Communication as a first class design constraint for reconfigurable systems

Prof. Simon Moore

UNIVERSITY OF CAMBRIDGE
Computer Laboratory
Computer Architecture Group

FRESH Workshop, Imperial, July 2016

But first...

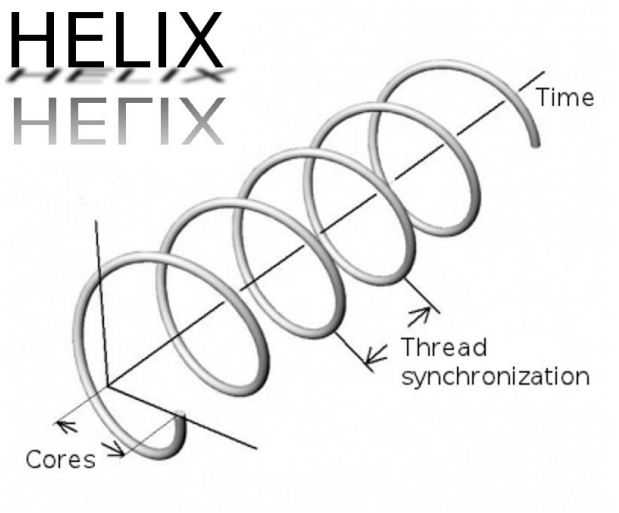
Quick introduction to the
Computer Architecture Group
at the University of Cambridge

Parallelism

Timothy Jones
http://www.cl.cam.ac.uk/ tmj32/

Compilers & Architecture

- Thread-Level
- Data-Level
- Memory-Level



lowRISC & Loki

Robert Mullins
http://www.cl.cam.ac.uk/ rdm34/

- lowRISC – open source system-on-chip platform for research & startups

- Loki – shape-shifting manycore

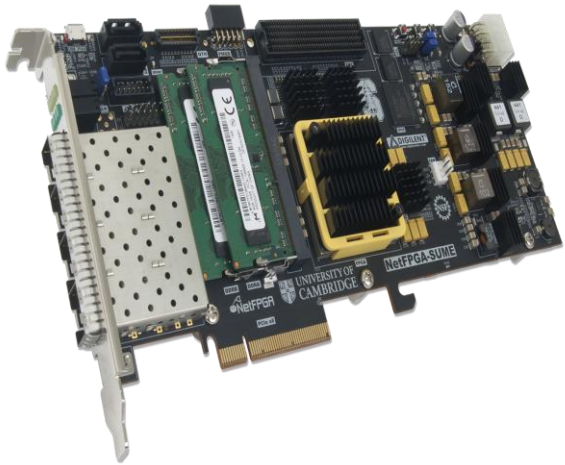
lowRISC
OPEN TO THE CORE

From Systems Research Group

Network software meets network hardware

Andrew Moore
http://www.cl.cam.ac.uk/ awm22/

- One language for all network hardware, firmware, and software www.naas-project.org
- Open Hardware and 100Gb/s Research Reality www.netfpga.org
- Useful Measurements - Merging Cause and Effect www.metrics-itn.eu
- Datacenter evaluate thine self - Emulating 1 million machines selena-project.github.io
- NEW: SSICLOPS - secure (fast) clouds for everyone www.ssiclops.net (coming soon)
- NEW: ENDEAVOUR - exploring Software Defined Networking for Internet-wide switches



Secure Heterogeneous Many-core Computers

Simon Moore
http://www.cl.cam.ac.uk/ swm11/

- CHERI processor adds fine grained memory protection and sandboxing
- Clang/LLVM support
- FreeBSD OS port
- Security eval.
- Cross research group
  Comp Arch + Security + SRG
- FPGA prototyping & chips

