Curriculum Vitae (23 December 2021)

Randal Burns

randal@cs.jhu.edu

ORCID: https://orcid.org/0000-0002-2924-1997

Current Position

Professor and Head, Department of Computer Science, Johns Hopkins University

Education

Ph.D. 2000 University of California, Santa Cruz, Computer Science
Advisor: Professor Darrell D. E. Long
Dissertation: Data Management in a Distributed File System for Storage Area Networks.

M.S. 1997 University of California, Santa Cruz, Computer Science
Advisor: Professor Darrell D. E. Long
Thesis: Differential Compression: A Generalized Solution for Binary Files.

B.S. 1993 Stanford University, Geophysics
Conferred with honors and distinction

Employment History

2020–present **Johns Hopkins University**, *Head*, Department of Computer Science, Baltimore, MD.
2015–present **Johns Hopkins University**, *Professor*, Department of Computer Science, Baltimore, MD.
2017–2021 **Gigantum, Inc.**, *Founder, Board of Directors*, Washington, DC.
2019–2020 **Johns Hopkins University**, *Chair*, Department of Computer Science, Baltimore, MD.
2018–2019 **Johns Hopkins University**, *Interim Chair*, Department of Computer Science, Baltimore, MD.
2015–2016 **Los Alamos National Lab**, *Visiting Scientist*, Data Science and Scale Team, Los Alamos, NM.
2008–2015 **Johns Hopkins University**, *Associate Professor*, Department of Computer Science, Baltimore, MD.
2012–2015 **Johns Hopkins Applied Physics Laboratory**, Research and Exploratory Development Department, Temporary On Call Employee, Laurel, MD.
2001–2008 **Johns Hopkins University**, *Assistant Professor*, Department of Computer Science, Baltimore, MD.
1999–2001 **International Business Machines**, *Research Staff*, Almaden Research Center, San Jose, CA.
2000–2001 **University of California**, *Lecturer and Research Fellow*, Santa Cruz, CA.
1996–1999 **International Business Machines**, *Research Associate*, Almaden Research Center, San Jose, CA.
1994–1995 **Dialogic Corporation**, *Software Engineer*, Sunnyvale, CA.

Professional Memberships

USENIX, ACM (Senior Member), IEEE (Senior Member)

Awards and Honors

2020 Member, Information Science and Technology study group, DARPA
2018 Distinguished Alumnus Award, University of California Santa Cruz
2017 Best Presentation, ACM Symposium on High Performance Distributed Computing (HPDC)
2015 Kavli Fellow
2014 Kavli Fellow
2012 Member of the Defense Science Study Group, Class of 2012-2013
2009 Johns Hopkins University Dean's Leadership Award
2009 Best Paper Award, USENIX Conference on File and Storage Technologies (FAST)
2003 IBM Outstanding Innovation Award
2003 National Science Foundation CAREER Award
2002 Department of Energy Early Career Principal Investigator Award
2001 IBM Corporate Accomplishment for Adaptive Differential Backup in the Tivoli Storage Manager
2001 IBM Third Plateau Invention Award
2000 IBM Interoperability in Heterogeneous Environments Special Patent Incentive Award
2000 IBM Second Plateau Invention Award
1998 IBM First Plateau Invention Award
1997 IBM Internet/Network Computing Special Incentive Award
1995 University of California Regents Fellowship

Publications

Journals

J-1 T. M. Tomita, J. Browne, C. Shen, J. Chung, J. L. Patsolic, B. Falk, J. Yim, C. E. Priebe, R. Burns, M. Maggioni, J. T. Vogelstein. Sparse Projection Oblique Randomer Forests. Journal of Machine Learning Research (JMLR), 21(104), pp. 1–39, 2020.

J-2 A. S. Charles, B. Falk, N. Turner, T. D. Pereira, D. Tward, B. D Pedigo, J. Chung, R. Burns, S.S. Ghosh, J. M. Kebschull, W. Silversmith, and J. T. Vogelstein. Toward Community-Driven Big Open Brain Science: Open Big Data and Tools for Structure, Function, and Genetics. Annual Review of Neuroscience, 2020. DOI: 10.1146/annurev-neuro-100119-110036

J-3 R. Burns. Creating an Active Learning Environment using Reproducible Data Science Tools. ACM eLearn, July 2020. DOI: 10.1145/3409311.3403400

J-4 J. Pulido, D. Livescu, K. Kanov, R. Burns, C. Canada, J. Ahrens, and B. Haman. Remote visual analysis of large turbulence databases at multiple scales. *Journal of Parallel and Distributed Computing*, 2018. DOI: 10.1016/j.jpdc.2018.05.011

J-5 J. T. Vogelstein, *29 authors*, and R. Burns. A community-developed open-source computational ecosystem for big neuro data. *Nature Methods*, 2018. DOI: 10.1038/s41592-018-0181-1

J-6 D. G. C. Hildebrand, R. M. Torres, W. Choi, T. M. Quan, A. W. Wetzel, G. S. Plummer, R. Portugues, I. H. Bianco, O. Randlett, S. Saalfeld, A. Baden, K. Lillaney, R. Burns, J. T. Vogelstein, W.-K. Jeong, J. W. Lichtman, and F. Engert. Whole-brain serial-section electron microscopy in larval zebrafish. *Nature*, 18;545(7654):345-349, 2017. DOI 10.1038/nature22356

J-7 G. Kiar, K. J. Gorgolewsk, D. Kleissas, W. Gray Roncal, B. Litt, B. Wandell, R. A. Poldrack, M. Wiener, R. J. Vogelstein, R. Burns, and J. T. Vogelstein. Science In the Cloud (SIC): A use case in MRI Connectomics. *Gigascience*, 2017. DOI 10.1093/gigascience/gix013

J-8 P. Johnson, S. Hamilton, R. Burns, and C. Meneveau. Analysis of geometrical and statistical features of Lagrangian stretching in turbulent channel flow using a database task-parallel particle tracking algorithm. *Physical Review Fluids*, 2(014605), 2017. DOI https://doi.org/10.1103/PhysRevFluids.2.014605

J-9 D. Zheng, D. Mhembere, V. Lyzinski, J. T. Vogelstein, C. E. Priebe, and R. Burns. Semi-External Memory Sparse Matrix Multiplication for Billion-Node Graphs. *IEEE Transactions on Parallel and Distributed Computing*, 28(5), 2016. DOI 10.1109/TPDS.2016.2618791

J-10 J. Graham, K. Kanov, X. I. A. Yang, M. Lee, N. Malaya, C.C. Lalescu, R. Burns, G. Eyink, A. Szalay, R. D. Moser, C. Meneveau. A Web services accessible database of turbulent channel flow and its use for testing a new integral wall model for LES. *Journal of Turbulence*, 17(2), 181–215, 2016. DOI:10.1080/14685248.2015.1088656

J-11 K. Kanov, R. Burns, C. Lalescu, and G. Eyink. The Johns Hopkins Turbulence Databases, an Open Simulation Laboratory for Turbulence Research. *Computing in Science and Engineering*, 17(5), 10–17, 2015. DOI: 10.1109/MCSE.2015.103

J-12 N. Kasthuri, and K. J. Hayworth, and (22 more). Saturated reconstruction of a volume of neocortex. *Cell*, 162(3):648–661, 2015. DOI: 10.1016/j.cell.2015.06.054

J-13 K. M. Harris, J. Spacek, M. E. Bell, P. H. Parker, L. F. Lindsey, A. D. Baden, J. T. Vogelstein, and R. Burns. A resource from 3D electron microscopy of hippocampal neuropil for user training and tool development. *Nature Scientific Data*, 2(150046) 2015. DOI: 10.1038/sdata.2015.46

J-14 N. Weiler, F. Collman, J. Vogelstein, R. Burns, and S. Smith. Synaptic molecular imaging in spared and deprived columns of mouse barrel cortex with array tomography. *Nature Scientific Data*, 1(140046), 2014. DOI: 10.1038/sdata.2014.46

J-15 R. Burns, J. Vogelstein, and A. S. Szalay. From cosmos to connectomes: the evolution of data-intensive science. *Neuron*, 83(6), pp. 1249–1252, 2014, DOI 10.1016/j.neuron.2014.08.045

J-16 G. Eyink, E. Vishniac, C. Lalescu, H. Aluie, K. Kanov, K. Bürger, R. Burns, C. Meneveau, and A. Szalay. Flux-freezing breakdown in high-conductivity magnetohydrodynamic turbulence. *Nature*, 497(7450), 466–469, 2013. DOI 10.1038/nature12128

J-17 H. Yu, K. Kanov, E. Perlman, J. Graham, E. Frederix, R. Burns, A. Szalay, G. Eyink, C. Meneveau. Studying Lagrangian dynamics of turbulence using on-demand fluid particle tracking in a public turbulence database. *Journal of Turbulence*, 13(12):1–29, 2012. DOI: 10.1080/14685248.2012.674643

J-18 G. Ateniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. Peterson, and D. Song. Remote Data Checking Using Provable Data Possession. *ACM Transactions on Information and Systems Security*, 14(1), 2011. DOI 10.1145/1952982.1952994

J-19 R. Burns and Z. Peterson. Security Constructs for Regulatory-Compliant Storage. *Communications of the ACM*, 53(1), 126–130, 2010.

