Center for Research in Intelligent Storage (CRIS) & Our New Direction: Hyperconverged Infrastructure

November 2018

CRIS Motivation: Evolution of Computing

Before 1980

- Data sits almost exclusively in datacenters
- Data and compute centralized
- Business-focused

1980—2000

- Data and compute are distributed
- Datacenters expand role in managing data
- Quick expansion in entertainment

2000 to Today

- Datacenters expand to cloud infrastructures
- Compute continues to be distributed; data begins to contract
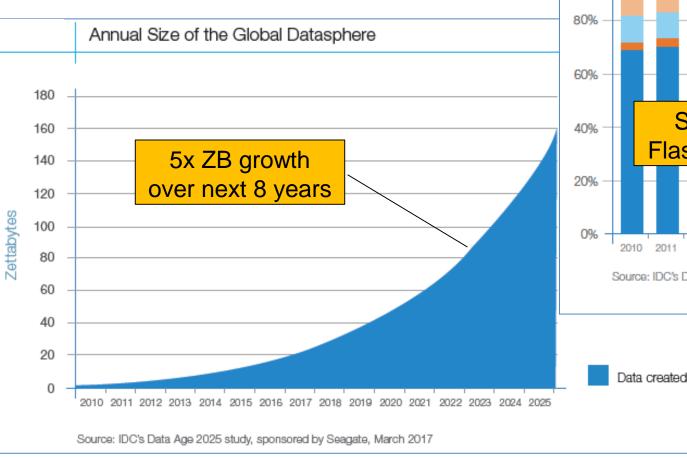- Add social to the mix

Source: Data Age 2025: The Evolution of Data to Life-Critical
David Reinsel John Gantz John Rydning | March 2017
IDC White Paper Sponsored by Seagate © 2017 IDC

CRIS — Center for Research in Intelligent Storage

UNIVERSITY OF MINNESOTA — Driven to Discover®   TEXAS A&M UNIVERSITY   TEMPLE UNIVERSITY®

CRIS Motivation…

Annual Size of the Global Datasphere

5x ZB growth over next 8 years

Source: IDC's Data Age 2025 study, sponsored by Seagate, March 2017

Byte Shipment Share by Storage Media Type

Substantial shift to Flash over next 8 years

DRAM
Optical
Tape
Flash
HDD

Source: IDC's Data Age 2025 study, sponsored by Seagate, March 2017

Source: Data Age 2025: The Evolution of Data to Life-Critical
David Reinsel John Gantz John Rydning | March 2017
IDC White Paper Sponsored by Seagate © 2017 IDC

Center for Research in Intelligent Storage

3

UNIVERSITY OF MINNESOTA
Driven to Discover®

TEXAS A&M
UNIVERSITY

TEMPLE
UNIVERSITY®

CRIS Mission & Research Portfolio

Mission: *push the boundaries of file, memory and storage systems by exploring and developing new technologies and techniques to improve the usability, scalability, security, reliability, and performance of storage systems.*

Research Portfolio:

- **New Storage Technologies:** Flash Memory based SSD, NV-RAM, Shingled Write Disks, Kinetic Drives, Object Storage Devices/Systems, and Active Storage Devices including more computing power in the devices

- **New Storage Hierarchies:** multi-level caching/prefetching, data allocation/migration, and tiered storage, differential storage hierarchy

- **Cloud Storage and Big Data:** OpenStack, Key-Value Store, Hadoop, Spark, Docker, Access Hint, Cloud Storage, and SDS

- **Tools:** Workload Characterization and Workload Generation, High Fidelity Replayer, etc.

CRIS Background & Operations

- NSF I/UCRC established in 2009
- Currently seven industry & three university members
  - University of Minnesota
  - Texas A&M
  - Temple
- Operating budget: $1,100,000
  - $650,000 in IAB fees
  - $460,000 in NSF fees

