For Cloud Professionals, part of the On Cloud Podcast

David Linthicum, Managing Director, Chief Cloud Strategy Officer, Deloitte Consulting LLP

Episode 5: Enterprise Blockchain - what tools exist today, what impact will it have on the future enterprise and how will it evolve?
Duration: 0:30:26

Operator:
Welcome to On Cloud, the podcast for cloud professionals, where we break down the state of cloud computing today and how you can unleash the power of cloud for your enterprise. Now here is your host David Linthicum.

David Linthicum:
Hey guys welcome back to the podcast, and this week we have a special treat for you. We got some hot Blockchain discussion going on here and joining me is Val Bercovici who is actually the Founder and CEO of PencilData, and I guess you guys are somehow involved with Blockchain - is that correct Val?

Val Bercovici:
Yes and more specifically "Enterprise Blockchain".

David Linthicum:
Enterprise Blockchain that makes sense because it doesn't really make sense that you have a Personal Blockchain thing but I guess Blockchain is one of those technologies that's going to touch us all whether we know it or not. So, tell us about your business? How did you come to be the founder of PencilDATA, what are you involved in these days and what kind of cool stuff you're working on?

Val Bercovici:
Okay thanks for the question really great opportunity to chat about this. So I have a long term relationship with data so to speak I like to say I've had a front row seat to the evolution of data across 25, 30 year career arc, starting with simple network databases if I'm dating myself and mainframe style databases all the way to modern real-time streaming datasets, and along the way I've had the pleasure of seeing relational databases be democratized by companies like Microsoft, SQL Server I had the pleasure of helping Eric Baldish and Rob Brydon and team kind of exit the remaining engineers, Hadoop Engineers, out of Yahoo that Doug (inaudible) and Cloudera didn't get, and spent a long time actually as CTO of NetApp, SolidFire and a storage career, and when the time to move on from enterprise storage came about two years ago I wanted to learn about the hottest industry trends, and it actually wasn't Blockchain back then two years ago if you recall, the hype was just starting on Blockchain but certainly Artificial Intelligence and machine and deep Learning were all the rage, and I knew nothing about it beyond the headlines.

So I got recruited into a really cool AI Startup called "Peritus.ai" that's Latin for "expert" and the notion there was to really fill the gap in Autonomous Data Centers where provisioning and orchestration are highly automated today with tech support when software bugs emerged or other break fix situations emerged you're back into this non-deterministic human centric process. So long story short I think confirmation bias being what it is I learned that Machine Learning is a very data hungry set of applications if you will, and there were enormous challenges and therefore business opportunities around catering to data scientists with better data management, data protection, governance and so forth and that's what we found at PencilData on the back of and our original use-case and we've encountered many, many, dozens more fascinating ones but the original use-case was to help data scientists get access to better data, more accurate into timely data and ultimately help them fit their models better than sort of the current tide of all models with public datasets, and Public IP in terms of neural networks and it's been a fun journey along the way - we learned a lot.

David Linthicum:
So what does it mean in terms of - as an Enterprise IT Consumer so in another words I'm - you're going to come to me and explain to me what you guys do you? You just kind of crafted kind of a general idea of some of the issues that you saw, so what specifically would be the use-case is it going to be a financial application? Is it going to be a - is it going to be a merchandising application, manufacturing application?

Val Bercovici:
So great question, out here in the west is a major bank, this bank has presumably top level security infrastructure in place. It has strong auditing controls in place particularly so you wonder how could "data tampering" incidents continue to happen fraud effectively continue to happen and the only way it can happen is actually logically by not - sales people becoming more technical and becoming hackers overnight it's by those sales people or sales leaders perhaps with Administrators, Database Administrators, or IT or Network Administrators, Sys admins as we commonly know them because it's the administrators often they have that full authority to see all data, to change perhaps all data, but also change access logs to try and cover their tracks.

