**Project Completed by:**

**Selections from the**

**EDUCAUSE Security Professionals Conference 2018**

**August 21, 2018|12:00 - 4:30 PM EASTERN**

**Selections from the**

**EDUCAUSE Security Professionals Conference 2018**

**August 21, 2018|12:00 - 4:30 PM EASTERN**

>>> And we're back. Our next session is GULP, do you really know who your users are? Join us for this talk is Joel Rosenblatt, director of computer and network security at Columbia university. Joel, we are delighted to have you join us, please, begin.

>> Thank you very much. This talk is -- I'm changing gears for everybody. We're going to talk about some technical stuff now as opposed to thought processes. But hopefully, it's going to be useful to you. Any way, GULP is a program that we developed at Columbia, and to start off with, let me give you some background on what our environment looks like. We're a large research university in New York city. We have a totally decentralized management structure, so there's no central I.T. for the entire university. Every department does their own I.T. thing. It's fairly large. We have about 250,000 network nodes around New York city. About 100,000 addresses, MAC addresses that we see online, on an average every day. No central computer support, we have a help desk that you can call, that basically triages the call and then directs it out to where it's supposed to go. One of the things that we don't allow on our network is we don't allow sniffing traffic, or scanning machines, so we're not allowed to see what you're doing our network, we can see where you're going, but we can't see what you're actually doing there. One of the things that basis of this entire program was the idea of free love, I'm going to explain that in a little bit. But we give out IP addresses on demand, there's no preregistration required, there's no MAC registration, there's no knack program, you get an IP address, that IP address is registered, you can run a server on it if you like. We have open Wifi and we let you do pretty much -- anybody can connect to the network. We have no corporate firewall. There's no border firewall on our network. Anything you want to do, you can pretty much do. And we have just for context, we have about 130,000 e-mail accounts live right now. That are active. As far as why we developed this, well, when we have -- a large research university typically, what you want to do is you want to be able to give access to people as they come on campus. We see about almost 20,000 different MAC addresses every day that come on to our network. And those are guests, they're visitors, people that are coming to campus for various reasons. So one of the things that we wanted to do was we wanted to offer a pain-free use of our network for visiting people. We don't want to have them preregister, we didn't want to have to set up, you know, get their MAC address and put it into a system somewhere where they can connect to the network if they had registered properly. We needed to reduce the overhead, because we tried doing the registration deal and we had a whole room of people just registering MAC addresses, and turned out that it was a lot of mistakes, people would come on campus, for a conference, 100 people coming in for a conference and 20 of them can't get on the network and it turned out to be a very big deal. One of the things we said, how do we fix this problem? And the solution we came up is free love. The idea of free love is we allow all computers, whether public or private, wired or wireless, in the residence halls or libraries, anywhere on campus, to connect directly to the network and thereby the world without further adieu. There's a paper that's in the EDUCAUSE library, you can read about it, it was -- who was the I guess, the CIO of our Columbia, way back in the -- I'm going to say in the 80s, and that's when we started doing "free love" allocation. Once you have free love, you have a new problem to solve. So how do you know who is using an IP address? What MAC address corresponded to what computer? Or who was -- which IP address was being used by a particular person? So now you have to solve these problems because that comes up sometimes from law enforcement or just problems with the network, or just problems. You always want to know who is doing things. So you had a -- sorry, I'm on the wrong button here. Come on. The new solution is GULP. And GULP is the grand unified logging program. So I'm sure that everybody that's listening to this program has logs. You have logs from computers. Computers are really good at logging things. They are very good at telling you what happened. They generate millions and millions and millions of logs. And most of you have those logs squirrelled away on drives somewhere, and if I asked you to go find them, you would have to go find the right guy to find the right log to find the right drive, to find the right thing, the right privileges to look this stuff up. And why do you do it that way instead of processing the stuff in advance? Most people wait until they actually need a log or need an answer before they look at the logs. So we said, well, why don't we take those logs and actually process them? So that's what we did. We read in the logs, we have from pretty much any authenticated service at the university. We suck in the logs. And we suck in the logs even from our door swipes, so we have a Linh L system, if you walk around campus and get into buildings, swipe your ID card, those are little computerings, they have IP addresses. We suck in those logs, too. We take in the logs from our DHEP servers, our art caches and MAC addresses with IP addresses. GULP is basically a huge correlation engine, it takes all of the information that we get from all of our logs, dumps it interest a database and correlates it. And we can track using GULP the actions or the presence of any IP address, MAC address or UNI, our user ID's, and even if the UNI, even if the connection is not on the Columbia campus. So if somebody is logging in from somewhere in Europe, if they authenticate to a system, I know the IP address. I can actually see people flying across the country as they connect to different IPS, the MAC address changes and we can watch them flying around the world in almost real time. As they connect to different