Hello.
Forced data localization has no significant security advantages.

We'll start with the takeaways from our research on forced data localization, the exciting new government plan that unites Canada, China, and Vietnam in Internet governance strategies. First: it doesn't significantly help security over simply not doing it, just like a cardboard TSA station. Or a real one.

https://www.flickr.com/photos/kentwang/21492768295
Requiring data to stay in small geographic areas has significant impacts to data integrity and availability.

The second is that countries are small, and natural disasters in the age of global climate change are big---so confining data to a tiny country, even one as big as the US or Canada, destroys your business’ data either temporarily or permanently. Either one will likely destroy your business as well.

https://www.flickr.com/photos/nudelbach/15671098451
Forced data localization imposes many additional costs on businesses, including both storage/cloud costs and personnel costs.

Finally, forced data localization has tons of added costs, both technological and personnel.

https://www.flickr.com/photos/michaelandannabel/16664159481
WHO?

WHAT? WHY?
  WHY THEY SAY THEY DO IT
  WHY THEY DO IT
  SAFE HARBOR

PLACES

THINGS

PEOPLE
JAMES ARLEN

• Director, Risk and Advisory Services @ Leviathan
• CISA, CRISC, CISSP
• Extremely Proud Hamiltonian (just ask him)
• #HamOnt
BRENDAN O’CONNOR

- Tame Lawyer @ Leviathan
- IANACL, but I’m from Montana, which is kind of like the US’ Manitoba
- Purple

Brendan bio summary
“Funding for these papers was provided by Google. All opinions in this study reflect the independent views and analysis of the author(s) alone.”
WHO?

WHAT? WHY?
  WHY THEY SAY THEY DO IT
  WHY THEY DO IT
  SAFE HARBOR

PLACES

THINGS

PEOPLE
So, um. What's forced data localization? Do we care? We have a lot of other things to worry about.

So we opened the presentation by telling you that what we're talking about doesn't work. So what is this and why should you, as technologists, care? Well, you may have heard of this before, but you'll definitely be hearing a lot more about it soon.
"If a company holds data on a citizen of our country, that data must be held within the country."
(Or the EU)

—FORCED DATA LOCALIZATION

This is a broad definition, but the key idea is geographic restriction. It can take many forms; you might be familiar with the EU Data Protection Directive (you may have heard of the epic slap across the face the US got from the Europeans last week; if not, we’ll get there soon), but this also exists in either proposed or enacted form in China, Vietnam, Brazil, Russia, and Canada, among others. Canada is a weird case. (James)
Let us assume for a moment that the government is not out to get us and destroy everything we love. (This is a good assumption here in Canada, at least in most years; in the US it’s a really bad assumption.) Why would they do this?

One reason is regulation. It turns out to be tricky to apply your country’s laws outside of your country. ("Tricky" is a very technical legal term in public international law. It also applies when the US loses a case at the World Trade Organization but keeps doing the same thing.) If you keep everything in your borders, your domestic laws clearly apply in full force, whether that means privacy protection, warrant, or destruction laws.

https://www.flickr.com/photos/diversey/15966987237
The second reason is a traditional winner: protection of local industries, in this case, data centers. As we will see later, most countries have no public cloud providers, so the only data center options are likely the local telecoms---the same telecoms with enormous lobbying power to enact rules like this. (Until very recently, Canada was one major country with no public cloud providers.)

https://www.flickr.com/photos/icanchangethisright/4879511760
A third reason that goes to the heart of why this is such a political issue right now is, well, the ongoing set of disclosures about all the special friends the US intelligence service have made over the last few years. It’s a very populist response to push data localization in response to these revelations; it makes people feel like they’re sticking it to the US, who so richly deserves it. This was particularly true for Brazil, who realized that domestic Internet connectivity often routed across the US, and pushed not only data localization laws (eventually defeated) but a massive new South American fiber building project that is still ongoing.