From Security Research Group

OSes, ISAs, Program Analysis & Security

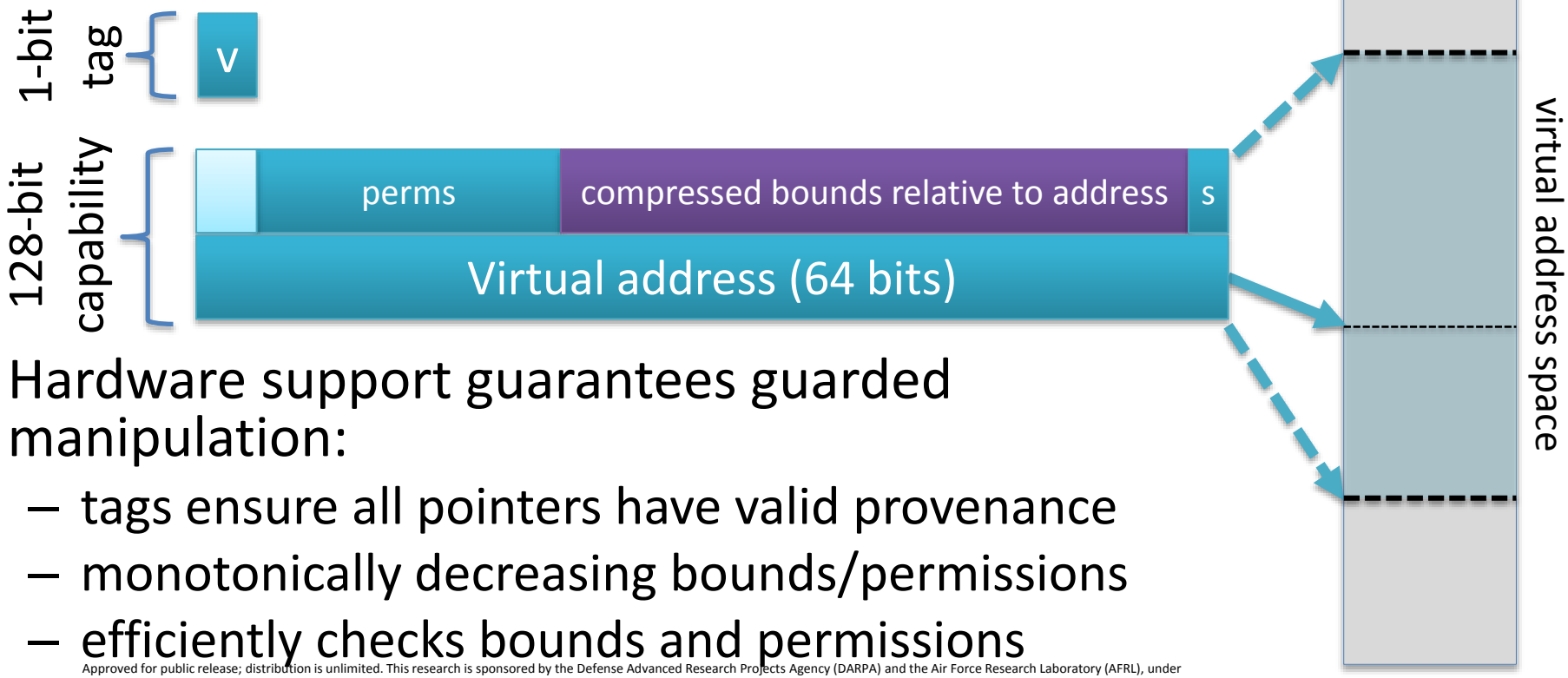
Robert Watson

http://www.cl.cam.ac.uk/ rnw24/

- CHERI: revisiting RISC for the age of Risk
  - with Comp. Arch. group
  - processor ISAs for security + rest of the stack
- Capsicum: POSIX + capability protection
  - initially on FreeBSD but ported to Linux by Google
- Network-stack specialization
  - clean-state network-stack & storage designs

CHERI – Secure Processor

- **Principle of least privilege** is key to security
- CHERI processor provides **capabilities** – a new hardware data type, the tagged and bounds-checked pointer:

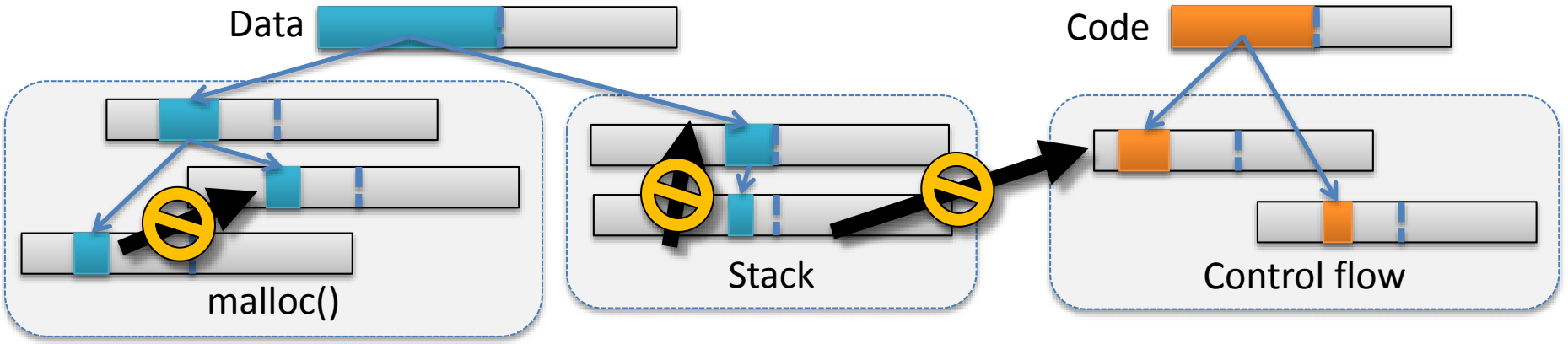


- Hardware support guarantees guarded manipulation:
  - tags ensure all pointers have valid provenance
  - monotonically decreasing bounds/permissions
  - efficiently checks bounds and permissions

Approved for public release; distribution is unlimited. This research is sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory (AFRL), under contracts FA8750-10-C-0237 ('CTSRD') and FA8750-11-C-0249 ('MRC2'). The views, opinions, and/or findings contained in this article/presentation are those of the author(s)/presenter(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

CHERI – Secure Software

- CHERI capabilities: HW support for **least-privilege pointers**
  - Enforce bounds, permissions, provenance, and monotonicity
  - Mitigate C-language software vulnerabilities



➞ Efficient, fine-grained, compiler-driven **memory safety**

➞ Highly efficient and scalable **software compartmentalisation**

Approved for public release; distribution is unlimited. This research is sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory (AFRL), under contracts FA8750-10-C-0237 ('CTSRD') and FA8750-11-C-0249 ('MRC2'). The views, opinions, and/or findings contained in this article/presentation are those of the author(s)/presenter(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Publications cross traditional boundaries

The CHERI capability model: Revisiting RISC in an age of risk. Jonathan Woodruff, Robert N. M. Watson, David Chisnall, Simon W. Moore, Jonathan Anderson, Brooks Davis, Ben Laurie, Peter G. Neumann, Robert Norton, and Michael Roe. **ISCA 2014**.

Beyond the PDP-11: Processor support for a memory-safe C abstract machine. David Chisnall, Colin Rothwell, Brooks Davis, Robert N.M. Watson, Jonathan Woodruff, Simon W. Moore, Peter G. Neumann and Michael Roe. **ASPLOS 2015**.

CHERI: A Hybrid Capability-System Architecture for Scalable Software Compartmentalization. Robert N. M. Watson, Jonathan Woodruff, Peter G. Neumann, Simon W. Moore, Jonathan Anderson, David Chisnall, Nirav Dave, Brooks Davis, Khilan Gudka, Ben Laurie, Steven J. Murdoch, Robert Norton, Michael Roe, Stacey Son & Munraj Vadera. **IEEE Security and Privacy 2015**.

Clean Application Compartmentalization with SOAAP. Khilan Gudka, Robert N.M. Watson, Jonathan Anderson, David Chisnall, Brooks Davis, Ben Laurie, Ilias Marinos, Peter G. Neumann, and Alex Richardson. **ACM CCS 2015**.

