kubernetes. We now have 346 members and 31 projects, and it's definitely just much more complicated. There's a lot more things going on. There're many more services that we're offering to our projects and to our members. So it's definitely taken some scaling, and I think many of your listeners have experience with startups and know that scaling can sometimes be a little awkward. But it's also been a really fun ride.

[00:31:40] JM: The first two projects to graduate from the CNCF were Kubernetes and Prometheus, and a third project that recently graduated is Envoy, the service proxy created at Lyft. Could you use Envoy as an example case study for how projects get shepherded through the CNCF?

[00:32:00] DK: Oh! Absolutely, and I think it really is a great example and a particularly compelling one. So Envoy was originally created by Matt Klein Lyft and it was designed to solve a very specific set of issues within Lyft. Fascinating, Lyft does not use Kubernetes internally today. So when Envoy was developed, it was not developed with Kubernetes in mind or certainly not solely with Kubernetes in mind. Like a lot of sort of successful fast-growing companies that have come up over the last 5 or 10 years, they wound up deciding that they needed a container orchestrator and built their own, and that's the pattern that we've seen I guess in China at Alibaba and Tencent. We've seen it at Yelp, and Spotify, and many others.

One of the sort of background trends that's happening right now is that a lot of those companies are saying, "Hey, we're having to invest all these resources just to maintain our internal system. Why don't we go ahead and begin the investment to migrate over at Kubernetes?" I believe Lyft is going through that process as well.

But anyway, Matt developed that software. The one way of thinking about Envoy, a very limited one, is that Apache as a Web server provided ingress to your backend applications. Ingress,

meaning connection to the internet for – And that was sort of the dominant standard for a decade. Then nginx has in a lot of ways planted it and become the dominant standard in this much more dynamic capabilities and higher throughput than Apache.

In a lot of ways, Envoy is really targeting that same trend. The space is often called service proxy, in which it's looking at not just say static webpages out to the internet, but the idea that you've split your application up into microservices. You have dozens or hundreds or thousands of different services that all need to communicate to each other. Then Envoy allows them to do that and for you to track it and secure those connections and provides a lot of other functionality as well as connecting the ones that you want to out to the internet.

Anyway, one other fascinating part about Matt that if you want to link in your show notes, he wrote a really interesting medium blog post 18 months ago where when he came out with Envoy and a ton of people were interested in it, he got a bunch of proposals from venture capitalists who said, "You should quit Luft and go do a startup," and they would fund him to create an Envoy startup.

He thought very hard about it and then came back and said, "The issues with Envoy is it needs to be used so tightly integrated into the rest of your cluster and software and microservices that it just doesn't make sense for this to be provided as a software as a service or as an open core, or one of the other kinds of standard business models.

So he both was interesting in having it be open source and having all the functionality be open source, not just an open core model. Then he was also interested in having the most widespread usage and adoption possible. So he and Chris Aniszczyk from CNCF were able to convince Lyft to contribute to CNCF. When that contribution happened, it's not like it stopped being Matt's project. He and the other maintainers of Lyft, Constance and I think there are seven or eight of them now, are still the ones who are reviewing the pool request and in charge of the project, and everything that they were before, it's just the kinds of services that they might've gotten from Lyft or not have gotten like some marketing help are now being provided by CNCF.

So that came in as an incubating project and has matured over time and just gotten tons of incredibly high-quality pool requests and contributions, different features and bug fixes and such

over the last year and a half. I believe now, at the very least, Amazon, Microsoft and Google are all spinning out commercial. Amazon announcer theirs last week, or seriously considering it. As well as just dozens and dozens of other big-name organizations that have adapted it.

So the aspiration for graduation as supposed to it being this sort of incredibly challenging milestone that it's just a very natural step in the evolution of a project and just more of a signifier or a signpost to say, “Yeah, this project is incredibly widely used. It’s been very carefully vetted and is highly respected, and now it is a marketing signaling mechanism to say that we think enterprises of all sizes and sort of aptitudes for openness to risk should really be adapting it.”