- IAB Meetings
  - Spring & Fall (2 Days), Summer & Winter (2 Hours)
  - Propose, review & direct research projects
  - Assess & advise on center direction & operations
  - Network with industry members, PI's & students
- Workshop (1 Day)
  - Spring
  - Industry & academia present major trends, challenges & research needs
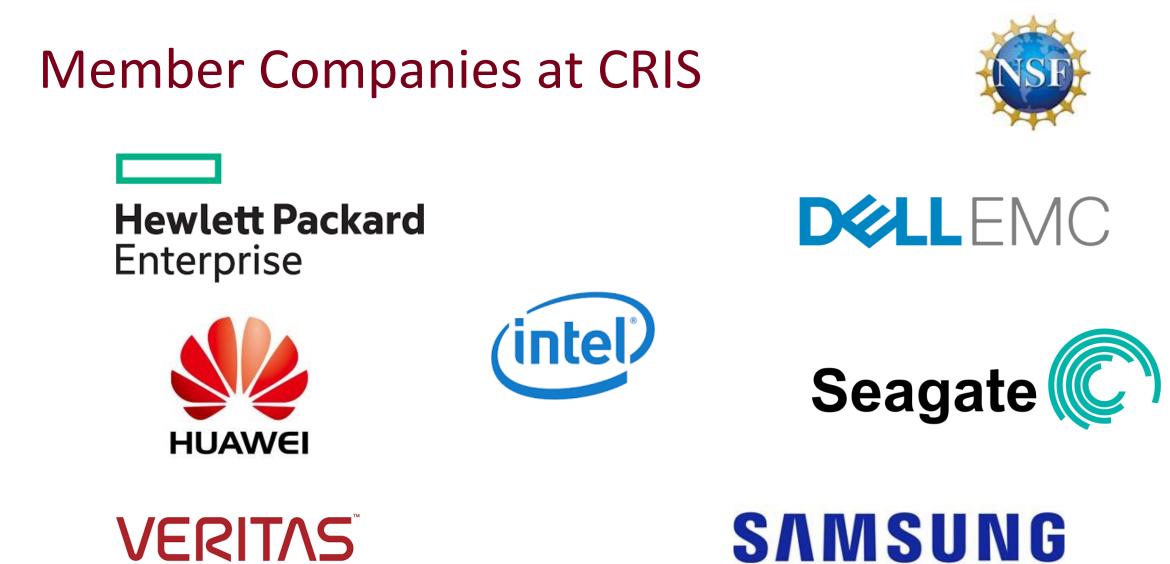 - Open to public

CRIS Website: cris.cs.umn.edu

Member Companies at CRIS

Potential New Sponsors: Alibaba, AT&T, Century Link, Cisco, Microsoft, Micron, SKHynix, ARM and Bank of America

Why Join CRIS?

- Supplement your company's pre-competitive research effort in developing & transferring knowledge
- Support development of & access for recruiting future data storage scientists & engineers
- Collaboration with a PI & student to define & direct a research project (bi-weekly calls)
- Leverage – one membership ($50K for Large/$15K for Small Companies) provides access to ≈$1M in research project results
  - 16 projects in 2018
- Non-exclusive, royalty free license for all IP
- Attend:
  - Bi-annual IAB meetings to review all projects, assess/influence CRIS direction & network with industry colleagues
  - Bi-annual virtual interim IAB meetings to review/recommend new research proposals & assess/influence CRIS direction
  - Annual workshop on storage trends

CRIS Members: "IT Environment is Transforming"

Goal: Data Processing → Information Retrieval → Knowledge Generation & Decision Making

+

White-Box Effect (Learned from Cloud Computing)

+

Open Source Effect

Drives CRIS Members' Interest in Hyperconverged Infrastructure

Hyperconverged Infrastructure (HI) - Seamless integration of compute, network & storage:

- Key Internet building block

- Consolidates compute, storage & network for big data applications
  - Increased efficiency
  - Broad scalability
  - Improved agility
  - Reduced costs

- Research spaces: Interactions between compute, network & storage AND new building blocks/components:
  - New architectures, Non-Volatile memory, VM & Containers for server compute
  - New optical networks, 5G cellular system, NFV (Network Functional Virtualization) & software-defined network for switches & routers
  - Hierarchical & Hybrid Storage, Software-Defined Storage (SDS), Data Archive Systems

Center for Hyperconverged Infrastructure (CHI)

Proposal

- Building upon the I/UCRC Center for Research in Intelligent Storage (CRIS)

- CRIS established working relationships with Intel, IBM, HPE, SGI, Cray, Salesforce, Seagate, Veritas, NetApp, Dell/EMC, HGST, LSI, Samsung, Huawei & SK Telecom

  - They identified hyperconvergence trend & recommended this change

- All current CRIS members committed to join

- Thirteen PI's: Temple (4), Texas A&M (4) & U of MN (5) specializing in IT systems, computing & networking architectures & data storage systems

- Same operating model & member benefits as CRIS

CRIS Director, Proposed CHI Director & Research

David Hung-Chang Du, proposed CHI Director, is the Qwest Chair Professor of Computer Science and Engineering at the University of Minnesota, Minneapolis. He is the Center Director of the NSF multi-university I/UCRC Center of Research in Intelligent Storage (CRIS). He is an IEEE Fellow and a Fellow of Minnesota Supercomputing Institute. His research areas include mass storage systems, high-speed and sensor networks, cyber security and parallel/distributed processing.

