Maximizing RAN performance for best application QoE

A conversation with John Reister, Vice President of Marketing and Product Management Vasona Networks

By Frank Rayal

Frank Rayal. Hello, and welcome to this conversation with John Reister from **Vasona Networks**. I'm Frank Rayal. This conversation is part of a Senza Fili report on deployments of heterogeneous networks and the latest solutions from the wireless ecosystem that will make HetNets a reality.

Today we're speaking with John Reister, Vice President of Marketing and Product Management at Vasona Networks. Vasona is a pioneer in cell-level traffic management.

John, I would like to start off by asking you for a brief introduction to Vasona Networks.

John Reister. Vasona is focused on the mobile RAN and trying to get the best possible performance out of that RAN in terms of providing high-quality subscriber experience. We're trying to get to that "Be all you can be" and enable the network to make full use of the capacity that it has and improve the quality of service. We do that by going down to the cell level and looking at each individual cell and understanding the contention that's going on within the cells. We also enable the assessment of how much congestion there is in the RAN, and allow resources to be used best and to give each application and each subscriber the quality that they need for that particular application they're using.

Frank Rayal. What exactly have you developed? Can you drill deeper into some of the use cases to provide an example of what cell-level traffic management is?

John Reister. One of the challenges with mobile networks is that you end up with multiple subscribers in a cell. It's a combination of things that go on. First of all, the demand on the capacity is very dynamic. You get these spikes that last one, two or three seconds as the bandwidth bursts up. So you get very fluctuating demand.

And then you have capacity as the UEs move around within each cell and as they do handovers into and out of the cell. The actual bandwidth that's available to each UE changes. So you can have a subscriber that's sitting stationary in a park watching a video on his tablet, and he's across the street from the antenna and getting very fine bandwidth. Then somebody hands over into the cell from farther away. Even though that first subscriber hasn't moved, his bandwidth can suddenly change. So demand changes, and you've got capacity changes.

And then the third piece is that the applications have different needs in terms of quality. It's not only how much throughput they need, but how much packet loss and how much latency they can tolerate. So what we're trying to do is to look at the contention on a cell-by-cell basis. If there's no contention, that's great. We can just observe and monitor what's happening. But if there is contention, we're going to detect that and take corrective action in terms of adjusting the queuing

Sponsored by

vasona networks®

and prioritization, and some other techniques, to ensure that the applications that have a high need for delivery can get through in time. Applications that are running in the background – for example, somebody's got their phone in their purse and it's synchronizing – can be delayed by some time, to enable that video or browsing experience to be the best it can be.

Frank Rayal. John, where is the solution placed in the wireless network?

John Reister. It's placed between the core and the RAN. So if you think in terms of 3G, it's going in between the RNC and the SGSN. That's the place in the network where you can actually see which subscribers are in which cells and when they're doing a handover from one cell to the other. You also can see all the data traffic, whether it's a roaming subscriber, a VPN subscriber or a regular subscriber. You can see all that traffic, and you can manage it from there.

Frank Rayal. And the solution operates on a subscriber by subscriber basis, in terms of optimization?

John Reister. It does. I mean, we're looking at the contention of a cell. So we want to see all subscribers within that cell. But you are correct that we're looking at each and every individual subscriber: what kind of bit rate they're able to get and how well the network is serving each subscriber. We're able to classify the applications that the subscribers are using into types, and then, based on that type of application, we try to give the package the appropriate treatment for that type. For example, streaming obviously has higher demands than browsing, and browsing has higher demands than a background download application.

Frank Rayal. In terms of the capacity gains, that is done on which segment of the wireless network? Does it, for example, include the backhaul, or is it on the radio access side?

John Reister. It's everything from where we're located down. So it's going to be the backhaul that runs from the RNC in 3G going out to the base station. And then, of course, the air interface is the important one, where most of the congestion occurs.

Frank Rayal. In LTE, they pushed the entire Medium Access Control layer to the base station. And that, of course, includes the scheduler, which is the part responsible for optimizing the traffic flow on the air interface. How does your solution fit within those concepts? – because the scheduler is on the base station trying to optimize traffic to the UE.