[00:36:47] JM: Nginx has been used for so many different purposes. It's used as a load balancer to the front your entire infrastructure, or it's also been used – Like I spoke to Kong recently, and Kong uses nginx as their sidecar proxy for deploying a service mesh. What have people built on top of Envoy? Has it been used for as many kinds of purposes as nginx?

[00:37:13] DK: Oh, definitely. Yeah, it's really in exactly the same space. I will say that nginx remains a fantastic software and tons of people use that in production on Kubernetes clusters and other cloud services. But probably the thing that Envoy is best known for in terms of enabling sort of next-generation software is Istio. So Istio has over the last year have gotten a ton of excitement and interest around this concept of service meshes.

Again, Matt Klein wrote a really good medium post that I’d recommend that you link to, where he described Envoy and that role of a service proxy as being the data plane that's allowing all these different services to communicate together. Istio as being the control plane, which in more sort of traditional systems a decade or two ago would essentially be a role played by the system administrator who would be manually logging on to different services and moving things around, or configuring them, or pushing new versions, doing blue-green deployments and all that kind of thing. The idea of Istio is to try and automate a huge amount of that functionality. So I don't know that everybody who uses Istio realizes this, but istio in fact is built on top of and relies on Envoy.

[SPONSOR MESSAGE]

[00:38:37] JM: This episode of Software Engineering Daily is sponsored by Datadog. Datadog integrates seamlessly with container technologies like Docker and Kubernetes so you can monitor your entire container cluster in real-time. See across all of your servers, containers, apps and services in one place with powerful visualizations, sophisticated alerting, distributed tracing and APM.

Now Datadog has application performance monitoring for Java. Start monitoring your microservices today with a free trial, and as a bonus, Datadog will send you a free t-shirt. You can get both of those things by going to softwareengineeringdaily.com/datadog. That's softwareengineeringdaily.com/data.

Thank you, Datadog.

[INTERVIEW CONTINUED]

[00:39:32] JM: As I've talked to some other people who are leading projects that are involved with the CNCF, I'm starting to get a picture for how the Kubernetes ecosystem not – It doesn't just improve the deployment of your containers, but there's a whole vision around improving administration. So, for example, this idea of having the control plane layer gives you a place to push out things to your sidecar proxies, that might be Envoy, and those can communicate with your different services. Then you have projects like Spiffy for defining the workload identifiers. Then you have projects like – If I'm getting this right – No. I think Spiffy and Spire are for identifying the workload objects, and then you have open policy agent, which can be used to define policies and then you could use something like a control plain like Istio to push out those policies, and you could have some caching system built around your pods. So you see just a whole vision for how people can manage infrastructure more intelligently. Can you talk about that overall vision for how administration can be improved by this stack?

[00:40:54] DK: Yeah. I think you're also a Hacker News reader, and there's at least an article or two and sometimes many more every week on Kubernetes, and you can never get to more than 10 posts or so before someone has to chime in and just say, "Yeah, nobody should be using Kubernetes unless at your Google scale, and it's just so complicated and it has so many of these new concepts and so much overhead for you to learn."

Really, for almost anybody, you should just be able to get by with a couple bash scripts. Of course, if you are a batch expert and you just have a small website or something, and you can even probably get containers running and just have a couple batch scripts that move back and forth between them. But I do think that Kubernetes is an expression of this larger trend of computing that's been called DevOps SRE or the pets to cattle. There's been a kind of an intriguing process that took place on the public technical oversight committee mailing list over the last year where we wanted to create a new definition for cloud native and what we meant by that.

What's pretty interesting about the definition we came up with – I mean, it's five sentences long. So I'm going to read it to you, but is that it doesn't include the word Kubernetes anywhere in it and it doesn't even include the word orchestration. But I do feel like it's trying to get across this idea that it is, and I apologize for the buzzword implosion, but it is a paradigm shift in how you think about computing and how you think about your application and then how you deploy it and work with it.