CRIS/CHI Research

- Cloud computing & big data support
- Management at scale across all components including data distribution over the Internet & quick response to data independent of location
- Efficient use of software systems like Key-Value Store, Hadoop, Spark, Docker, access hints, cloud storage & software defined storage
- Exploration of internal architectures of servers, switches, routers, & storage systems all the way to silicon
- Using AI & machine learning approaches to reduce cost & improve performance of large hyperconverged infrastructures like the Internet

11

CRIS Co-PI, Proposed CHI Co-PI & Research

Paul Gratz is an Associate Professor in the Department of Electrical and Computer Engineering at Texas A&M University where he is a member of the Computer Engineering and Systems Group. He leads the CAMSIN research group; which focuses on efficient and reliable design in the context of high performance computer architecture, processor memory systems and on-chip interconnection networks. He is a co-PI for the Center for Research in Intelligent Storage (CRIS).

CRIS/CHI Research

- Implications of hyperconvergence on compute & storage components
- Novel systems software & software/hardware rearchitecting for non volatile memory (NVM)
- Novel device interfaces/organizations for efficient systems software
- Data center architectures for NVM
- New & novel CPU & systems architectures cognizant of emerging NVM & non-volatile storage
- Techniques & mechanisms to support virtualization & containerization of persistent memory
- Hardware support & acceleration for hybrid memory composed of emerging NVM & DRAM
- Techniques to leverage speculation driven, forward looking memory management, rather than demand driven, backward-looking policies
- AI & machine learning approaches to direct & manage speculation
- Workload/application dependent positioning of boundary between edge compute & centralized compute, in an effort to improve overall performance & reduce network traffic

Center for Research in Intelligent Storage

UNIVERSITY OF MINNESOTA
Driven to Discover®

TEXAS A&M UNIVERSITY

TEMPLE UNIVERSITY®

CRIS Site Director, Proposed CHI Site Director & Research

Krishna Kant, proposed CHI Site Director, is a professor in the Computer and Information Systems (CIS) department at Temple University, Philadelphia and site director of CRIS. He is an IEEE Fellow and carries 37 years of experience in academia, industry and government. His recent research interests include data driven analysis of storage systems performance, energy management in data centers, automated mechanisms for configuration management, resilience in high performance computing, edge computing, security in cyber physical systems including smart grid and smart transportation, and body area communications.

CRIS/CHI Research

- High speed decentralized storage hierarchies
- End-to-end integration of complex storage hierarchies
- Automated configuration management & diagnosis
- Integrated management & diagnosis in multiparty environments
- Automatic end-to-end energy & resource management
- Holistic security & privacy protection during system design time

Industry Testimonials

"HPE have been directly engaged in many of the research projects of the CRIS since its inception. Since then HPE have sponsored research in a wide range of storage-related fields, including archival storage, scalable distributed indexing and non-hierarchical file systems, large-scale distributed storage systems, file systems for next-generation storage devices, and data deduplication. We also have partnered with CRIS with particular interest in cross-cutting issues such as security and reliability in file and storage systems. Over the years we hired many bright CRIS research students and universities. It's one of our goals in HPE and CRIS is to build a bridge between the industry and the academia to enrich the technologies we all use. Storage is changing with the emergence of Hyperconvergence and cloud, also machine learning and AI. CRIS is well suited for practical research and is led by very knowledgeable staff. Many of the staff members are actively engaged in the high tech industry. This brings value to the CRIS center and to the industry alike. We continue to see tremendous value in CRIS, Now CHI, hence we continue to fund its research projects. We highly recommend the CHI center of research and we stand by its research staff and its projects and we look forward to continue to partner with CHI for many years to come for a better world we build together."

Respectfully, Ayman Abouelwafa 3PAR-CTO   HPESD

Industry Testimonials…

"For Seagate, participation in CRIS has offered an opportunity to work with students and leverage the expertise of the PIs to address technical challenges we face. As the students work on their projects, the exchange between industry sponsor and student can also be a recruitment opportunity. Additionally, we get a view into the trends and issues facing the wider storage industry through the interaction with other industry members, outside of traditional customer relationships. CHI looks to expand the what was built with CRIS, augmenting storage with compute, networking, and the software that enables storage systems. Seagate is fully supportive of this realignment, as we believe the broader focus will help maintain active industry participation as well as continue to offer insight into industry trends. Plus, developing a broader range of expertise within the center can only help as we look to solve tomorrow's challenges."