Of course, as we know from those same documents, it’s not like the US’ tapping is limited to US fiber. Data localization merely means that the US and its allies will come to your country to take your data. But it does help someone else:

https://theintercept.com/2014/02/10/new-photos-of-nsa-and-others/
All the *other* spy agencies. This is GCHQ, of course, but these laws help agencies with and without 23 movies made about their agents. Not every country has the worldwide networking reach of the US, so for countries who want extremely robust domestic intelligence gathering capabilities (also known as hack all the things, creep all the creeps), this is an efficient, if unsubtle, method of assisting that.
Heck, they’re even over in the expo area recruiting! Right now! (I thought about taking their photo, but they seemed to dislike that idea. Though they were very polite about their dislike.)

https://www.flickr.com/photos/defenceimages/7985695591
Of course, in that case and every other, the reason *given* is to keep the citizenry safe. So there's that.

https://www.flickr.com/photos/squiremorley/9032306448
So on that note, let's talk about the EU Data Protection Directive.

Those of you who secure data from the EU know about this: it's the stringent set of privacy guidelines the EU uses for PII. If you keep data in the EU, you follow EU rules; if you take data out of the EU, you have to go through a process of attesting that you will have at least the same protections. You can actually do this through several different means: you can have specific contract provisions provided to users, you can do certain things with EU or per-country regulators, or your country can have a treaty with the EU itself.

https://www.flickr.com/photos/tdlucas5000/15070603679/
Until recently, most people didn’t think much about the EU as a forced data localization area, because though the DPD is fairly stringent, the EU had a series of “Safe Harbor” agreements allowing export (with loose conditions) to the US and Switzerland. (The EU recognizes that unlike the US, Canada enshrines privacy in law, and therefore recognizes it as sufficient to protect EU citizens’ privacy automatically.)

So: EU-US Safe Harbor. The US Department of Commerce managed it, and what was required for compliance was a company to annually self-certify that it promised not to breach the privacy of EU citizens. If that sounds unbelievably lax, you’re right; the compliance regime was so useless that companies would often simply stick the logo on their website and not bother to do the paperwork, or---and this is my favorite---forget to fill out the renewal for a few years and just tell everyone they were Safe Harbor certified. In fact, a prominent US security consultancy got whacked (whacked is a very technical legal term, but another would be “was subjected to a twenty-year consent decree”) for doing exactly that just a few months ago.
Someone else thought this was too lax: the European Court of Justice. In a case brought by an Austrian privacy activist, they ruled that given NSA surveillance and the lack of any real enforcement, EU-US Safe Harbor didn't meet the threshold of the EU's human rights guarantees to its citizens. They therefore invalidated the program as a whole, effective immediately.

https://www.flickr.com/photos/puisney/1674586821/
Well, Safe Harbor is dead. For US businesses, that means that the individual company will have to negotiate data protection requirements with 22 countries' privacy regulators. If that sounds like a mess, yes, it is. It also puts the rest of the world on notice that the DPD has more teeth than was previously thought.

From the perspective of the US, then, the EU is now a forced localization state.

Oh, it'll also deal a large blow to the EU and US economies---but that was probably coming anyway when the US Second Circuit Court of Appeals gives the US Department of Justice the right to be international police without restraint in “In re Warrant to Search a Certain Email Account Controlled & Maintained by Microsoft Corp.,” AKA the US v Microsoft in re Ireland case.

But wait, there’s more!

https://www.flickr.com/photos/dcjohnc/32727942
This issue is bigger than one court case. The US Second Circuit Court of Appeals is currently deciding a case helpfully titled “In re Warrant to Search a Certain Email Account Controlled & Maintained by Microsoft Corp.”

What’s actually going on there? Well, the DOJ is trying to use antiterrorism laws to force Microsoft to turn over the contents of an email account. Ordinarily, they would—but in this case, the email account is of an EU citizen, so it’s held by Microsoft Ireland. The DOJ’s response? “Same logo, give us the data.” Microsoft pointed out that the US has a shiny treaty with Ireland to send warrants. The DOJ said “we don’t need no stinkin’ treaties.” Microsoft went to court.

The court held Microsoft in criminal contempt.

The EU privacy commissioner has informed Microsoft that if it complies with the warrant—even if its alternative is to send its CEO to federal prison—it will no longer be able to operate in Europe.

…So that’s gone well.

(And by the way, it’s not just antiterrorism laws the DOJ uses like this. They used Sarbanes-Oxley—the law we passed after one of our energy companies tried to eat our face—to seize fish. It went to the Supreme Court. Being a lawyer is awesome.)

