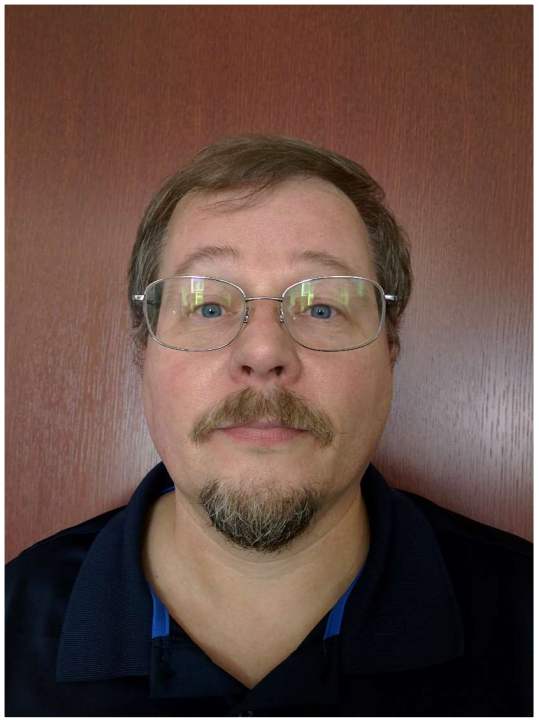

CI/CS WORKSHOP
THE COMMUNITY TOGETHER
ResearchSOC
CI CoE PILOT

NCSA Cyberinfrastructure Professional Internship Program: **Goals, Methods & Results**

Daniel LaPine, CII Division Lead, NCSA
Dr. Volodymyr (Vlad) Kindratenko, Sr. Research Scientist, NCSA
Dr Luisa Rosu, Research Associate, I-Stem (UofI)

CI/CS WORKSHOP ResearchSOC | CI CoE PILOT

Primary Investigator (PI)



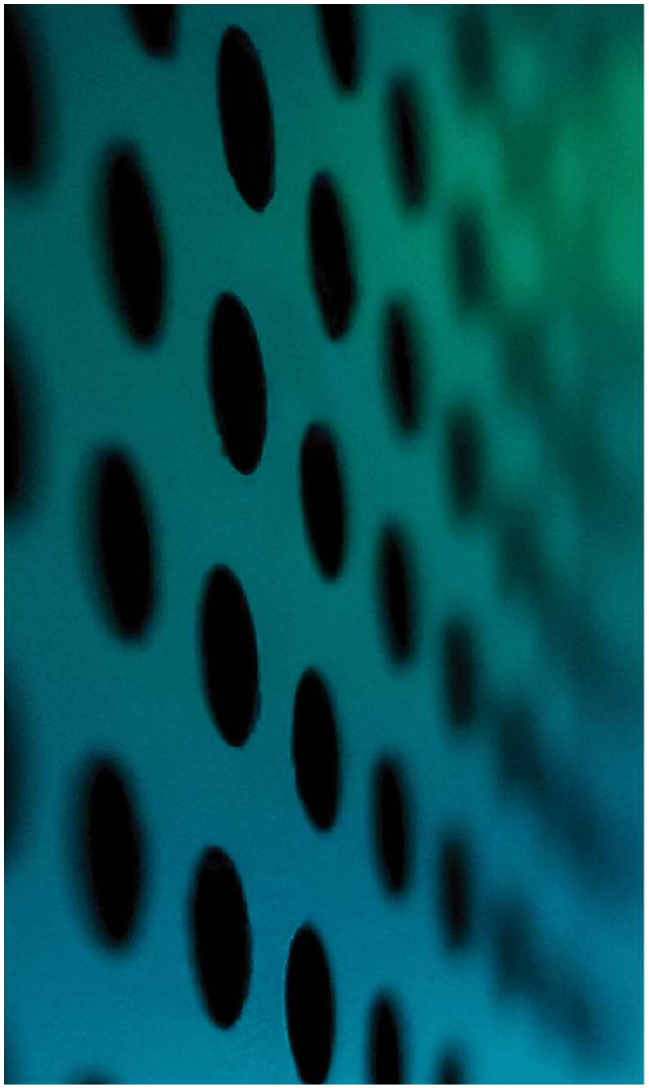
- Mr. Daniel Lapine
- NCSA employee since 1998
  - undergrad student - 2001
  - Systems Engineer - 2009
  - XSEDE L3 for Technology Investigation Service - 2015
- Technical Program Manager
  - Scientific Computing Services division - 2020
  - Continuous Improvement & Innovation division - Now
- BS in CS from University of Illinois College of Engineering
- 13 years military service USMCR/USAF
  - Machine gunner, Airborne CryptoLinguist, Instructor for flight and simulator

