Constellations Podcast

SPECIAL EPISODE 3/3 – Big Data for Satellites, Deep Learning and Growing Security Threats

Speakers: Loic Coulet, Technology & Products Manager, Kratos, Bob Potter, VP Technology, Signal and Ground Systems, Kratos, Jay Finnigan, Director of Network Products, Kratos, Anthony Semiao, Chief Solutions Architect, Kratos– 36 minutes

Bob: Thank you so much for staying this long. I know that we are the last session, and that we are the people who are between you and your lunch, and your travel plans home. So, but hopefully we'll find out, create a panel that's of interest. Because big data really is I think something that's coming if it's not already here. And it really is the future. A lot of our discussions this week have been circling around cloud and data analytics and so on. So hopefully it'll be of interest. So I'd like to introduce the panelists.

On this panel we have Jay Finnigan from our network management group. You can come on up. Anthony Semiao who's also from our network management site, more on the federal side. Our Federale. And Loic who is our product manager for our big data solution, Skyminer.

Thank you guys.

Anthony: Bob who got promoted so he doesn't have to answer any questions.

Bob: Yes. So let me start then with one of the questions. So the big data analytics market is growing like crazy, at least where I live in Silicon Valley. There's a big data company on virtually every street and every town. It's growing like crazy . I'll ask the panelists, what do you think are some of the key drivers in the growth of this big data?

Jay: I think, personally, we see big data impact in our lives more and more every single day. I think about when I browse something on the web and a minute later I'm on Amazon and that's the first thing I see, with those advertisements being pushed to me. We're all recognizing it in our online presence and our social networking, just how intelligent the applications and services that we're using are becoming. That's all driven by machine learning, big data, and artificial intelligence.

It's proliferating to our personal lives and it's coming into our professional lives. Everyone these days knows what big data kind of means. But 5 years ago we didn't talk about it too much in this industry. But now we talk about it every single day. So I think it's proliferating through all aspects of our lives, personally and now professionally.

KRATOS

Bob: Thank you. Loic.

Loic: From my point of view there are several aspects. First of the all is the increase of the data usage. Data usage has increased exponentially during the past 15 or 20 years. Internet, video, image data, various kinds of data. So there was a need for this data storage. New technical solutions, cloud computing, hives of storage. And then the need to process all of the data and to extract useful and accurate information out of it.

There has been tremendous activities and machine learning with the new rise of neural networks and deep learning and so much thought has been added to this topic that now big data is everywhere, machine learning is everywhere. We have machines that become already intelligent on some very narrow aspects but that beat the humans in every aspect. Google for example with Go, artificial intelligence, is very impressive proof of that.

Everyday this artificial intelligence is bringing challenges in day-to-day life and applications. It might be that it's only the beginning. Next step would be to figure out how far this artificial intelligence will go. Autonomous vehicles already exist and already are driving, especially in California. That's just exploding like crazy and I believe it's only the beginning of this era of machine learning, artificial intelligence. That I say this would be a big job to the next 10 to 15 years. Before maybe machines take over humans.

Anthony: You mean I might be able to retire someday.

Bob: Google are designing their own chips by the way, you know that right?

Loic: There is concern about that because artificial intelligence is a threat and has many improvements as well. Taxicabs, what do they become when autonomous vehicles are driving everywhere?

Bob: My kids actually figured out how mess with Waymo because autonomous vehicles have to obey all the rules, right? So in California we have a rule you have to be 3 feet from anybody on a bicycle. So they know that when a Waymo is behind them, they can start to move out of the gutter a little bit and it messes with the vehicle behind them. So they have lots of fun with that.

Loic: Yes. But it's very important to keep the human in the loop because we also share the concerns and the human has always to be in the loop and artificial intelligence has to be a tool. That's my professional opinion.

Bob: Thank you. Anthony.