Fast Protection-Domain Crossing in the CHERI Capability-System Architecture. Robert N. M. Watson, Robert Norton, Jonathan Woodruff, Alexandre Joannou, Simon W. Moore, Peter G. Neumann, Jonathan Anderson, David Chisnall, Nirav Dave, Brooks Davis, Khilan Gudka, Ben Laurie, A. Theodore Markettos, Ed Maste, Steven J. Murdoch, Michael Roe, Colin Rothwell, Stacey Son, Munraj Vadera. **IEEE Micro Journal 2016** (to appear).

A Consistency Checker for Memory Subsystem Traces. Matthew Naylor, Simon W. Moore, Alan Mujumdar. **FMCAD 2016** (to appear).

We also publish in FPGA venues

Rapid codesign of a soft vector processor and its compiler. Matthew Naylor and Simon W. Moore. **FPL 2014**

Interconnect for commodity FPGA clusters: standardized or customized? Paul J. Fox, A. Theodore Markettos, Simon W. Moore and Andrew W. Moore. **FPL 2014**

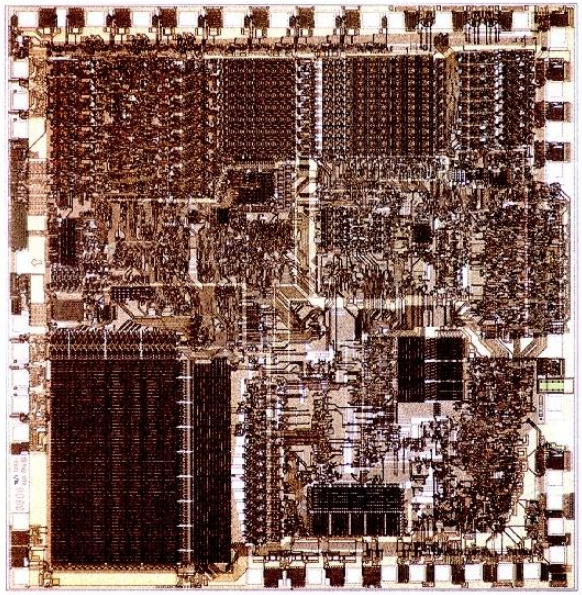
Managing the FPGA Memory Wall: custom computing or vector processing? Matthew Naylor, Paul J Fox, A Theodore Markettos and Simon W Moore. **FPL 2013**

Bluehive - A Field-Programmable Custom Computing Machine for Extreme-Scale Real-Time Neural Network Simulation. Simon W. Moore, Paul J. Fox, Steven J.T. Marsh, A. Theo Markettos, Alan Mujumdar. **FCCM 2012**

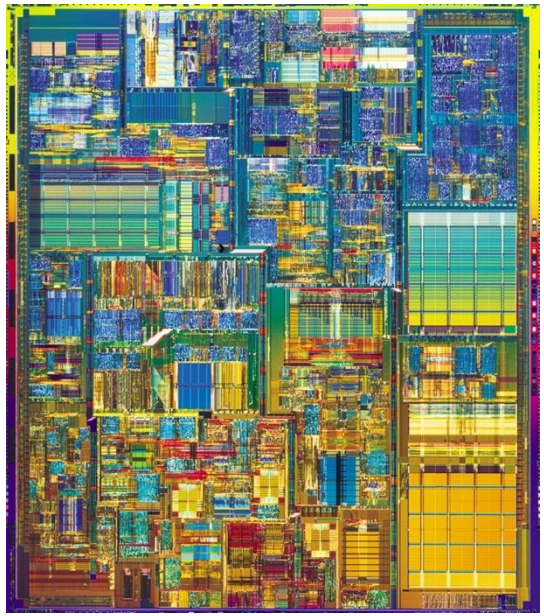
Back to the main talk…

Communication as a first class design constraint for reconfigurable systems

Wires stopped scaling 40 years ago



Intel 8086 (1978)
29,000 transistors
3μm process



Intel Pentium-4 (2000)
42,000,000 transistors
0.18μm process

Power of Computation vs. Communication

technology node	130nm CMOS (2006)	45nm CMOS (2008)
transfer 32b across-chip	20 computations	57 computations
transfer 32b off-chip	260 computations	1300 computations