J-20 R. Musaloiu-E., J. Cogan, K. Szlavecz, A. Szalay, J. Gray, S. Ozer, C.-J. M. Liang, J. Gupchup, and R. Burns. Wireless Sensor Networks for Soil Science. *International Journal of Sensor Networks*, 7(1-2), 53–70, 2010. DOI: 10.1504/IJSNet.2010.03185

J-21 A. Batsakis, R. Burns, A. Kanevsky, J. Lentini, and T. Talpey. CA-NFS: A Congestion-Aware Network File System. *ACM Transactions on Storage*, 5(4), 15:1–15:24, 2009. DOI 10.1145/1629080.1629085

J-22 Y. Li, E. Perlman, M. Wang, Y. Yang, C. Meneveau, R. Burns, S. Chen, A. Szalay, and G. Eyink. A Public Turbulence Database Cluster and Applications to Study Lagrangian Evolution of Velocity Increments in Turbulence. *Journal of Turbulence*, 9(31), 1–29, 2008. DOI: 10.1080/14685240802376389

J-23 B. Ball, D. Brady, M. Brooks, R. Burns, B. Cuker, D. DiToro, T. Gross, M. Kemp, L. Murray, R. Murphy, E. Perlman, M. Piasecki, J. Testa and I. Zaslavsky. A Prototype System for Multi-Disciplinary Shared Cyberinfrastructure—Chesapeake Bay Environmental Observatory. *Journal of Hydrologic Engineering*, 13(10), 960–970, October 2008. DOI: 10.1061/(ASCE)1084-0699

J-24 A. Batsakis and R. Burns. NFS-CD: Write-Enabled Cooperative Caching in NFS. *IEEE Transactions on Parallel and Distributed Systems*, 19(3), 323–333, 2008. DOI: 10.1109/TPDS.2008.14

J-25 Z. Peterson and R. Burns. Ext3cow: A Time-Shifting File System for Regulatory Compliance. *ACM Transactions on Storage*, 1(2), 190–212, 2005. DOI: 10.1145/1063786.1063789

J-26 C. Wu and R. Burns. Adaptive and Tunable Randomization for Load Management in Shared-Disk Clusters. *ACM Transactions on Storage*, 1(1), 108–131, 2005. DOI: 10.1145/1044956.1044962

J-27 R. Burns, L. Stockmeyer, and D. Long. In-Place Reconstruction of Version Differences. *IEEE Transactions on Knowledge and Data Engineering*, 15(1), 973–984, 2003. DOI: 10.1109/TKDE.2003.1209013

J-28 M. Ajtai, R. Burns, R. Fagin, D. Long, and L. Stockmeyer. Compactly Encoding Unstructured Inputs with Differential Compression. *Journal of the ACM*, 49(3), 318–367, 2002. DOI: 10.1145/567112.567116

J-29 R. Burns, R. Rees, and D. Long. Efficiently Distributing Data in a Web Server Farm. *IEEE Internet Computing*, 5(4), 56–65, 2001. DOI: 10.1109/4236.939451

J-30 R. Burns, R. Rees, L. Stockmeyer, and D. Long. Scalable Session Locking for a Distributed File System. *Cluster Computing Journal*, 4(4), 295–306, 2001. DOI: 10.1023/A:1011860527389

J-31 B. Reed, E. Chron, D. Long, and R. Burns. Authenticating Network Attached Storage. *IEEE Micro*, 20(1), 49–57, January 2000. DOI:10.1109/40.820053

Refereed Conferences

C-1 B. Wheatman and R. Burns. Streaming Sparse Graphs using Efficient Dynamic Sets. IEEE Conference on Big Data (BigData), 2021.

C-2 M. Madyhasta, K. Lillaney, J. Browne, J. Vogelstein, and R. Burns. PACSET (Packed Serialized Trees): Reducing Inference Latency for Tree Ensemble Deployment. SIGKDD Conference on Knowledge Discovery and Data Mining, 2021.

C-3 B. Choi, R. Burns, and R. Huang. Understanding and Dealing with Hard Faults in Persistent Memory Systems. European Conference on Computer Systems (EuroSys), 2021.

C-4 M. Madhyastha, G. Li, V. Strnadova-Neeley, J. Browne, J. T. Vogelstein, R. Burns and Carey E. Priebe. Geodesic Forests. SIGKDD Conference on Knowledge Discovery and Data Mining, 2020.

C-5 K. Lillany, V. Tarasov, D. Pease, and R. Burns. Agni: An Efficient Dual-access File System over Object Storage. ACM Syposium on Cloud Computing, 2019.

C-6 J. Browne, T. Tomita, D. Mhembere, J. T. Vogelstein, and R. Burns. Forest Packing: Fast, Parallel Decision Forests. SIAM International Conference on Data Mining, 2019.

C-7 K. Lillaney, D. Kleissas, A. Eusman, E. Perlman, W. Gray-Roncal, J. T. Vogelstein, and R. Burns. Building NDStore Through Hierarchical Storage Management and Microservice Processing. IEEE eScience, 2018, DOI: 10.1109/eScience.2018.00037

C-8 D. Zheng, D. Mhembere, C. Priebe, J. Vogelstein and R. Burns. FlashR: Parallelize and Scale R for Machine Learning using SSDs. ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, 2018.

C-9 D. Mhembere, D. Zheng, C. Priebe, J. Vogelstein, and R. Burns. knor: A NUMA-optimized In-memory, Distributed and Semi-external-memory k-means Library. *High Performance Distributed Computing*, 2017.

C-10 S. Hamilton, P. Lindstrom, J. Patchett, C. Meneveau, and R. Burns. Extreme Event Analysis in Next Generation Simulation Architectures. *International Supercomputing Conference*, 2017.

C-11 K. Kanov and R. Burns. Particle tracking in open simulation laboratories. *Supercomputing*, 2015.

C-12 W. Gray Roncal et al. VESICLE: Volumetric Evaluation of Synaptic Inferfaces using Computer vision at Large Scale. *British Machine Vision Conference*, 2015.

C-13 Z. Liu, N. Ivkin, L. Yang, M. Neyrinck, R. Burns, G. Lemson, and A. Szalay. Streaming algorithms for Halo finders. *eScience*, 2015.

C-14 D. Zheng, D. Mhembere, H. Wang, C. Priebe, A. Szalay, J. Vogelstein, and R. Burns. FlashGraph: Processing Billion-Node Graphs on an Array of Commodity SSDs. *Conference on File and Storage Technologies* (FAST), USENIX, 2015.

C-15 K. Kanov, C. Lalescu, and R. Burns. Efficient evaluation of threshold queries of derived fields in a numerical simulation database. Extending Database Technology (EDBT), 2015.

C-16 D. Zheng, R. Burns, and A. S. Szalay. Toward Millions of File-System IOPS on Low-Cost, Commodity Hardware. *Supercomputing* (SC) IEEE, ACM 2013.

C-17 R. Burns, W. Gray Roncal, D. Kleissas, K. Lillaney, P. Manavalan, E. Perlman, D. Berger, D. D. Bock, K. Chung, K. Deisseroth, L. Grosenick, N. Kasthuri, M. Kazhdan, J. Lichtman, R. C. Reid, A. S. Szalay, J. T. Vogelstein, R. J. Vogelstein. The Open Connectome Project Data Cluster: Scalable Analysis and Vision for High-Throughput Neuroscience. *Scientific and Statistical Databases Management Conference* (SSDBM), 2013.

C-18 D. Crankshaw, B. Falck, A. S. Szalay, R. Burns, T. Budavári, and J. Wang. Inverted Indices for Particle Tracking in Petascale Cosmological Simulations *Scientific and Statistical Databases Management Conference* (SSDBM), 2013.

C-19 K. Kanov, R. Burns, G. Eyink, C. Meneveau, and A. Szalay. Data-Intensive Spatial Filtering in Large Numerical Simulation Datasets. *Supercomputing* (SC) IEEE, ACM 2012.

C-20 O. Khan, R. Burns, J. Plank, and C. Huang. Rethinking Erasure Codes for Cloud File Systems: Minimizing I/O for Recovery and Degraded Reads. *Conference on File and Storage Technologies* (FAST), USENIX, 2012.

C-21 K. Kanov, E. Perlman, R. Burns, Y. Ahmad, and R. Burns. I/O Streaming Evaluation of Batch Queries for Data-Intensive Computational Turbulence. *Supercomputing* (SC), IEEE, ACM, 2011.