Unfortunately that’s not the only instance of this happening. In 2004, the Privacy Commissioner for British Columbia pointed out that since BC privacy rules require that citizens are notified if their information is drawn into a criminal investigation (against someone else), and USA PATRIOT says it’s illegal to notify someone if their information is drawn into a criminal investigation (against someone else)....
Yesterday, the US House of Representatives passed this bill (introduced in March, as you see). It purports to give the citizens of certain countries (including, in theory, the EU and Canada) citizens judicial oversight and the right to sue US government agencies for misuse of their data. But it only does so if:

- The disclosures were “willful” (this means a bunch of things under US law, but to sum up, our Constitution now has a “good faith” exception—the “whoopsies” clause for human rights—and my guess would be that this will be similar) AND
- the Attorney General thinks the citizen’s home country is giving the US all the data the US wants to fight “crime” anyway, AND
- the citizen’s home country never “impedes the transfer of information (for purposes of reporting or preventing unlawful activity) to the United States by a private entity or person” AND
- the head of the agency that misused the data CONSENTS to be allowed to be sued, in advance, in writing, to the Department of Justice, AND
- If all this doesn’t make you happy as an EU or Canadian citizen, too bad, judges can’t review these determinations.
- Oh, and if they classify it, nothing makes them have to respond to you. So the Snowden revelations still didn’t happen for purposes of this law—even though those are one thing the European Court of Justice specifically talked about. So that’s going well.

http://src.bna.com/Em
So now we have a good overview of what data localization is, and how it comes about. Let's turn to our research questions for this project.
WHO?

WHAT? WHY?
  WHY THEY SAY THEY DO IT
  WHY THEY DO IT
  SAFE HARBOUR

PLACES

THINGS

PEOPLE
We begin with places---where we put data, and what effect centralizing data has.

https://www.flickr.com/photos/archer10/2590434432
The first question in our analysis of place is where the cloud is; since cloud computing is just another term for “other people’s computers,” it turns out that borders are very important in determining whose computers we should use or avoid.

To figure out where it is, however, first we need to state precisely what we are looking for; we need to define the cloud. I think most of you have a general sense that it’s not really cloud computing if you’re racking and stacking boxes you own, but where’s the line?

https://www.flickr.com/photos/davedehetre/4787216846
Cloud Computing has five essential qualities:

- On-Demand Self-Service Provisioning
- Broad Network Access
- Resource Pooling
- Rapid Elasticity
- Measured Service

-NIST / CLOUD SECURITY ALLIANCE

Thankfully the Cloud Security Alliance has a nice five-part definition that it stole from NIST.

I've highlighted the first in Orange, and that's because for our work, we added a bit of specification to this element. We define self-service to mean that anyone with a credit card should be able to sign up for the cloud service provider and use them immediately; no business licensing negotiation, no confidentiality on pricing.
We focused on compute capability---while there are great services that meet this definition for, for instance, storage *only* (such as Backblaze B2 or RSync.net), we think computation is the critical core. Given that, and our definition, we find just seven providers.
Now that we’ve defined the cloud, and we have some names for who to call if we need a cloud in a hurry, where can we look? This next slide surprised us. And we work in this every single day. But... The Earth isn’t really cloudy all over.

https://www.flickr.com/photos/gsfc/7394700302
“The Earth is, at best, partly cloudy.”

12 countries: US, Brazil, Ireland, the UK, the Netherlands, Belgium, Germany, Singapore, Australia, Taiwan, Japan, and Hong Kong. If you’d like to play with this map, it’s at ValueOfCloudSecurity.com.

And one thing to note on this map is the bright pink dot: that’s DigitalOcean’s TOR1, the first public cloud datacenter to be located in Canada. It was announced after we did this work, and we congratulate DigitalOcean on being the first cloud provider to recognize that, as the Arrogant Worms put it, Canada is really big.

Still, thirteen countries total: not very much. And it turns out that many of the countries yelling loudest about data localization have none at all. Interesting.
Of those seven providers, we have an interesting split between those who make it more expensive to compute in certain areas and those who don’t. For HP, we have no idea if they would or would not charge with regional differences, because their two regions are both in the US. But the biggest providers do all charge with regional differences.
PLACES - COST

HOW MUCH?