numbers derived from work by W. Dally, Stanford

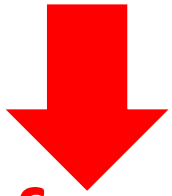
Observation

Electronics technology scaling favours

transistors over wires

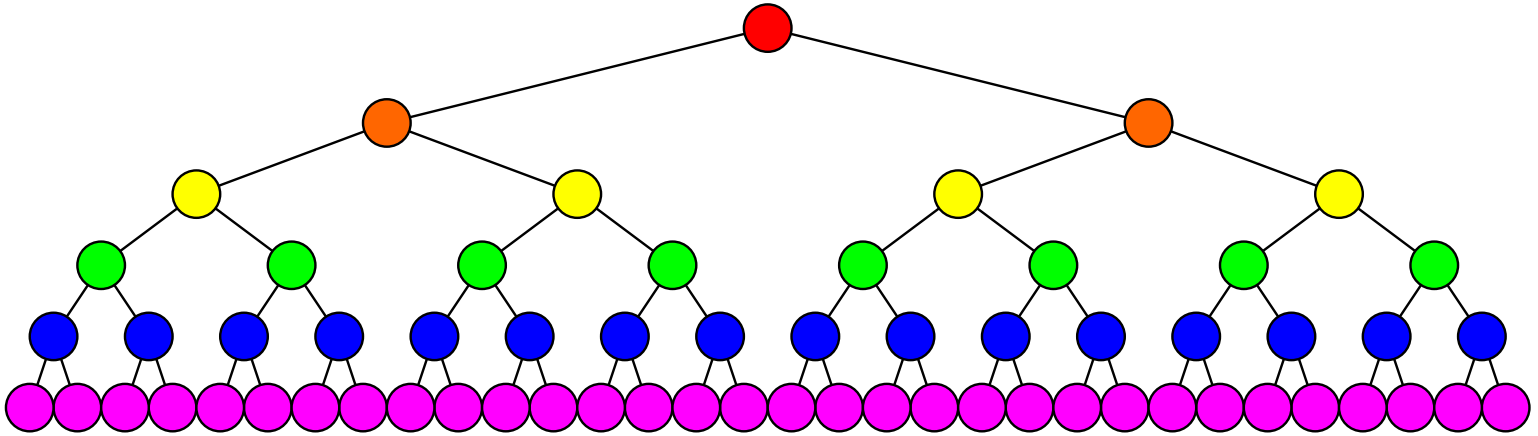
Time to Rethink Complexity

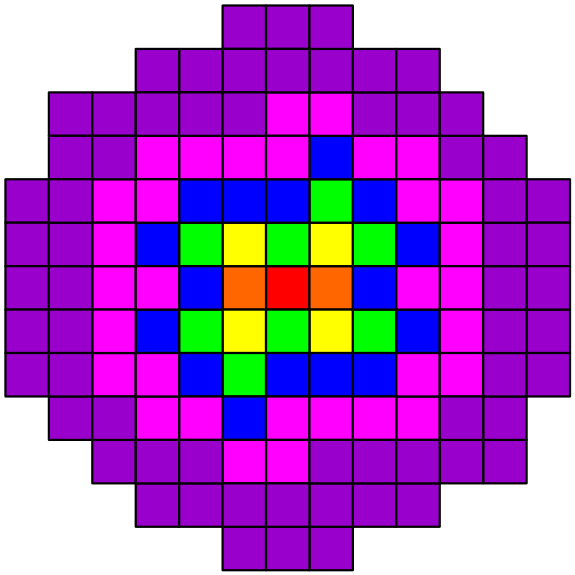
~~computational complexity~~



complexity of communication

Example: embedding a binary tree in 2D





Binary Tree Traversal

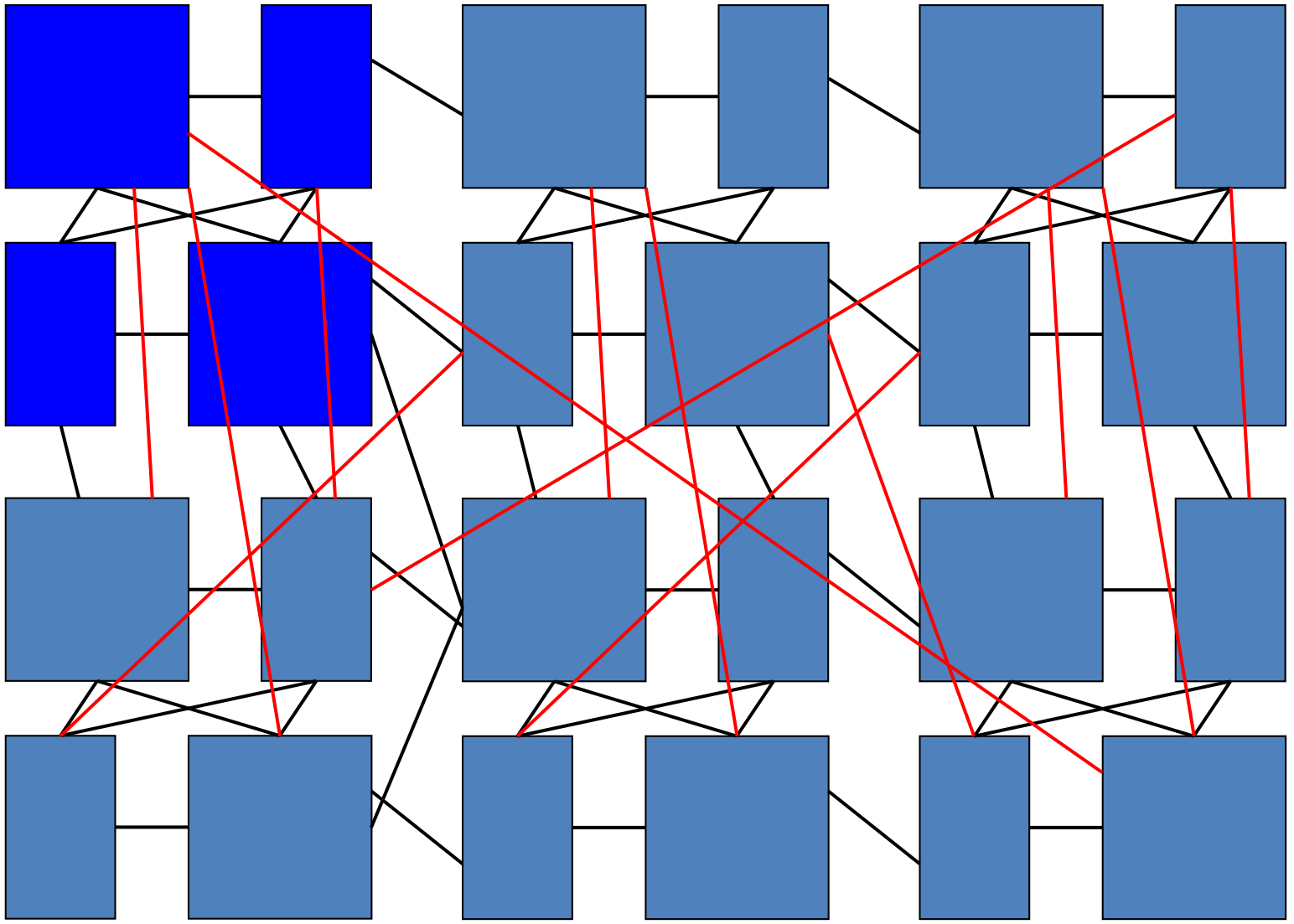
No Communication Cost: $O(\log n)$

Random Placement: $O(\sqrt{n}.\log n)$

Optimal Placement: $O(\sqrt{n})$

work by Dan Greenfield, University of Cambridge

Rent's rule – predicting complexity of communication



Generalizing Rents Rule

- Rents rule can also be applied to communication scaling for:
  - networks-on-chip
    (i.e. dynamically routed communication)
  - communications in software
    (communication between registers, caches and DRAM)
  - communication in mammalian brains
- Work of Dan Greenfield (one of my PhD students)

Rent's rule for Mammalian Brains

OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Efficient Physical Embedding of Topologically Complex Information Processing Networks in Brains and Computer Circuits