Anthony: It's hard to add more to what the two gentlemen have already said. But I think if you look at big data and analytics overall, it's an industry that feeds itself. Right? Every time we come up with a new algorithm or extract a piece of useful information, somebody goes, "Ah ha. And if I combine it with this other piece of information, I have yet another piece of information." And that becomes useful. The data kind of feeds onto itself and then it just drives technology.

Before, everything had to have a relational database, 20 years ago. If it wasn't a relational database, it wasn't right. Now it's like that's the old thing. If you've got a relational database in big data, you're too slow. So we've got to do ... been in memory databases and all these things. As soon as we come up with something that we can't process fast enough, well we'll just put another terabit processor in the cloud and crunch it with that and come up with yet the next piece of machine learning or technology. It constantly feeds onto itself.

I agree with Loic, that we don't know if this is the beginning or the infancy, but the end is certainly way, way out in the future. The folks that can look at patterns and datasets and derive the next set of algorithms are the key to the workforce when it comes to looking at data for the next many, many years.

Bob: Okay. So on that thought, what do you see as the biggest opportunities for big data analytics with satellite operatives in the next 5 to 10 years?

Anthony: I'll start even though I come from the network management side and the terrestrial side, I think all the data that the satellite industry around imagery can bring to bear on earth observation type technologies is a huge place to look. I was reading somewhere in preparation for this, that the folks that are in the natural world, if you will, ... mining, oil and gas and then on the government side ... the folks that consume the majority of the imagery data, they're looking to grow in this area by like 60% year on year. It's obscene.

I think when you take what the satellite industry can offer in terms of imagery and combine that with other datasets, we go back into that feeding problem where we're just going to find new pieces of information we never thought before.

How many people really don't know where they're going? Or when they get there, they're surprised at what they find when they reach their destination. I know on the way here this morning, I drove an hour and 40 minutes from my home in Virginia over this way, and by the time I got here I felt like I had toured all of Annapolis. Because Google knew where I was going, I was using Google maps and they're like, hey you want to visit this place and hey you want to visit that place. So they were feeding me all the way here while I was sitting on 495.

Bob: Thank you. Loic?

Loic: The different use cases and I believe in different challenges depending on the application. Satellites that provide image data have different concerns than communications and so on. Big data there are some challenges in image data with IO and IO resolution, faster and faster image providing and processing of the data. The storing the data for a larger amount of time here is already a challenge for them. More than complexity in size of the data, from my software engineer perspective, complexity of systems is becoming probably the new challenge. HGS satellites, android off beams, and dozens and thousands of services everywhere, all over the place. And you have to optimize for this complexity. And so people that work in optimization problem in algorithmic, it's always, always very complicated. Because you end up with a number of possibilities that is next to infinity. This is where I believe the real challenge is, is to touch this complexity. To have systems that are smart and efficient and powerful enough to address this complexity and to provide the errorless to the human beings to the operators that are behind and operating the systems.

Jay: I'll add a little bit more to what Loic was saying. I think we're going to see an explosion of big data and machine learning. Very specifically when it comes to orchestration of the networks that you'll all be operating in and providing to your customers. As we virtualize the spacecraft and we virtualize the ground, we'll be virtualizing the network as well.

Big data and machine learning will be crucial there for planning and provisioning and monitoring those very complicated network that will span the terrestrial and space.

Bob: Okay. So, picking up on the point from Anthony where we talked about imagery. Imagery seems to be one of the big growth areas in satellite right now. So do you see a time when basic imagery and analysis will be a commodity? We'll actually get it on our phones and make use of it?

Loic: I can start. It already exists.

Bob: Yeah. Give us an example right? It's already there.

Anthony: The downside is they haven't found out how to make good money on it. Right? They give it away for free. The data imagery is ... what do you want to look at? Do you want to look at rainfall patterns? Do you want to look at ... you name it you can get that information. I think the key is how you exploit that in the future to make even more money off of it than what they're doing today.

Bob: So one of the things with big data is ... people have concerns about security and how do we keep that data secure? So what do you think are the biggest security challenges for satellite operators and what are the security challenges they face with respect to big data?

