

Director's Welcome

CERES comes to life! After many months of hard work by our faculty with a set of bold, visionary partners, together we launched the CERES Center for Unstoppable Computing in January 2016. As the CERES Director, it's exciting to the many interesting projects and burgeoning faculty-company researcher relationships take off quickly! This early energy and success is a great validation of the potency of our Unstoppable Computing vision, the extraordinary corporate partners and faculty, and an indication of great things to come! Here are a few highlights of what we've accomplished in just eight months since our inaugural research summit:

- Two dynamic Research summits! January 2016, July 2016 over 60 researchers at each, and with strong representation from our Partners and Members, growing from 14 (January) to 20 (July).
- Launched 8 new projects in data analytics, machine learning, software bug finding and tolerance, trustworthy cloud computation, and

data management. These projects are compelling and garnered enthusiastic support from our terrific partners and members. Thanks Huawei, Argonne, NetApp, and Citadel!



Its early days for CERES, but we're all looking forward to what these projects will grow into! Now, we're already into the intern hiring season for next summer; all partners have received a list of CERES students considering summer internships; many students will settle their summer plans as early as December-January!

At the July summit, we agreed to shift the bi-annual research summits to a September – March schedule, so our next Research Summit will be in March 2017!

Andrew A. Chien, CERES Director
Hyde Park, October 2016

CERES Launch and First Group of New Projects, March 2016

In January 2016, we launched CERES in a gala event at the University of Chicago's 53rd Street 11th Floor complex in Hyde Park. Dean Rocky Kolb (Physical Sciences) kicked off the festivities and welcomed over 50 researchers, including 14 guests from our partners, a dozen faculty, and thirty graduate students and postdocs. With vision talks from Ian Foster, Hank Hoffmann, and Haryadi Gunawi in a strong program of 14 by visitors and CERES faculty,



combined with a huge poster session (over 30 posters presented!), what an exciting technical day!

Welcoming our sponsors and building connections were critical priorities, but the main course was 7 "project pitches" from which we selected 4 strong CERES projects to launch. We describe each, and key progress to date:

- 1. Accelerating Data Analytics (Chien/Elmore)**
Objective: Achieve 2x to 5x performance increases for in-memory data analytics systems, using mainstream software stacks on novel data processing hardware. July 2016 Progress: (1) Defined a set of database loading and format

Center Launch and 1st Research Summit

checking benchmarks, defined a set of compression/decompression benchmarks, (2) Defined a simulation framework for the performance studies. Design includes a scaleup model to get to large-scale experiments (first software emulation, later FPGA), (3) Initiated implementation of the simulation/performance framework, (4) Completed full architecture definition of Unstructured Data Processor (UDP), and (5) Initial implementation of several application kernels enabling initial performance assessment as promising (20-40x single core performance at 1/10th the energy).

2. Limplock-Free Clouds (Gunawi/Lu)

Objective: Build datacenter distributed systems that are free from "limplock" performance glitches where the system suddenly and unexpectedly slows dramatically and is "locked" in slow mode. July 2016 Progress: (1) Invented and implemented path-based speculative execution (PBSE) as a limplock-free replacement of the default Hadoop MapReduce straggler algorithm. We have covered the datanode-map path for Hadoop-HDFS and Hadoop-QuancastFS and we are now working on the map-reduce and reduce-datanode paths. (2)



For data center distributed systems in general, we

have identified a common cascading patterns such as an expensive loop inside a lock critical section used by multiple threads. We are refining the above pattern and investigating others.

3. POETdroid (Hoffmann)

Objective: Extend mobile battery life through application-aware resource management. July 2016 Progress: (1) open source release of portable energy monitoring library with C, Java, and rust bindings, (2) Evaluation of open source navigation apps as test cases, (3) Development of in house navigation app (existing open source projects unsuitable for many reasons - most are not compatible with latest Android)

4. Secure Outsourced Data Analytics (Feldman)

Objective: Design and implement a data analytics framework offering users of third-party analytics significantly greater confidentiality, integrity, and correctness guarantees.