So this is the CNCF cloud native definition version 1.0. Cloud native technologies empower organizations to build and run scalable applications in modern dynamic environments, such as public, private and hybrid clouds. Container service meshes, microservices, immutable infrastructure and declarative APIs exemplify this approach. These techniques enable loosely coupled systems that are resilient, manageable and observable combined with robust automation. They allow engineers to make high-impact changes frequently and predictably with minimal toil. The Cloud Native Computing Foundation seeks to drive adoption of this paradigm by fostering and sustaining an ecosystem of open source vendor-neutral projects. We democratize state-of-the-art patterns to make these innovations accessible for everyone.

[00:43:39] JM: This whole space is getting so big, and like you said, hyper scale. I know that the KubeCon Seattle that is coming up next week sold out pretty quickly, and it's like 7,500 people or something. Are you anticipating a day where the KubeCon CloudNativeCon appendage gets broken up into KubeCon and then separately Cloud Native Con? Like maybe these are two distinct environments.

[00:44:06] DK: I hope not. I mean, anything could happen. I feel like that change would really hurt the community in a lot of ways. One of the reasons is that not all that many of the talks at the conference are just about Kubernetes or just about Kubernetes internals. They are often talks about, “Here’s how we use Kubernetes in this new environment, or here’s how we use Kubernetes with this new open source project that’s part of CNCF yet, but we think is really intriguing, or here's this edge case we encountered.”

Some of the talks are just about all those other projects, but one of the sort of underlying concepts of the Cloud Native Computing Foundation, of our parent, the Linux Foundation, and I would say of open source in general is that most interactions in the world are not zero-sum games. Now, it is the case that we only have a certain number of rooms at KubeCon CloudNativeCon, and so almost by definition if a talk is going to be solely about Prometheus, then it cannot be solely about Kubernetes.

But I think that's a really sort of artificial and constrained way of looking at it, and that most of the attendees in particular are not just focused on one project to the exclusion of all others. They’re really focused on their business and the technology problems they’re running into and they're looking for the set of solutions or the ecosystem of solutions that they can learn about and then implement themselves. So I do think this concept of a positive sum game that by combining together we’re able to have larger crowds, more funding, more sponsors and therefore things like more diversity scholarships, more fun events and other kinds of stuff is a really promising one. Now, that said, I totally agree that KubeCon CloudNativeCon is now going to be a very large conference. So it’s actually 8,000 that we have sold out.

So one of the pieces that we are doing is also offering a set of side conferences, and so for example, on Monday will be the first ever EnvoyCon, as a pre-day conference before to KubeCon. In the last couple of years, we've sponsored the Prometheus team, has run PromCon in Munich and they’re planning to do that again in 2019, and there’ll also be like a GRPC Con and such.

So I definitely don't want to come off as strident on the subject. I mean, I feel like there're a lot of different things that we’re trying and in different modes that we want to experiment with. I mean, another huge part is the Meet Up South Air, where we now have over 150 different Meetup

groups all around the world. They've had more than 1,600 meetings in the last three years. So that's just a way that very widely gets a lot of this messaging out there. Then we are looking at some other kinds of conferences that we could do that might reach different groups and different regions.

[00:47:11] JM: I think what you're encountering is something that AWS Reinvent incidentally walked into, which is both the fact that there is massive expansion in people who are interested in software engineering, and also the appeal of the cloud in terms of how accessible it is. AWS obviously pioneered that. So in some sense, it makes sense that they walked into it. But KubeCon CloudNativeCon is kind of extension of that, because it's like, "Okay. Here it's not just kind of talks that are anchored around AWS services. As inspiring as AWS services are, this is more about the entire cloud environment, which encompasses a whole lot of companies and also a whole lot of open source projects." So it seems like the scope of where it could grow to is quite large.