Co-PI

- Dr. Volodymyr (Vlad) Kindratenko
- Senior Research Scientist (NCSA)
  - Computer and Data Sciences
- Leads the Innovative Systems Lab 2.0
  - applied R&D at NCSA
- Adjunct Associate Professor ECE
  - Teaches freshman and sophomore computer engineering classes
- Many, many publications
  - http://www.ncsa.illinois.edu/assets/php/directory/contact.php?contact=kindr

External Evaluator

- Dr. Luisa Rosa
- Research Associate (I-STEM)
  - CORE Early Career Fulbright U.S. scholar 2014
- Has conducted and participated in evaluations
  - iFoundry, a College of Engineering (CE) curriculum incubator
  - Strategic Instructional Initiative Program (SIIP) in CE
  - NCSA's CIP Intern program
- PH.D in Continuing Teacher Education (Uofl)



Goals

- National Science Foundation (NSF)
- National Center for Supercomputing Applications (NCSA)

NSF Award Goals

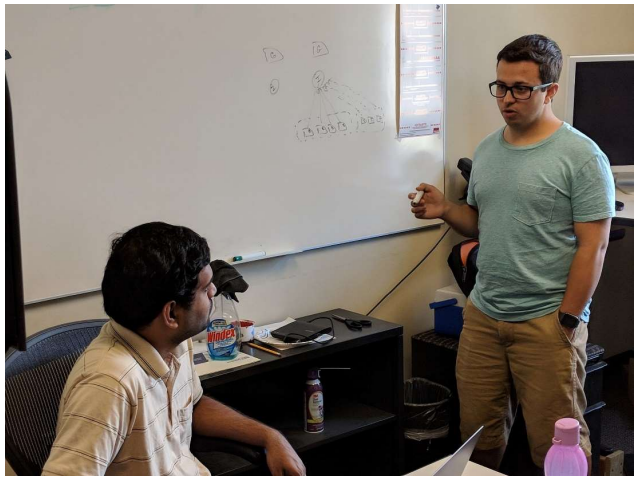
- Solicitation by NSF's Office of Advanced Cyberinfrastructure (OAC)
- The award goal was to address the shortage of a workforce with the specialized skills needed to operate and support advanced cyberinfrastructure
- NSF recognized a lack of trained and experienced staff were available to operate HPC and advanced CI.
- Wanted programs that could expand that pool and be widely applied

Workforce Issues

- Insufficient experienced operators available to support HPC and Big data
- Lack of diversity in the workforce that did exist
- Lack of training programs (in 2017) to bridge the gap between classwork and advanced CI operations
- HPC & ACI operations have a limited career draw for general CS graduates

Our NCSA Proposal

- Provide training and experience using On-the-Job training methods
- Treat the participants as full time, paid interns for a semester
- Accept a wider range of candidates for training than those who follow the “standard path” to systems engineering