C-22 X. Wang, A. Das, C. Olston, and R. Burns. CoScan: Cooperative Scan Sharing in the Cloud. *Symposium on Cloud Computing* (SOCC), ACM, 2011.

C-23 E. Givelberg, A. S. Szalay, K. Kanov, and R. Burns. MPI-DB, A Parallel Database Services Software Library for Scientific Computing. EuroMPI, 2011.

C-24 X. Wang, E. Perlman, R. Burns, T. Malik, T. Budavári, C. Meneveau, and A. Szalay. JAWS: Job-Aware Workload Scheduling for the Exploration of Turbulence Simulations. *Supercomputing* (SC), IEEE, ACM, 2010.

C-25 E. Perlman, R. Burns, M. Kazhdan, R. Murphy, and B. Ball. Organization of Data in Non-Convex Spatial Domains. *Scientific and Statistical Databases Management Conference* (SSDBM), 2010.

C-26 A. Batsakis, R. Burns, A. Kanevsky, J. Lentini, and T. Talpey. CA-NFS: A Congestion-Aware Network File System. (Best Paper Award at) *Conference on File and Storage Technologies* (FAST), USENIX, 2009.

C-27 X. Wang, R. Burns, and T. Malik. LifeRaft: Data-Driven, Batch Processing for the Exploration of Scientific Databases. *Conference on Innovative Data Systems Research* (CIDR), ACM, 2009.

C-28 M. Bolitho, M. Kazdhan, R. Burns, and H. Hoppe. Parallel Surface Reconstruction. *International Symposium on Visual Computing* (ISVC), 2009.

C-29 T. Malik, X. Wang, D. Dash, A. Chaudhary, R. Burns, and A. Ailamaki. Adaptive Physical Design for Curated Archives. *Scientific and Statistical Database Management Conference* (SSDBM), 2009.

C-30 A. S. Szalay, G. Bell, J. Vandenberg, R. Burns, D. Fay, J. Heasley, T. Hey, M. Nieto-Santisteban, A. Thakar, C. Van Ingen, R. Wilton, and A. Wonders. GrayWulf: Scalable Clustered Architecture for Data Intensive Computing. *Hawai'i International Conference on Systems Sciences* (HICSS), 2009.

C-31 T. Malik and R. Burns. Workload-Aware Histograms for Remote Applications. *International Conference on Data Warehousing and Knowledge Discovery* (DaWaK), 2008.

C-32 R. Curtmola, O. Khan, R. Burns, and G. Ateniese. MR-PDP: Multiple-Replica Provable Data Possession. *International Conference on Distributed Computing Systems* (ICDCS), IEEE, 2008.

C-33 A. Batsakis, R. Burns, A. Kanevsky, J. Lentini, and T. Talpey. AWOL: An Adaptive Write Optimizations Layer. *Conference on File and Storage Technologies* (FAST), USENIX, 2008.

C-34 X. Wang, R. Burns, A. Terzis, and A. Deshpande. Network-Aware Join Processing in Global-Scale Database Federations. *International Conference on Data Engineering* (ICDE), IEEE, 2008.

C-35 G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song. Provable Data Possession at Untrusted Stores. *Computer and Communication Security* (CCS), ACM, 2007.

C-36 E. Perlman, R. Burns, Y. Li, and C. Meneveau. Data Exploration of Turbulence Simulations using a Database Cluster. *Supercomputing* (SC), ACM, IEEE, 2007.

C-37 M. Bolitho, M. Kazhdan, R. Burns, and H. Hoppe. Multilevel Streaming for Out-of-Core Surface Reconstruction. *Eurographics Symposium on Geometry Processing* (SGP), ACM, 2007.

C-38 X. Wang, T. Malik, R. Burns, S. Papadomanolakis, and A. Ailamaki. A Workload-Driven Unit of Cache Replacement for Mid-Tier Database Caching. *Database Systems for Advanced Applications* (DASFAA), IEEE, 2007.

C-39 Z. Peterson, R. Burns, G. Ateniese, and S. Bono. Design and Implementation of Verifiable Audit Trails for a Versioning File System. *Conference on File and Storage Technologies* (FAST), USENIX, 2007.

C-40 T. Malik, R. Burns, and N. Chawla. A Black-Box Approach to Query Cardinality Estimation. *Conference on Innovative Data Systems Research* (CIDR), ACM, 2007.

C-41 T. Malik, R. Burns, N. Chawla, and A. Szalay. Estimating Query Result Sizes for Proxy Caching in Scientific Database Federations. *Supercomputing Conference* (SC), Best Student Paper Nominee, ACM, IEEE, 2006.

C-42 C. Wu and R. Burns. Improving the I/O Performance of Clustered Storage Systems with Adaptive Request Distribution. *High-Performance Distributed Computing* (HPDC), IEEE, 2006.

C-43 Z. Peterson, R. Burns, J. Herring, A. Stubblefield, and A. Rubin. Secure Deletion for a Versioning File System. *Conference on File and Storage Technologies* (FAST), USENIX, 2005.

C-44 A. Batsakis and R. Burns. Cluster Delegation: High-Performance, Fault-Tolerant Data Sharing in NFS. *High-Performance Distributed Computing* (HPDC), IEEE, 2005.

C-45 T. Malik, R. Burns, and A. Chaudhary. Bypass Caching: Making Scientific Databases Good Network Citizens. *International Conference on Data Engineering* (ICDE), 94-105, ACM, 2005.

C-46 R. Burns. Fastpath Optimizations for Cluster Recovery in Shared-Disk Systems. *Supercomputing Conference* (SC), Best Paper Nominee, ACM, IEEE, 2004.

C-47 C. Wu and R. Burns. Achieving Performance Consistency in Shared-Disk Clusters. *High-Performance Distributed Computing* (HPDC), 140-149, IEEE, 2004.

C-48 D. Rasch and R. Burns. In-Place Rsync: File Synchronization for Mobile and Wireless Devices. *USENIX Annual Technical Conference, FREENIX track*, 91-100, USENIX, 2003.

C-49 C. Wu and R. Burns. Handling Heterogeneity in Shared-Disk File Systems. *Supercomputing Conference* (SC), ACM, IEEE, 2003.

C-50 G. Whittle, J.-F. Paris, A. Amer, D. Long, and R. Burns. Using Multiple Predictors to Improve the Accuracy Using Multiple Predictors to Improve the Accuracy of File Access Predictions. *Mass Storage Conference*, 230-240, NASA, IEEE, 2003.

C-51 A. Amer, D. Long, and R. Burns. Group-Based Management of Distributed File Caches. *International Conference on Distributed Computing Systems* (ICDCS), 525-532, IEEE, 2002.

C-52 A. Amer, D. Long, J.-F. Paris, and R. Burns. File Access Prediction with Adjustable Accuracy. *International Performance, Computing and Communication Conference* (IPCCC), IEEE, 2002.

C-53 R. Burns and W. Hineman. A Bit-Parallel Search Algorithm for Allocating Free Space. *International Symposium on Modeling, Analysis, and Simulation in Computer and Telecommunication Systems* (MASCOTS), 302-310, IEEE, 2001.

C-54 R. Burns, R. Rees and D. Long. An Analytical Study of Opportunistic Lease Renewal. *International Conference on Distributed Computing Systems* (ICDCS), 146-153, IEEE, 2001.

C-55 R. Burns, R. Rees, and D. Long. Safe Caching in a Distributed File System for Network Attached Storage. *International Parallel and Distributed Processing Symposium* (IPDPS), 155-162, IEEE, 2000.

C-56 R. Burns, R. Rees, and D. Long. Semi-Preemptible Locks for a Distributed File System. *International Performance, Computing and Communication Conference* (IPCCC), IEEE, 2000.

C-57 R. Burns and D. Long. In-Place Reconstruction of Delta Compressed Files, (Extended Abstract). *Conference on the Principles of Distributed Computing* (PODC), 267-275, ACM, 1998.

C-58 R. Burns, and D. Long. Efficient Distributed Backup and Restore with Delta Compression. *I/O in Parallel and Distributed Systems* (IOPADS), 27-36, ACM, 1997.

C-59 R. Burns, and D. Long. A Linear Time, Constant Space Differencing Algorithm. *International Performance, Computing and Communications Conference* (IPCCC), IEEE, 1997.

Refereed Workshop and Short Papers

W-1 K. Lillaney, D. Pease, D. Tarasov, and R. Burns. The Case for Dual-access File Systems over Object Storage. Workshop on *Hot Topics in Storage* (HotStorage), USENIX, 2019.

W-2 H. Weng, D. Zheng, R. Burns, and C. Priebe. Active Community Detection in Massive Graphs. Workshop on Mining Networks and Graphs: A Big Data Analytic Challenge (SDM-Networks), SIAM, 2015.