So these inside jobs now are one of the main sources of data breaches, data loss, data tampering and fraud and if you actually listen to some real security experts, at the very top of the list if they would tell anyone to conduct in terms of a security review and what top priority is for improving the security of - it would be in monitoring, supervising, and then foreseeing policies for the administrators in your network, in your environment.

So that is one of the top used-cases we address actually is what I like to call that "Biggest Single Point of Failure" in Enterprise Security today in particularly on Data Integrity our specialty and that is the inside jobs that unfortunately are conducted by administrators that are blackmailed, bribed or very often nowadays have their credentials stolen, published, and sold in the dark web, and then leveraged, exploited by bad actors.

David Linthicum:
So this is the problem that Chainkit solves, your past product?

Val Bercovici:
This is one of many problems Chainkit solves and so one of the use-cases I think that hopefully explains this in a general sense as possible would be in the legal vertical. So in that legal vertical there are companies now that are clearly delivering digital transformation in innovating very ordinary mundane but important things like process serving. So we all seen the movies hopefully we haven't experienced this ourselves where a law firm you know files some kind of subpoena or some of the kind of requests that we have to appear in court, hire a courier, the courier has got them in an envelope on a bicycle they roll up to your door, knock on it ask you who you are to verify who you are and say you've been served. And so it's been a long time now where that entire process could be digitized and effectively that could be delivered via email today. But you have to have the right hooks, you have to be able to trap the right events not the least of which of course is authenticate the PDF or PDF Files in question to make sure that is the actual process or subpoena or what have you that was served, and its lifecycle. So it's origin in a law office, it's its origin in an email transaction if you will from the law firm's email server routed through the Internet received by the recipient's email server, routed into that recipient inbox opened obviously by the recipient when they provide their credentials based on challenge response, and then of course acknowledging that last important critical event which is the attachment itself being opened and then read.

And there were startups that are our partners customers of ours that consume the "Chainkit Platform" that not only trap all those events so "Microsoft, Azure, and Office 365" is able to write this kind of App but it's really when they sell it to their end customers these law firms particularly that some law firms buy it side not sight unseen, but buy it right away first call close where they love the efficiency of the new process, it's digitally transformed it's lower cost, it's more traceable, but other law firms particularly with experience in court hesitate to purchase such a solution because if you were to challenge the veracity, the authenticity of digital evidence in court if you put up a real technical expert on the stand the very first thing they'll ask is what was the chain of custody for this email? Or the chain of custody for that PDF file? And if it just boils down to an administrator or just a narrow group of administrators once again you have that single point of failure scenario where a technical expert can explain how an administrator has full authority to change everything and can't be bribed because they're human or their credentials can be stolen.

So the real use-case here for something like Chainkit to bring a decentralized consensus or a decentralized trust, not a centralized authority and beyond the point of any reasonable doubt prove irrefutability that yes that that PDF file in question as presented in court is the same one it's proven authentic by thousands of nodes in a Blockchain that are independent not just one or two centralized admins. Similarly for all the lifecycle events of the routing ultimately into the recipient's inbox and reading of the attachment, all of those events are trapped by "Office 365" and "Azure" but now they're verified by the decentralized entry on a Blockchain it says yes this event at this timestamp actually did happen and there's no real practical way for someone to tamper with that record or that log event. Those are the kinds of the very Mainstream Enterprise Use-Cases we see over and over again across a pleasantly surprising amount of industry verticals for us as a business opportunity.

David Linthicum:
So aren't you still challenged in order to kind of explain the - basically the verification process of this so you have the multi-entity verification of these systems so therefore people can't typically commit fraud on top of them, but if you take it to court isn't this a very complex kind of an explanation to make to laymen that are really looking to adopt this – and the same thing kind of goes in the Enterprises as well everybody kind of wants the notion of Blockchain but doesn't really understand how it works and really understand how the security in the validity of the various things that we are looking to validate kind of take place?

