

CCC Spring Council Meeting 2018

Attending: Greg Morrisett, Nadya Bliss, Daniel Lopresti, Shwetak Patel, Nina Mishra, Kevin Fu, Keith Marzulo, Cynthia Dwork, Mark Hill, Beth Mynatt, Sampath Kannan, Liz Bradley, Ben Zorn, Jen Rexford, Maja, Andy Bernat, Peter Harsha, Ann Drobnis, Helen Wright

Visitors: Kim Hazelwood, Amin Vahdat, Chris Ramming

Amin Vahdat, Google “Networking Challenges for the Next Decade”

- Example of Google research model
 - Driven by novel application requirements to solve problems that others might not be aware of
 - Impact can be huge
 - Getting from point A to point B, with the current challenges with networking
- Software defined networking
- Protocols can become more robust if you solve from higher more robust systems
 - But that means you have to give up the style of the protocols
 - Make the whole system very complex
- Open Source
 - If we come up with the approach but don't share the information, we limit ourselves from the beginning
 - Perhaps would have been better off if Open sourcing from the beginning
 - Tensorflow - good product, can run Open Source, but we feel it runs best on our hardware → gives us a competitive advantage
 - gRPC - ours is the same as open source, we lose competitive advantage, but have broader community who can help make it better
 - Google software Innovations are driven by unprecedented demand for scale, bandwidth, reliability
 - Centralized metadata to scale?
 - Yes and no, it has never been fully decentralized
- Ben- Formal methods are at play to design the code at Google?
 - Not enough, we have a strong formal methods team but proving the spec is challenging
 - If you update on a daily basis you are also introducing new bugs
 - “Spanner” Google's approach to building a highly reliable infrastructure
- Distributed systems and network codesign go hand and hand
- Google publishes a subset of networking papers
 - Mark- Why publish? News for the competitors?
 - Advance the state of understanding
 - Reproducibility
 - Motivation/gratification on part of the engineers
 - We learn a huge amount about our systems in the process of publishing these papers

Task Force Discussions

- II Task force
 - Trying to organize a workshop around disaster recovery and public safety
 - Ben - Michael Dunaway is very interested in being involved
 - AI in the context of public safety and recovery seems like a popular way to focus
 - Pre-reads: Cricis report co-written by Robin Murphy and NIST public safety supercluster
 - Cricis report had a strong ask, but I don't know that it went anywhere
 - Trying to organize a steering committee
 - Dan - Cricis report was huge. Is this from one workshop?
 - Want to know what the follow-up was from the Cricis report before we move forward
 - Dan - something that would take place in the fall
 - Ben- How do we bring researchers in, esp. those who aren't necessarily working in this area?
 - Ben - need someone to bring in connections from the AI side
 - Thoughts?
 - Liz - Might be nice to map what you guys have with the gaps in the Cricis report/landscape
 - Liz - what about NGOs?
 - Ben - they're another player in this. So many different communities that have to come together to talk about this
 - Dan - Cricis might have not succeeded because it was so all-encompassing. This will be more focused
 - Liz- Social media amplifies false information. 70% of retweeted tweets are false
 - Ben - Dealing with video data at scale. There are a lot of drones, but they don't actually have time for someone to look at the videos
 - Beth - need to get Robin involved to do gap analysis from report recommendations
 - Beth - reminds me a bit of when we did the inclusive access workshop. We needed AI expertise, but those folks didn't see themselves as inclusivity researchers. May need to approach this similarly
 - Nina - I don't know of any public repositories for this data
 - New class of real-time algorithms that could be designed to help in these situations
 - If you're trapped and everyone has a phone, but no one has internet, what can you do in that situation? Logistic apps?
 - Nadya - does FEMA have data sets?