W-3 D. Zhang, R. Burns, and A. Szalay. A Parallel Page Cache: IOPS and Caching for Multicore Systems. Workshop on *Hot Topics in Storage* (HotStorage), USENIX, 2012.

W-4 P. Stanton and R. Burns. APR-Quad: An Update Efficient Authenticated Dictionary for Spatial Data. Workshop on Security and Privacy in GIS and LBS, ACM, 2011.

W-5 E. Givelberg, A. S. Szalay, K. Kanov, and R. Burns. An Architecture for a Data-Intensive Computer. Workshop on Network-Aware Data Management, IEEE, ACM, 2011.

W-6 R. Burns. Forget Locality (The Radio Edit). Workshop on *High-Performance Transaction Processing Systems* (HPTS), 2011.

W-7 O. Khan, R. Burns, J. Plank, and C. Huang. In Search of I/O Optimal Recovery from Disk Failures. Workshop on *Hot Topics in Storage* (HotStorage), USENIX, 2011.

W-8 P. Stanton and R. Burns. I/O Efficient Search of Large Social Networks. *Network Science Workshop*, Network Science Center, United States Military Academy, 2010.

W-9 B. Hong, R. Curmola, G. Ateniese, and R. Burns. Remote Data Checking for Network Coding-based Distributed Storage Systems. *Cloud Computing Security Workshop*, ACM, 2010.

W-10 N. Walfield, P. Stanton, and R. Burns. Practical Protection for Personal Storage in the Cloud. *EUROSEC: European Workshop on System Security*, ACM, 2010.

W-11 P. Stanton, B. McKeown, R. Burns, G. Ateniese. FastAD: An Authenticated Directory for Billions of Objects. *Workshop on Hot Topics in Storage and File Systems* (HotStorage), ACM, USENIX, 2009.

W-12 S. Richardson, B. Olson, J. Dymond, R. Burns, S. Chandrasegaran, J. Boeke, A. Shehu and J. Bader. Automated Design of Assemblable, Modular, Synthetic Chromosomes. Workshop on Parallel Computational Biology, 2009.

W-13 E. Perlman, R. Burns, and M. Kazhdan. Organizing and Indexing Non-Convex Regions. Demonstration Paper in *International Conference on Very Large Data Bases* (VLDB), 2008.

W-14 R. Curtmola, O. Khan, and R. Burns. Robust Remote Data Checking. *Workshop on Storage Security and Survivability* (StorageSS), ACM, 2008.

W-15 R. Burns and G. Ateniese. Currency and Correctness of Content in Object Storage Networks. Whitepaper at the *DARPA Distributed Object Storage and Retrieval Workshop*, 2008.

W-16 X. Wang, T. Malik, R. Burns, D. Dash, and A. Ailamaki. Automated Physical Design in Database Caches. *Workshop on Self-Managing Database Systems* (SDBM), IEEE, 2008.

W-17 J. Gupchup, R. Burns, A. Terzis, and A. Szalay, Model-Based Event Detection in Wireless Sensor Networks. *Workshop on Data Sharing and Interoperability on the World-Wide Sensor Web* (DSI), ACM/IEEE, 2007.

W-18 X. Wang, R. Burns, and A. Terzis. Throughput-Optimized, Global-Scale Join Processing in Scientific Federations. *Workshop On Networking Meets Databases* (NetDB), USENIX, 2007.

W-19 A. Batsakis, R. Burns, T. Talpey, A. Kanevsky, and J. Lentini. Enhancing the Linux Memory Architecture to Support File Systems over Heterogeneous Devices. Position paper, *Linux File Systems and IO Workshop*, USENIX, 2007.

W-20 S. Ozer, A. Szalay, J. Gray, A. Terzis, R. Musaloiu-E., K. Szlavecz, J. Cogan, and R. Burns. Data Analysis Tools for Sensor-Based Science. Demonstration paper, *Conference on Embedded Networked Sensor Systems* (SenSys), ACM, 2006.

W-21 R. Burns, A. Terzis, and M. Franklin. Design Tools for Sensor-Based Science. *Workshop on Embedded Networked Sensors* (EmNetS), IEEE, 2006.

W-22 Z. Peterson and R. Burns. Building Regulatory-Compliant Storage Systems. Refereed project highlight, *Conference on Digital Government Research* (dg.o), 2006.

W-23 R. Burns, Z. Peterson, G. Ateniese, and S. Bono. Verifiable Audit Trails for a Versioning File Systems. *International Workshop on Storage Security and Survivability* (SSS), ACM, 2005.

W-24 Z. Peterson and R. Burns. Limiting Liability in a Federally Compliant File System. *PORTIA Workshop on Sensitive Data in Medical, Financial, and Content Distribution Systems*, 2004.

W-25 A. Batsakis and R. Burns. NFSv4 as the Building Block for Fault Tolerant Applications. *Workshop on NFS Extensions for Parallel Storage* (NEPS), 2003.

W-26 R. Burns, R. Rees, and D. Long. Consistency and Locking for Distributing Updates to Web Servers Using a File System. *Performance Evaluation Review*, 28(2), 15-21, Workshop on the Performance and Architecture of Web Servers (PAWS), ACM, 2000.

W-27 R. Burns and I. Narang. Version Management and Recoverability for Large Object Data. *Workshop on Multimedia Database Management Systems* (MMDBMS), 12-19, IEEE, 1998.

Invited Papers

I-1 Neuro Cloud Consortium. To The Cloud: A Grassroots Proposal to Accelerate Brain Science Discovery. *Neuron*, 2016. DOI:10.1016/j.neuron.2016.10.033

I-2 D. Mhembere. W. G. Roncal, D. Sussman, C. E. Priebe, R. Jung, S. Ryman, R. J. Vogelstein, J. T. Vogeltstein, and R. Burns. Computing Scalabal Multiavriate Global Invariants of Large (Brain-) Graphs. IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2013.

I-3 W. Gray Roncal, Z. H. Koterba, D. Mhemebere, D. M. Kleissas, J. T. Vogelstein, R. Burns, A. R. Bowles, D .K. Donavos, S. Ryman, R. E. Jung, L. Wu, V. Calhoun, and R. J. Vogelstein. MIGRAINE: MRI Graph Reliability Analysis and Inference for Connectomics. IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2013.

I-4 Y. Ahmad, R. Burns, M. Kazhdan, A. Szalay, and A. Terzis. Scientific Data Management at the Johns Hopkins Institute for Data Intensive Engineering and Science. *SIGMOD Record*, 39(3), 18–23, 2010.

I-5 J. Gupchup, A. Sharma, A. Terzis, R. Burns, and A. Szalay. The Perils of Detecting Measurement Faults in Environmental Monitoring Networks. *International Conference on Distributed Computing in Sensor Systems* (DCoSS), IEEE, 2008.

I-6 R. Burns, L. Stockmeyer, and D. Long. Experimentally Evaluating In-Place Delta Reconstruction. *IEEE Mass Storage Conference*, NASA, IEEE, 2002.

I-7 B. Reed, E. Chron, D. Long, and R. Burns. Authenticating Network Attached Storage. *IEEE Symposium on High Performance Interconnects*, IEEE, 1999.

Other

O-1 J. T. Vogelstein et al. (35 authors). To the cloud! A grassroots proposal to accelerate brain science discovery. *Neuron* 92(3), 2016.

O-2 A. S. Szalay et al. (23 authors). GrayWulf: Scalable Clustered Architecture for Data Intensive Computing. *Supercomputing*, Winner of the HPC Storage Challenge, 2008.

O-3 A. S. Szalay, G. Bell, J. Vandenberg, A. Wonders, R. Burns, D. Fay, J. Heasley, T. Hey, M. Nieto-SantiSteban, A. Thakar, C. van Ingen, and R. Wilton. GrayWulf: Scalable Clustered Architecture for Data Intensive Computing. Microsoft Research Technical Report MSR-TR-2008-187, 15 September 2008.

O-4 R. Burns. Data Management in a Distributed File System for Storage Area Networks. *Ph.D. Dissertation*, University of California, Santa Cruz, 2000.

O-5 R. Burns. Differential Compression: A Generalized Solution for Binary Files. *Masters Thesis*, University of California, Santa Cruz, 1997.

Patents

P-1 Producer/consumer locking system for efficient replication of file data (with A. Goal, W. Hineman, and R. Rees), US6925515, 02 August 2005.

P-2 System for dynamically evaluating locks in a distributed data storage system (with R. Rees), US6917942, 12 July 2005.

P-3 System and method for managing authentication and coherency in a storage area network (with D. Long), US6792424, 14 September 2004.

P-4 Lease based safety protocol for distributed system with multiple networks (with A. Goel, D. Long, and R. Rees) US6775703, 10 August 2004.