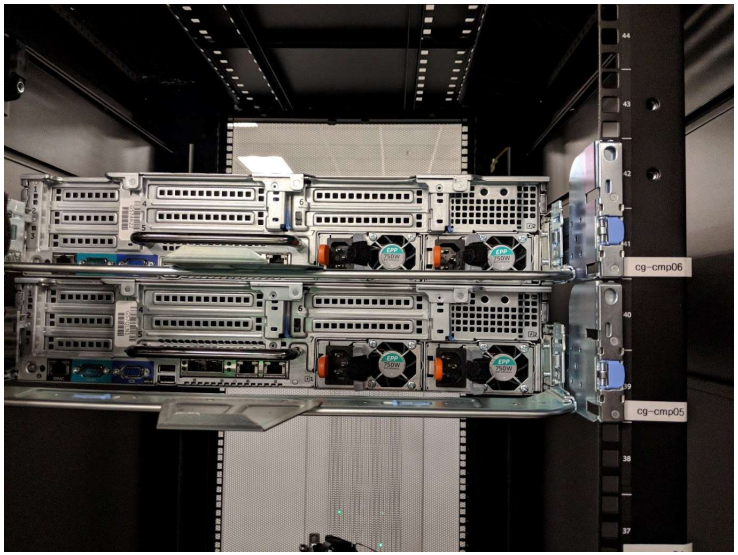
MENTORING

Official Program Goals

- Provide interns with the opportunity to learn about CI operations at a major center with an emphasis towards **On the Job training** (OJT)
- Encourage access to CI operations for a larger array of potentially overlooked candidates, with a focus on inclusion
- Generate and refine the program and training materials for dissemination

Unofficial Program Goals

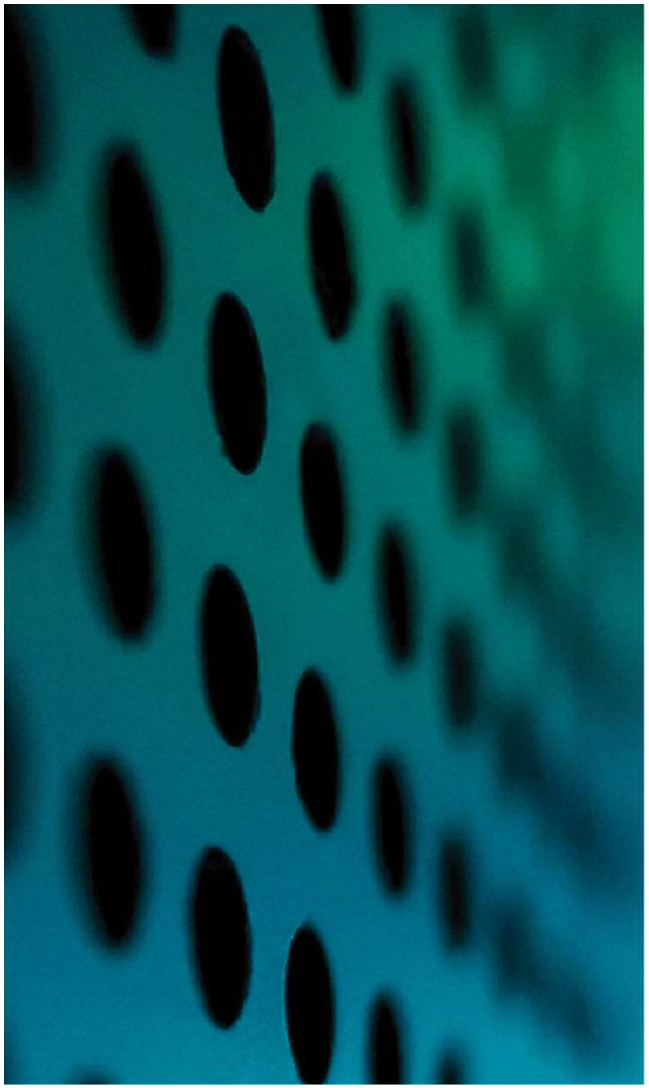
- Get them hands-on experiences with HPC hardware as much as possible
- Teach the interns why we do things and when services are required or not
    - The difference between “I can do this” and “Should I do this?”



HARDWARE PROVISIONING

Inform Career Decisions

- Show them how a “real world ACI facility” operates with deadlines & project managers and everything.
- Give the interns an opportunity to learn what being “an operator” means
- Gives us the first chance at hiring some promising new employees