Val Bercovici:
Great question it really kind of cuts to the heart of it. It's a very often as you can imagine - I would I kind of describe this market as being definitely in its early stages and much like any technology we do have to spend a lot of time on education to make sure people understand exactly one the answers to the question you ask. Like the conversation naturally starts with Bitcoin how is it different? Bitcoin isn't Bitcoin speculative, risky, used by criminals, etcetera. So one of the first things we do is essentially educate people on the notion of cryptocurrencies, but also many other applications of Blockchain particularly in Enterprise settings and it's the underlying technology here that's all the time on mainstream interviews about Bitcoin is Bitcoin is an example of Blockchain but it isn't just Blockchain you know Blockchain has many, many, other things and that's where we sort of pivot the discussion towards using the Pareto Rule where 20% particularly in an Enterprise context of Blockchain transactions do involve some kind of financial transaction, a store and exchange of value such as Bitcoin but 80% are really all about the underlying capability of Blockchain which is to use that decentralized mechanism to verify really anything digital, and the ability now to use that decentralized approach where you can't really even identify much less code or steal the credentials of a 1,000 independent nodes that are required to have consensus to be able to add blocks to the chain to agree on the fact that yes this particular transaction is what it is, we expend a lot of energy in a Bitcoin case to prove that fact occasionally depending on how fast we do that, we get the reward for confirming that fact, but when you try to break it down to it's not Bitcoin and it's not just as femoral or mysterious thing called Blockchain it is a group of independent data centers of independent owners of infrastructure working together, solving a puzzle, proving that data is what it is, people start to understand that yes this is one of those

things where I've essentially hedged my risk enormously by not depending on one entity to verify something, but depending on multiple people to verify something and knowing that all of them as opposed to just one of them one have to be compromised for what it is and verifying to be wrong. So once you really break it down into its component parts people get more comfortable and familiar with it particularly once you take cryptocurrency off the table and therefore 80% of all the corresponding hype around cryptocurrency.

David Linthicum:
Alright so what is Zero Trust Security in the context of this?

Val Bercovici:
Zero Trust is fantastic it's instead of a business level or business processes case such as legal process serving, what excites me as a propeller head ultimately are the infrastructure level integration opportunities because clearly that's higher volume from my business, and so with Zero Trust and we're talking about innovators such as "John Kindervag" who coined this over at Forrester and is now a Palo Alto Networks, and Dr. Chase Cunningham who still at Forrester, the notion of Zero Trust quite simply I think is finally what I like to call an adult or a matured conversation around security. It's no longer about preventing the tax by putting up a castle wall I think we all have to acknowledge our castle walls have been breached almost any organization or individual I can think of has basically been breached, and a lot of our personal data certainly a lot of enterprise data reading the headlines every day is on the dark web, so now it's all around not necessarily trusting Zero Trust identities that we see on our networks or access points that access our network and particularly the most popularly for Zero Trust today it's actually about not trusting a lot of internal network traffic presuming there is an advanced persistent threat or other form of malware already acting inside our network exfiltrating data just tampering with data, corrupting data, holding data ransom via encryption and so forth. So we have to assume that's happening and therefore we see a lot of effort now, a lot of recognition for Zero Trust, and a lot of effort being put into not trusting things continuously verifying key things, but very few people have extended that concept on to the data itself. So we're still doing Zero Trust and continuous verification of things around the data, but it's very important now to really assume I think in a legal context, a technical expert in court will pretty much tell any organization you've got to assume your data - there's a reasonable doubt that your data has been compromised now what are going to do about that? And so what we're proposing is Zero Trust extended which is the ability now to extend your current Zero Trust Frameworks which you hopefully have on to being able to continuously which means efficiently at a low cost, at low latency, verify all your key data, verify all your key database transactions all your files and file updates, all your objects and objects stored and so forth, all of your key milestone events in a log that perhaps log the progress of a regulated business processor, highly insured business process anything where you do have to explain it sometime the road you will have to defend it, that's where you want to continuously verify that data, and at the very least manage like Zero Trust for traditional network security, be able to instantly detect, have the earliest possible warning that something has been breached, that something isn't following an approved business process and that you're detecting now for the first time that there's a problem you have to address intermediate.