P-5 Data placement and allocation using virtual contiguity (with D. Long and R. Rees) US6651147, 18 November 2003.

P-6 System for managing asset access in a distributed storage system (with A. Goel and R. Rees), US6571276, 27 May 2003.

P-7 System and method for allocating storage space using bit-parallel search of bitmap (with W. Hineman), US6510505, 21 January 2003.

P-8 Decentralized remotely encrypted file system (with E. Chron, D. Long and B. Reed), US6405315, 11 June 2002.

P-9 System and method for differential compression of data from a plurality of sources (with M. Ajtai, R. Fagin, D. Long, and L. Stockmeyer), US06374250, 14 April 2002.

P-10 Continuous availability and efficient backup for externally referenced files (with I. Narang), US06088694, 7 October 2000.

P-11 Method for generating and reconstructing in-place delta files (with D. Long), US06018747, 25 January 2000.

P-12 Secure array of remotely encrypted storage devices (with E. Chron, D. Long, and B. Reed), US05931847, 3 August 1999.

Contracts and Grants

National Science Foundation. "Frameworks: Advanced Cyberinfrastructure for Sustainable Community Usage of Big Data from Numerical Fluid Dynamics Simulations." Co-PI with PI C. Meneveau, A. Szalay, T. Haine, and T. Zaki. NSF OAC-02103874, $3,992,109, 7/1/2021–6/30/2026.

National Science Foundation. "NeuroNex Innovation Award: Towards Automatic Analysis of Multi-Terabyte Cleared Brains." Co-PI with PI J. T. Vogelstein, F. Engert, C. Priebe, and M. Miller. NSF IOS-1707298, $800,000, 9/1/2018–8/31/2020.

DARPA. "Lifelong Learning Forests." Co-PI with PI J. T. Vogelstein HR001117S0016-L2M-FP-054,$1,846,041, 02/01/2018–01/30/2020.

National Institutes of Health. "Sensorimotor processing, decision making, and internal states: towards a realistic multiscale circuit model of the larval zebrafish brain." Co-PI with PI F. Engert and J. T. Vogelstein. $952,570, 9/01/170–8/31/20.

National Science Foundation. "Computational Infrastructure for Brain Research: EAGER: BrainLab CI: Collaborative, Community Experiments with Data-Quality Controls through Continuous Integration." PI with M. I. Miller and J. T. Vogelstein. ACI-1649880, $294,599, 01/01/17–12/31/19.

National Science Foundation. " BIGDATA: IA: Democratizing Massive Fluid Flow Simulations via Open Numerical Laboratories and Applications to Turbulent Flow and Geophysical Modeling" Co-PI with PI C. Meneveau, A. Szalay, G. Eyink, and T. Haine. OCE-1633124, $952,570, 11/1/2016–10/31/2019.

DOE Los Alamos National Laboratories. "Advance in-situ analysis capabilities of numerical simulations", PI. 89049-001-11/329030, $139,147, 1/1/2015–5/13/2016.

DARPA. "From RAGs to Riches: Utilizing Richly Attributed Graphs to Reason from Heterogeneous Data." Co-PI with J. Vogelstein (PI), and C. Priebe. DARPA N66001-15-C-4041, $1,975,863, 03/01/15–05/01/18.

National Institutes of Health. "Synaptomes of Mouse and Man", PI with S. Smith (PI), E. Lein, W. Seeley, J. Trimmer, W. Weinberg, J. Vogelstein, M. Chevillet, and G. Sapiro. NIH/NINDS 1R01NS092474-01, $1,311,966 to JHU (of ∼$8.7M), 09/30/2014–06/30/2019.

National Institutes of Health. "Neural circuits in zebrafish: form, function and plasticity", Co-PI with F. Engert (PI) and J. Lichtman. NIH/NINDS 1U01NS090449-01, $1,736,376, 09/30/2014–07/31/2017.

National Science Foundation. "BIGDATA: F: DKA: Collaborative Research: Clustering Algorithms for Data Streams." Co-PI with PI V. Braverman, A. Szalay, T. Budavari, and B. Van Durme. IIS-1447639, $1,000,000, 9/1/2014–8/31/2017.

DARPA. "Scalable Brain Graph Analyses using Big-Memory, High-IOPS Compute Architectures." PI with J. Vogelstein, C. Priebe, and A. Szalay. DARPA N66001-14-1-4028, $435,226, 05/01/2014–11/31/2015.

National Science Foundation. "CIF21 DIBBs: Long Term Access to Large Scientific Data Sets: The SkyServer and Beyond." Co-PI with PI A. Szalay, S. Salzberg, A. Thakar, and C. Meneveau. ACI-1261715, $7,603,720, 10/1/2013–09/30/2018.

National Institutes of Health. "BIGDATA: Small DCM: ESCA DA Computational infrastructure for massive neuroscience image stacks." co-PI with P. Mitra, J. T. Vogelstein, and R. J. Vogelstein. NIDA 1R01DA036400-01, $907,194, 03/15/2013–03/14/2016.

National Institutes of Health. "CRCNS Data Sharing: The EM Open Connectome Project." PI with G. Hager, M. Kazhdan, J. Lichtman (Harvard), J. T. Vogelstein, and R. J. Vogelstein. NIH/NIBIB 1RO1EB016411-01, $830,128, 09/01/2012–08/31/2015.

National Science Foundation. "Interdisciplinary Scientific Data Management." co-PI with PI A. Szalay, C. Meneveau, K. Svalecz, and A. Terzis. OCI-1244820. $1,000,000, 10/01/2012–09/31/2014

National Science Foundation. Amendment to AST-0428325 "Exploring the Lagrangian Structure of Complex Flows with 100 Terabyte Datasets." co-PI with PI A. Szalay, G. Eyink, and S.-Y. Chen. AST-0939767, $432,478, 9/1/2009–8/31/2012.

National Science Foundation *Cyber-Enabled Discovery and Innovation* (CDI) Type-II. Database enabled multiscale simulations and analysis of fluid turbulence. Co-PI with PI C. Meneveau, S. Chen, G. Eyink, and A. Szalay. CMMI-0941530, $1,899,469, 09/01/2009–08/31/2013.

National Science Foundation. *High-End Computing University Research Activity* (HECURA). "CRAM: A Congestion-Aware Resource and Allocation Manager for Data-Intensive High-Performance Computing." PI with J. Griffin. CCF-0937810, $495,000, 12/15/09–12/14/12.

National Science Foundation. *Major Research Instrumentation* (MRI). *Acquisition of 100TF Graphics Processor Laboratory for Multiscale/Multiphysics Modeling*. Co-PI with PI M. Robbins, L. Graham-Brady, A. Szalay, and R. Dalrymple. CMMI-0923018, $381,518, 08/01/09–07/31/12.

National Science Foundation. *Sustainable Digital Data Preservation and Access Network Partner (DataNet)*. "Data Conservancy: A Digital Research and Curation Virtual Organization." NSF OCI-0830876, Co-PI with PI S. Choudhury and 12 others. $20M, 09/04/09–09/03/14.

Johns Hopkins University. *Provost's Discovery Grant.* "Data Intensive Scalable Computing at JHU (JH-DISC)." Co-PI with PI A. Szalay and 6 others. $505,914, 11/2008–10/2011.

National Science Foundation. *National Science, Mathematics, Engineering and Technology Education Digital Library* (NSDL), NSF 07-538. "Archival Introspection and Maintenance Metadata." PI with S. Choudhury, T. DiLauro, and J. L. Griffin, DUE-0734862, $99,075, 12/15/2007–11/30/2008.

National Science Foundation. *Cyberinfrastructure for Environmental Observatories: Prototype Systems to Address Cross-Cutting Needs* (CEO:P), NSF 06-505. "A Prototype System for Multi-Disciplinary Shared Cyberinfrastructure—Chesapeake Bay Environmental Observatory (CBEO)." co-PI with William P. Ball, Dominic M. Di Toro, Thomas Gross (PI), William M. Kemp, Michael Piasecki, ATM-0618986, $2,149,906, 10/1/2006–9/31/2009.

Library of Congress (administered by the National Science Foundation). *Digital Archives*, NSF 04-592. "Securely Managing the Lifetime of Versions in Digital Archives." PI with A. Rubin and G. Ateniese, IIS-0456027, $300,000, 7/1/2005–6/31/2008.

National Science Foundation. *Information Technology Research*, NSF 04-12. "Exploring the Lagrangian Structure of Complex Flows with 100 Terabyte Datasets." co-PI with E. Vishniac (PI), G. Eyink, S.-Y. Chen, and A. Szalay, AST-0428325, $2,165,860, 9/1/2004–8/31/2009.

