RSSAC Packet Sizes WP Teleconference Call

May 11 I 17:00 - 18:00 UTC

Attendance: Duane Wessels, John Bond, George Michaelson, Kevin Wright, Robert Story, Andrew McConachie, Mario Aleman

Notes:
John: In the last meeting we only had Duane and myself.

George: So the meeting before that we had the code just published. And people we're considering using it.

John: I looked at Geoff's code and used it in a Scapy script. I tested it on our infrastructure and I can confirm that it acts in a default behavior. We use NSD and Knot but I suspect BIND would also have this behavior. The MTU will be set to the IPv6 minimum for that interface. In the normal world that 1240. Unless NSD or BIND unless they use their own configuration they will use MTU 0f 1240 regardless of Linux kernel config.

George: To recap MTU is hardset by at IPv6 min in the code for both NSD and BIND.

John: I looked at the code of BIND and it also does this. I looked at BIND and it is hard to read so I'm a little less confident. In Knot it will go with an MTU of 1500 because L does not change their config. We also do not set EDNS advertise size so we will fragment if on NSD above and on Knot above 1500.

George: If we imagined a table with roots as rows and behaviors as columns. We are in a position to fill in the blanks for known code levels. ICANN is sitting standard settings.

John: And on places that are not running NSD or BIND it will fall back to iface behavior. And on Knot it will do what the kernel does. The kernel does some PMTU discovery and knot follows that.

George: The fallback to PMTU has small differences on 1500 but it might attempt to probe on the path.

John: PMTU doesn't happen on the other systems, from my reading it says don't even bother. So on NSD it does not do PMTU and on the knot ones it does. It might change in the future. We said using the socket option it is no longer a good idea and instead should use PMTU.

George: That's very informative. Thanks. Your scapy script, how comfortable are you with that being generally available?

John: It's on github and already available. I'm not sure when to transfer it to the RSSAC-Caucus github. I'm happy for it to be available.

George: So now we have 2 ways to measure MTU settings, and some measurements as well. So basically for one of the root letters we can be mostly authoritative for the 7 questions in the SoW?

John: Yes, I do think we have the information for one root letter. But it wouldn't go into the document, but now we have a way to determine it, a methodology for finding it. Specific config options determine what we are looking for here.[Lists a few questions for operators]

George: That's a wonderfully simple list of things to ask people to confirm.

John: Yeah, I think those questions should be simple to answer. And it gives the information we want.

Duane: So just so I'm clear, we're talking about asking the operators to answer these questions. Whereas we also discussed in the past just doing the measurements instead of just asking.

George: Yes.

Duane: My concern is that the last time we asked this it was really hard.

George: I'm not convinced that eventhough it's easy to describe, we can do the measurements correctly. Certain things like PMTU are very conditional.

John: I think the problem there is vantage points. We could try something like ATLAS or NL-NET.

George: I would say on like NLNET we could try it at the lower level. Scriptable dig maybe, but then that becomes a statement like, using dig instead of classic DNS. But they do use python, so if there was a way to do Python. We only need packets on the wire.

John: Scapy uses raw sockets so we'd need root. If we wanted to do it with dig you'd need libpcap to see if they're coming in. You need to be root no matter what. That's the problem.

George: I don't always like this behavior, but we are a group of people with a mission. We don't need to boil the ocean. Kind of in that process committee thing. What can we usefully do here to properly document the problem space and get that to people who can make a decision. Here are a series of questions that we can get to people. Duane said it's really hard to get answers, but we could still ask. We might get some answers.

Choice 2 we could try active tests, and the problem is that we need root access, which it tough.

Duane: I think asking the questions is OK, but we need to give lots of detail. Like, do your hosts use MTU discovery. Like on Linux you run this test.

George: I absolutely agree. John, you've indicated in the past that people aren't comfortable with running code.

John: I think it depends on how complicated the code is. I actually struggled to run that script and it would pull out the information that I needed. BSD and Linux. But if we can say this is how Knot, BIND and NSD works, and as long as you have not modified the kernel or daemon code, we should be able to figure it out.

George: I could live with that. If 85% of the roots can answer this then we could say that these configurables answer most of the questions.

Duane: So if we continue down this path we need to write down these questions. But I think we need to expand on these questions.

John: I agree.

Duane: can we do this on the list. Someone take a stab on the questions.

Robert: John you said you'd already taken a stab at a script.

John: It might be lost. I'll have to look around.

George: OK small steps. Let's try to get these questions captured. We should look into how easy it is to answer these questions. I like the questions, they're like, did you modify the code. Asked the right way, if we know you didn't change anything you are in box A, or box B, etc.

The more complicated questions are where people have modified things like the kernel or code. That would be tougher. We put that forward as a model and we ask about the applicability of that.

Duane: The thing we need to keep in mind is that even within an operator there will be varying behaviors. So we'll need some way to capture that.