systems. We keep the data for 28 days. And then we purge it. And that's in our log retention policies. You want to have an accurate log retention policy for law enforcement reasons, if somebody comes and asks you for a log, you need to produce that log. And unless you have a policy that says they automatically get deleted at a certain time, you really don't want to have logs around that you arbitrarily delete at different times, that will get you in trouble with the law enforcement people. So now, let's talk a little bit about what the work flow is. So if you have logs, you basically read the logs. You suck in the -- you have the user authentication, you have the time stamp, you have the IP address, and you take that information and you put it into a database. Then you want to use your DHEP and art cache to associate the IP addresses that are in your logs with actual devices. So you use that to associate them with MAC addresses. You throw in your network information and if you have a door swipe and you have machines that are authenticating people that are not necessarily computer logs, if you put that information in there, you have a physical association of where somebody was with an IP address with a time stamp, it turn out that that information is incredibly useful for figuring out if people's ID's have been compromised. Typically, if somebody logs in to a building and in another country, they're not the same person. It's very useful information. One of the things that you will notice is that there's a difference if you're doing network security versus public safety. Turn out that GULP is really useful for both, because they ask the same question, they just ask it in completely different ways. So in a network security way, you would want to know which machine used an IP address at 3:00 because there was a problem. Public safety people are going to say who used an IP address at 3:00? In network security, you're going to say, which -- which machine was using that MAC address which connected to the network yesterday? Because maybe the machine was compromised, and you want to find out what the machine was so you can go have it fixed. Public safety is going to ask you, did a person named John Doe log into the network yesterday? Really, the same question, because if they logged into the network, they were using their laptop, they -- you really just finding a device. And in network security, how many different MAC addresses used that jack or in that room? So if you had a -- let's say a wireless or you had a public jack somewhere where people could plug in, that's how many MAC addresses were using it. In public safety, they're going to ask how many people used that IP address? And when? I've gotten asked that question when somebody used a kiosk in a library to send threatening mail to -- turned out they were sending it to the white house and I got a call on the Sunday morning from the secret service, they wanted to know who was in that library at that time because the white house got a bomb threat? And that was just something that we had to figure out and GULP let us do that kind of research. Here's some -- this is what GULP actually looks like, it's a little hard to read. But I think if you put your nose up against the screen, you might be able to see. This is a GULP for me. And you can see -- let's see, it tells you who I am. And then, it shows you various places that I logged in from. And if you look on the lower part of the screen, you'll see on the right side, there's the three arrows and those are the kind of the interesting ones. One of them is the first one is the log in from my desk top, which is actually what I'm sitting in front of now talking to you. The next one is a door swipe, which is the building that I'm in. It's called stewed Studabaker. And optimum online.net, which is my home. You can see here that that -- it works -- gullP tells you where I was at various times, it tells you kind of what I was doing. And it tells you even if I wasn't on campus, I happened to log in from home to check e-mail. Now, one of the things that I did here was GULP, you can GULP by IP address. So if you go back to the previous slide, it says that I logged in from this 68.197 address, my opt much address from home, and I put in a GULP for that IP address. Now, if you look down at the UNI column here, the first column in the lower part of the slide, you'll see that it has somebody named Audrey 0 and then JLR9. JLR9 is me, but who the heck is Audrey 0, Audrey 0 is my wife. Because she works at the university, too. And she logs in from home and checks her e-mail a lot more than I do, apparently. And that's another cool thing about GULP. Because GULP will show you everybody that used that IP address. Not just one person. So now we have an IP address that was used by two different people. And that gives you the ability to associate people. So for example, if you have students that are living in the same dorm room, they may have an access point in the dorm room, they're going to show up on the same IP address. There are times where that has become very relevant doing investigations. They say, well, I don't know that other person. And then you look and you say, well, how come you log on from the same IP address and that IP address happens to be an apartment. How come they're in the same apartment with you? There are interesting things you can figure out from this stuff. I did a GULP for Audrey 0 to show that it looks a little bit different than mine, because she lives -- she works in a different building than me, she has logged in from different machines and different locations, and it just lets you do this kind of investigation. So now, I have a poll that I wanted to put up. And here it is. You can click on all of them. These are things that you might have been asked to do as a security person. Check all the ones that you've ever been asked to do. I'll let this go for a bit. And look at that. So everybody has been asked to do all of these things. Huh? How surprising. Okay. Wow. Uh-huh. We're getting there. Okay. So you get the idea. People in our positions at universities are asked to do a lot of different things. And the reason I had this poll put up is because if you think about doing these things, how easy is it for you to get these answers? Can