National Science Foundation. *Science and Engineering Information Integration and Informatics*, NSF 04-528. "Bypass-Yield Caching for Large-Scale Scientific Database Workloads in the World-Wide Telescope." PI with A. Thakar, IIS-0430848, $631,863, 10/01/2004–9/30/2007. Linked award: IIS-0431008, PI Anastassia Ailamaki, Carnegie Mellon University, $298,000.

National Science Foundation. *Concept Development Toward a Collaborative Large-Scale Engineering Analysis Network (CLEANER)*, NSF 03-607. "CLEANER with focus on the Chesapeake Bay." co-PI with Y. Amir and B. Ball, BES-0414372, $22,783, 8/1/2004–7/31/2005. Linked awards: BES-0414429, PI Dominic Di Toro, University of Delaware, $9,999. BES-0414347, PI William Kemp, University of Maryland Center for Environmental Sciences, $5,245. BES-0414214, PI Thomas Gross, Chesapeake Research Consortium, $31,932.

National Science Foundation. *Faculty Early Career Development (CAREER) Program*, NSF 02-111. Advanced Computer Research Program, Division of Advanced Computer Infrastructure. "CAREER: Interoperation among Heterogeneous Global-Scale Storage Systems." PI, CCF-0238305, $406,973, 5/1/2003–4/30/2008.

Department of Energy. *Early Career Investigator Program*, Notice 02-16. "Global-Scale File Sharing for High-Performance Cluster Computing." PI, DE-FG02-02ER25524, $318,227, 4/15/2002–4/14/2005.

Other Grant Activity

DARPA, Phase II, Small Business Innovative Research (SBIR) D2-2060, "Tools for Sharing and Analyzing Neuroscience Data" $1M, 06/18–12/20.

Allen Institute for Brain Science. "Processing and Visualization of Array Tomography Data in the Cloud." $104,625, 12/1/2017–6/30/18

DARPA D17PC00174, Phase I and Option, Small Business Innovative Research (SBIR) SB163-001, "Tools for Sharing and Analyzing Neuroscience Data" $149,731, 05/08/17–04/20/2017.

Subcontract for IARPA Machine Intelligence from Cortical Networks (MICrONS) project. Johns Hopkins University Applied Physics Labs. $150,0000, 09/20/2016–03/31/2018.

Subcontract of "National Center for Applied Neuroscience (NCAN)." Johns Hopkins University Applied Physics Labs, PI J. Vogelstein. $34,912, 06/01/2013–09/30/2013.

Subcontract of "National Center for Applied Neuroscience (NCAN)." Johns Hopkins University Applied Physics Labs, PI J. Vogelstein. $34,912, 06/01/2013–09/30/2013.

Child Mind Institute Endeavour Training Fellowship. Mentor for Awardee Joshua Vogelstein. $75,000. 02/11/2014–02/11/2015.

National Institutes of Health. "CLARITY: Fully-Assembled Biology" Consultant to PI K. Deisseroth (Stanford). NIMH 1R01MH09964701 , $150,000 of $5.65M, 08/31/2012–08/31/2017.

Subcontract of "Reverse Engineeering Neural Microcircuits." Internal Research and Development Grant, Johns Hopkins University Applied Physics Labs, PI J. Vogelstein. $19,304.

National Science Foundation. CI-Fellows Sub-Award for Ragib Hasan. PI R. Hasan and R. Burns (mentor), $140,000, 10/02/09–10/01/10.

National Science Foundation *Major Research Instrumentation* (MRI). "Development of Data-Scope—A Multi-Petabyte Generic Data Analysis Environment for Science." Senior personnel with PI A. Szalay, OCI-1040114, $2,087,760. 01/01/2011–12/31/2012.

Corporate Grants and Gifts

IBM Corporation. "ObjectFS: An Open-Source File Systems for Cloud Objects." $25,000, 9/15/2017.

Amazon Web Services. Education grants for AWS cloud compute resources.

- *Parallel Programming*, $12500, February 2015.
- *Parallel Programming*, $10500, February 2014.
- *Data Intensive Computing*, $3200, September 2013.
- *Parallel Programming*, $9800, February 2013.
- *Data Intensive Computing*, $1600, September 2012.
- *Parallel Programming*, $5400, February 2012.
- *Parallel Programming*, $7500, September 2011.

nVidia CUDA Center of Excellence at Johns Hopkins University, Co-PI with PI Szalay, August 2011.

NetApp. "Reducing Memory and I/O Interference for Virtualized Systems and Cloud Computing." Research Gift. Randal Burns, $40,000, 8/30/2010.

NetApp. "Server-Directed Adaptive Data Destaging for NFS (on High-Performance Interconnects)." Research Gift. Randal Burns, $45,000, 11/13/2008.

Network Appliance. " Optimizing Client-Server Data Transfer through Memory Management." Research Gift. Randal Burns, $29,192, 10/19/2006.

International Business Machines. *Shared University Research* (SUR) program. Randal Burns, $362,000 (Equipment), 2002.

Program Committees

2022	Conference on File and Storage Technologies (FAST), USENIX
2021	Symposium on Cloud Computing (SOCC), ACM
	Interntaional Conference on Big Data (BigData), IEEE
2020	Conference on File and Storage Technologies (FAST), USENIX
	Systems and Storage Conference (Systor), ACM
2019	USENIX Annual Technical Conference (ATC), USENIX
	Principles and Practices of Parallel Programming (PPoPP), ACM
2016	USENIX Annual Technical Conference (ATC), USENIX
	Conference on Networking, Architecture, and Storage (NAS), IEEE
2015	Conference on File and Storage Technologies (FAST), USENIX
	Workshop on Hot Topics in Storage Systems (HotStorage), USENIX

Conference on Innovative Data Systems Research (CIDR), ACM

2014 *Program Chair*, Systems and Storage Conference (Systor), ACM
Programmable File Systems Workshop (PFSW), ACM

2013 Conference on Very Large Data Bases (VLDB), ACM
Conference on Innovative Data Systems Research (CIDR), ACM
Scientific and Statistical Database Management Conference (SSDBM), LNCS
Symposium on Cloud Computing (SOCC), ACM, IEEE, USENIX
Conference on Mass Storage Systems and Technology (MSST), IEEE
Parallel Data Storage Workshop (PDSW), ACM,IEEE

2012 Conference on File and Storage Technologies (FAST), USENIX
Conference on Very Large Data Bases (VLDB), ACM
International Conference on Data Engineering (ICDE), IEEE
Scientific and Statistical Database Management Conference (SSDBM), LNCS
International Systems and Storage Conferfence (SYSTOR), ACM, IEEE, USENIX
Workshop on Hot Topics in Storage Systems (HotStorage), USENIX
Workshop on Management of Big Data Systems (MBDS), IEEE
Conference on Mass Storage Systems and Technology (MSST), IEEE

2011 Conference on Very Large Data Bases (VLDB), ACM
Conference on Innovative Data Systems Research (CIDR), ACM
Scientific and Statistical Database Management Conference (SSDBM), LNCS
International Systems and Storage Conferfence (SYSTOR), ACM, IEEE, USENIX
International Workshop on Storage Network Architecture and Parallel I/O (SNAPI), IEEE

2010 *Program Chair*, Conference File and Storage Technologies (FAST), USENIX
Storage Challenge Committee, Supercomputing (SC), ACM, IEEE

2009 Conference on Innovative Data Systems Research (CIDR), ACM
Scientific and Statistical Database Management Conference (SSDBM), LNCS

2008 Workshop on Computing with Massive and Persistent Data (CMPD), IEEE
File System and Storage Technology Conference (FAST), USENIX
International Workshop on Storage Security and Survivability (StorageSS), ACM

2007 File System and Storage Technology Conference (FAST), USENIX
Conference on Mass Storage Systems and Technology (MSST), IEEE

2006 Conference on Mass Storage Systems and Technology (MSST), IEEE

2005 Supercomputing Conference (SC), ACM, IEEE
Security in Storage Workshop, IEEE
Conference on Mass Storage Systems and Technology (MSST), IEEE

2004 Conference on Mass Storage Systems and Technology (MSST), IEEE

2003 Security in Storage Workshop, IEEE
Conference on Mass Storage Systems and Technology (MSST), IEEE

2002 File System and Storage Technology Conference (FAST), USENIX
Security in Storage Workshop, IEEE
Conference on Mass Storage Systems and Technology (MSST), IEEE

2001 International Conference on Distributed Computing Systems (ICDCS), IEEE
Conference on Mass Storage Systems and Technology (MSST), IEEE