- Greg - someone who was doing modeling and simulations for flooding
 - Dan - that's Michael Dunaway
 - Nina - Perhaps people would be willing to donate their data so that researchers/responders could use it
 - Keith - SAVI: 5 years ago program funded with Japan (after Fukushima)
 - Beth: Recent EAGR solicitation around Crisis response
- Cybersecurity
 - Kevin - Embedded security workshop: we have submitted a proposal to the USENIX board to have a visioning workshop in August
 - Commitment from a number of interesting speakers
 - Participants will submit a position paper to qualify for an invitation
 - Sam Fuller has confirmed as the keynote. Former VP of research at Deck. Now his main interest is sensor security for IoT and autonomous vehicles etc.
 - Arranged for a large number of international researchers. Commitments from UK, China and Korea
 - Haven't sent out government invites yet, but have a short list of key program managers
 - Should get good attendance since it's in Baltimore
 - I've already started to get a trickle of emails from hearing about the workshop
 - I'm happy to circulate the proposal if you're interested
 - Ben - there was a report that came out at the end of the Obama admin about cybersecurity. What's the current state of cybersecurity on the federal level?
 - Beth - this is a key point for OSTP. One of the things in August is how to create channels to get more uptake of these ideas
 - Nadya- more emphasis on implementation than research, but I do think there is a robust research profile
 - Interesting thing that's happening at DHS: moving cyber from S&T to NPP (?). Some concerns in the research community that it will dilute the research component, but I'm not sure that will happen. Want cybersecurity to be centralized rather than scattered around S&T
 - Greg - one could ask if the CCC should be involved in this topic. It feels like shouting into a hurricane. All the agencies are working on it. We need to do something really different and unique, ex. Synergy with II
 - Kevin - formal methods is still maturing and discussing something related to 5G
 - Greg - I got someone from MIT to work on the formal methods blog post
- Privacy and Fairness
 - Held a workshop on fairness in Philly March 18-19th

- Cynthia - I loved the workshop. Workshop that was very heavily concentrated on the science and the math rather than the problems of fairness. 50 plus participants
 - Nice that there is an intellectual community that is starting to gel
- Sampath - Who knew fairness could be so complicated?
 - There are questions about the definitions of fairness. Ex. not just about a kid not getting into one college, but all of his college applications
 - You can be fair to separate groups but there is intersectionality and you might not be fair in that way. Ex. women vs. AA vs AA women
 - Fair representation is you take someones data and convert it into a form where you cannot determine demographics from the data
 - Fairness and transparency, fairness and merit
- Liz - organizers are preparing a draft of the paper and then the task force will generate a blog post and CACM article
 - Ann - make it more DC friendly for a general audience
- Second workshop being planned on intersectionality and fairness. Third workshop being discussed on intersection between privacy, fairness, and economics
- Liz - privacy and fairness seems to be bubbling up. Lots of incoming faculty at Boulder in this area
- Cynthia - never seen the level of general societal concern for fairness compared to other things
 - Board meeting of the Siemens theoretical computer science institute and a large chunk of the board meeting was devoted to fairness
- Beth - I was in a separate meeting with the Hewlett foundation that reflected similar trends
- Beth - Adversarial system approach: funding for breaking systems, not just building them
 - A few folks who discussed looking at this from a sociotechnical standpoint. This may intersect legal frameworks. Systems that are legally fair must intersect social system
- Sampath - Solon Barocas was pointing out ?
- What are the legal mechanisms around fairness
- Ben - a lot of data for certain classes of individuals and not others
- Sampath - we have this view in formulations of fairness of agents to enforce fairness. Ex. landlord and Airbnb. Landlords aren't necessarily fair but Airbnb wants to enforce fairness in it's landlords
- HTF
 - Planning workshop on live long learning. As technology gets integrate into workforce, look at how to integrate technology into training with AR/VR etc.
 - 40 advanced technology centers that are near major industries. Looking at ways to observe how manufacturing is evolving and how technology can play a role
 - Shwetak - I visited one for nano, that showed how to rivet a bolt with VR