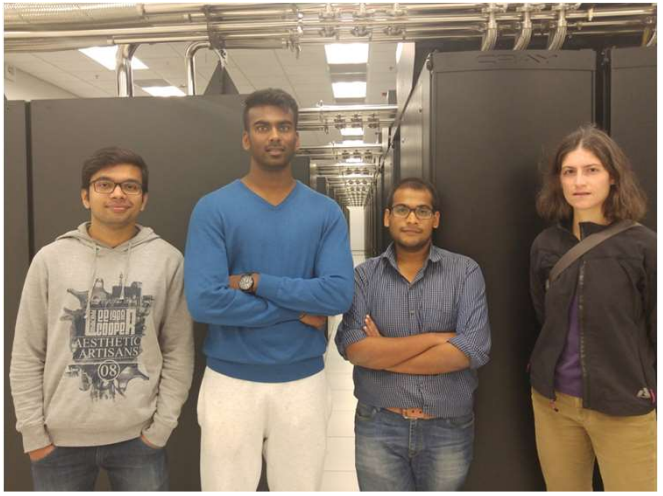


Methods

- Overview
- Intern selection, HR and Support
- Scheduling
- Training & Special projects

Program Overview

- In 2017 NCSA established a pilot internship program for cyberinfrastructure professionals (CIP)
- The CIP program was funded for 3 years by the NSF's Office of Advanced Cyberinfrastructure
- Received a no cost extension to 2021



FALL 2017 INTERNS

NCSA

- Well established national supercomputing center (30 years of operation)
- Large array of HPC operations and user support
  - 10 operational CI focus groups
  - 6 major HPC clusters in operation
  - Participation in XSEDE, OSG, Midwest Big Data Hub and collaborations with international partners

NCSA Capabilities

- NCSA knows and conducts advanced cyberinfrastructure operations
- Have multiple and varied ACI groups
- NCSA operations staff is experienced and capable in multiple areas
- Ability to support multiple interns, up to five per session, one per group

NCSA Limitations

- NCSA is an institute of the University of Illinois (U of I)
- Large number of complex systems in operation at all times
- Not overstaffed with engineers

FALL 2018 INTERNS

Intern Selection

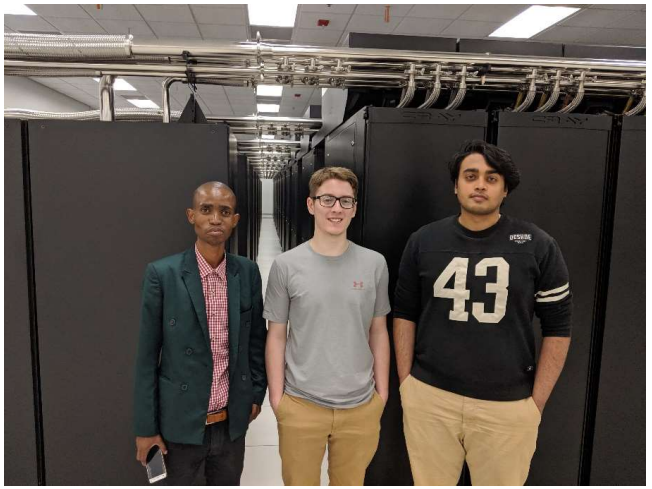
- The program is continually advertised to a broad community, from the U of I students to a national audience through emails, flyers and direct recruiting
- Applications are via the program website, and are open for at least 1 month, about 2 months before the start of the session
- Applicants provide information and their CV's via the website and that is reported to the program officers for review at the end of the period
- Interns must have at least 60 credit hours of college work and must intend to work in CI operations
- Applicants are screened and interviewed to pick the most appropriate candidates

Intern HR

- Interns are provided a monthly stipend, sufficient to cover living expenses in the local area for the 15 weeks.
- Interns are considered fellows of a college on campus, not employees
- Interns will receive U of I and NCSA identities as do staff
- Interns are expected to work 40 hours a week and attend all group meetings

Intern Support

- Interns report to an assigned mentor or the group lead on a daily basis
- The program PI acts as an ombudsman for general issues with NCSA HR handling any HR issues
- Interns may attend any events for staff that NCSA or the campus hosts, such as colloquia, lunch and learn meetings, professional training, etc



SUMMER 2019 INTERNS

General Program Schedule

- Host the program twice a year, usually fall and spring
- Sessions are concurrent with university's fall and spring schedules
- Duration is 15-16 weeks
- Intent for 5 students per session (50% of available operational groups)

Schedule Breakdown

- One week of initial training (35 hours over 5 days)
- One week of single day group embedding
- Selection and assignment to a group for the extent of the internship
- On the Job training with an operational group and assigned mentor
- Last five weeks have a group project assigned to all the interns
    - Group projects are operational efforts using CI skills learned and practiced
    - 25% time effort towards group project goals for the interns

