From:	(b) (6)
To:	Alagic, Gorjan (Assoc)
Cc:	Perlner, Ray A. (Fed); Apon, Daniel C. (Fed); Dang, Quynh H. (Fed); Moody, Dustin (Fed); internal-pqc; Dang, Thinh H. (Fed)
Subject:	Re: Kyber"s response discussion tomorrow ?
Date:	Friday, June 5, 2020 11:46:07 AM

sorry nano-seconds!

On Fri, Jun 5, 2020 at 11:41 AM Daniel Smith <(b) (6) > wrote:

> Ray's intended comment:
>
> "The Kyber team thinks it is impressive/ infeasible to build micro-SD cards to cover New York City; this is nowhere near the level of security that we are asking."
>
> Agreed.
>
> My interpretation of Ray's comment:
>
> "We are defining level I security on a razor's edge."
>
> My extremely brief reaction and reply is, "Yes, that is exactly what we are doing." Given our observations of technological progress and the open possibilities of breakthrough technologies in conjunction with much greater wealth, there is no way that 10000 years is anywhere near a lower bound.
>
> That all being said, I think that it is fairly reasonable to incorporate these issues into the complexity, but very clearly not at the level of New York City sized things and not in a very asymmetric way across the candidates. A team has justification for their claims that we will find convincing or they do not.
>
> As far as throwing misleading analogies out there, let's put this in another perspective. Micro-SD cards are approximately 1 mm thick. So we only need a block about 92 meters on a side of memory as dense to access to have 2^89 bits of memory. At 92 meters, it takes at most about 300 micro-seconds (the physical limit) to communicate. Why not put the memory in that perspective?
>
> Maybe they should have put it in a 1-dimensional perspective. Then we get something around 70 million km, so about half the distance from the earth to the sun. Is that more impressive? How are we supposed to take these discussions seriously?
>
> On Fri, Jun 5, 2020 at 11:33 AM Alagic, Gorjan (Assoc) <gorjan.alagic@nist.gov> wrote:
>
> > I'm probably the least qualified person to be talking about concrete bits of security. But let me do it anyway...
> >
> > One take on what Ray is saying is something like this. If one buys the argument "don't worry about memory attacks that require NYC-sized microSD cards because that's absurd" then one might also buy the argument "don't worry about attacks that require a solar panel the size of North America and 10,000 years because that's absurd." But clearly we

don't buy the latter argument, because we *do* want bits of security to be that high. So maybe we shouldn't buy the former argument either.

Now, why do we buy the latter argument? Is it because we're actually worried about people covering North America with solar panels and computing for 10,000 years? (But not, say, covering North America *and* South America with solar panels and computing for 50,000 years?) Of course not. It's because there's a lot of uncertainty in our estimates of the actual concrete complexity of these algorithms, and because we have to account for decades of future progress in theory, algorithms, hardware, etc. And these things are really hard to understand and estimate, so we give ourselves a large margin.

Now, can we be at least somewhat scientific about how large this margin should be, both in cases of time complexity and space complexity? Probably. Should we be open about how we're doing these assessments and how we're selecting these margins? I think so.

Also: I agree with DanielS that we need to be even-handed in applying our assessments to all schemes, to whatever extent that is possible.

-Gorjan

From: Perlner, Ray A. (Fed) <ray.perlner@nist.gov>
Sent: Friday, June 5, 2020 11:10 AM
To: Apon, Daniel C. (Fed) <daniel.apon@nist.gov>; Daniel Smith (b) (6) ; Dang, Quynh H. (Fed) <quynh.dang@nist.gov>
Cc: Moody, Dustin (Fed) <dustin.moody@nist.gov>; internal-pqc <internal-pqc@nist.gov>; Dang, Thinh H. (Fed) <thinh.dang@nist.gov>
Subject: RE: Kyber's response discussion tomorrow ?

Just as an aside. I'm supposed to be impressed by "**A planar sheet of terabyte micro-SD cards the size of New York City (all five boroughs, 800 km^2 ~ 2^49.5 mm^2) would hold 2^89 bits.**"?