David Linthicum:
So what would be you know kind of the demarcation line for data that exists in Enterprise to be to kind of qualify for Zero Trust and you applying your technology to it. Obviously it's not every piece of data that we're tracking in the Enterprise but it's some of the data or most of the data and if so what are the patterns of the data which kind of lead it to leveraging your technology to in essence put the Zero Trust security into it?

Val Bercovici:
Great question and we tend to actually start that discussion not necessarily at the technology level which we'd love to but really more at the industry vertical level. So very often we focus on regulated industries or highly insured industries essentially any business that has to explain what happened in its databases, what happened on its storage systems, what's happening in its business processes. And so we start with a business problem first, we identify are you regulated, do you conduct frequent internal audits, would you like to conduct internal audits more efficiently? Have you not been able to audit a process that needs explanation and is part of regular regulatory review or a surprise regulatory audit - those are the kind of questions we start with and yes very often that does translate down into databases, data storage and databases, data stored in file systems, so it is impractical I think in this first generation of the - to just verify every transaction all the time but auditors typically take a look at weekly scans of data or monthly or at the very least quarterly scans of data and at the cadence by which you perform an internal audit is a really, really, good cadence by which you want to take a look at a set of transactions or some key transactions and verify by them, and see whether things are on the rails or off the rails with regards to your data authenticity, your data integrity but this now extends very deeply into IT operations itself. So you want to find out if a network configuration is changed particularly now in the era of Cloud Native and software define networks if it changed was it part of a regular cadence, was it a part of a documented emergency patch update, or was it some kind of unauthorized network change that's permitting malware or a persistent threat to do it's nasty thing an excellent trader or ransomware your data. So this really extends quite excitingly once you understand the business requirements down to a lot of infrastructure layers and the granularity isn't super fine yet, but the granularity is certainly it can be done multiple times a day, multiple times an hour and in some cases multiple times a minute without

really impacting the operations.

David Linthicum:
So how does this affect the retail user, I mean that's a we're going to have a lot of people where is a public podcast a lot of people who aren't necessarily into technology they're just trying to track where technology is going in the retail consumers and thinking about Blockchain and you know, multi-identity authentication, the ability to leverage encryption services things like that so what should they look for in terms of their lives improving over the next couple of years?

Val Bercovici:
So I would say the irony of Zero Trust is that by doing a good job of not trusting your data and continuously verifying it you're actually able to deliver the value of business or even personal trust to your end-users your end consumers and so forth. So I think the ultimate impact is business trust in folks who buy services and products from be it Amazon, be it your regulated utility and you establish that trust through transparency as a business delivering that product or service. I actually wrote a blog about a year ago pre Facebook and Cambridge Analytica that I hash tagged data responsibility and fortunately or ironically that very same day Ginni Rometty of IBM had opened up for keynote at Davos on the exact same topic and her IBM policy team used the exact same hash tag data responsibility and I think that's the heart of it in my mind.

The only way you can establish more transparency and build that trust and maintain that trust with your customers is being as transparent as possible, not just with regulators, not just with insurances companies but with your customers and proving in an irrefutable way that you are managing their personal data in a responsible way and that you are doing everything to the best of your abilities to prevent attacks and to prevent data breaches. They will continue to happen but when customers see that you've tried to do the best possible thing, to apply the best possible best practice, that trust builds as also almost a joint empathy there. What's happening today and I'm surprised still in 2018 this is happening but hopefully this will start to change in 2019 is very often when vendors whether they're social media vendors, whether they're Enterprise IT vendors, whether the people that sell your coffee it doesn't really matter, when they are forced now to disclose in their minds prematurely that a data breach has happened but in regulators minds such as the EU and GDPR already two weeks is too late, when they're forced to disclose that you often see we think this number of customers these number of customers are affected and we were hoping it's only this amount of data not their credit card information, not where they were last Tuesday in terms of check-ins or GPS location, not what their kids names are, there's a lot of vague claims or a lot of vague statements made by these vendors and I personally think that's not acceptable in this day and age, because I think you and I are both old enough to remember the "Enron Scandal" and how that email dataset became one of the most queried datasets in terms of machine learning, but also in terms of establishing precedents for how to deal with data in a legal context so every email backup archive vendor a couple years later came out with a legal hold feature for email so that when these kinds of things happen in the future it's very easy now for companies to satisfy the necessary subpoenas and prove they're providing all the emails and that they are authentic and so forth to lawyers and to regulators.