Example Schedule

Fall 2019 runs Tuesday Aug 27th through Dec 13th

Training Schedule (week 1- week2) Sessions in Green are all new NCSA staff invited

Day & Rooms	Tuesday (Aug 27th) 2000	Wednesday (Aug 28th) 2100	Thursday (29th) 2000	Friday (30th) 3000	Tuesday (Sep 3rd) 3000
0900	~~Pre-eval (Dr. Rocu)~~ Program Overview Daniel Lapine	Storage HW survey Daniel Lapine	Andrew Loftus LSST Travel Report - Aug 2019 Base Data Center Build-Out in Chile	Comp Ops (room 2000) 0910	NPCF Tour Chit Khin Meet at NCSA
1000	NCSA Building Tour Daniel Lapine	Project Management Overview Katherine Kendig	NERD David Wheeler	Former CIP student final report Daniel Lapine	NPCF cont'd
1045	Break	Break	Break	Break	Break
1100	Desktop support intro	Intro to Linux clusters Daniel Lapine	DAV Roberto Sisneros	Elastic Computing Stephen Squaire	(TBD) Chit Khin
1200	Lunch	Lunch	Lunch	Lunch	Lunch
1300	NCSA Communications Daniel Lapine	IRST James Eyrich	AVL Stuart Levy	HAL Intro with Slurm Overview Dawei Mu	ISL Volodymyr Kindratenko
1400	ENV Modules Daniel Lapine	IRST James Eyrich	AVL Stuart Levy	Monitoring Yan Zhan	Systems Overiew Timothy Bouvet
1445	break	break	break	break	break
1500	ITS content	linux and command line Daniel Lapine	ISDA Jong Lee	IDDS Group Overview Kimberly Blum	A day in the life of a sysadminJacob Rundall
1600	vim Daniel Lapine	linux and command line cont'd	Storage James Glasgow	Admin- Group selection process	R and analytics Yifang Zhang

One-Day embeds (tentative) (confirmed in blue)

	4th (Weds)	5th (Thurs)	6th(Fri)	9th (Mon)	10th (Tues)
Cyan	IRST	ISL	Systems	Nerd	SET
Danjin	ISDA	DAV	Genomics	Systems	Dora/
Gary	ISL	System	DAV	SET	IRST

Initial Training week

- Training presentations cover three general areas
  - Introduction and group overviews
  - Technical topics
  - Professional development
- This material is available for others

FALL 2019 INTERNS

Group Embedding and Selection

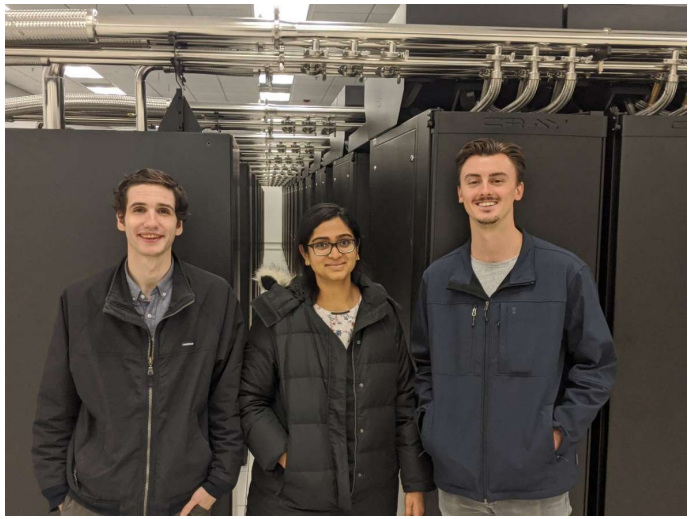
- Before the second week, students note groups of interest
- The PI coordinates group availability and creates a scheduled visit
- Each intern then spends one day for a week with different groups
- By the end of the second week, interns generate a preference list
    - For all the sessions (6) and the participants (17), only once was student not able to get their first choice
- The interns then spend the rest of their time in the session working with the assigned group and their mentors