Anthony: I'll jump in there. I think the biggest challenges will be focused around the number of threat records that the operators will start to see. I think when you provide a fixed capacity, in the past your user base or your customer base was relatively small. As you start to provide end-to-end services and bring on many more end users, the number of entry points into your networks and operations is going to explode. If we look at the Telcos, if we look at some of the social network or other consumer big data companies like Google and so on, they have very large and focused operations on security. It's one of the biggest things that they focus on.

The satellite industry's going to have to take that very seriously and make large investments in analyzing the attack surface that will continue to increase over time.

Bob: Okay. Anybody else?

Loic: I can say if you want a bit more about the computer science aspects of the security. It turns out that once I spoke to a security officer in bank trading. It's not typical but what this person said is that security in computers simply does not exist. Everything is interconnected and it is an illusion of security that you can build as many fences as you want. It's never secure. Especially that's for the last 50 years. The computer industry has focused mainly on performances, features, and security was not so much a concern years ago. Now everything is interconnected. Everything speaks to everything and the complexity of the networks and of all of the systems does not allow a human being to monitor even a single machine properly.

I think it needs a change of paradigm in the world computer industry. We have attacks on side channels, we have attacks on security at the CPU exhibition event and every day there are new kinds of threats that nobody ever thought about until recently. And those threats were existing for the case and now it shows that the world computer industry rethink that from scratch and think security as a holistic approach. What is a secure computer or secure processor? A secure operating system? A secure network, a secure crowd infrastructure? And rather than building fences, think security at the core of the system. And I hope it will be able to try to answer these questions and to avoid getting too much of extinguishing the fire around the place.

Anthony: As Bob said, when he introduced me, I spend a lot of time with our federal customer base. Our federal customer base loves the word encryption. We encrypt everything and then we encrypt it again. And if we're lucky we encrypt it a third time. The problem with big data and encryption is, what does encryption do? Slows it down. What is the enemy of big data processing? Slowing things down. So the techniques ... this goes to exactly what Loic was saying ... the techniques we have today just aren't innovative enough. Especially around things like securing massive amounts of data and using things like encryption. Because it defeats the purpose of processing big data.

The other thing that I think we're going to see a lot of ... that isn't about what fences you build or how well you secure things ... people with all the proliferation of end points and things that are generating data, it becomes very easy to generate fake data. So if someone wants to really toy with your data analysis, they don't even have to really break in. They just have to act like a sensor on your network and generate amounts of data that's an anomaly to your overall dataset. And they're going to skew your results, they're going to skew your algorithms. The innovation has to be around, how do we find patterns in all this information that is anomaly? And pull it out and allow people to look at those patterns and move them to the side or get them out of the core piece of the information they're looking at.

Bob: So pulling up, I've got a question from the audience on that. Does the data need to be secure or just anonymized?

Anthony: I think it's both, right? There are datasets that need to be not only made anonymous ... look at the healthcare industry here in the United States and what we call PII, personal information or integral information or whatever the acronym stands for ... that data when you look at it as a whole, if it’s about my medical history, it is off-limits in a big data set as a chunk. But there's probably key pieces of information in that data that aren't necessarily the fact that I have a medical condition or that I'm on these medications that are ... if I was anonymous, there's probably really good information in there ... that these drugs help these conditions, so forth and so on. So it has to be anonymous but at the same token, if I go back to my federal customers base and we've got a dataset that is talking about the location of all of our intelligence operations and operatives and our data sensors, that has to be not only anonymous but it has to be secure. Because we can't let other state actors get into that information. So it's both and it's a difficult problem.

Loic: I agree with you especially anonymous data. As a dataset goes out and there are more indicators, it's not anonymous anymore. It's not anonymized. For example, your cell phone may report anonymous usage of your position and from this position one can infer who you are because you have been to your work but at a point in time, then back home. Even if it does not know who are, it

will cross-correlate this information with data from credit card information so they know, not only who you are but what is your bank. As the data goes out and you have more indicators, it's not anonymous anymore because you can correlate the facts and you can find whether it's a person, a human being, or a system, you can still find a lot of information. So it has to be secured and anonymized.