Fast forward what 10, 15 years now and you're telling me what we're not able to do that for all of the databases we use today whether they're sequel or no sequel and all the data streams they use today I think that's unacceptable - I know there's solutions to this and that's what I propose in my blog which is a suggested framework using Open Source tools for example the Apache Beam Framework for processing data combined with very modern and cloud-native approaches such as ISTIO who are actually processing that data and running complex business rules to do it, and I jokingly said let's start a project called Bistio which combines the best of these things Open Source Tools, is able to persist any key data stream or dataset either metadata or the entire dataset on a Blockchain so that when the lawyers and regulators and investigators come down and look to review what happened both inside and outside the company, there's no doubt as to the veracity or authenticity of the full dataset that they're looking at, and all these vague claims that are made during these data breach disclosures can be highly targeted, highly specific and ultimately they deliver certainty to customers and certainty to business which is a very much desired state for any business.

David Linthicum:
So the stuff is pretty complex pretty cryptic where people early kind of insiders in the whole Blockchain world so where is the stuff going specifically for you what do you think you guys are going to be working on in 2019, 2020? What are the hot topics in this particular space and where are the Enterprises going to be making the investments?

Val Bercovici:
So the thing is like any new technology at first you're absolutely right it's cryptic, it's complex, it's certainly poorly understood and the opportunity in the early days I can certainly see for the next three years for us is to emphasize on demystifying getting more importantly from a pragmatic standpoint and making it drop-dead simple to deploy, so that's been our early focus and talking to customers as customers gave us a fantastic set of use-cases including autonomous driving and medical images and in auditing robotic assembly during

the assembly line, and then for the assembly of complex industrial IoT devices auditing their operations. So the use-cases are really fascinating but over and over again when we would ask the question of when would you buy this kind of solution from us it did come down to help me make it simple, help me make it easy to integrate and implement and that is really the focus right now, it's not necessarily about just educating and demystifying the technology - is it's basically saying if you have a requirement to audit to verify your data to verify your business process to verify software releases today, let's make that integration as simple as five minutes or less of production. And so that's going to be I think the focus for the next few years at least is in bringing the value of Blockchain, bringing the value of decentralized and consensus trust into existing business apps as opposed to what's happening in the Blockchain world today which is having to rewrite apps from scratch just to utilize one of potentially 2,000 different Blockchain's, bringing the value of existing popular Blockchain such as Ethereum or very (inaudible) Hyperledger Blockchains with a IBM, Fabric, Intel, Sawtooth, we're having some fun discussions with VMWare on Project Concord bringing that value to Enterprise to existing apps and processes existing datasets and making it as simple as possible.

David Linthicum:
So where should the Enterprises go for information besides this podcast of course but what are some of the better books out there, the blogs to go to, what are you guys publishing, where is kind of the beginning the intermediate and the advanced information on Blockchain and you know how it's going to be morphed sometime in the future and there is how do I get on - how do I start surfing that wave?

Val Bercovici:
Yes great question I've been a big fan of Open Source quite some time now, and I believe one of the values of Open Source hasn't just been a lot of recoup and code an intellectual property available to us at a low cost or free in some cases, but also some of the educational not vendor neutral so to speak educational aspects of it. So Hyperledger you Hyperledger Foundation and a lot of the resources they publish and of course there, their social media feeds is a great starting point for really understanding what's available where the immediate business value is and it's a great way to basically go and get away from the cryptocurrency hype and speculation which tends to drown out a lot of information sources that people are looking for. Once you've exhausted data, once you've found what you need to there, absolutely look to follow our social media handles are LinkedIn and Twitter feeds at PencilDATA since Chainkit has proven so popular with customers, it has its own brand, has its own Twitter feed for example, lot of good signal to noise ratio there because we don't publish often, and of course I would say folks that are focused at a high profile level on the commercial on the enterprise opportunity such as ConsenSys they're doing a really good job and they were aligned recently with Hyperledger on educating the enterprise market on educating folks on the actual opportunity there.