Bryan Wyatt, Seagate Logic Design Engineer, VLSI Engineering & CRIS IAB Chair

Industry Testimonials…

"Veritas has been working with CRIS for many years, and benefits from the collaboration greatly. The HyperProtect project, for example, helps us understand the backup telemetry better and uses the information for intelligent data protection."

Weibao Wu, Veritas Distinguished Engineer

Next Steps…

Upcoming Events to Learn More:

Hyperconverged Infrastructure Workshop: April 30, 2018; U of MN, Minneapolis, MN

Industrial Board Meeting: May 1-2, 2018; U of MN, Minneapolis, MN

Meet / Discuss in Depth:

David Du: du@cs.umn.edu

Narasimha Reddy: reddy@tamu.edu

Paul Gratz: pgratz@tamu.edu

CRIS Website: cris.cs.umn.edu

Krishna Kant: kkant@temple.edu

Dave Aune: dbaune@umn.edu   952 807 6853

ADDITIONAL MATERIAL

CRIS Site Director, Proposed CHI Site Director & Research

Narasimha Reddy, proposed CHI Site Director, is a J.W. Runyon Professor in the department of Electrical and Computer Engineering at Texas A&M University and site director of CRIS. He is also the Associate Dean for Research with the Texas A&M Engineering Program and the Assistant Director of Strategic Initiatives & Centers with the Texas A&M Engineering Experiment Station. His research interests are in Computer Networks, Storage Systems, Multimedia Systems, and Computer Architecture.

CRIS/CHI Research

- Implications of hyperconvergence on compute & storage components
- Novel systems software & software/hardware rearchitecting for non volatile memory (NVM)
- Novel device interfaces/organizations for efficient systems software
- Data center architectures for NVM
- New & novel CPU & systems architectures cognizant of emerging NVM & non-volatile storage
- Techniques & mechanisms to support virtualization & containerization of persistent memory
- Hardware support & acceleration for hybrid memory composed of emerging NVM & DRAM
- Techniques to leverage speculation driven, forward looking memory management, rather than demand driven, backward-looking policies
- AI & machine learning approaches to direct & manage speculation
- Workload/application dependent positioning of boundary between edge compute & centralized compute, in an effort to improve overall performance & reduce network traffic

CRIS Research Portfolio, Fall 2018

Mission: push the boundaries of file, memory and storage systems by exploring and developing new technologies and techniques to improve the usability, scalability, security, reliability, and performance of storage systems.

Research Pathway	Research Project Title	Students / Advisors	Sponsors	Start Date
NEW STORAGE DEVICES	UMN - New Storage Devices	Baoquan Zhang, Fenggang Wu, Xuchao Xie, Hebatalla Eldakiky / Dr. David Du	Seagate	Nov 2015
	TX A&M - Data Redundancy for Key-Value Device	Mian Qin / Dr. Narisimha Reddy & Dr. Paul Gratz	Samsung	Mar 2017
	UMN - Database Application Accelerators in Solid-State Disks	Jinfeng Yang / Dr. David Lilja	Intel	Aug 2016
NEW STORAGE HIERARCHIES	TX A&M - Virtualization of NVM	Jaemin Jung, Chih Chieh Chou/ Dr. Narisimha Reddy & Dr. Paul Gratz	HPE	Sep 2015
	UMN - Storage System Design with Hybrid Memory	Ming-Hong Yang / Dr. David Du	HPE	Jun 2017
	TX A&M - Peripheral Memory Management Unit	Fei Wen, Mian Qin / Dr. Narisimha Reddy & Dr. Paul Gratz	Dell-EMC	Nov 2015
	Temple - Provenance Directed Data Prefetching	Dusan Ramljak, Jit Gupta / Dr. Krishan Kant	HPE	Oct 2015
	Temple -Implementation of Intelligent Caching in Huawei's Storage System	Dusan Ramljak, Jit Gupta / Dr. Krishna Kant	Huawei	Oct 2017
	Temple -Storage System Heat Analysis and Prediction	Lu Pang, Anis Alazzawe, Tanya Roy Dusan Ramljak / Dr. Krishna Kant	Dell EMC	Oct 2018
CLOUD & BIG DATA	UMN - Enhanced Kubernetes Support on Storage	Hao Wen / Dr. David Du	HPE	Apr 2017
TOOLS	UMN - HyperProtect: An Automatic Data Protection	Yaobin Qin / Dr. David Lilja	Veritas	Jan 2017

CRIS Research Portfolio, Fall 2018…

Mission: push the boundaries of file, memory and storage systems by exploring and developing new technologies and techniques to improve the usability, scalability, security, reliability, and performance of storage systems.