Jay: Okay. I agree with both of the other panelists here. I spent a number of years working in Kratos' services network where we have a lot of RF centers around the globe and we're monitoring quite a bit of spectrum. And even those datasets, having them anonymized ... we can still over time recognize patterns. The patterns in the anonymous data give a very good indication about what the activities were. So we can glean from the very anonymous data that very specific tactical things were occurring in locations and periods of time. So anonymization isn't sufficient in itself. Because in the larger dataset you can draw many conclusions that the individual stats observed data didn't indicate. But you can learn very valuable strategic tactical information out of the patterns that you can discover.

Bob: I would agree. In a sense it's already there right? Because I know that when I'm browsing the web I like to go to the BBC website to look at the news. Because I believe that they are a news entity. Fair, free and impartial. But I'm getting ads for Prop 8 in California. So they already know that I live in California. What the hell do I want to know abut kidney dialysis for? But those ads appear in the middle of my news scripts. So it's already there right?

Following up on that sort of thing is another question from the audience. Do you think that human intervention in networks poses a higher security threat than AI? Humans in the loop makes things more secure or less secure?

Anthony: Both. Both. It is without a doubt a threat. I work in one place here where 83 ... I think the total ... probably a little bit higher now ... we'll say 90 people have access to the door to get into the room where the PC is and they need three manners of tokens and a password. They're still worried about the security. And it's because the person in the middle, if you can be bought off, you are a threat. What's the first question on the United States application for a secret clearance? It's, "Have you ever filed bankruptcy or do you have financial problems." So yeah, the human in the loop is ... both the person securing it and both the security threat. I don't know that there's a way around it though.

Loic: I agree that the human is always the weakest spot, but AI is not reliable. So you cannot have AI without human in control. At some point, you have to trust humans, you have to find trustworthy persons because there's no other way to have a trustworthy system. AI poses other issues. They can be completely wrong. They can be perfectly right 99.999% of the time and one time over ten

thousand, it could be completely wrong and nobody knows why. So you need to have this human being to be able to be in control at the end. Especially as AI become more intelligent, you have deep learning and nobody understands how the model works because it's getting too complicated inside. It just works. It just works, but nobody knows how and why it works inside. So you need someone in control. Someone in charge.

Bob: It's funny because ... I'll just pass this anecdote. We were in a meeting last week with some really smart people in Silicon Valley, and they actually told that they don't actually write code anymore. They actually have the machines write the code and they just have these machine learning things. Incredibly smart people and they don't write a line of code anymore.

And it's going to wind up in our cars.

Anthony: So built by the lowest bidder.

Bob: Yes. Yeah technically ... lowest price. Yes.

So another ... I think maybe by way of explanation of some of the phrases we've been using. There's a question from the audience, "What is deep learning and what is the implications of satellite industry?

Loic: I can take that one. Deep learning is the new era of neural networks for computer science. Neural networks are trying to replicate in a very simplified way because there are some external factors. You have neurons and neurons are interconnected with other neurons. By learning, the connection between the neurons will change. You have some connections that will become stronger, some of them weaker.

Neural networks started in the 1950s. It's very old computer science area, artificial intelligence, and it never took off. From time to time, it went up then fell down because it was not working well enough. For instance, about 15 to 20 years, the rise of big learning changed those things. They discovered that bringing dense layers of artificial neurons and chaining those dense layers all together- it's why it's called deep learning because it is a deep layer after layer, a deep capacity of learning. It used to have very good results. The learning process is very long. They have to feed the mothers with millions of data. If you want to recognize a dog you have to feed it with millions of pictures of dogs and cats and whatsoever. But then the model recognizes exactly what is a dog and what is not a dog, what is a cat, what is another animal. And it works very, very well. There are different way of tutoring the different layers, if you want to recognize shapes, if you want to learn from sound, if you want to reproduce ... because using the deep learning, you can also do inference of information and

the creation of information. Or you can construct an image. Building new pictures out of pieces of information. They do that very well. The only concern is that nobody really knows how it works inside. But it works.