John Reister. The scheduler does not look at the application types. So certainly in 3G, that is very true. In LTE, there are some concepts with QCI and

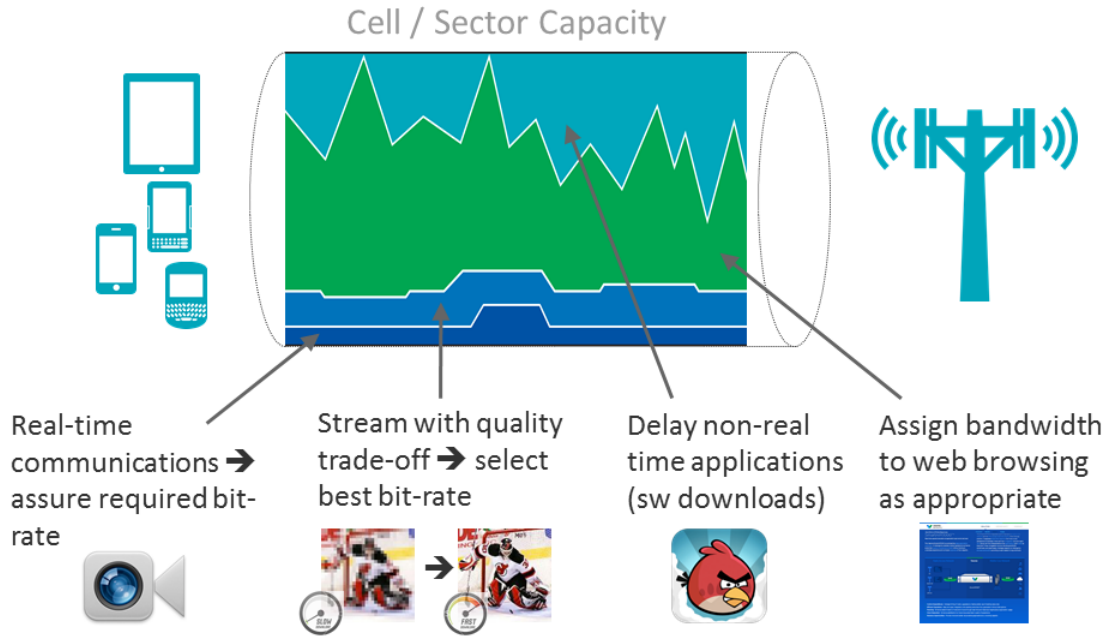

Figure 1. Vasona's application edge controller identifies the state of the RAN to maximize performance.

different bearers, but I think the fact is that, today, virtually all the applications are over the top and they are all going into the default bearer. So essentially what we are doing is we're watching what the scheduler is doing. We are able to detect exactly what is happening between us and the subscriber. We are looking at how many packets are in flight between us and the subscribers. So we know how the packets are backing up in the RAN and how the queuing is going – the roundtrip times and the latency. Then we are essentially managing all that to make sure that the quality perceived by the subscriber is as good as it can be.

I don't want to only focus on the corrective action, though. A huge part of the value we provide is without taking any corrective action. We're actually watching what's going on, and we can read out on where the congestion points are and which applications tend to trigger congestion: time of the day, handovers, etc. We determine which cells are having issues, and we can do this at a very detailed level. The monitoring data that we collect on the cells allows operators to figure out where to prioritize their investments, whether it is to do cell splits, or add more channels, or introduce small cells into the mix.

Frank Rayal. Were you able to quantify some of the benefits of your solution in terms of capacity or any other way?

John Reister. We look at the quality that is there without us taking corrective action. We can monitor that. Then, after we enable corrective action, we look at the resulting quality improvement. And I can talk a little bit about those improvements. But then when we look at those improvements, we ask ourselves, how much more capacity would the network have needed to achieve that without us taking corrective action? It's like we're making the network behave as if it has more capacity than it really does.

So for example, for video streaming delivery, we can improve the proportion of videos that are delivered without interruption by 20%. If you look at the videos that today get interrupted with pixilation or freezes, first without us taking corrective action and then with us taking corrective action, we improve that by 20%.

In addition, we improve the actual bit rate of the video that the subscriber is experiencing. When you look at a YouTube video, operators will adapt the bit rate somewhat to match the network. We can actually provide around 15% higher bit rate for the video. We've looked at, say, 5,000 videos delivered with our solution turned off and then 5,000 videos with our solution turned on. And when we compared the bit rates that the subscribers are able to get on average, we've improved them by 15%. So not only are we reducing the interruptions, they're getting more