External Professional Activities

Member, DARPA Information Science and Technology advisory group, 2020–present.
Visiting Professor, National Space Researh Institute, Brazil, 2020.
Nominee and Participant, CCC Leadership in Science Policy Institute (LiSPI) Workshop, Nov. 2019.
Participant, National Academies Workshop on Neuroscience Data in the Cloud, Oct. 2019.
Panelist and Site Review, NIH Other Transactions, 2017.
Participant, Chan Zuckerberg Science Workshop: Computational Tools for Microscopy, Feb. 2017.
Participant, GE Kavli Brain Trust VI: Building an International Brain Station! October, 2016.
Participant, Kavli/NSF Global Brain Workshop, 2016.
Organizing Committee, Kavli Indo-American Frontiers of Science Symposium, National Academies of Science, 2015.
Big Data Session Chair, Chinese-American Kavli Frontiers of Science Symposium, National Academies of Science, 2014.
Defense Science Study Group, Class of 2012–2013.
Program Chair, International Systems and Storage Conference (SysStor), ACM, 2014.
Steering Committee, USENIX Conference on File and Storage Technologies, 2010–2017.
Editorial Board, ACM Transactions on Storage, 2004–2012.
Program Chair, Conference on File and Storage Technologies, USENIX, 2010.
Agency for Science Technology and Research (A*STAR) Visitor, Singapore, September, 2010.
Academic Advisory Board, Association for Computing Machinery, Washington DC Chapter, 2008–2011.
Academic Advisory Board, StorageNetworking.org, 2003–2008.
Participant, Institute for Information Infrastructure Protection Forum on *Technology for Cyber-Physical Systems*, Committee on Homeland Security and Governmental Affairs, United States Senate, *Chairman:* Joseph Lieberman. 2008.
Panel Organizer and Moderator. "Scientific Data Management: An Orphan in the Database Community?" International Conference on Data Engineering (ICDE), IEEE, 2008.
Vice Program Chair, International Performance, Communications, and Computing Conference, IEEE, 2002.
Moderator of Usenet newsgroup for operating systems research, *comp.os.research*, 1996-2002.

Internal Professional Activities

Non-Tenure Track Faculty Committee, 2020-2021.
Whiting School of Engineering, PhD Fellowship Committee, 2020-2021.
ECE Head Search Committee, 2020.
Conflict Review Committee, Whiting School of Engineering, 2019–present.
Steering Committee, Kavli Neuroscience Discovery Institute, 2015–present.
Steering Committee, Science of Learning Institute, 2012–2017.
Oversight Committee for the Moore Foundation Integrated ICU Project, 2012–2014.
Leveraging Data to Knowledge Strategic Committee, Whiting School of Engineering, 2012.
Steering Committee, Insitute for Data Intensive Science and Engineering (IDIES), 2009–2014.
Homewood High Performance Compute Cluster Committee, 2008-2016.
Chair, Faculty Search Committee, Department of Computer Science, Johns Hopkins University, 2008-2013.
High Dimensional Data and Intensive Computing Campaign Group, 2011.
eScience Task Force, 2007.
WSE/KSAS Academic Computing Advisory Committee, 2005–2013.
Director of IT, Dept. of Computer Science, Johns Hopkins University, 2004-present.
Committee Member, Dept. of Computer Engineering, Johns Hopkins University, 2004-present.
Curriculum Committee, Dept. of Computer Science, Johns Hopkins University, 2004-2012.
Faculty Search Committee, Johns Hopkins University, 2002-2004, 2006-2007.
Library Advisory Committee, IBM Research, 1999-2001.
Liaison, Sponsored University Research (SUR) Grant, IBM Corp., 1999-2001.
Organizer, Storage Software Seminar Series, IBM Almaden Research Center, 2000-2001.

Consulting

2016–2017 **Los Alamos National Laboratories**, Data Science at Scale Team, Technical Consultant.

2012–2017 **Stanford University**, Deisseroth Lab, Consultant for Data Management of CLARITY Neuroimaging Data.

2012-2013 **Axiom Law**, Technology and Analysis Consultant.

2006-2007 **Sterne, Kessler, Goldstein, and Fox**, Intellectual Property: Expert Witness.

2002-2003 **Keysec Corp.**, Bethesda, MD. Technical Consultant, Storage Architecture and Security.

2002 **Jenner and Block**, Intellectual Property: Expert Witness.

2001-2002 **Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo**, Intellectual Property: Expert Witness.

Course Instruction

Parallel Programming, Johns Hopkins University
- Fall 2021 (82 enrolled: 39 undergraduate and 43 graduate)
- Fall 2020 (87 enrolled: 49 undergraduate and 38 graduate)
- Spring 2020 (55 enrolled: 37 undergraduate and 18 graduate)
- Spring 2019 (78 enrolled: 46 undergraduate and 32 graduate)
- Spring 2018 (148 enrolled: 78 undergraduate and 70 graduate)
- Spring 2017 (88 enrolled: 25 undergraduate and 63 graduate)
- Fall 2017 (91 enrolled: 27 undergraduate and 64 graduate)
- Spring 2016 (80 enrolled: 30 undergraduate and 50 graduate)
- Spring 2015 (115 enrolled: 46 undergraduate and 69 graduate)
- Spring 2014 (86 enrolled: 33 undergraduate and 53 graduate)
- Spring 2013 (91 enrolled: 33 undergraduate and 58 graduate)
- Spring 2012 (45 enrolled: 22 undergraduate and 23 graduate)
- Fall 2011 (69 enrolled: 25 undergraduate and 44 graduate)
- Fall 2010 (78 enrolled: 34 undergraduate and 44 graduate)
- Fall 2009 (35 enrolled: 9 undergraduate and 26 graduate)
- Fall 2008 (69 enrolled: 18 undergraduate and 51 graduate)

Data-Intensive Computing, Johns Hopkins University
- Fall 2013 (23 enrolled)
- Fall 2012 (13 enrolled)

Advanced Topics in Data-Intensive Computing, Johns Hopkins University
- Fall 2018 (21 enrolled)
- Fall 2017 (17 enrolled)
- Fall 2014 (12 enrolled)

Data Organization: Storage and External Memory Systems, Johns Hopkins University
- Spring 2011 (18 enrolled)

External Memory Algorithms and Data Structures, Johns Hopkins University
- Spring 2010 (8 enrolled)

Storage Systems, Johns Hopkins University
- Fall 2007 (22 enrolled: 6 undergraduate and 16 graduate)
- Spring 2006 (41 enrolled: 17 undergraduate and 24 graduate)
- Fall 2004 (42 enrolled: 23 undergraduate and 19 graduate)
- Fall 2002 (56 enrolled: 30 undergraduate and 26 graduate)

Transaction Processing Systems, Johns Hopkins University
- Spring 2009 (11 enrolled: 1 undergraduate and 10 graduate)
- Spring 2007 (18 enrolled: 7 undergraduate and 11 graduate)
- Spring 2005 (21 enrolled: 6 undergraduate and 15 graduate)
- Fall 2003 (21 enrolled: 8 undergraduate and 13 graduate)

Distributed Database Systems, Johns Hopkins University
Spring 2002 (22 enrolled: 6 undergraduate and 16 graduate)
Intermediate Programming, Johns Hopkins University
Spring 2005 (40 enrolled)
Spring 2004 (48 enrolled)
Advanced Storage and Transaction Processing, Johns Hopkins University
Fall 2006 (6 enrolled)
Spring 2003 (11 enrolled)
Database Systems, University of California, Santa Cruz
Winter 2001 (80 enrolled)
Advanced Operating Systems, University of California, Santa Cruz
Spring 2000 (14 enrolled)

Other Course Instruction (Seminars, Lecture Series, Co-Teaching)

Parallel Programming for Data-Science, National Space Research Insitute, Brazil.
Short course (8 hours) offered to 40 students at the graduate level.
Digital Preservation, Johns Hopkins University
Spring 2008 (8 enrolled, 4 lectures, 4 panels)

Keynote and Distinguished Lectures

July 2021 "External Memory Systems for Data Science and Machine Learning." *Distinguished Lecture*, IBM Research—Storage Virtual Speaker Series.

Aug. 2013 "Data-Intensive Computing for Neuroscience: The Open Connectome Project." INCF NeuroInformatics Congress.

Apr. 2013 "Big Data Computing for Connectomics: Spatial Databases, Scalable Analytics, HPC, NoSQL, Cloud, and Much Much More." Stanford Cracking the Neural Code (CNC) Annual Symposium.

Invited Lectures

Jan. 2020 "External Memory Systems for Data Science and Machine Learning", Data Science Seminar, National Space Research Institute (INPE), Brazil. *Host:* Rafael Santos.

Jan. 2020 "Johns Hopkins Turbulence Databases (JHTDB): Immersive Numerical Laboratories", Centro de Previsão do Tempo e Estudos Climáticos, Brazil. *Host:* Enver Ramirez.

Jan. 2020 "Reproducible and Shareable Data Science in Distributed Clouds", Data Science Seminar, National Space Research Institute (INPE), Brazil. *Host:* Rafael Santos.

Oct. 2017 "Embracing Open Source in Big Data Neuroscience", Kavli Frontiers Symposium on Neurotechnology.