- Brazil: 54% cost savings using cloud outside Brazil
- European Union: 36% cost savings
- Schengen: 10% more expensive
- India: No Public Cloud
- Indonesia: No Public Cloud
- Russia: No Public Cloud
- Canada: Just DigitalOcean
So the summary here is that it costs a lot of money to force localization. You end up with single-source suppliers, which ends poorly, and in the vast majority of the world with no public cloud, we’re back to the old, extremely capital-intensive, method of doing Internet computing: buying hardware. This is real microeconomic costs of the forced localization.

There are also macroeconomic costs: the European Centre for International Political Economy found that forced localization laws destroy (remove without giving back value) between 0.7% and 1.1% of GDP as a whole. That’s a devastating impact on the scale of a national economy.

...but we’re security people; we don’t write the checks. Are there problems for data security in forced localization?

https://www.flickr.com/photos/doctorow/13446520523
In Calgary, the primary data center building is owned and run by Shaw. Like any building, it had its share of problems caused by age, including a persistent, but very slow, drip from a sprinkler pipe in the back of a battery room that supplied backup power to the data center. Over time, the calcium in the water built mineral formations.

On July 12, 2012, water dripping onto these formations shorted a battery, causing an explosion. (This picture is, sadly, not of the explosion, just from steam systems in the building being vented.) Since this tore through the power distribution center, it knocked the primary systems offline as well.

https://www.flickr.com/photos/mdaines/4438382520/
How bad was it? This is Calgary now.

No, but it was in fact really, really bad. This, as it turned out, was either the host for or in the critical path of a huge number of systems, all of which failed badly. An abbreviated list includes:

- Ambulance dispatch (all of the digital systems, fallback to voice)
- 3 hospitals’ EHRs
- Alberta Justice
- Alberta Health Service
- Checker Yellow Cabs
- Bell
- 3 radio stations
- Calgary Parking Authority
- Calgary Fire Department
- ATB Financial
- IBM
- 911 Service
- (Ironically) 311, the service you call to figure out why 911 isn’t working

Note that this is still the primary data center facility in Calgary---and that until very recently, Canada had no public cloud providers. Why would a public cloud help? In a
There have actually been a large number of outages that have taken out single Amazon Availability Zones—US-East-1 was famous for years for being uptime for the whole Internet. Turns out if you run a large chunk of the whole Internet, everybody notices if you misconfigure a router. In 2011, though, a single lightning strike near Dublin, Ireland managed not only to shut down a Microsoft Azure datacenter, but also shutdown an AWS Ireland datacenter—and corrupt Elastic Block Store devices, which took a "huge" amount of time to fix.

https://www.flickr.com/photos/usfwsmtnprairie/14538796137/
Joplin refers to a particularly devastating area tornado in Joplin, MO and surrounding regions of the American South. For reasons I will leave as an exercise to the audience, there aren't very many cloud data centers in the rural South, and Joplin didn't hit anything in that category. However, this is changing fast; both Google and Microsoft have located cloud data centers in major tornado zones (Iowa and Virginia), and so it’s entirely possible that the entire Internet will notice the next weather pattern.

https://www.flickr.com/photos/dvids/5864127016/
Moving upward in scale, you may remember the Internet being shut down for nearly all of Egypt a few years back, and the dramatic response by the modem (both phone and ham radio) teams of the Telecomix Crypto Munitions Bureau. A lack of distribution meant the primary Internet in Egypt was easy to cut off, and the distributed nature of Telecomix helped to put it back together.

https://www.flickr.com/photos/adoyle/7952517908
Syria losing its whole-country Internet a few years back looked at first like another instance of a government shutting down communications. However, subsequent document releases point to a possibility of it being a screwup by a non-Syrian party that had penetrated the core routers. Whoops!

https://www.flickr.com/photos/christiaantriebert/7955551210
When you think of the Internet, you’re told to think of a globally interconnected web with no single points of routing failure. Neal Stephenson's Mother Earth Mother Board (which you should all read, if you haven’t done so already) pointed out that wasn’t particularly true—and decades later, it’s still not. There are plenty of US-EU links, of course, but look at North Africa. Notice how it all looks like one dense line?