Yes. It's a big number, but the computational complexity we're asking for even in category 1 is HUGE. I did some calculations. Supposing you didn't just blanket new York city with solar panels operating at 20% efficiency, but the entire continent of north America. That would get us about 2^51 W. I did a quick Google for a high-end bitcoin miner, and found Antminer S17+ advertising 73TH/s at 2920W. Assuming 1 Hash is a double SHA2 operation, i.e. 2^19 bit operations, this corresponds to 2^53.5 bit operations per second per watt. Putting this together with our power budget, that's 2^104.5 bit operations per second. There are 2^25 seconds in a year. Thus category 1's classical security requirement of 2^143 bit operations comes out to the computational capacity of an array of top end bitcoin miners powered by tiling North America with solar panels running for 10000 years.

From: Apon, Daniel C. (Fed) <daniel.apon@nist.gov>
Sent: Friday, June 5, 2020 3:41 AM
To: Perlner, Ray A. (Fed) <ray.perlner@nist.gov>; Daniel Smith (b) (6) ; Dang, Quynh H. (Fed) <quynh.dang@nist.gov>
Cc: Moody, Dustin (Fed) <dustin.moody@nist.gov>; internal-pqc <internal-pqc@nist.gov>; Dang, Thinh H. (Fed) <thinh.dang@nist.gov>
Subject: RE: Kyber's response discussion tomorrow ?

I felt that the key point in the Kyber Team's response was

"We agree that … 141 [is] smaller than 143, but at the moment we do not consider this to be a sufficient reason to modify the Kyber-512 parameter set.
…
The additional memory requirement of this attack strongly suggests that Kyber-512 is more secure than AES-128 in any realistic cost model.
…
A planar sheet of terabyte micro-SD cards the size of New York City (all five boroughs, 800 km^2 ~ 2^49.5 mm^2) would hold 2^89 bits.
"

I still feel we should do our own internal analysis at the start of the 3rd Round.

I'm utterly opposed to letting DJB's eleventh-hour protestations influence absolutely anything whatsoever.

--Daniel

From: Perlner, Ray A. (Fed) <ray.perlner@nist.gov>
Sent: Thursday, June 4, 2020 3:02 PM
To: Daniel Smith (b) (6) ; Dang, Quynh H. (Fed) <quynh.dang@nist.gov>
Cc: Moody, Dustin (Fed) <dustin.moody@nist.gov>; internal-pqc <internal-pqc@nist.gov>; Dang, Thinh H. (Fed) <thinh.dang@nist.gov>
Subject: RE: Kyber's response discussion tomorrow ?

Here are my current thoughts on the matter:

I am open to the idea of using a more realistic model of computation than the basic gate model. However, a lot of the ideas in I've seen in the literature seem too pessimistic (as in they reckon attacks as being harder than they should be. – at least in the long term)

DJB's favored model, for example, assumes the computation must be implemented by only nearest neighbor interactions in a 2 dimensional grid. This has some justification, in that trying to violate these assumptions clearly costs more than the basic gate model assumes, but

1. Today's Supercomputers generally use a meaningfully 3 dimensional arrangement of processors (although the processors themselves are 2 dimensional)
2. Long distance connections needing high performance are implemented by fiber optic cables, and sending a bit through a kilometer of fiber optic cable, while more expensive than sending the bit across a single AND gate, clearly costs less than sending it through a kilometer of densely packed AND gates (which is how DJB's favored model would treat it.)

NTRU's "local" model seems in practice to be even more extreme, simply ignoring any algorithm that hasn't explicitly been implemented locally

Hard limits on the total memory size have also been proposed. I think the smallest numbers I could really convince myself were commensurate with an adversary actually capable of threatening the appropriate security level were 2^100 for levels 1 and 2, 2^150 for levels 3 and 4 , and 2^180 for level 5.

One could perhaps adjust the RAM model to cost random access queries to a memory of size N at N^(1/3) in terms of depth and (log(N))^2 in terms of gate count and require all other gates to be local. (I think I might actually be ok with that, keeping in mind that if the whole thing can be implemented locally, you don't need to make RAM queries, no matter how large the computation is.)

The other worry though is that things like memory cost are much more susceptible to being optimized away by incremental improvements, which the first iteration of a new attack rarely includes. But there are a lot of smart lattice people, so maybe I can be convinced they've thought about this stuff enough that there is no room for further improvement. I'm not convinced yet, though.