- Still in the midst of getting the advisory committee set up
- Other areas of focus:
 - Physical and mental health
 - Augmenting human performance
 - Voice input - Siri, Echo etc.
- Beth - the accelerators are areas that will be getting increased awareness from the NSF
 - Put it in a blog
- Peter - Office of integrated activities - human technology frontier and big data.
 - Beth - \$30 M set aside
- Shwetak - 90% of it went toward community colleges
- Beth - call that's out now is to prepopulate for FY19. That is my guess
- Ann - I know there are communications people within NSF who hadn't even heard the term accelerator when it was announced
- AI
 - Ann - hasn't done much
 - Been hearing for a year that partnership on AI is getting up a going. Hired an executive director in the fall, we're meeting with her tomorrow
 - Maja - I'm now seeing opposition about AI within our ranks, it's unsettling to tear ourselves apart
 - Is AI good for or going to ruin civilization
 - People wanting to dismiss it is irresponsible
 - Cynthia - there are things that are not good for the field. Are you trying to put together a more thoughtful discussion?
 - Maja - leading minds are speaking independently, but every once in a while they are pitted against each other and then the NY Times runs "crazy scientists fight"
 - Beth - all of this stuff happening within the press, can we write a level setting paper that corrects the record
 - Ben - AI is being integrated into software and then the software starts doing things I don't necessarily understand
 - Cynthia - does anyone now what AI Now is or does? I'm on the board but can't get them to tell me
- PMLC
 - 3 workshops in development
 - Digital computing and quantum computing are both in May
 - Tom Conte led workshop in the fall around thermodynamic computing
 - Beth - I was on the CSTB report on quantum review. At least 6 months until it comes out
 - Once we reach the trough of disillusionment, what will the role of government be?
 - Mark - in the best case quantum computing is doing regular computing and finding

- Greg - factoring requires millions of qubits. Only a few algorithms that are currently viable
- Ben - none of the things you can do are commercial viable?
 - Greg - quantum chemistry basically, but the costs of the machines every high
- Sampath - when I was at NSF I wrote a position paper that got \$30 million for quantum computing. One of the arguments is China and Canada are spending so much money
- Kevin - I visited one of these quantum physical labs. Physicist love to talk about qubits and laser, but they are so ideal it's hard to run algorithms. More like physical computing

Communications and Outreach plan

- Plan to be vetted and signed by the end of June
- Got informal feedback from NSF last week. Overall positive set of reactions
- This is associated with the increased budget for CCC. We don't have to do less of something to fund this
- Nadya - increasing structured communications. Not going to be reactive - get calls from reports when there is a breach etc.
 - Really target sustained relations with key reporters that we want to communicate with annually
 - The conversation is a venue we're discussing. Want to hit Wired, Science New Yorker
- Greg - we're putting reports on ArXiv now. II papers are all up there
- Greg - amplifying signals of CCC council members which helps with perceived CCC expertise
 - Nadya - if you do something cool that's relevant, highlight that
 - Organizing people by areas, not just task forces
- Beth - what's the actionable piece for agency. I feel like for II we should have produced a powerpoint deck a while ago
- Ben - what are the right resources
- Nadya - getting a large audience at a generic CCC presentation at a university won't get a big audience but if one of us show up for a discussion that will get more interest
- Greg - lot's of younger faculty don't know CCC, CRA, or that they are related
- Greg - I Schools are a resource we haven't really tapped into
 - Snowbird will give us a chance to connect with department heads
- Ben - what's the next step if they hear about the CCC at Snowbird or whatever? All of this is invite only right? What would they do next?
- Beth - when I've had discussion it's what would a successful visioning proposal look like and what do CCC Council members do
- Sampath - to play devils advocate I don't think CCC will ever be a household name among CS researchers. It's a one-of for a workshop, not a day to day concern
- Ben - you say fairness is catching on. Involve junior researcher more

- Nadya - a lot of the things that are in here are currently being done, but it's ad hoc. This will make it more organized
- Greg - engaging with other agencies besides NSF. Might make sense to go around the table- who do you know at DARPA, at DOE, etc?
- Nadya - much easier to maintain connections once they are established (and will connect us to Leo)
- Ann - all of the embassies have science liaison. Attempted to do a workshop with the British one once, it didn't work out due to politics
- Beth - like that we made a weak attempt at an evaluation plan. Want more focus on outcomes
 - Other suggestions include email list or more social media. Suggestions are not necessarily evidence based
- Nadya - was there any conversation with the NSF about establishing a baseline
 - Beth - yes, we did

Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective

Kim Hazelwood, Facebook

- Moore's Law -> LeCun's Law
 - Number of citations from LeCun's 1998 paper
- Machine Learning Executive Flow
 - Data, Features, training, Eval, Inference
- Infrastructure Research Challenges
 - Data
 - Offline Training
 - Online Inferences
- Does Facebook Use Machine Learning?
 - Yes--- for example NewsFeed, shows the most relevant stories to the user
 - Also, Ads, Search, Facer, Content Understanding (Lumos)
- What machine learning model does Facebook use?
 - Mostly ranking with MLP, which is newsfeed, ads, and search
- How often do we train models?
 - Relearning models every month (haircut) verses clicking on an add that retrains itself based on your searches
 - No single answer, retrain at all different times and it takes all different levels of time
- How much compute does inference consume?
 - Different levels of problems (100x, 10x, and 1x problems)
- Does facebook design hardware?
 - Yes, since 2010 facebook released through open compute
- Does Facebook design hardware for AI/machine learning?
 - Buying Nvidia cards but designing their own chases
- Facebook is in the Ads business, not tied to the hardware
 - Facebook is not secret about its hardware

- Facebook AI Ecosystem
 - Deep learning frameworks- Caffe2 and pytorch
 - ONNX (Open Neural Network Exchange) - new consortium on the model and agreement of what it should look like
- FB Learner Platform
 - AI workflow
- How do we scale this?
 - 2 billion users
 - US and Northern Europe data centers
 - There are a lot of challenges/opportunities at the same time
- What is the ethics of this?
 - My advertiser is not my friend
- There is good that comes out of this too
 - Safety check
 - Connecting with people who normally wouldn't be
 - You need to have a business model otherwise it wouldn't survive, we need ads since people aren't going to pay for this
- The ethics and privacy team do push back a lot
- Scaling Challenges/Opportunities: Data
 - Data quality and potentially quantity
- Scaling Opportunity- free compute
 - Leveraging CPUs
- Disaster Recovery
 - Losing the entire datacenter
 - Compute diversity become critical
- Key Takeaways
 - Facebook AI
 - Lots of Data
 - Wide variety of models
 - Full stack challenges
 - Global scale
- What is facebook looking forward in terms of the computing research community?
 - Having the right focus on the right problems
 - Am I solving the right problem?
 - Seeing clustering the doesn't have the right logic to it
 - What fraction of the overall resources is machine learning?
 - Relationship between industry and academia research- based on HPCA paper

Chris Ramming

- Software defined infrastructure
- NSF/VMware Partnership on Software Defined Infrastructure as a Foundation for a Clean- Slate Security
- Processing storage and communication

- 1981- growth of non-virtualized, horizontal compute industry
- 1995- provocation of programmable networks
 - Virtual infrastructure
 - Idea that infrastructure should be programmable
- 1997- DARPA launches active networks program
 - Active networks
 - Moves one packet from one input to one output
 - Focus was on obscure programming language instead of performance
- 1999- revival of compute virtualization
- 2002- Server consolidation emerges as virtualization driver
 - Intel amplifies planetlab signal
 - An overlay network testbed
 - Network slicing- ability to take a whole network instead of just look at some components of it
- 2012- VMware acquires Nicira for \$1.26 Billion
 - From processor virtualization to software defined hybrid cloud
- Private cloud vs. public cloud tensions
 - Control of data
 - Ability to span clouds has become a monument to access cloud
- 2014- The rise of the edge
 - Internet of things begins to motivate a three-tiered model
 - IoT gateways
- In process Mobile Edge
- In process: Software-defined Telecom
 - Architectural approach in which compute, network, and storage components of the physical infrastructure are all virtualized
 - Where are these properties is going to be useful?
- SDI and security
 - SDI provides a powerful set of tools for revising infrastructure security challenges
 - What is the software defined architecture that is easy to security?
- Why does industry care about the joint public/private partnership with Govt?
 - VMware could not be where it is today without this relationship
 - Industry players notice when there is a strong relationship, one way to provide input to the community
- Greg- How do we not increase more software out there when we put out a response to malware attacks?
 - Chris- What if we verify software instead of building hardware.
- Ben- If verify is a solution because it is cheaper than software, then why hasn't a ton of money been invested in this?
 - Chris- We have always been comparing verified software to other software? That might be the difference.

