CRNCH Rogues Gallery: A Community Core for Novel Computing Platforms

E. Jason Riedy, Jeffrey Young

School of Computational Science and Engineering
Georgia Institute of Technology

3 November 2017
Outline

Why Do We Need Rogues?

Current Rogues

- Emu Chick
- 3D Stacked Memories and FPGAs
- Neuromorphic Hardware
- Infrastructure and Invitation

(Some) Enabled Research

- Streaming Graph Analysis
- Sparse Tensors
- Sparse Matrix Operations
Why do we need rogues?

Rogue

Someone who goes their own way, who breaks away from the crowd.

- Current architectures are hitting limits on manufacturing, heat dissipation, memory latency...
- What happens when novel prototypes hit reality?
- Designers need feedback, a software ecosystem, and trained students.
Apps: Massive+scale data analysis

Cyber-security Identify anomalies, malicious actors
Health care Find outbreaks, population epidemiology, similar patient association
Social networks Advertising, searching, grouping
Intelligence Decisions at scale, regulating markets, smart & sustainable cities
Systems biology Understanding interactions, drug design
Power grid / Smart cities Disruptions, conservation, prediction

Irregular data access. Changing data.
Introducing the CRNCH Rogues Gallery

CRNCH Rogues Gallery
A physical & virtual space for hosting novel computing architectures, systems, and accelerators.

Host / manage remote access for novel architectures to

- kick-start software ecosystems (Kokkos),
- leverage real applications to train students, and
- provide rapid feedback to architects.

Amortize effort and cost of trying novel architectures. Break the “but it’s too much work” barrier.

http://crnch.gatech.edu/rogues-gallery
Why here and now?

• Past (and present): Experience with hosting, software, training for novel architectures:
  • STI Center for the Cell Broadband Engine Processor;
  • NVIDIA Center for Excellence;
  • multiple Intel Parallel Computing Centers; and
  • multiple NSF & DARPA projects.

• Developed expertise, now sharing efforts.

• Current local- to national-scale applications:
  • Machine learning: ML@GT
  • Health care analytics, clinical focus: CHAI
  • Data-driven science & discovery: NSF South Big Data Hub, FLAMEL, IDEaS, ATL Data Science for Social Good

• Combining novel architectures & applications.

• And enabling new combinations for others!
Current Rogues

Emu Chick

3D Stacked Memories and FPGAs

Neuromorphic Hardware

Infrastructure and Invitation
Emu Chick

- "Migratory Memory Side Processing" to exploit weak locality.
- Data for graph edge attributes, documents / medical records, etc. reside nearby even if accessed irregularly.
- Moving threads to data on reads means all accesses are local, common case needs to tolerate less latency.
3D Stacked Memory and FPGAs

- Rogues Gallery includes traditional FPGA platforms (Arria 10, Ultrascale+) and stacked memory variants
- Enable “near-memory” and memory-centric processing.
- FPGA platforms enable prototypes of non-traditional accelerators like Automata, sparse data engines, etc.
- Current work is supported in part by Micron hardware donation.
Neuromorphic systems

- Field-Programmable Analog Array (FPAA) System-On Chip, designed in the lab of Dr. Jennifer Hasler.
- Uses analog technology supported by digital components to achieve unprecedented power and size reductions.
- Near-term industry products include TrueNorth and Intel Loihi
Flexible Infrastructure

- Always WIP. Will adapt to individual needs.
- Front-end node for general access.
- Slurm for overall queuing and access control.
- Controls access-limited development tools.
  - NDAs: Groups, file permissions
  - Limited licenses: VMs and slurm resources
- Working towards sensitive data, e.g. health care.
  - Swap VLANs for resources? (SDN)
(Some) Enabled Research

- Streaming Graph Analysis
- Sparse Tensors
- Sparse Matrix Operations

Only a sampling, more projects are ongoing & interested!
Streaming graph analysis

Streaming graph frameworks:

- **STINGER**: 3M upd/s in 1.1B vertex, 17B edge graph on 4×Haswell
- **Hornet**: 120M upd/s in 50M vertex graph on P100

DynoGraph: wrap above for evaluation on other architectures!

- **Emu Chick simulator**
- **Emu Chick hardware**!

Adding dynamic community detection, incremental PageRank, concurrent analysis, and more.

Support: IARPA, NSF, DARPA, Intel, IBM, NVIDIA, Oracle, ORNL
Sparse tensors

- ParTI: Supports fast, essential sparse tensor operations on CPU and GPU.
  - Rogues Gallery supports extensions to Emu and FPGA+3D memories
  - Incorporates open data sets like FROSTT

Support: NSF, DoE
Sparse matrix operations

- Sparse primitives provide a key point for optimizing data movement.
- We are looking to optimize sparse primitives with 3D memories for HPC libraries like SuperLU.
  - Scatter and gather are key but other primitives are also useful in graph analytics.
Others are joining!

Other local researchers are looking to use the Rogues:

- Multilingual topic modeling with NMF
- Fraud detection in health insurance
- Anomalies for computer network security

We welcome external users and their applications!

CRNCH Rogues Gallery connects researchers with novel architectures and architects with upcoming applications.

Let us host / manage your neat stuff!

http://crnch.gatech.edu/rogues-gallery
External Image Credits

- Oscar Wilde: public domain, obtained from Wikipedia
- Edna St. Vincent Millay: public domain, obtained from Wikipedia
- Dread Pirate Roberts: copyright 20th Century Fox, used for identification
- Mary Jackson, Katherine Goble Johnson, Dorothy Vaughan (Hidden Figures): copyright 20th Century Fox, used for identification
- Malcolm Reynolds: copyright Universal Pictures, used for identification
- Rogue One: copyright Walt Disney Studios Motion Pictures, used for identification
- The Story of Karrawingi, the emu (cover): copyright estate of Leslie Rees, used for identification