George: And that goes to, we can't answer all questions. I'm skeptical that we can write a magic script to test all these things. I wanted to believe there was a simple text, now I'm not sure.

Duane: I agree it's hard to do that test.

George: Risk analysis is going to be pretty important in this.

Duane: What do you mean?

George: If we propose to the root-ops a simple matrix with defined behaviors, we should put a caveat that says there might still be ways to contradict those answers.

Duane: That will be necessary. Still this is an area where the operators have been picked on recently. But that should not stop us.

George: We might as well say, this is what we think.

Kevin W: If we end up publishing this data, will that give any adversaries any better attack vector to attack the RSS?

Duane: I can't think of anything, but that doesn't mean it won't happen.

George: I think knowledge of knowing which softwares are being run(BIND, Knot) I don't think that would really motivate anyone to attack. The wider knowledge about a set of v6 behaviors exist. I think it's a more general problem about how open we are about how systems work. How open should we be about how our systems run? This is maybe just another part of the risk analysis.

John: I suspect that someone will make that case. But from a practical PoV the tools we've created are already in the public domain. From my testing the tools give the same information, also from spot tests. I get pretty consistent results. I guess someone will make that case, but the info is public if you look.

Kevin W: That helps answer my concern.

George: Unless there is something else, I think we're done with this call.

Duane: I wanted to talk about the last RSSAC meeting. I gave an update to the last RSSAC meeting. I said there was a bit of apathy towards this, but today's meeting gave me a good feeling. How would you feel about starting to write a document and putting some deadlines around that?

George: Who is going to be lead author? Someone needs to take ownership. I think it's completely reasonable for the RSSAC to ask if this WP is going anywhere.

Duane: This WP needs to decide if we want to take on this work and start wriing a document. I can support either way, but I do think we need to start making progress.

George: So who is going to take the pen?

Robert: I can start to put something together.

George: I don't want a big document. I think we should try to find a simple plan English response to this. We haven't said authoritatively we should say it. The rest of it I feel John has already done a good job. We've moved up a layer in the stack and go the moment of these questions.

Duane: I see your point. My only advice would be to we should try to basically everything that was asked in the scope. I think RSSAC finds it frustrating when a WP doesn't stay in scope. We can also explain that something is not the right question.

George: I'm saying I don't get a sense that the root ops expect to be asked just to answer these questions. John has come up with a script, but RSOs don't want to just run a root level script. We don't have a straightforward way to answer that. But we do have some good questions based on what John suggested that we can ask. So for one of the questions we can just ask, what software do you run and did you modify it? I think John's approach about asking a higher level question is the shorthand to get us there.

Duane: I'm looking at the scope and you're talking about item #1. But there are also these others things like defining a registry, etc.. Those things as well are supposed to go into our doc, and we need to do it.

George: I think the problem here is that seek advice from relevant people, but the relevant people can't really answer them. What is the specific reason that would provide that they should be all the same. I don't think we are in a position to say that they should be all the same. I think I probably say in the report that we don't know. I'd even go further and say we don't even know if it is an important question.

John: I think it is an important question, and I think the diversity is important. Half of the roots are setting the TC bit and half are fragmenting. Some people don't like UDP and some people don't like fragments. Because of this diversity we've been able to respond to more people.

George: I think that's defensible, but it's a statistical position. I'm stuck with only one Anycast server.

John: Then you're using the RSS correctly. I've worked with 2 orgs with the roots and I do think that is by design. I use a rarely used Linux system just for the diversity. I think all the operators look at this diversity and make decisions based on it. It's not an accident.

Duane: I think the WP has these open questions, and once we have a draft it goes to the wider caucus and they can comment on it.

George: The variance question is item #4. I'm proposing the direct questions are too difficult to answer, so we have indirect questions. Then there will be this grid for the registry. I think we should promote variance. The hard part is between 3 and 4. 1 and 2 is not that difficult, we can do that. I guess you're also saying with respect to variance, why do we think that?

Duane: Yes, the justification. What I want to avoid. What I've seen happen is that the WP is asked a specific question. And what comes back is that we don't really know.

George: and you're trying to set a deadline. Just out of interest, what would the deadline be?

Duane: They're tough. I think we're at least 6 months away from what we think is a finished document.

George: I think we should try writing. And we can get words easily for 1 and 2. 3 and 4 I'm not sure.I think we need a F2F meeting to discuss 3 and 4. Can we meet in Montreal and put forward an outline position?

Duane: At least someone could make a presentation to the Caucus and we could ask for input from the Caucus? Could you do that George?

George: I would happy to do that.

John: I will not be there.

Duane: I will not be there.

Robert: I will not be there.

George: I will make a presentation in Montreal about the progress. Pointed code, questions to ask, what do we want to do about variants.

ACTION ITEMS: George to present at Montreal. Andrew to send a starter Google doc, schedule a call for IETF 102 week, and send the SoW to the WP.