- Nadya- what is going to drive the market towards more secure systems? Our result is nothing- it costs us money when breaches are known, but not if they aren't known. There is no reason to building a security if you are in the business to repair attacks.
- Greg- Is there a fourth leg to virtualizing the user?
- Mark- what is the relationship between DARPA and NSF?
 - Interesting interplay with other organizations
 - Need both for the research community
 - Super valuable to do joint solicitations with NSF
- Ben- How do you summarize the benefit?
 - Chris- These ideas germinate 1-2 years, the acquisition happens and is expensive
- Sampath- DARPA and SRC
 - Chris-no need to separate
- Beth- what advice would you give companies who are trying to establish relationships with NSF?
 - Chris- Process worked well
 - Keith- Agreed, NSF has a set of values that they need to maintain
- Chris- Lot of value of having joint participation if you can afford it
 - Reaching out in advanced
- Greg- non US based companies partnering with US government agencies?
 - Depends on the agency, money needs to flow correctly

Peter Harsha

- New appropriations numbers that we've learned. I have some good tidings
- Will cover President's requested FY 19 and final FY 2018 numbers
- Background and fiscal environment
 - Limits on nondefense spending chart
 - Force congress to stick within spending limits
 - No room to grow any of the discretionary accounts
 - Only really alienated the far right and far left
 - After a few government shutdowns, did manage to pass a budget
 - Increased nondiscretionary spending by 13%
 - President agreed to the budget, but OMB went back and revised the caps
- FY19 was drawn up without having the FY18 numbers to even consider
- Congress not committed to President's numbers. Mulvaney doesn't seem committed to it either
- Still details in the budget that will be instructive. Don't worry about the numbers but the programmatic changes
- NSF would be flat overall, 2% increase to research and related activities. SBM hardest hit, lose about 11%
- President's budget request goes program by program but OMB can go as granular as the like

- NIH - increase of \$1.4 billion, but because of consolidation, which include NIDLRR. CRA and a group of universities are opposed to this consolidation
 - In the FY18 there is no mention of consolidation. Haven't gotten the sense that Congress is on board with this
- DOE Office of Science - flat but huge increase in ASCR and Exascale. 39% and 28% respectively. ARPA-E would see reduction
- NIST - 34% cut overall, 17% cut to NIST labs
- DHS science and technology - 25% cut, transfer cybersecurity research out of S&T to NPPD
- Keith - what was the rationale for such a large cut for NIST?
 - Peter - mostly a cut to the manufacturing program. Seen as corporate welfare
- Defense S&T - basic research up 5%, applied down 4.4%, adv. Tech development down .5%, DARPA would increase 19%
- Beth - how likely is the manufacturing cut
 - Peter - unlikely that there will be much of cut in reality
- FY18 final omnibus
 - Overall science does well in this budget. Could have been much worse
 - NSF - overall NSF grows 3.9% vs. FY17 to \$7.71 billion. Increase of \$295 billion
 - R&RA highest budget since 09
 - DOE office of science grows 16.1% to \$6.26 B
 - ASCR grows 25/2% to \$810 million
 - ARPA-E grows to \$353 million (all-time high)
 - Defense S&T - 6.1 basic research increase 2.9%, 6.2 Advanced research up 7%
 - Other agencies - NASA up 6%, NIST labs up 5%, DHS S&T up 7.6%, NOAA up 6.7%
 - CS Education Funding - 2 Dept of ED grant programs have new guidelines specifically including CS Education efforts
 - Student support and academic enrichment - \$108 million boost to \$1.1B. Mostly for securing schools
 - Directs the program to support pre-K through grade 12 CS ed programs that address the enrollment and achievement gaps for underrepresented students
 - Education Innovation and Research carves out \$50 million for "innovative STEM education projects including CS ed"
 - Cameron Wilson at Code.org is constantly trying to figure out who has the knowledge at ED to lead these initiatives
 - Infrastructure spending - boosts infrastructure spending to \$21B
 - Includes \$265 million to increase and expedite rural broadband expansion within USDA
 - \$398 million to support "cutting-edge sciences at National Labs and other DOE sites"
 - Gains for science agencies across the board in FY18. Still room for continued growth in FY19, though much more modest