Bob: So hopefully that explains what we think of deep learning. So we'll move onto another question. What type of space or network data is best suited for big data analytics? What are your thoughts on that?

Anthony: Oh, all of it. Keeping the data around whatever you're sensing ... whatever information you're collecting could potentially be valuable in the future. The Googles, the Amazons, those very large corporations don't throw away a bit of data. You collect it and you keep it, because you never know what questions you're going to ask in the future. One of the principles of data science is to identify the question you're asking. I don't think any of us have all of the questions right now, we don't know what the new questions will be down the road. But if I need five years' worth of data to make a conclusion and I ask that question, I really do want to have the data around.

So I'd say all the data that you have access to, it is big data. You should keep it all.

Bob: I'm guessing everybody's not going to agree with that statement.

Right so, what type of companies do you think are positioned to take advantage of big data analytics in the satellite industry?

Anthony: I'll leave it to my satellite industry experts here.

Jay: I'll give the same answer ... I think all of them ... the big data, the analytics, the learning, the AI, it's here and it's in every single industry. It's in everything that we touch and do. I would suggest that we are in the infancy of it now. It's just going to continue to grow. It's in every aspect of our life and in every aspect of technology.

Bob: With that said, what do you think are the biggest impacts Kratos Solutions have for incorporating big data?

Anthony: On the network management side I think being able to look through the management data, the performance metrics, the changes of states, the failures in the network and not only understand patterns of reoccurrence but understand and predict impact on the network and operations is probably one of the greatest things we should be looking at. Some of the customers I work with have been trying to look for network signatures and combing through, manually combing through, tons and tons of information. Say, "Oh when I see

these three events happen in the core of the network, I know that I'm going to take an overall three-second hit in the network." That's just the microscopic view of it. But I think that's one of the things if we were to incorporate that type of intelligence into the way we treat our network management applications and say, "Listen, you are X amount of time ahead of the problem. But this is going to happen so here's where you need to be focused." That's invaluable data for all the people in this room.

Bob: Anyone else?

Jay: I think there's a certain number of things that we have done or we're looking at now. And not just us, our partners in industry. Bringing in weather information is a big topic of conversation now and predicting impacts based on weather. Using the tools to proactively notify customers that something's going to happen in a hour or three hours or six hours. Use that information to feed the orchestration and provision of services. I'm going to have an impact in three hours and I can take the opportunity now to reroute my network. To move my traffic somewhere else.

The biggest impacts in our industry I think will be for, as Anthony was saying, predicting failures, reducing the number of problems and incidents that occur in the networks. Big data and the algorithms and the orchestration around that is probably where I see the biggest impact.

Bob: But to take advantage of our big data analytics you have to have those underlying infrastructure with things like SpectroNet that will allow you to reroute-

Anthony: Absolutely.

Bob: Okay. I think the last question. I'm not quite sure I understand this completely. So what are your thoughts on cloud constellation space belt?

Anthony: I think I need to Google that. If I recall correctly space belt is the basically the folks that are putting cloud compute and storage up in space and using laser technology in between satellites to access that information. I had a brief conversation, this goes back probably 18-24 months now, with representatives of that company and they pitched out of what they were doing. And I was like, this is ingenious. From my ... I'll call it neophyte satellite industry experience; focused more on the terrestrial world. I don't see how this doesn't take off if they can make it work. I know the problem they were trying to tackle was that if a satellite operator brings data down into a country, they are now bound by the export and import rights of that country. If it's up there, there are no rules yet.

So it's like the new frontier. I think it's really neat and saying that my dataset sits in the sky where it belongs to no one, so to speak, or everyone.

Bob: Yeah, that's interesting isn't it? You get away from geophysical issues by putting it up there in the sky. So with that said, I think we're reaching the end of our time. I'd like to thank you all for your interest and for your questions. I thank you all for attending the conference and I'll leave it to James to close up.

James: Thank you.