University	New Storage Devices	New Storage Hierarchies	Cloud Storage & Big Data	Tools	2017 - 2018 Publications
U of MN	2	1	1	1	28 [1, 2, 3, 4]
Texas A&M	1	2			4 [1,2,3]
Temple		3			5 [1]
TOTAL	**3**	**6**	**1**	**1**	**35**

U of MN
1. 14 Conference Papers
2. 2 Journal Articles
3. 11 Drafted for Conference or Journal Submission
4. 1 Open Source Distribution (Linux HF Replay)

Temple
1. 1 Conference Papers
2. 1 Journal Articles
3. 3 Drafted for Conference or Journal Submission

Texas A&M
1. 2 Conference Papers
2. 2 Drafted for Conference or Journal Submission

Publications: UMN

Paper Title	Authors and Affiliations	Submitted Conference	Earliest Publication Date	Date Posted on CRIS Wiki
JoiNS: Meeting latency SLO with Integrated Control for Networked Storage	Hao Wen	MASCOTS 2018	June 2018	Sept 2018
ChewAnalyzer: Workload-Aware Data Management Across Differentiated Storage Pools	Xiongzi Ge, Xuchao Xie, David H.C. Du, Pradeep Ganesan and Dennis Hahn	MASCOTS 2018	June 2018	Sept 2018
Data Management Design for Interlaced Magnetic Recording	Fenggang Wu, Baoquan Zhang, Zhichao Cao, Hao Wen, Bingzhe Li, Jim Diehl, Guohua Wang and David H.C. Du.	HotStorage '18	July 2018	April 2018
Reducing Relational Database Performance Bottlenecks Using 3D Crosspoint Storage Technology	Jinfeng Yang and David J. Lilja	IEEE-BigDataSE	July 2018	NA
On Improving the Write Responsiveness for Host-Aware SMR Drives	Ming-Chang Yang, Yuan-Hao Chang, Fenggang Wu,Tei-Wei Kuo, and David H.C. Du	IEEE Transactions on Computers	TBD	NA
Improving Data Integrity in Linux Software RAID with Protection Information	Baoquan Zhang, Alireza Haghdoost	AHPAMA 2018	May 2018	NA
An NVRAM-based Burst Buffer Coordination System for Parallel File Systems	Ziqi Fan	HPC 2018	April 2018	NA
Key-value Pairs Allocation Strategy for Kinetic Drives	Hebatalla Eldakiky and David H.C. Du.	IEEE Bigdata Service 2018	March 2018	Jan 2018
NetStorage Replayer: A Cloud Server Performance Evaluation Tool that Synchronously Replays Network and Storage Traces	Bingzhe Li, Hao Wen, Farnaz Toussi, Clark Anderson, Bernard A. King-Smith, David J. Lilja, and David H.C. Du	FAST 2018	February 2018	November 2017
JoiNS: Guaranteed QoS with Integrated Control for Networked Storage	Hao Wen, Zhichao Cao, Yang Zhang, Xiang Cao, Ziqi Fan, Doug Voigt and David H.C. Du	FAST 2018	February 2018	November 2017
HintStor: A Flexible Framework to Study I/O Access Hints in Heterogeneous Storage Systems	Xiongzi Ge, Zhichao Cao, David H.C. Du, Pradeep Ganesan, Dennis Hahn	FAST 2018	February 2018	November 2017
ZoneTier: A Zone-based Storage Tiering and Caching Co-Design to Integrate Solid State Drives with Host-Aware SMR Drives	Xuchao Xie, Liquan Xiao, David H.C. Du	FAST 2018	February 2018	November 2017
ALACC: Accelerating Restore Performance of Data Deduplication Systems Using Adaptive Look-Ahead Window Assisted Chunk Caching	Zhichao Cao, Hao Wen, Fenggang Wu, David H.C. Du	FAST 2018	February 2018	November 2017
SmartRAID: A RAID-5 with Alternating Idle Shingled Magnetic Recording (SMR) Drives	Baoquan Zhang, David H.C. Du	TBD	TBD	November 2017
Evaluating Media Cache Cleaning and Improving I/O performance in Shingle Magnetic Recording (SMR) Drives	Baoquan Zhang, David H.C. Du	TBD	TBD	November 2017
Tier-code: An XOR-based RAID-6 Code with Improved Write and Degraded-mode Read Performance	Bingzhe Li, Soheil Mohajer and David Lilja	IEEE Cluster 2017	September 2017	April 2017
Hibachi: A Cooperative Hybrid Cache with NVRAM and DRAM for Storage Arrays	Ziqi Fan, Fenggang Wu, Dongchul Park, Jim Diehl, Doug Voigt, and David H.C. Du	MSST 2017	May 2017	March 2017
Evaluating Media Cache Cleaning and Improving I/O performance in Shingle Magnetic Recording (SMR) Drives	Baoquan Zhang, Xuechao Xie and David H.C. Du	HotStorage'17	July 2017	March 2017
SMaRT: An Approach to Shingled Magnetic Recording Translation	Weiping He, David H. Du	FAST 2017	February 2017	February 2017
Kinetic Action: Application Level Performance Analysis of Kinetic Drives	Manas Minglani, Jim Diehl, Xiang Cao, Bingzhe Li, David J. Lilja, David H. Du	FAST 2017	February 2017	October 2016
On the Accuracy and Scalability of Intensive I/O Workload Replay	Alireza Haghdoost, Weiping He, Jerry Fredin, David H.C. Du	FAST 2017	February 2017	October 2016
TDDFS: A Tier-aware Data Deduplication based File System	Zhichao Cao, Hao Wen, Xiongzi Ge, Jingwei Ma, David H.C. Du	FAST 2017	February 2017	October 2016
Virtual Persistent Cache: Remedy the Long Latency Behavior of Host-Aware Shingled Magnetic Recording Drives	Ming-Chang Yang, Yuan-Hao Chang, Fenggang Wu, Tei-Wei Kuo, David H.C. Du	FAST 2017	TBD	October 2016
Data Allocation of Large-scale Key-value Store System using Kinetic Drives	Xiang Cao, Manas Minglani, David Hung-Chang Du	IEEE BigDataService 2017	April 2017	October 2016
Performance Evaluation of Host Aware Shingled Magnetic Recording (HA-SMR) Drives	Fenggang Wu, Ziqi Fan, Ming-Chang Yang, Baoquan Zhang, Xiongzi Ge and David H.C. Du	HotStorage'17	July 2017	October 2016
Improving Data Integrity in Linux MD RAID with T10-PI	Baoquan Zhang, Raghunath Raja Chandrasekary, Alireza Haghdoost, Fenggang Wu, Lance Evansy, David H.C. Du	TBD	TBD	October 2016
TraceRAR: An I/O Performance Evaluation Tool for Replaying, Analyzing, and Regenerating Traces	Bingzhe Li, Farnaz Toussi, Clark Anderson, David Lilja, David H.C. Du	ISPASS 2017	April 2017	October 2016