- Some programmatic changes require a close watch through appropriations process
- Election year, so appropriations will stop in the summer. The soonest we'll see the resolution will be end of November
- II research briefing January 30th
 - Benefits of II and underlying critical research gaps
 - Almost everyone stayed the whole time, rare for these events
- Tax reform issue - Would increase cost to grad students. CRA and some other organizations came together to oppose the graduate students tuitions waiver. We were successful so costs will not go up

- Peter - issue with briefing is there is a capital cost to putting them on and getting people to show up. I tend to be conservative with holding briefings
- Privacy and fairness for research needs is a great area for CRA, privacy and fairness for algorithms is not
- Ann - should we do something around quantum for the fall after the quantum workshop?
 - Peter -- maybe
- Keith - Caucus structure in congress, is this something we are involved in?
 - Peter- yes involved with Robotics caucus. AI caucus is ramping up and we have been in touch with them. Widely varying in usefulness
- Dan - things on hold until the midterms elections
 - Peter - there will be progress until the summer but then it will shutdown until after the elections
 - Dan - is there something we should do to plan for that?
 - Peter- we have our Congressional visits day in the fall to give our final pitch for investments. Trick is getting bandwidth from the offices
- Peter - my way of assessing damage to the community people yelling at me

Jen Rexford Report

- Campus CIOs and Campus Administrators talk about access to data, cloud providers to care about market
- 5 Recommendations (from the Draft Report):
 - Articulate case for academic institutions to use the cloud
 - Articulate the “business case” for cloud providers to support academic users
 - Remove artificial costs that make cloud computing less attractive
 - Create support structures for academics transitioning to the cloud
 - Form a central entity to serve as a nexus between multiple cloud providers on one side and multiple academic institutions on the other
- Removing artificial costs, looking at ways to see how people can select the cloud offering and look at other ways to use the cloud
- How to lower the barriers further
- Voice of NSF/forming a relationship with agencies
- Sampath- For research and education

- Jen- cloud providers can help but will be more challenging
- Beth- Do you see a potential for universities to spring into this group of public/private partnerships?
 - Jen- A lot of faculty who run research groups, students may use the cloud for a lot of different uses
- Greg- What about portability?
 - Jen- Set of baseline offering across cloud providers, identifying a set of baseline capabilities, less bewildering
- Greg- Integration for third party systems, easier to not use the cloud
 - Jen- having a central entity is good for the community

Communications and Engagement Plan Feedback

- Liz- do we really want to be held to all of this? Can we set some low bars? What do we drop?
- Shwetak- can we prioritize some of this?
- Sampath- what is the end result that we want?
 - NSF wants us to be well known
- Beth- NSF wants us to be an NSF/CCC not a CRA/CCC
 - They want the community to know that they are doing something for the community
- Mark- many other fields are trying to start like CCC
- Beth- busy faculty members, companies, etc what other pieces of the pie do we need?
- Keith- What could CCC be known for? CRA is known for taulbee.
 - Beth- high impact pieces are needed
 - Like the SRC annual report
- Sampath -What are the resources that junior faculty need? We could use the funding as a teaser to bring people together
- Ben- Could we do some annual thing to bring junior people in to agencies (not policy folks)?
 - Peter- it should be OK
 - Ann- We are planning to do something like this this summer at the CI Fellows 2.0 meeting in August
 - Keith- NSF might not like that, since it is not reaching that they want specifically, it will need to be painted
 - Greg- we could be much more strategic about various areas, go to the SIGS with results of successful visioning activities
 - Maja - get visioning activities just for junior people, and match them up with funders
- Keith - Foundations - it is worth it to get to know them, as they work very differently from the agencies, and we should tap into their resources, as they have a different angle with the communities
- Ben- Good to have an inventory of the agencies and their topic interests and what they are caring about

AAAS

- Due Date April 19th for 2019 AAAS
- Possible topics
 - Greg might bring in some agriculture folks and CS folks- Greg
 - Fairness/privacy- Sampath/Ben

Workshop for Early Career Researchers

- Invite only
 - CI Fellows
 - Post docs
 - Early career visioning workshop folks
- Possible sessions
 - Start-ups
 - 50/50 gender split
 - Task force breakup?