David Linthicum:
I like to ask this question so say we go forward in time ten years and we always look back at the technology we can actually kind of figure out the value of the technology based on its ability to kind of solve problems and how much we're using today, so fast forward to 2028 how will we be remembering Blockchain will it be built in to the existing infrastructure or it would be like that was a huge boondoggle that we shouldn't gotten into of course I don't think you'd actually answer yes to that, but you know going forward where is the value going to be placed on Enterprises, how is the change the Enterprise going forward and what are kind of the data points that we gotten out of the utilization and Blockchain.

Val Bercovici:
I love that question because I've actually spoken at some rally conference a specifically about this, one of the things we didn't have if we look back 10, 20 years ago on the technologies we depend on today's we didn't have 2,000 options for networking, we didn't have 2,000 options for file sharing or 2,000 options for what a relational database should look like, and so we have this what I like to call perverse incentive today for people to innovate it this low, low scientific level of coming up with net new encryption schemes, and net new consensus algorithms when you know perhaps 100 others or at least a dozen other perfectly good ones already exist. So the whole token sale, you know initial coin offering hype in mainly of about a year eighteen months ago, I think still is continues to do damage to the industry certainly on the currency side but also on the enterprise side and ten years from now there's no doubt in my mind having seen this play out over and over again, that we will value all these Blockchain innovations today exactly how we value the BIOS on our laptops or servers or perhaps how we value the TCP/IP stack in our operating systems and our network adapters, that is to say essential for getting business done but not exactly game changing in terms of businesses that we operate or services we deliver today. So that's why I caution even Kelsey Hightower who does delivers fantastic keynotes to the Cloud-Native Community cautions people not to get too caught up in container hype because after a few years if those projects truly succeed they become boring they become things we take for granted and assume and that's absolutely where Blockchain technologies headed in ten years and having said that it will just like TCP/IP enabled this amazing explosion of business innovation and the world we live in today, and the mapping we used today to get to work, and get to meetings and so forth amazing new business services and other personal services will be enabled by Blockchain but the hype in the mania and the interests will certainly fade that the layer four or five level of the OSI stack outer space.

David Linthicum:
Yes I think it's very profound I think any technology that actually has value will be ultimately forgotten and really kind of the metric of success is that it's been so successful that it's boiled down the infrastructure the way in which we do computing, things like that and I think that's a great way to describe it. So you mentioned it some you mentioned it prior but where can we find you guys on the web and where can we reach out for more information on your company?

Val Bercovici:
We are at www.pencildata.com I particularly encourage the Slash Sign Up page because it is as easy as five minutes including looking at a three minute video to basically sign up and integrate the power of Blockchain Integrity into your existing app or workflow even on Excel macro in five minutes in production on the live Blockchain. So you can go to our website, you can go to social media it handles we're PencilDATA on LinkedIn, we're PencilDATA on Twitter, pencil_data actually on Twitter there is a long story about why the underscore is there which we can talk about some other time over a beer as well as of course Chainkit, www.chainkit.com or @Chainkit on Twitter.

David Linthicum:
Val it's great having you on the podcast. Hope to have you back at some point in the future.

Val Bercovici:
Very much a looking forward to it thanks for the really fun conversation this morning David.

Operator:
Thank you for listening to On Cloud For Cloud Professionals with David Linthicum. Connect with David on Twitter and LinkedIn and visit the Deloitte On Cloud blog at www.deloitte.com/us/deloitte-on-cloud-blog. Be sure to rate and review the show on your favorite podcast app.

Visit the On Cloud library

www.deloitte.com/us/cloud-podcast