Ray

From: Daniel Smith (b) (6)
Sent: Thursday, June 4, 2020 2:23 PM
To: Dang, Quynh H. (Fed) <quynh.dang@nist.gov>
Cc: Moody, Dustin (Fed) <dustin.moody@nist.gov>; internal-pqc <internal-pqc@nist.gov>; Dang, Thinh H. (Fed) <thinh.dang@nist.gov>
Subject: Re: Kyber's response discussion tomorrow ?

Hmmm...

They are calling us out explicitly to offer our position on this. It is a muddy issue in my mind.

I have a bit of a problem with saying, "We are secure because of other stuff that we can't measure really well." For other areas we have been requiring them to ignore memory costs even when that makes a difference for them.

A clear example comes to mind: GeMSS. For GeMSS they had a quite exhaustive analysis of known techniques applied to GeMSS. They quite conservatively used analyses and coefficients that are unrealistic even with zero cost of memory and memory access (which is why confusingly they chose to report some of the numbers as lower than the security bounds when actually they should be fine). When you consider the hidden polynomial factors or actual coefficients, the least costly attack (and the one they are basing the parameters on) is the direct algebraic attack. They are being super conservative and choosing a linear algebra exponent of 2 for dense linear algebra (I think that we can't use sparse techniques here because of the number of solutions (or the density after fixing variables)), but if we take memory into account, then the complexity is altogether different. If our metric is New York City, then this scheme should benefit fairly significantly.

On a historical note, Ray and I argued fairly extensively about this memory issue when we were drafting the CFP. I recall having discussions about the physical feasibility of converting Jupiter into atomic scale memory that violates causality with the speed of its access (sending replies and being set to different values before being asked to) leading up to the release of this document. The issue as I recall was allowing the community to address some complexity issues that had not been pinned down yet at the time and for the community to come to a consensus on how to address these things. Still, we need to have some standard metric for comparisons between schemes.

I think that it is entirely reasonable to address memory and memory access in a cost model. A problem occurs when we lack justification and when we lack consistency in how we apply restrictions in these analyses. Ray and I were arguing on the level of Jupiter and breaking the laws of physics, whereas Kyber is arguing on the level of the 5 boroughs.

I would be open to allowing teams to specify their cost model addressing memory (in communication with us and with clear justification and theoretical support), and to adjust parameters accordingly. This would need to take place extremely quickly, though, to not make analysis placed on a moving target.

The easiest way to handle the situation is exactly the opposite, though. That is to let the teams do what they are doing and then judge them by our own metrics. The downside of this approach is that there is plenty of room for bias and plenty of reason for skepticism in our choices if any parts of our community think that we are cutting corners unreasonably.

If we chose to allow memory access cost as part of the complexity analysis, there will be consequences. We may have to communicate with each team explicitly, but I think we should make it clear (if we go that route) that they should analyze the memory concerns with strong justification for **minimal** cost models that they can then incorporate. We also need to assess the feasibility of these models and the appropriateness of the bounds they suggest.

I think that we have plenty to talk about, but we'll follow your lead, Dustin.

Cheers,

Daniel

On Thu, Jun 4, 2020 at 1:47 PM Dang, Quynh H. (Fed) <quynh.dang@nist.gov> wrote:

> I think so. If more people think that a talk tomorrow would be good, then I would ask you to consider that.
>
> ---
>
> **From:** Moody, Dustin (Fed) <dustin.moody@nist.gov>
> **Sent:** Thursday, June 4, 2020 1:41 PM
> **To:** Dang, Quynh H. (Fed) <quynh.dang@nist.gov>

Cc: internal-pqc <internal-pqc@nist.gov>; Daniel Smith (b) (6) ; Dang, Thinh H. (Fed) <thinh.dang@nist.gov>
Subject: Re: Kyber's response discussion tomorrow ?

I think we can discuss via email.

I don't think we need to have a meeting tomorrow. Maybe on Tuesday.

Let me know if you think otherwise.

Dustin

From: Dang, Quynh H. (Fed) <quynh.dang@nist.gov>
Sent: Thursday, June 4, 2020 1:34 PM
To: Moody, Dustin (Fed) <dustin.moody@nist.gov>
Cc: internal-pqc <internal-pqc@nist.gov>; Daniel Smith (b) (6) ; Dang, Thinh H. (Fed) <thinh.dang@nist.gov>
Subject: Kyber's response discussion tomorrow ?

Hi Dustin,

Are we going to discuss Kyber's response tomorrow at 10 ?

Quynh.