On the Job training

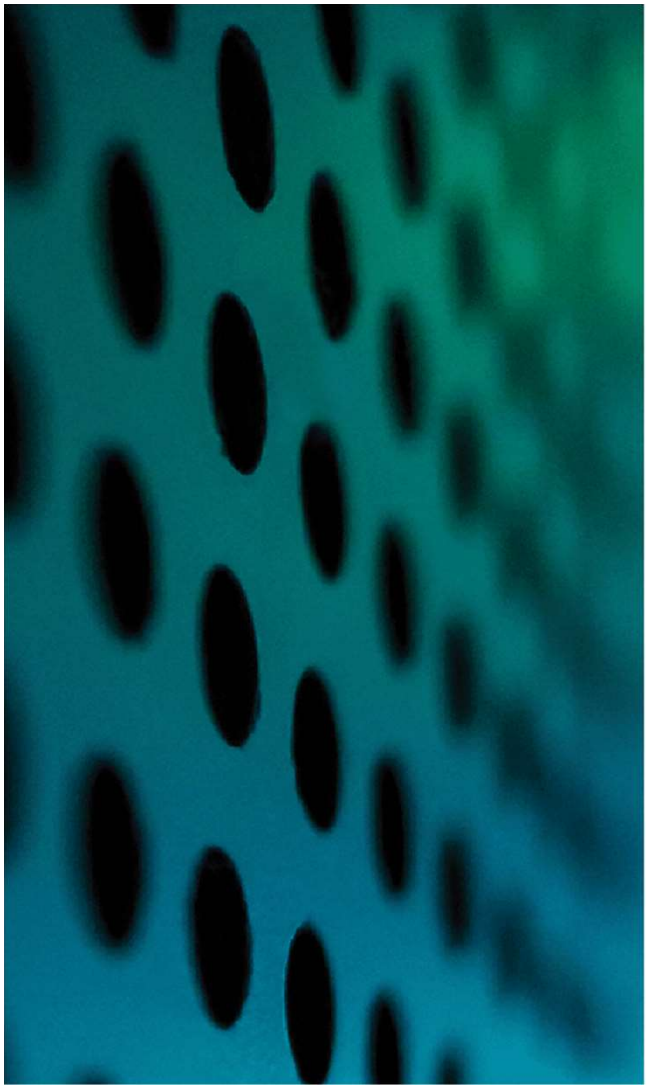
- Interns work within a focused operational group as they would if hired, with some limitations
- Focus on learning what operational HPC issues exist and why we operate in certain ways
- Attend operational meetings with the group members
- Take on assigned tasks and have expected timeframes and deliverables
- Have access to group and divisional documentation

On the Job training, cont'd

- Emphasis on working with cluster hardware installation and decommissioning for hands on HPC opportunities
- Have access to NCSA facilities with an accompanying engineer
- Don't have root or superuser access
- Aren't answering/responding directly to user tickets



SPRING 2020 INTERNS



Results

- Participants
- Evaluation
- Results
- Lessons
- Covid-19
- Future

Participants

- Six sessions over the first three years
- Application rates have varied over time
- Twenty accepted but three dropped out before starting
- Seventeen interns completed
    - Education levels ranging from just 60 Credit hours of study to those with MS degrees
    - At least one women or minority per session

Evaluation

- External Evaluation by Dr. Rosu annually before NSF reporting
- Pre and Post program student capability and knowledge assessment
- Survey of climate and career assistance during the evaluations
- Evaluation feedback is used to make program updates as possible
- In general:
  “Interns noted that they had opportunities to evaluate what working in their selected field might be like, and how it differed from their expectations, and how that would influence their future decisions”

Results

- Ten participants working in the CI field post their time in the program
    - Three of which were hired by NCSA
- Five continuing academic studies
- Two not working in the field
- Unexpected benefits
    - Training week is useful for other employees and new hires
    - Creating and giving overview presentations is good practice for CI professionals

Lessons

- ‘Intern’ is not a university pay status
  - Addressed over time with fellowship from a college on campus
  - Convincing university institutions that a paid internship on campus is useful
- Intern collaboration and socialization
  - Interns separated into to distinct groups without much interaction
  - Established separate communication channels for just interns