http://www.submarinecablemap.com/
In 2008, a boat traversing the Suez Canal dropped its anchor in the wrong place and took all of North Africa offline, as well as areas of countries as far away as Pakistan. One boat took much of the Arabic-speaking world off the Internet.

https://www.flickr.com/photos/94341334@N08/16518682613
Of course, sometimes widespread disasters actually have big, unavoidable causes. You are all familiar with the Japanese tsunami, which took out a large variety of Eastern Pacific datacenters (due both to the water and the ensuing power loss). Unfortunately, as we mention in the paper, the increasing number and force of Pacific typhoons means that scale of disaster is going to become more frequent. If your cloud strategy relies on one small chunk of the world---essentially, anything you can see on one side of the globe---these are going to get you.

https://www.flickr.com/photos/jetalone/5822881926/
In North America, we have our own insanely huge storms: hurricanes, and things that are like hurricanes. (SuperStorm just sounds like it was made up by CNN.) And when you see flooding in the streets of Manhattan, or read the blog posts by a former special forces operator defending a datacenter with a pump shotgun, remember: get around the world. Fail elsewhere.

https://www.flickr.com/photos/jasonahowie/8142908668/
That is ultimately the message. When we're young and play soccer (or football), we clump together because that's where we think the ball is. When we learn strategy, we learn that a whole bunch of bad things---like the other team scoring---can be minimized or prevented if we spread out over the whole area we have available. We can't do a reasonable lunar computing center yet (although I heard Google is working on it), but we have a whole planet: use it. It'll save your wallet, and it'll save your data. Between the two, that might be enough to save your startup idea of iPhones for plants or whatever the kids are doing these days.

https://www.flickr.com/photos/usaghumphreys/6205421569
WHO?

WHAT? WHY?
  WHY THEY SAY THEY DO IT
  WHY THEY DO IT
SAFE HARBOR

PLACES

THINGS

PEOPLE
Now we turn to the second focus of our inquiry: building datacenters. Or put another way, is it cheaper or more expensive to use the cloud for storage as compared with building storage datacenters to hold our stuff?

https://www.flickr.com/photos/kewl/8475764430
It turns out that this question is difficult to answer. But we have a theory.

https://www.flickr.com/photos/doctorow/14452760238
### METHODOLOGY

<table>
<thead>
<tr>
<th>THINGS</th>
<th>Small Consultancy</th>
<th>Mid-Sized</th>
<th>BigCorpInc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50 employees, 20 remote</td>
<td>500 FTE, 4 satellite offices</td>
<td>2500 FTE, 2 HQ, 6 satellite offices</td>
</tr>
<tr>
<td></td>
<td>300GB/user, 15TB total</td>
<td>In-office server room</td>
<td>Two datacenter</td>
</tr>
<tr>
<td></td>
<td>Uses NAS</td>
<td>150TB total storage</td>
<td>750TB storage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enterprise-class storage</td>
<td>Dual-Replicated SAN</td>
</tr>
</tbody>
</table>

We set up three mock corporations at three different sizes.
### THINGS

<table>
<thead>
<tr>
<th>METHODOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Small Consultancy</td>
</tr>
<tr>
<td>• PowerVault NX400</td>
</tr>
<tr>
<td>• 1.8GHz Xeon</td>
</tr>
<tr>
<td>• 2x500GB RAID1</td>
</tr>
<tr>
<td>• 8x3TB RAID5</td>
</tr>
</tbody>
</table>

| • Mid-Sized  |
| • Storinator Redundant NAS  |
| • 2x500GB RAID1  |
| • 90x4TB  |
| • Tandberg T160+ Tape Library  |

| • BigCorpInc  |
| • Petarack High-Availability SAN  |
| • 512GB RAM, replicated 750TB of storage  |
| • 5x Tandberg T160+ Tape Library  |

Then we built and priced out entire specific solutions. Including tape.
So we took a large number of cloud providers. Which plan we picked, and which providers we excluded, goes to specific details that I’ll leave to the paper (and I’ll give you the link to the papers at the end). But you can see that the range of prices is extreme—and not particularly correlated to value or to security features.
Here’s the costs, and cost comparisons, for each. An the numbers are by and large close—not for the mid-sized, but for small and large. But cloud is more expensive—at the way we priced out. Note that these costs don’t include depreciation and obsolescence of hardware after a few years.

