LSN Huge Data Workshop Agenda

Monday Afternoon   April 13, 2020

*All Times Eastern Time

Time	Session	Affiliation	Title
12:00pm EDT	**Workshop Begins**		
12:00-12:10 EDT	**Introduction**	(KC Wang, Jim Griffioen, Ron Hutchins, Zongming Fei)	
12:10-12:20 EDT	**NSF Welcome**		
	Deepankar Medhi	Program Director	
	Erwin Gianchandani	Deputy Assistant Director for CISE	
12:20-12:50 EDT	**Keynote**		
	Craig Partridge	Colorado State	Are Our Networks Trashing Our Files?
12:50-1:30 EDT	**Data Generation**	(Session chair: Jim Griffioen --- Zoom moderator: KC Wang --- Note taker: Zongming Fei)	
	Jim Kenyon	Michigan	Modern Research Imaging and X-omics
	Cody Bumgardner	Kentucky	Huge Data in Medicine Workshop 2020
	Soumya Rao	Missouri	Computational Challenges in Genomic Data Analyses
	Zachary Jacokes	Virginia	Computational Needs for Multimodal Explorations in Differential Autism Spectrum Disorder Phenotypes
	Yunsheng Wang	Kettering	Huge Data for Connected Vehicles
	Hongxin Hu	Clemson	Trustworthy AI for Huge Data Generation and Process From IoT Devices
1:30-1:50 EDT	**Break**	Please join a hallway conversation	
1:50-2:40 EDT	**Data Storage**	(Session chair:Zongming Fei --- Zoom moderator: KC Wang --- Note taker: Ron Hutchins)	
	David Halstead	NRAO	Next Generation Very Large Array Communications
	Frank Wuerthwein	UCSD	HL-LHC Data Challenges
	Khulud Alsultan	Missouri	Scalable Characteristic Mode Analysis: Requirements and Challenges
	Micah Beck	Tennessee	Location, Location, Location: The Exposed Buffer Approach to Problems of Data Logistics
	Ian Foster	ANL	Building a Global Research Data Platform: Globus Perspectives
	Seymour Douglas	Ithaka . org	The Informational Value of Another Column in Massive Data and Feature Stores
	Jian Wu	Old Dominion	Scholarly Very Large Data: Challenges For Digital Libraries
2:40-3:00 EDT	**Break**	Please join a hallway conversation	
3:00-4:00 EDT	**Huge Data Issues Discussion**	Ron Hutchins (leader)	Huge Data Issues Discussion, Tuesday's Breakout Preview, Other Suggestions and Topics

	Tuesday	**April 14, 2020**	
10:00-11:20 EDT	**Data Movement**	Session chair: KC Wang; Zoom moderator: Zongming Fei; Note taker: Jim Griffioen	
	Leon Reznik	RIT	From Data Communication to Delivery of Quality Data
	Wenji Wu	Fermilab	BigData Express: Toward Schedulable, Predictable, and High-Performance Data Transfer
	Shawn McKee	Michigan	Networking Areas of Interest for HEP
	Igor Sfiligoi	UCSD	Characterizing networking as experienced by users
	Celeste Anderson	Pacific Wave	Pacific Wave Support of Huge Data applications
	Hans Addleman	Indiana	EPOC: Roadside Assistance to the Rescue!
	Jacob Fosso Tande	UNC-Greensboro	Mitigating the dilemma of huge data flow at the University o fNorth Carolina, Greensboro
	Joe Mambretti	Northwestern	Next Generation Infrastructure for Big Data Science
	Engin Arslan	Nevada	End-to-End Parallelism for Distributed Science Workflows
	George Papadimitriou	RENCI	Leveraging Dynamic Resource Provisioning and In-network Processing for Scientific Workflows Operating on Huge Data
	Edmund Yeh	Northeastern	An SDN/NDN Integrated Big Data Ecosystem for Big Science
	Harvey Newman	Caltech	Update on Network Requirements and Computing Model R&D for the HL-LHC Era
	Violet Syrotiuk	ASU	Challenges with Petabyte-Scale Flows and Beyond
	Clifford Lynch	CNI	Shipping Computation to Very Large Data – an Old Idea Demanding Reexamination
11:20-11:40 EDT	**Break**		
11:40-12:30 EDT	**Breakout Session**		
	Room1	New Areas of Research Beyond Big Data (Facilitator: KC; Note Taker: Ron)	
	Room2	New Types of Data & Ways to Get Them (Facilitator: Jim; Note Taker: Zongming)	
12:30-1:00 EDT	**Lunch Break**	Please join a hallway over-lunch conversation	
1:00-2:20 EDT	**Data Processing and Security**	Session chair: Jim Griffioen; Zoom moderator: KC Wang Fei; Note taker: Zongming Fei	
	Mariam Kiran	ESnet	Debugging Bad Performance in Huge Infrastructure: Using ML and AI
	Liqiang Wang	Central Florida	Optimizing Big Data Applications Using Hybrid Program Analysis
	Richard Yang	Yale	Toward Collaborative Networks and Applications (CNA) as a Key Component for Huge Data Networking
	Zhili Zhang	Minnesota	5G, Edge Computing and Future Network Support for Huge Data
	Hao Che	UTArlington	STITCH: Orchestrating Sensing-as-a-Service at Scale
	Ling Liu	Georgia Tech	Improving Huge Data Analytics with In-Network Memory Computing
	Joaquin Chung	ANL	Leveraging Network Programmability and In-network Computing for Streaming Analysis of Huge Data
	Malathi Veeraraghavan	Virginia	Challenges with collecting, anonymizing, sharing and using high-speed network-traffic data
	Sheng Di	Chicago/ANL	Lossy compression for transferring, storing and analyzing huge scientific datasets
	Jon Calhoun	Clemson	Error-bounded Fixed-ratio Lossy Compression for Scientific Data
	Tony Luo	MissouriUST	Scalable Distributed Machine Learning with Huge Data for IoT and Scientific Discovery
	Saptarshi Debroy	CUNY	Multi-Cloud Performance and Security-driven Brokering for Bioinformatics Workflows
	Kaiqi Xiong	South Florida	Deep Learning for Adversarial Attacks in Large Scale Networks
	Lan Wang	Memphis	Toward a Name-Based Data-Centric Platform for Scientiↂc Research

2:20-2:40 EDT	**Break**		Please join a hallway conversation
2:40-3:40 EDT	**Breakout Session**		
		Room3	Collaboration across Disciplines (Facilitator: Ron; Note Taker: KC)
		Room4	Critical Research Infrastructure Needed Beyond Big Data (Facilitator: Zongming; Note Taker: Jim)
3:40-4:20 EDT	**Breakout Report**		Leaders from the Breakout Sessions will report on their discussions
4:20-5:00 EDT	**Key Observations Discussion**	Ron Hutchins (leader)	Discuss findings and key observations