Danielle S. Bassett[1,2,3,6*☯], Daniel L. Greenfield[4☯], Andreas Meyer-Lindenberg[5], Daniel R. Weinberger[6], Simon W. Moore[4], Edward T. Bullmore[3*]

1 Department of Physics, University of California Santa Barbara, Santa Barbara, California, United States of America, **2** Institute for Collaborative Biotechnologies, University of California Santa Barbara, Santa Barbara, California, United States of America, **3** Behavioral & Clinical Neurosciences Institute, Department of Psychiatry, University of Cambridge, Cambridge United Kingdom, **4** Computer Laboratory, University of Cambridge, Cambridge, United Kingdom, **5** Central Institute of Mental Health, Mannheim, Germany, **6** Genes, Cognition, and Psychosis Program, Clinical Brain Disorders Branch, National Institute of Mental Health, Bethesda, Maryland, United States of America

Abstract

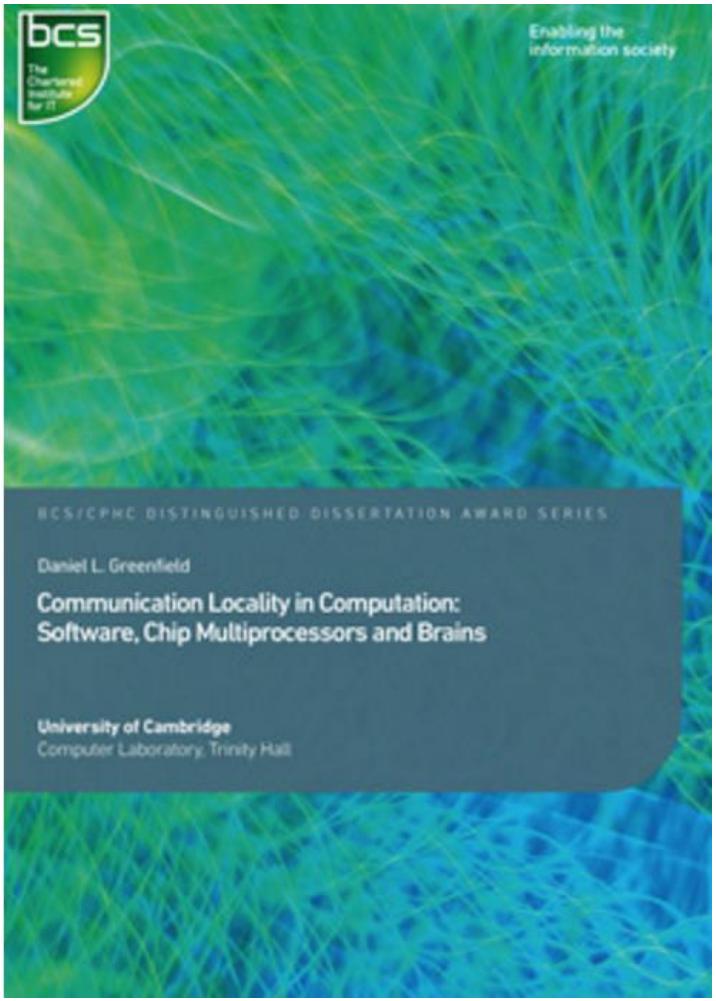
Nervous systems are information processing networks that evolved by natural selection, whereas very large scale integrated (VLSI) computer circuits have evolved by commercially driven technology development. Here we follow historic intuition that all physical information processing systems will share key organizational properties, such as modularity, that generally confer adaptivity of function. It has long been observed that modular VLSI circuits demonstrate an isometric scaling relationship between the number of processing elements and the number of connections, known as Rent's rule, which is related to the dimensionality of the circuit's interconnect topology and its logical capacity. We show that human brain structural networks, and the nervous system of the nematode *C. elegans*, also obey Rent's rule, and exhibit some degree of hierarchical modularity. We further show that the estimated Rent exponent of human brain networks, derived from MRI data, can explain the allometric scaling relations between gray and white matter volumes across a wide range of mammalian species, again suggesting that these principles of nervous system design are highly conserved. For each of