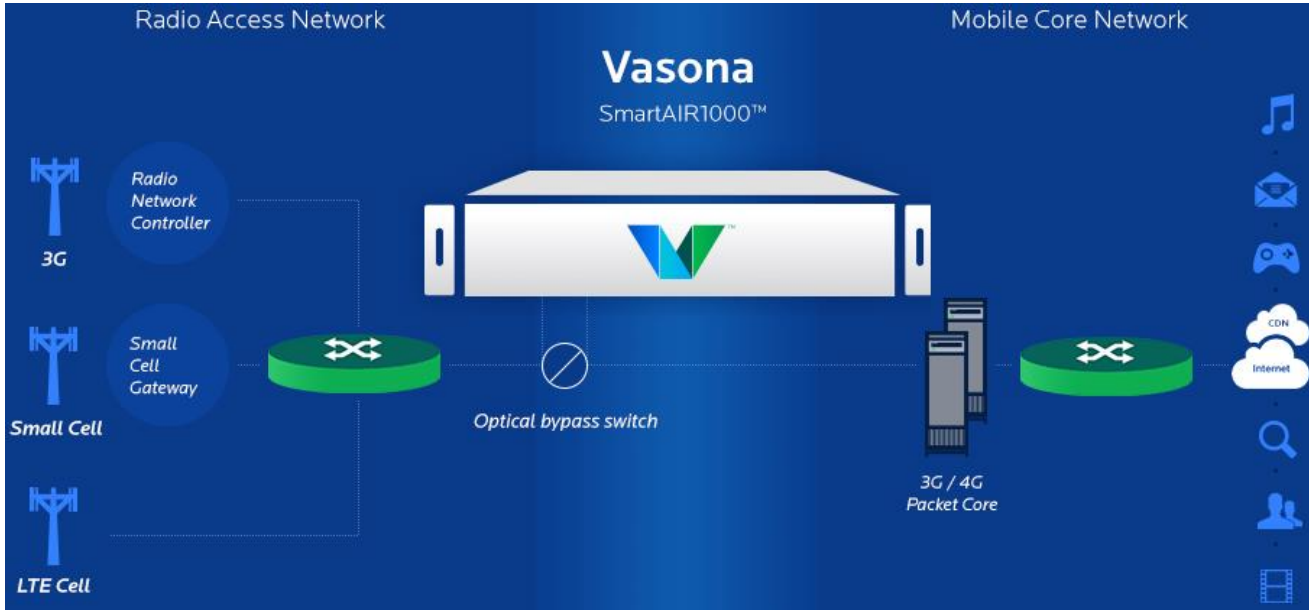

Figure 2. The Vasona SmartAIR1000™ deploys between the RAN and the mobile core network. Source: Vasona Networks

bits for their video. That generally corresponds into better overall quality.

Frank Rayal. The benefit would really depend on the type of traffic, right? So video may experience more benefit than other types of traffic.

John Reister. It does, so let me talk about browsing a little bit. What we've seen in recent trials with browsing are a couple of nice factors. One is that when we look at a browsing session, we are actually delivering around 25% more total bytes for the browsing session, but we're doing it in less time. So we have reduced the average time of the browsing session. I think in the most recent trial, we reduced the time of the browsing session from 38 seconds to 31 seconds, but we increased the number of bytes that were delivered on those sessions. In other words, the subscribers are getting the information they need, faster. They're getting more information, more bytes in a shorter time. They're not abandoning the session because they couldn't get the information or the browsing session was stalling out. So those are the kinds of metrics, and if we translate this into equivalent capacity, it's around a 35% to 40% improvement in equivalent capacity in the network. You would have to deploy 35% more capacity in a cell to be able to achieve those same results without the corrective action.

Frank Rayal. And the solution would work equally well with LTE, 3G networks, irrespective of the radio access network?

John Reister. Absolutely. The solution is agnostic to the type of the radio access network. We've actually seen that, in LTE, the spikes in congestion get even bigger. I think when people initially deploy LTE, they have a few months without any congestion, but the data that we've seen is that usage expansion and spikes are actually more significant. So managing those spikes in demand becomes even more important. Obviously, a 20% improvement in uninterrupted video is great. We can identify this as the best you can get. We can also show that these are the cells where you need more capacity. For example, there might be events occurring around a stadium or around an enterprise campus. We can look at the control plane and handovers. You don't want to put small cells where you have a very high degree of mobility – you're just going to exacerbate the load on the control plane as people have to jump in and out of a cell.

Frank Rayal. This is a good transition into the area of heterogeneous networks. How does your solution apply when small cells are deployed? Does it apply the same as in the case of macro cells?

John Reister. Absolutely. When you get a small cell, you can have multiple subscribers sharing that cell and contending with different applications. You want to be able to do the best job you can in getting those applications to where they're highly interactive, like video and browsing applications. You want to monitor and improve those experiences. Then to the extent that you have applications running in the background, you can prioritize the interactive ones over the ones running in the background. It's the same concept.

In fact, in small cells, congestion can be more pronounced. For example, if that small cell is placed in a location that tends to experience particular busy moments, like an event location or commuting center, our value can be even further enhanced by improving quality of experience at these busy times.

Frank Rayal. One aspect that I would also like to explore is the fit of your solution within the concept of self-organizing networks. Because in a way, you're doing some form of optimization. So how does all that fit together, if it does at all?

John Reister. I think we are a nice complement to self-organizing networks, but self-organizing networks focus more on controlling and forcing subscribers from one cell to another cell, effectively changing the size of cells. We focus on optimizing what happens inside each and every cell. If a UE is in the heart of a cell, SON solutions will not be able to provide any benefit. Vasona can improve this UE situation several-fold. So, the idea is to let SON optimize the border of the cells, and let Vasona optimize the core of each cell.