Lessons Cont'd

- Advertising and applicant pool
  - Over time applicant pool has declined from ~25 per session to ~3
    - Requirement to work 40 hours a week and not attend school is onerous for current students
    - Tried a summer session but insufficient time to impart as much knowledge
    - Better general economic conditions favor direct to employment career path
    - We are trying to offer the internship to underqualified job applicants

COVID-19 Effects

- Spring 2020 session was in progress during the event
  - State and campus orders for shelter in place implied all staff to shift to remote work as of March 16th 2020
  - Critical and essential operations needs applied to some staff for on-site visits but not the interns
  - No decline in intern outcomes reported or noted due to the shift, save for lack of hands on hardware post March
- Fall 2020 session planned to be remote, with possible shift to some on site operations as conditions allow

Program Future

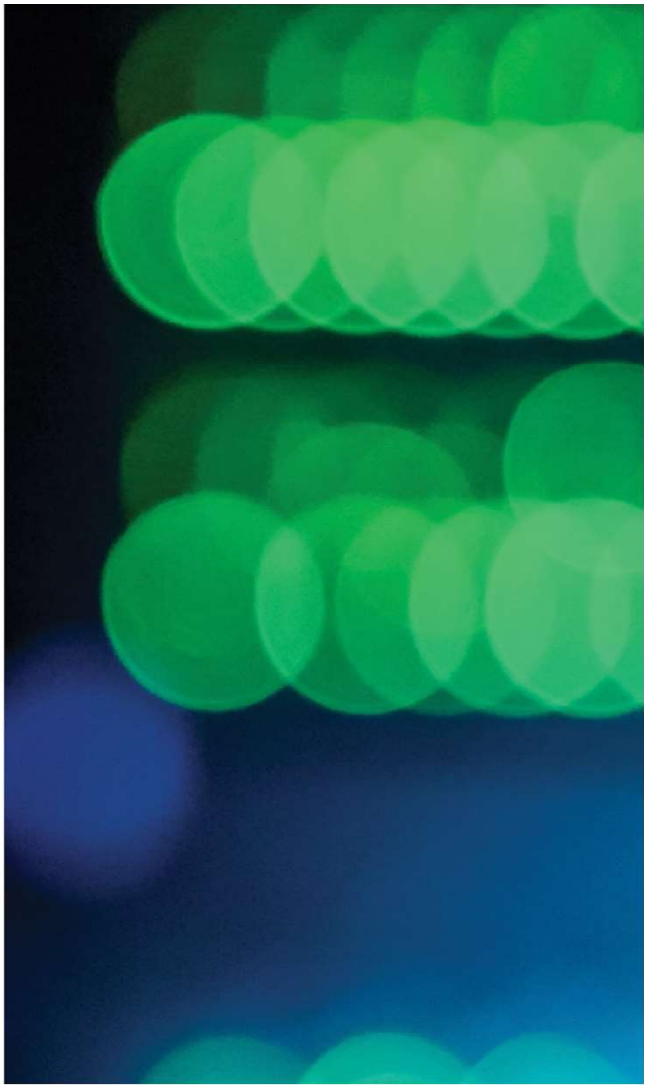
- Program extension into a fourth year
- Need to get wider exposure of the program materials and methodology to other institutions and national HPC centers
- Possible program expansion to support professionals already employed
    - With support from industry, government or other institutions, could offer some degree of OJT for new Industry staff as a training service offering.
    - Could offer value in exposure to Advanced CI operations and techniques

Acknowledgments

- This work is supported by the National Science Foundation's Office of Advanced Cyberinfrastruction (OAC), grant No. 1730519, as well as the University of Illinois at Urbana-Champaign.
- We would like to thank NCSA HR and the ICI directorate for their support, Olena Kindratenko with her assistance for student outreach to attract new participants and Dr. Luisa Rosu and her team for thorough evaluation and reporting of participant experiences

Contact Info

- Program Website
  http://www.ncsa.illinois.edu/enabling/cybersecurity/cip
- Dan LaPine's email: lapine@illinois.edu
- Vlad Kindratenko's email: kindrtnk@illinois.edu




Questions?
CI/CS WORKSHOP
ResearchSOC
CI CoE PILOT


Thank you!
CI/CS WORKSHOP
ResearchSOC
CI CoE PILOT