Publications: Texas A&M

2017 - 2018

- "SPaN: Speculative Paging for future NVM and SSD," Viacheslav Fedorov, Jinchun Kim, Mian Qin, A. L. Narasimha Reddy and Paul Gratz. *In the International Symposium on Memory Systems (*MEMSYS'17*), October 2017.*

- "Kill the program counter: Reconstructing program behavior at the last level cache," Jinchun Kim, Elvira Teran, Paul V. Gratz, Daniel Jimenez, Seth Pugsley, and Chris Wilkerson. *In the 22nd International Conference on Architectural Support for Programming Languages and Operating Systems* (ASPLOS), April 2017.

- "KVRAID: A Space Efficient, Low Write Cost Data Redundancy Scheme for KV SSDs", Mian Qin, A. L. Narasimha Reddy, Paul V. Gratz, Yang Seok Ki *

- "vNVML: An Efficient Shared Library for Virtualizing and Sharing Non-volatile Memories", Chih Chieh Chou, Jaemin Jung, A. L. Narasimha Reddy, Paul V. Gratz, Doug Voigt *

* Under review for the 17th USENIX Conference on File and Storage Technologies (FAST '19)

Publications: Temple

2017 - 2018

- "Modular framework for Data Prefetching and Replacement at the Edge", Dusan Ramljak, K. Kant, submitted to Edge Computing 2018
- "Understanding Impact of Workload Phase Change on Caching Characteristics", Dusan Ramljak, K. Kant, submitted to Mascots 2018
- "Belief-Based Storage Systems"; Dusan Ramljak, K. Kant. *HotStorage '17, JULY 10–11, 2017, SANTA CLARA, CA*
- "Data Driven Dynamic Adaptation of Caching Configuration Parameters in Storage Systems", Dusan Ramljak, Madhurima Ray, Spencer Melnick, Andrew Posmonteir, Krishna Kant, under review.
- "Adaptive Combined Cache For Managing Performance-Endurance Trade-off", Tanaya Roy, Dusan Ramljak, Krishna Kant, in preparation.