Mar. 2017 "Data Intensive Applications: The OpenConnectome Project", Keystone Symposium on Molecular and Cellular Biology, X2 Connectomics.

Nov. 2016 "NeuroData for KNDI: Storage and Analytics Services", Kavli Neuroscience Discovery Institute Kickoff meeting.

Oct. 2016 "Storing and Accessing Untrusted Storage: Big Data Security in the Cloud and Networked Systems", JASON Fall meeting, 2016.

July 2016 "BrainLabCI: Collaborative, Community Experiments with Data Quality Controls through Continuous Integration", NSF Workshop on Open Data Ecosystems for Neuroscience.

Nov. 2015 "NeuroDataViz: Cloud, Caching, and Mobile Vizualization of Big Brain Data", Data Science at Scale Seminar, Los Alamos National Lab, *Host:* David Rodgers.

Apr. 2015 "Neuro-Visualization at the Network Edge", CodeNeuro New York City, *Host:* Jeremy Freeman.

Apr. 2015 "Analytics Infrastructure for Richly-Attributed Brain Graphs", DARPA SIMPLEX Kickoff Meeting, *Host:* Reza Ghanadan.

Mar. 2015 "The Case of Scaleup Graph Analytics", DARPA GRAPHS PI Meeting, *Host:* Reza Ghanadan.

Oct. 2014 "The Hazards (Perceived and Real) of Big Data", KAVLI Frontiers of Science, Chinese-American Symposium.

May 2014 "FlashGraph: Processing Billion Node Graphs on an Array of SSDs." DARPA GRAPHS JHU Kickoff Meeting. *Host:* Reza Ghanadan.

Nov. 2013 "Open-Science, Data-Intensive Computing Architectures foor Vision and Analysis of Brain Data." DARPA Collaborative Brain Data Exploitation Workshop, *Host:* Reza Ghanadan.

July 2013 "The Open Connectome Project: A Big Data Architecture for the BRAIN Initiative." Information Science and Technology Seminar, Los Alamos National Lab, *Host:* Jim Ahrens.

Feb 2013 "BIG Data Computing for MICrONS: Storage Management and Analysis Engines." IARPA Machine Intelligence from Cortical Networks (MICrONs) Workshop. *Host: Jacob Vogelstein*.

Feb. 2012 High-Throughput Data-Intensive Computing: Shared Scans in Scientific Databases and the Cloud, Computer Science Seminar, George Washington University, *Host:* Howie Huang.

Oct. 2012 "EM Open Connectome Project Volume Databases." Workshop on Scaling Up EM Connectomics, HHMI Janelia Farm Research Campus, *Host:* Davi Bock.

Oct. 2012 "EM Open Connectome Project: Overview and Architecture." Lichtmab Lab, Neuroscience Seminar, Harvard, *Host:* Jeff Lichtman.

June 2011 High-Throughput Data-Intensive Computing: Shared Scans in Scientific Databases and the Cloud, Seminar, Microsoft Research, *Host:* Cheng Huang.

Mar. 2011 "The Open Connectome Project", Data-Intensive Research: Statistical Databases Workshop, Johns Hopkins University, *Host:* Alex Szalay.

Feb. 2011 "Democratizing Data Intensive Computing", Applied Physics Laboratory, Johns Hopkins University, *Host:* Jacob Vogelstein.

Sep. 2010 "Congestion-Aware Resource Management," Data Storage Insitute, Agency for Science Technology and Research (A*STAR), Singapore, *Host:* Khin Mi Mi Aung.

Sep. 2010 "NFS-CD: Write-Enabled Cooperative Caching in NFS," Data Storage Insitute, Agency for Science Technology and Research (A*STAR), Singapore, *Host:* Khin Mi Mi Aung.

Sep. 2010 "Engineering the Billion-Object Authenticated Dictionary," Data Storage Insitute, Agency for Science Technology and Research (A*STAR), Singapore, *Host:* Khin Mi Mi Aung.

Apr. 2010 "Engineering the Billion-Object Authenticated Dictionary," Computer Science Seminar, Brown University, *Host:* Ugur Cetintemel.

Oct. 2009 "TurbulenceDB: A Data-Intensive Architecture for the Analysis of Multiscale Fluid Simulations", Los Alamos Computer Science Symposium, *Host:* Francis Alexander.

Sept. 2009 "Engineering the Billion-Object Authenticated Directory," Database Seminar, University of Waterloo, *Host:* Ken Salem.

Aug. 2009 "CRAM: A Congestion-Aware Resource and Allocation Manager for Data-Intensive High-Performance Computing," Report on NSF HECURA funded project at the *High-End Computing File System and I/O Workshop* (HEC FSIO).

Sept. 2008 "Remote Data Checking: Auditing the Preservation Status of Massive Digital Data Sets on Untrusted Stores," Mass Storage Technology Conference, *Host:* Jean-Jacques Bedet.

July 2008 "Remote Data Checking: Toward Verifying Remote Computation," DARPA Distributed Object Storage and Retrieval Workshop, *Host:* Joshua Alspector.

June 2008 "The *Store Everything* Model for High-Performance Computing," JASON Summer Study, *Host:* Dan Meiron.

Mar. 2008 "Querying the Cosmos across the Globe: A Reduction in Scale?" Purdue University, Computer Science Seminar, *Host:* Ahmed Elmagarmid.

Feb. 2008 "Querying the Cosmos across the Globe: A Reduction in Scale?" University of California, Santa Cruz, Computer Science Seminar. *Host:* Neokolis Polyzotis.

Oct. 2007 "Auditing Long-Term Archives Built on Untrusted Storage Systems," University of Massachusetts, Computer Science Seminar. *Host:* Gerome Miklau.

Sept. 2007 "Network-Aware Join Processing in Global-Scale Database Federations," University College Cork, Computer Science Seminar. *Host:* Cormac Sreenan.

July 2007 "Securely Auditing the Contents of Untrusted Storage Systems," Network Appliance, Seminar. *Host:* James Lentini.

Nov. 2006 "Caching and Result-Size Estimation for Scientific Database Federations," University of California Berkeley, Database Group Seminar, *Host:* Joe Hellerstein.

June 2006 "Design Tools for Sensor-Based Science," Baltimore Ecosystem Study, Quarterly Research Meeting, University of Maryland, Baltimore County, *Host:* Claire Welty

May 2006 "Securely Implementing Regulatory Policy in Versioning File Systems," Department of Computer Science, University of Maryland, *Host:* Peter Keleher.

Feb. 2006 "Bypass-Yield Caching: Making Scientific Databases Good Network Citizens," Computer Science Seminar, University of Notre Dame, *Host:* Douglas Thain.

Dec. 2005 "Securely Implementing Regulatory Policy," Library of Congress, *Host:* Bill Lefurgy.

May 1998 "An Algorithm for Delivering Software to Network Attached Devices," Graduate Engineering Seminar, Department of Computer Engineering, Santa Clara University. *Host:* Linda Seiter.

Postdoctoral Fellows

2015–2017 Da Zheng, PhD, Johns Hopkins University.

2009–2011 Ragib Hasan, PhD, University of Illinois Urbana-Champaign.
Assistant Professor, *University of Alabama Birmingham*

PhD Students

2019 James Browne, Johns Hopkins University
United States Military Academy, West Point, NY

2019 Kunal Lillaney, Johns Hopkins University
Amazon, Seattle, WA

2019 Disa Mhembere, Johns Hopkins University
Kyndi, San Mateo, CA

2017 Stephen Hamilton, Johns Hopkins University
Army Cyber Institute, United States Military Academy, West Point, NY

2016 Da Zheng, Johns Hopkins University
PostDoctoral Researcher, *Johns Hopkins University*, Baltimore, MD

2015 Kalin Kanov, Johns Hopkins University
Bloomberg, New York, NY

2013 Osama Khan, PhD, Johns Hopkins University
Twitter, San Francisco, CA

2012 Eric Perlman, PhD, Johns Hopkins University
Data Director, *Kavli Neuroscience Discovery Institute, Johns Hopkins University*, Baltimore, MD

2011 Paul Stanton, PhD, Johns Hopkins University
United States Army, Fort Meade, MD

2011 Xiaodan Wang, PhD, Johns Hopkins University

Salesforce, San Francisco, CA

2009 Alexandros Batsakis, PhD, Johns Hopkins University
TeraData, Sunnyvale, CA

2007 Tanu Malik, PhD, Johns Hopkins University
Research Assistant Professor, *University of Chicago*

2006 Zachary Peterson, PhD, Johns Hopkins University
Assistant Professor, *California State University, San Luis Obispo, CA*

2006 Changxun Wu, PhD, Johns Hopkins University
Google, Mountain View, CA

current Meghana Madhyastha, PhD, Johns Hopkins University

current Brian Choi, PhD, Johns Hopkins University

current Brian Wheatman, PhD, Johns Hopkins University