you do this in real time at your university? Can you just -- if somebody calls you on the phone, can you tell them when you're on the phone, ask them, ask to find a missing student, or do any of these kinds of things? Well, with GULP, you can do this in real time. So let's go on here. So no one has Stein a student for ten days -- seen a student for ten days, can you tell me anything? You stick it into GULP and here's the procedure, you look up the I had of the missing person -- ID of the missing person, you analyze the records for the location and times and you figure out from talking to public safety if that information matches up. So we can actually, if they were on Wifi, track them around the campus, because as they connect to different access points, we get different access -- we get different GULP entries. If they are out of -- off campus and connected from a hotel somewhere or off -- went home, they went somewhere, we can actually figure that out. If they're flying around, we can watch them fly around. Turns out that this is typically incredibly useful, because usually, you get these calls when their parents call and say, you know, Joey has not called me, I don't know where he is, you know, he calls me every day and he didn't call me for the last couple of days, you know, you use GULP and figure out where they are. So another question that we get asked all the time is somebody gets an anonymous e-mail from yahoo, can you tell me who sent it? Well, typically, that's a hard question to answer. Because if it's a fake e-mail address that they just made up, how do you figure out who sent it to you? Well, if you stick it into GULP, you have the IP address that the anonymous e-mail came from, because you look through the headers, and unfortunately, that doesn't work well with gmail, because gmail takes out the originating IP address, but most of the other vendors still leave it in. If you take it, you pop it into GULP, you see what GOM carbon monoxides up. For the most part, a lot of -- what can come up. For the most part, most of the people who send the anonymous nasty e-mails, they'll go to the trouble of figuring out, I can't use my real e-mail address, you'll know who it is. Let me create this fake yahoo address and I'll send it from that. But as soon as they finish sending the bad e-mail, they log in with the regular e-mail address and guess what, same IP address, we know who it is. Done this millions of times. It works really well. Next question is, I've got a call from law enforcement, someone is applying for credit cards, using identity of employees, can you help. We got this call and we put it into GULP and turns out that that they were creating credit card applications from a particular IP address, they were getting the information from someone who was temporary employee, working in one of our departments. They were creating -- they were basically getting credit cards under different people's names and using them to buy stuff. And then selling the stuff to make money. Currently that, person is in jail and because we put it into GULP, we found the -- they did the stupid thing of logging in that IP address using the e-mail address that we had given them. And we figured out who it is. Any way. Another thing is, this is an interesting one. If you have a department that maybe does their own thing and they put up machines, had one the other day, somebody -- somebody's machine was compromised, I sent an e-mail to the department saying you have a compromised system, it's doing SSH scanning. And they came back to me, I don't know where this machine is, we have no idea who this. So took the IP address of the machine that was in question, put it into GULP, and I said, well, here's the person that's using that machine, maybe that will help you, and sure enough, yes, that's a new professor, he brought a server with him from wherever he came from, and I know where he is. He's at the 14th floor of this building here. And they went over, and they talked to the professor, and sure enough, his machine that he brought from wherever he came from, was hacked. And it was busy scanning our network, looking, you know, doing SSH scans. So handy, handy thing, it took me two seconds to do and I got an answer that actually worked. GULP data mining. So now, this is something that GULP does for us that we use on a daily basis. Use it to discover compromised passwords, you can use it to satisfy audit requirements. And you can use it to expose MAC spoofers. The external auditors are going to ask you things like how do you monitor for log ins for employees logging into sensitive systems, put it into GULP, and you know who logged in. Even know what hour they logged in. You can build a filter for that. MAC spoofers, if you check the MAC address, you'll see the same person is using a million MAC addresses, because they checked their e-mail from different MAC addresses and you can actually figure out who they are and have a word with them about that. Compromised passwords. This is the last thing. One of the things we did, we built a system that looks at all log ins, correlates the log ins that looks for multiple log ins from the same person, 72 hours from different ISP's or countries, and an example is here. You can see this -- if you look down at the line with the arrow next to it, that person logged in from Taiwan, United States and China and line L means they logged into a building somewhere. This is a GULP of that particular log in, that account. You can see here, the lines that were important. Line L means they logged interest a building, they logged in from Taiwan, they logged in from China, they did that in five hours, and that's kind of China to five hours, I don't think they were doing that physically. Any way. So GULP is a really good tool for bringing together all different kinds of information. If you have "free love" or a management environment, it tells you the actual answer. And you can use it all the time. If you want to build your own GULP kit, send me an e-mail, and I will just make sure you put in the subject line GULP or GULP kit, and I will send you some code to build your own. And we have questions? If you have questions, send them to me, and here's my e-mail address so you can send me that e-mail if you really want GULP. Okay. Questions?