[00:48:07] JM: I'd also emphasize the private cloud aspect of it. I mean, our — and hybrid cloud for that matter. Our aspiration is that if you as an enterprise are operating and maybe you have a commercial distribution on your own hardware and in your own data center and then you're also operating in a couple of different public clouds, that you can go to each of the conferences of each of those vendors, but it should also be feasible to come to KubeCon be able to book meetings with some of the top architects from each of those companies. So we do think there's some role for us to play as that vendor-neutral, multi-cloud, hybrid-cloud kind of environment.

[00:48:50] JM: I know we're running up against time, but since we're on the subject of multiple vendors, I talked to Bassam from Upbound earlier this week and we were talking about Kubernetes Federation and the vision of multi-cloud, and I think people want to have access to cloud resources across any cloud, because I think the clouds are going to get more and more differentiated. Obviously, they'll overlap, but you want to have to choose the services that you use based off of what cloud your monolith is anchored to.

Explain what Kubernetes Federation is and why it's important, and I guess what your vision is for what a multi-cloud company would look like.

[00:49:34] DK: Sure. I mean, I think the first thing that I would just say is that if you're looking for absolutely rock solid deployment hosting running of your software and not just of modern microservices, but of the monoliths and everything, then Kubernetes is a fantastic choice, and we have dozens of hundreds now of examples of companies doing that. Federation, I would say is really on the leading edge of new development and just isn't quite there yet. So, in particular, the multi-cluster group in Kubernetes started with a Federation V1 approach and then they went back and based on some of the learnings and some of the challenges they run into have gone forward on a Federation V2.

So I think there is an aspiration that a lot of end-users have today that just like they can build a bunch of containers and write a bunch of YAML and say, “Okay, this container needs to run on at least four machines, and this one needs a GPU every time it runs, and this Envoy sidecar needs to be deployed in every pod and those sorts of levels of interaction that they would love to be able to say, and I want you to decide whether it should be running on my private cloud or which of these public clouds based on the current spot pricing or based on the deal that I have or based on the need for some special service that they’re offering.

I think my sort of serve high-level understanding of the space right now is that functionality is not quite fully baked yet. So I think there're a lot of great work going on in the multi-cluster sig that definitely encourage people to get involved in. Those are all public meetings and interactions, mailing lists and GitHub repos. Then, unsurprisingly, you also have folks like Upbound that are saying, “We think that we could actually implement this best at a – I’m going to say higher layer than Kubernetes. It’s a little misleading whether it would be better as a metaphor to describe as a lower layer, but I’ll definitely say a different layer than Kubernetes, that we have this new software cross plane and we think that that can be a tool for provisioning your clusters in multiple different clouds and then ensuring that they’re run correctly and such.

I mean, I think Bassam would say that cross plane is very new software. It's promising. A lot of people who looked at it were quite intrigued. I don't think people are running it in production yet, but I'm thrilled to just see the level of ferment and innovation that is going on in that space, that I don't think it's likely that most enterprises are going to be interested or even able to standardize on just a single cloud provider.

I mean, obviously, if you're just a little startup and you're just getting started today, then you very likely just want to pick one cloud provider and focus on building out your product and finding product market fit and not worried about saving the last penny here and there, and it's quite likely in fact that a hosted Kubernetes service would be a great choice for that. But once you get to certain scale, and specially for, say, the global 5,000 enterprises, where one of things that we see again and again is so many of them are the result of acquisitions or the result of mergers and are constantly doing new acquisitions. So even if they would like to be single cloud in some principle, in reality they're not.

So in that environment, like I think it's great that there is this innovation going on. Again, I would sort of describe it as hill climbing as different enterprises and startups and such trying to find the right set of approaches that make that work, but what's great is that it's really taking place mostly today in an open source ecosystem where everybody's going to be able to get the advantage of their work.

[00:53:23] JM: You run several companies before and you've worked in technology for more than 20 years. Can you close off up by giving us a perspective on how company building today from a technological point of view? How does that compare to what it was like in the 90s?