Students 2017 -2018

Graduated

Student Name	University	New Organization
Ziqi Fan	U of MN	Samsung
Xiongzi Ge	U of MN	NetApp
Xiang Cao	U of MN	Grand Valley State University
Manas Minglani	U of MN	AMD
Jinchun Kim	Texas A&M	Apple

Working on CRIS Research Projects

	Student Name	Student Level & Year	University	# Projects
1	Fenggang Wu	PH, 4th Yr	U of MN	1
2	Bingzhe Li	PH, 4th Yr	U of MN	1
3	Zhichao Cao	PH, 4th Yr	U of MN	1
4	Chia-Wen Hsieh	PH, 1st Yr	U of MN	1
5	Hao Wen	PH, 4th Yr	U of MN	1
6	Baoquan Zhang	PH, 3rd Yr	U of MN	1
7	Hebatalla Eldakiky	PH, 1st Yr	U of MN	1
8	Yaobin Qin	PH, 2nd Yr	U of MN	1
9	Jinfeng Yang	PH, 2nd Yr	U of MN	1
10	Ming-Hong Yang	PH, 2nd Yr	U of MN	1
11	Dusan Ramljak	PH, 4th Yr	Temple	3
12	Tanaya Roy	PH, 2nd Yr	Temple	1
13	Madhurima Ray	PH, 3rd Yr	Temple	1
14	Lu Pang	PH, 1st Yr	Temple	1
15	Pavana Pradeep	PH, 1st Yr	Temple	1
16	Jit Gupta	PH, 1st Yr	Temple	1
17	Anis Alazzawe	PH, 3rd Yr	Temple	1
18	Jaemin Jung	PD	Texas A&M	1
19	Chih Chieh Chou	PH, 2nd Yr	Texas A&M	1
20	Mian Qin	PH, 2nd Yr	Texas A&M	3
21	Fei Wen	PH, 2nd Yr	Texas A&M	1

CRIS — Center for Research in Intelligent Storage

UNIVERSITY OF MINNESOTA Driven to Discover®   TEXAS A&M UNIVERSITY   TEMPLE UNIVERSITY®

About CRIS

The Center for Research in Intelligent Storage (CRIS) is a partnership between universities and industry, featuring high-quality, industrially relevant fundamental research, strong industrial support of collaboration in research and education, and direct transfer of university developed ideas, research results, and technology to U.S. industry to improve its competitive posture in world markets. Through innovative education of talented graduate and undergraduate students, CRIS is providing the next generation of scientists and engineers with a broad, industrially oriented perspective on engineering research and practice. CRIS currently consists of three universities: University of Minnesota, Temple University and Texas A&M

Collaborative Sites

CRIS at Temple University

CRIS at Texas A&M

Center Mission

The goal of the Center for Research in Intelligent Storage (CRIS) is to push the boundaries of file, memory and storage systems by exploring and developing new technologies and techniques to improve the usability, scalability, security, reliability, and performance of storage systems.

The Center is established in the University of Minnesota a university with strong ties to the storage system industry: Minneapolis has long been a center of the storage industry in the United States. Research projects in the Center are supported by industrial members, and the Center encourages frequent participation by industry employees in the projects: participation in weekly meetings via phone call, email discussions, and face-to-face meetings. The Center encourages students to engage member companies via summer internships, facilitating technology transfer and building strong ties that can result in full-time employment after graduation.

Support

The Center is primarily supported by its industrial membership and by funding from the National Science Foundation. The Center is part of the Industry & University Cooperative Research Program (I/UCRC) at NSF.

CRIS

Center for Research in Intelligent Storage

UNIVERSITY OF MINNESOTA
Driven to Discover®

TEXAS A&M
UNIVERSITY

TEMPLE
UNIVERSITY®

Hyperscale Data Centers & Storage Systems Workshop

May 1, 2018   DTC, Walter Library, Room 402

08:00 Continental Breakfast - Networking

08:30 Introductions and Welcome — Dr. David Du, Qwest Chair Professor, Computer Science and Engineering, CRIS Director

08:45 Towards Super-Datacenter Computing — Dr. Mike Marty, Google

09:35 Health Care & Cognitive Computing — Bill Rapp, IBM Distinguished Engineers, Watson Health

10:30 Coffee Break – Networking

11:00 UnitedHealthcare - Optum HyperscaleData Centers — Joel Carlson & Luke Burns