>> Thank you, Joel. So over the next few minutes, Joel can respond to your questions and comments. So please be sure to post any questions for Joel in the chat pod. We do have one already. And that is, have you had any faculty push back on privacy grounds, ours would get cranky about physical tracking.

>> Turns out that they don't get cranky at all, because they call me and ask me, somebody stole my machine, I want it back. Where is it? And I use GULP to figure out where it is. They want -- when they call them and tell them that their password has been compromised, they thank me right away. They said, oh, that's great. Thank you very much. And they change their password. We went through general counsel's office, we got permission to do this. We're not doing anything that we couldn't do without GULP. So if you think about it, you can do all of these things yourself, it just takes you three programmers and six days of time. I can do it in under 30 seconds. But I'm just using the information that you already have. I'm not creating anything special. I just am reading the logs. So no. Nobody has given me a hard time about it. We have been doing this for at least 15, maybe 15 years or more. And I have not had problems with any of the faculty with this. Okay. Next question.

>> -- analyzing the GULP data?

>> None. It's all done automatically.

>> All right.

>> So I get a look at a report once a day, it takes me about 5 minutes to look at the record and decide which ID's may be comp MIEed and that's about it -- compromised and that's about it. We are not using 802.1X anywhere. We have two different kinds of Wifi, we have free Wifi, open Wifi, and then we have registered Wifi if you want to use it. It's optional. But we don't have any other kind of NAK or MAC registration stuff. Let's see. Do you use GULP for proactive monitoring? It's actually, it's incident-based stuff. If I get a question, we use GULP to find it. I'm not really interesting in finding out where people are, or what they're doing, couldn't care less, but if there's a problem, we go right to GULP to figure out where it is and what the problem is about. And how to fix it. Or how to answer the question. It's a really great investigative tool. I can tell you that public safety is probably one of my -- is my biggest customer for this, because they get the phone calls, especially around finals time, they get a lot of phone calls from people saying, you know, I can't find my son, my daughter, whatever it is, can you help us? First thing they did is they call me and say, you know, can you check this ID and make sure that they're CHAK checking their e-mail and they're fine. It has happened where we found out that somebody wasn't fine. We have used GULP, we found that they were logged in the last time they were logged in was from a hotel room somewhere, and the public safety people went there, and it turned out that the person had jumped off a bridge. And they figured it out because the laptop was there with a picture of the bridge on the screen. I doesn't happen very often. It happens once or twice in the time I've been here, which is 40-plus years. But it has happened. And GULP did help us figure out what did happen to that person. So --

>> Joel, our thanks again for sharing your insight. If anybody has additional questions, for Joel, they can ask it in the chat pod of our online room. We need to move on to our next content. So we are going to go silent for a few minutes while we get ready for the next presenter. Thanks, again, Joel.

>> You're welcome, thank you, everybody. Bye. [ Please stand by for webinar to resume]