July 2016 Progress: (1) Based on explorations of side channels that Intel SGX doesn't mitigate, the untrusted cloud provider must be given some information about the sensitive data (e.g., the sizes of intermediate result sets), (2) We have refined our approach to focus on devising a query language that supports a wide variety of data analytics jobs but whose execution can be made independent of the sensitive data, given access to differentially private statistics about the sensitive data so that it can optimize the amount of padding and cover traffic without learning enough to compromise the privacy of individual records in the data set.

Contact Project PI's for more information!



Systems Research at UChicago on a Roll!

The Systems Group in Computer Science adds terrific new faculty -- Michael Franklin (Berkeley, Data Management and Data Science) and Blase Ur (CMU, Security and HCI). With these additions, we have tripled in size to 15 faculty, and are rapidly building a top-flight reputation in the systems, architecture, software and data management research communities (>35 top conference papers in the past 18 months, and >75 energetic, ambitious PhD students).



Computer Science to get new Space and lead UChicago Data Science Initiative!

Dean Rocky Kolb announced on June 28, 2016 that Computer Science would move to a renovated Crerar building in June 2018, to accommodate continued rapid faculty and student growth expected to double in size (again!) in the next five years. CS will be joined in Crerar by researchers from the Computation Institute. This new location will be the powerful new center of mass for computing, information, and UChicago's new initiative in Data Science, that incidentally is led by Computer Science Professor Michael Franklin!

2nd Research Summit and more new CERES Projects!, July 2016

We held our second CERES Research



Summit on July 27-28 at UChicago's Knapp Center for Biological Discovery, a leading genomics

center. The meeting surpassed expectations with 67 participants, 20 visitors, 15 talks, 28

posters, and three terrific keynotes: Dr. Ziang Hu (Huawei) – *Key Technology Challenges in the Cloud*, Robert Grossman, *Genomics Challenges in 10 Petabyte Systems*, and Michael Franklin, *Future Directions for Data Science and Computing at UChicago*.



With 8 project pitches, we selected four additional research projects for the CERES portfolio that began in September. They are:

1. COBE: Cascading Outage Bugs Elimination (Lu/Gunawi)

Objective: Build detection and runtime strategies for handling cascading outage bugs (“CO bugs” in short), a new class of bugs that can cause simultaneous or cascades of failures to many or all nodes/components of datacenter software infrastructure.

2. Control-theoretic Automatic Config Mgmt (Hoffmann/Lu)

Objective: Modern software systems all have many configuration options. These configuration options could severely affect the performance and energy consumption of software systems and are also difficult to set manually. This project aims to design a new configuration framework that allows automated and performance/energy-optimal

configuration using a combination of control theory and software engineering techniques.

3. Efficient Data Management for IoT and Data Streams (Elmore/Chien)

Data systems for supporting high ingest rates and complex queries (SQL or dataflow) lack automatic and efficient use of storage and data placement strategies. The goal of EDIS is to build a series of prototypes to evaluate efficacy of automatic encoding, ingestion, and placement strategies.

4. Machine Learning Investigations for Predictive Oncology (Ghosh)

Objective: Evaluate different feature selection methods and classification algorithms for predictive oncology. Specifically, we are interested in comparing and contrasting different feature selection methods against known biological pathway information. This will improve the ability to predict patient outcomes for a given drug regimen.



8 CERES Students are Graduating (MS and PhD's)! (contact CERES to get information on their technical expertise and career interests!) at ceres.cs.uchicago.edu

25 CERES Students are looking for Summer Internships! (contact CERES to get information on their technical expertise and career interests!) at ceres.cs.uchicago.edu



CERES Partners and Members (Thanks!)



Argonne

NATIONAL LABORATORY



NetApp™



CITADEL

CERES Mission Statement

The advance of computing has produced extraordinary capabilities and cost-effectiveness, driving the proliferation of computing into every aspect of society (commerce, government, entertainment, and social activity). However today's computing systems are fragile, failing without warning, suffering intermittent outages, and requiring expert management or repair. CERES mission is to bring together leadership companies with top faculty and student researchers at the University of Chicago Department of Computer Science to create new foundations for "Unstoppable Computing". Our new approaches reduce fragility and complexity of computing systems, but also increase their efficiency and lifetime. Our research activities span data management and analytics, cloud software, storage, programming systems, computer architecture, high-performance, and security. CERES researchers collaborate with our partners to create ground-breaking new integrated perspectives, capabilities, and robust solutions.