11:50 Lunch

1:00   Storage Research for Solving Big Data Problem — Dr. David Du

1:50   Network Function Virtualization for Cloud Computing - HydraNF: Accelerating Service Function Chains with Parallelism — Dr. Zhi-Li Zhang, Qwest Chair Professor and McKnight Distinguished University Professor, Computer Science and Engineering

2:40   LANL Data Centers — Dr. Josip Loncaric, HPC Technology Futures Lead, Los Alamos National Laboratory

3:30   Coffee Break – Networking

4:00   Salesforce Data Centers Trends & Challenges — Dr. Deepak Kenchammana-Hosekote, Chief Architect, Server-Storage Integration, Salesforce

4:50   Summary and Wrap Up — Dr. David Du and Dave Aune

CRIS — Center for Research in Intelligent Storage

UNIVERSITY OF MINNESOTA Driven to Discover®   TEXAS A&M UNIVERSITY   TEMPLE UNIVERSITY

CRIS Motivation

- Past decade witnessed tremendous advances in computing, wired and wireless communication and storage technologies. Remarkable cost reductions have made large computing and storage capacity available to increasing numbers of consumers.

- Unprecedented Internet connectivity enables vast amounts of data creation and availability to satisfy demand of many new applications, resulting in increasing challenges to manage this huge amount of data to be available at our fingertips at anytime from anywhere. NSF I/UCRC Center Research in Intelligent Storage (CRIS) was established to address these challenges.

- A complete solution to meet this challenge calls for new storage architectures, different storage systems design, new data models, new information access methods and new ways to deliver information. Industry has recognized the importance of this new computing and communication environment. Many new products for storage/file systems, storage management, and ways of preserving data have evolved. However, most of these are incremental and more research and development are needed. Much more interaction between academic and industrial researchers should be encouraged.
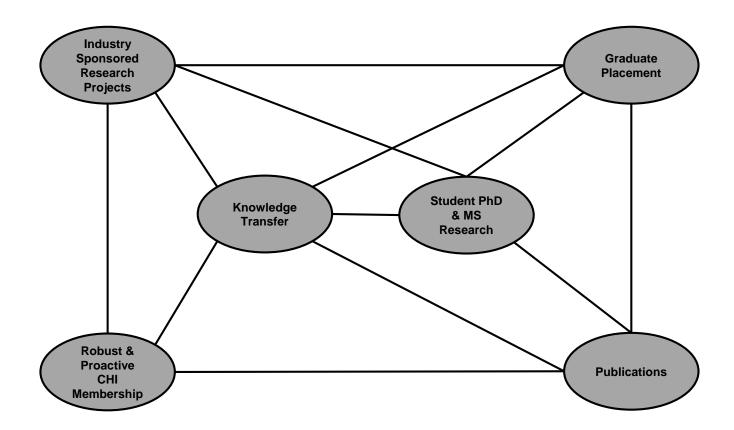
CRIS Activity-System Map

Activity-system maps, … show how an organization's strategic position is contained in a set of tailored activities designed to deliver it. In organizations with a clear strategic position, a number of higher-order strategic themes (in dark grey) can be identified and implemented through clusters of tightly linked activities (in light grey).

<u>**What is Strategy?**</u> By Michael Porter   COPYRIGHT © 2000 HARVARD BUSINESS SCHOOL PUBLISHING CORPORATION
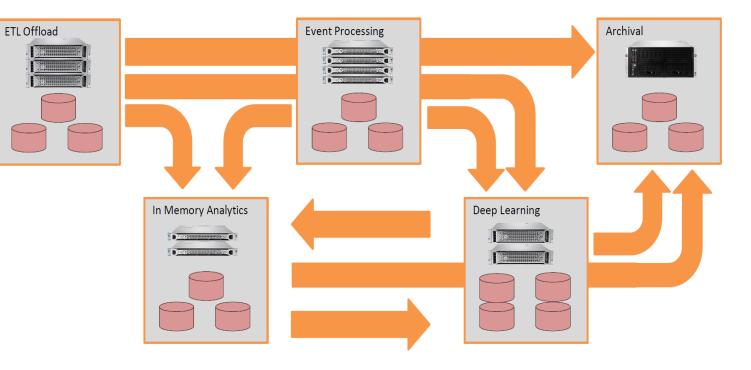
29

CHI Motivation: Typical Data Journey

- Data collected & transformed to different formats & offloaded to large scale distributed storage systems

- Simultaneously, through IoT and other event monitoring capabilities, collected data & real-time streamed data based on current events will be delivered to a large memory-based computing system to be analyzed (in-memory processing).

- Deep learning based AI & machine learning approaches will assist data analytics to support optimal decisions

- The original data as well as the analytic results are to be archived for future uses