On the other hand, this deals with single replications of data (except for BigCorp). So it’s cheaper to do one centralized datastore, but as we now know, one centralized datastore is likely to cause you to have a sad the next time a tornado wipes out your datacenter. When you start doubling these costs (or tripling them), having someone else handle replication looks increasingly good.
WHO?

WHAT? WHY?
   WHY THEY SAY THEY DO IT
   WHY THEY DO IT
   SAFE HARBOUR

PLACES

THINGS

PEOPLE
We’ve talked about places, we’ve talked about things, now we’re going to move to the third component of security: people. And we all know that in order to have true security, you need people in black hats with projectors pointed at their faces. (This photo courtesy of the amazing HackerDojo photoshoot.)

https://www.flickr.com/photos/brianlug/6870005158
So setting aside projectors in faces, this is an important question. At every level of infrastructure, we need people that actually defend that infrastructure. And we need a lot more of them.

https://www.flickr.com/photos/15609463@N03/9271442721
With more than one million cybersecurity positions unfilled worldwide, currently-identified security needs couldn’t be met if every employee at GM, Costco, Home Depot, Delta, and Procter & Gamble became security experts tomorrow.

This is a terrifying statistic; the one million comes from the Cisco security report. There are 25,000 needed in the US civil service alone (and as recent events have demonstrated, they’re really quite desperately in need).
So: we don’t have nearly enough people. Not even close. How are people trying to solve this?
https://www.flickr.com/photos/atin800/5269172393
Most of the Global North and BRICS countries are trying to educate security experts with formal university programs—mostly Masters and PhD programs, a few bachelor's degrees. The problem is that these programs don't scale—the UK’s GCHQ-designed program will start to graduate 66 PhDs per year. In 2017. The EU’s TEMPUS (Trans-European Mobility Program for European Studies) program noted that while there are bachelor's degrees available, “Master degrees are essential for providing a cybersecurity workforce with advanced capabilities.” So there's a clear bias there. And higher education is great, don't get me wrong—but it's not essential to defend the castle.

https://www.flickr.com/photos/deanhochman/8314450897
Ha ha ha ha ha! We could do apprenticeships, but nobody does—look around this room. How many of you are hiring senior security people right now? How many of you are hiring novices with a real interest, but no experience or degrees?
https://www.flickr.com/photos/erin_costa/6270627111
Importing security experts en mass is one way that some countries are solving this issue. The US and Canada have, for certain people, Trade NAFTA Status which allows somewhat-free flow of experts across the border. The US has its H1B program which hires 85,000 people with advanced training per year from around the world—of which a majority are computer scientists, of which a majority are security people. The EU has—well, had—free migration between EU countries. All of these programs allow certain countries to get more security experts, if and when other countries lose them. It’s organized brain drain, which will eventually rise to the level of a national security issue for the countries that are losing experts.

https://www.flickr.com/photos/lac-bac/9936332974
So if we can’t educate, we don’t apprentice, and importation only works for a few countries (and not sufficiently for those), what do we do?

We need to circle the wagons. When people use Amazon, or Azure, or Google Cloud, it’s entirely untrue to say they don’t need to worry about security; we do security consulting for cloud customers all the time. It is true, however, to say that huge chunks of their infrastructure are secured for them. Most of the world has gotten beyond this “the cloud is dangerous” idea; the rest of the world is going to have to do so quickly, because when we centralize data into fewer (not one, or even seven, providers, are enough—but not hundreds of thousands) repositories, we are better able to spread the few humans we do have more effectively. It’s not a solution, it’s a stopgap while we try desperately to increase the size of the industry as a whole.

https://www.flickr.com/photos/68387974@N02/13573579563
And yes, we’re hiring too. If you do technical security work, come talk to me or drop me an email. We have people spread all over Canada (including several here in Toronto) and the US.
So we’ve talked about places, and things, and people, and why we care anyway. Data localization laws don’t help security, they don’t defend against the NSA, and they stretch the far-too-limited resources we have available to defend our data. We need a better solution.

https://www.flickr.com/photos/madely87/11903869904/