[00:53:40] DK: Oh! I mean, big picture, it just seems awesome. I mean, there's so many just super annoying things that we needed to worry about 25 years ago. I mean, just as a random example. I set up the first music store on the web in 1994, and I have this memory of I can't even quite remember why we need to recompile something, but we wanted to use [inaudible 00:54:05], and there's a problem on it and we are using Ultrix at the time. This makes me really sound old. We needed to get the Sun compiler and we're willing to pay for it and there's no Sun salesperson that was awake in the U.S. We had to reach out the Japan office, and it was this incredibly expensive international phone call and just ridiculousness like that that's so time-consuming, memorable I guess in retrospect, that just today would be a home brew command on a Mac or an on apt-get command on Linux and you would get the software that would just run.

Now, that said, I mean I think things have improved dramatically. I would also say, certainly, one of the big takeaways I got from the keynotes in the sessions I attended in Austin in Copenhagen

was a consensus among almost all the leaders in the Cloud Native community that the ultimate winning solution was very likely to be Kubernetes underneath, but that the standard way that most developers were going to be deploying their applications on top of it was not what people were are doing today, which is generally building a Docker container and then writing some YAML and pushing that up with kube-cuttle.

That there is now – There was no consensus on what the correct approach would be. Lots of different people were proposing lots of different technologies and different ways of integrating with CICD and other kinds of things, and I expect in Seattle next week to hear many more ideas on that. But I genuinely do think that we're standing on the shoulders of giants and that Kubernetes is all based on the work of Linux and is in many ways it's just exposing features and capabilities of Linux or containers are.

Of course, Linux is based on UNIX and a lot of the work in computer science that had gone before that. I think there're hundreds of opportunities left for making developers' lives easier. In reality, software development today remains incredibly frustrating activity for a lot of the time and people talk about what a small percentage of their week they really feel like they get in that flow mode of writing useful code and seeing it interact and move forward the way they want, versus just annoying setting up this IDE or setting up this container or Googling this strange error message and finding the answer on Stack Overflow and all the other piece. So I definitely feel like we're very much in the middle of a long story right now.

[00:56:41] JM: Dan Kohn, thanks for coming back on. It's always a pleasure to talk to you.

[00:56:44] DK: Yeah! I hope you have a great time in Seattle, and I will give a quick pitch. We're sold out there, but we would love to see many of your listeners in Barcelona May 20th-23rd of 2019. We think it's going to be just also an amazing event.

[00:57:03] JM: Wonderful. All right. Well, hopefully I'll see you in Seattle.

[00:57:05] DK: Great. Thanks, Jeff.

[END OF INTERVIEW]

[00:57:10] JM: This podcast is brought to you by wix.com. Build your website quickly with Wix. Wix code unites design features with advanced code capabilities, so you can build data-driven websites and professional web apps very quickly. You can store and manage unlimited data, you can create hundreds of dynamic pages, you can add repeating layouts, make custom forms, call external APIs and take full control of your sites functionality using Wix Code APIs and your own JavaScript. You don't need HTML or CSS.

With Wix codes, built-in database and IDE, you've got one click deployment that instantly updates all the content on your site and everything is SEO friendly. What about security and hosting and maintenance? Wix has you covered, so you can spend more time focusing on yourself and your clients.

If you're not a developer, it's not a problem. There's plenty that you can do without writing a lot of code, although of course if you are a developer, then you can do much more. You can explore all the resources on the Wix Code's site to learn more about web development wherever you are in your developer career. You can discover video tutorials, articles, code snippets, API references and a lively forum where you can get advanced tips from Wix Code experts.

Check it out for yourself at wicks.com/sed. That's wix.com/sed. You can get 10% off your premium plan while developing a website quickly for the web. To get that 10% off the premium plan and support Software Engineering Daily, go to wix.com/sed and see what you can do with Wix Code today.

[END]