What we can do is we can provide great information that informs those decisions. So in our deployments where we have a lot of small cells packed into highly dense areas, we've seen handovers to be extraordinarily high, roughly a factor of nine, so almost a whole order of magnitude higher than what we'd characterize as

a normal or regular urban network. We've done some trials where the cells were packed in very tightly. When the UEs are constantly flip-flopping back and forth between cells, you end up with a massive amount of control traffic as they have to reattach, transition from idle to active, and establish context. There's a lot of control that has to happen before getting bearer-plane data. All that control actually poses a burden on the cell and the signaling gateway. We can look at that data, monitor it and identify where you have those issues. That's a complementary improvement that we can make to those networks.

Frank Rayal. Where do you see taking this product in the future? What's your vision and roadmap for this solution?

John Reister. When we look at what operators have done in pricing their data, it is very elementary, where you pay for gigabytes. If we look at other industries, for example the cable industry, you get your basic cable service and then you can buy the sports package and you can buy the premium package – there are lots of add-ons.

When we look at our technology, we can tie into the back office and we can allow the operator to have a much more robust or richer way of creating and pricing their services. For example, you could have a streaming package. That would mean that when you accessed video streaming services, those bytes could be rated differently. When there is congestion on the cell, you would have the highest quality and the least congestion if you subscribed to that package.

Another example is to look at usage patterns. Think about what the utility industry does with smart meters, where they're looking at usage over time and changing the rating based on the usage pattern. You can start to characterize subscribers as to what kind of quality they get as they move around and when they're in small cells. You can start to manage the services to match up with that subscriber behavior: segment your subscribers based on how much time they spend in their home cell and small cells versus how much time they are moving around from cell to cell. You can look at the quality of their experience and manage it around how they behave.

So the goal is to substantially improve the quality of the subscriber experience while minimizing the capital that the operator has to invest to achieve that. That's the best form of return on investment, where the return is your subscriber and customer happiness.

Frank Rayal. What do you see from your involvements with the operators when it comes to challenges in HetNet deployments? What do you see as being some of the key areas now that are critical to HetNets?

John Reister. HetNets do have their challenges. Interoperability is really important. Then, I think, figuring out where the HetNet investments should be made. As I said, if you deploy small cells incorrectly and you put them in a place where there's too much movement around, you can end up creating more problems than you solve. So I think it's sort of getting that intelligence and information about the network, and then when you put the cells in, you are able to monitor and manage what's going on to ensure that your investment is behaving in the way it's supposed to.

Frank Rayal. John, thanks for this conversation. It was very interesting. I would like to conclude by thanking the viewers for watching this conversation with John Reister from Vasona Networks.

This conversation is part of the Senza Fili report on HetNet deployments and the latest solutions from the wireless ecosystem that will make HetNets a reality. The report can be downloaded from the Senza Fili website at www.senzafiliconsulting.com.

Acronyms

3G	Third generation
4G	Fourth generation
CDN	Content delivery network
GPRS	General packet radio service
HetNet	Heterogeneous network
LTE	Long term evolution
QCI	QoS class identifier
QoS	Quality of service
RAN	Radio access network
RNC	Radio network controller
SGSN	Serving GPRS support node
SW	Software
UE	User equipment
VPN	Virtual private network

About Senza Fili

Senza Fili provides advisory support on wireless data technologies and services. At Senza Fili we have in-depth expertise in financial modelling, market forecasts and research, white paper preparation, business plan support, RFP preparation and management, due diligence, and training. Our client base is international and spans the entire value chain: clients include wireline, fixed wireless and mobile operators, enterprises and other vertical players, vendors, system integrators, investors, regulators, and industry associations.

We provide a bridge between technologies and services, helping our clients assess established and emerging technologies, leverage these technologies to support new or existing services, and build solid, profitable business models. Independent advice, a strong quantitative orientation, and an international perspective are the hallmarks of our work. For additional information, visit www.senzafiliconsulting.com or contact us at info@senzafiliconsulting.com or +1 425 657 4991.

About the interviewer

Frank Rayal is founding partner at Xona Partners, a boutique management and technology advisory firm specializing in telecom, media and technology. He is a telecom industry professional with more than 20 years of experience working with network operators and system vendors to develop and deploy innovative wireless solutions. Frank co-founded small-cell backhaul pioneer BLiNQ Networks. He held senior product management, marketing and business development positions at Ericsson, Redline, and Metawave. He holds a BS in electrical engineering from Case Western Reserve University, Cleveland, Ohio, and an MASc in electrical engineering and an MBA from the University of Toronto, Canada. He is a senior member of IEEE, and a member of Professional Engineers Ontario.