CPHC/BCS Distinguished Dissertation Prize 2011

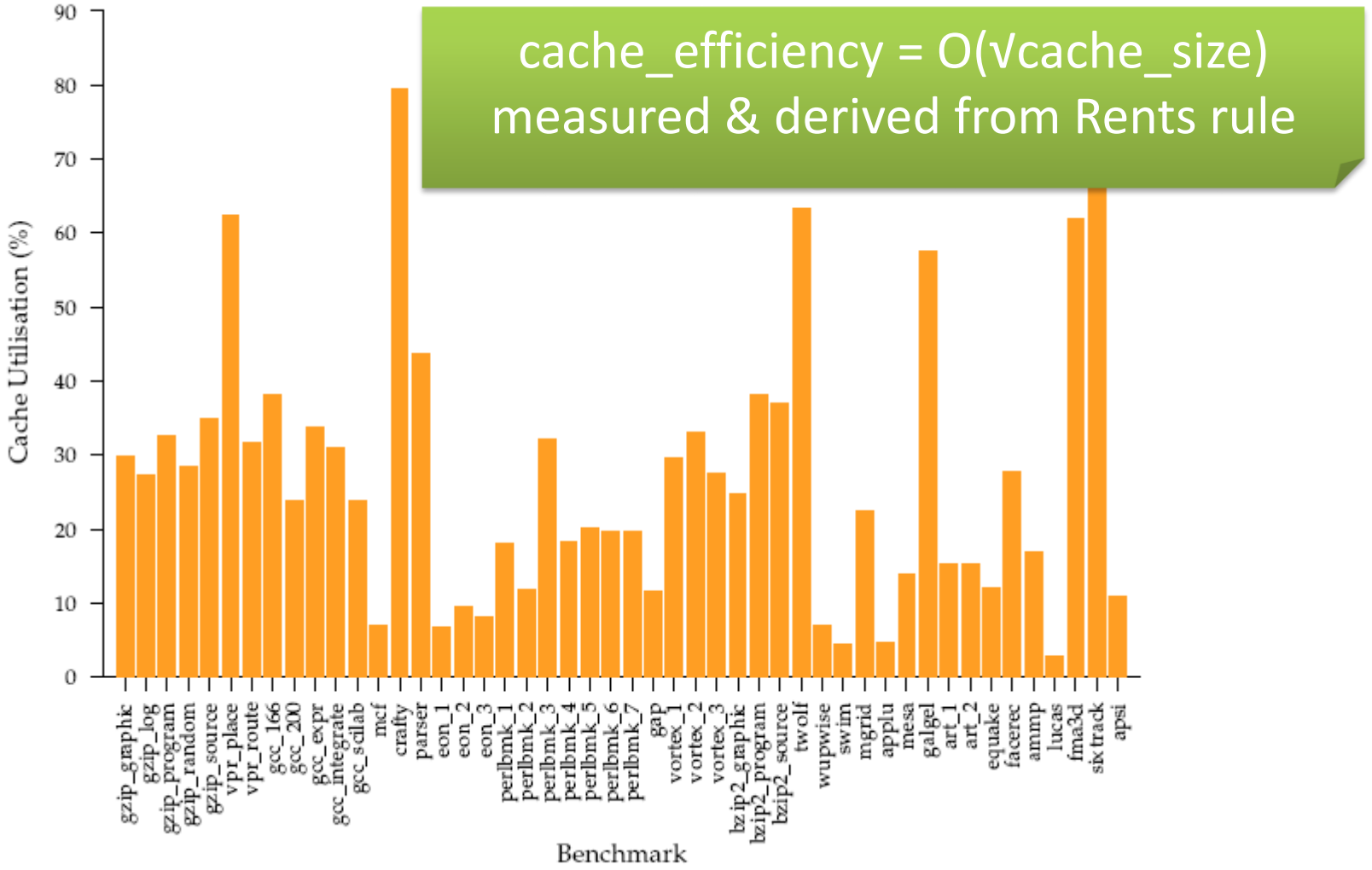
Communication Locality in Computation: Software, Chip Multiprocessors and Brains

Daniel Greenfield, 2011.

Supervised by Simon Moore



Level-2 unified cache utilisation

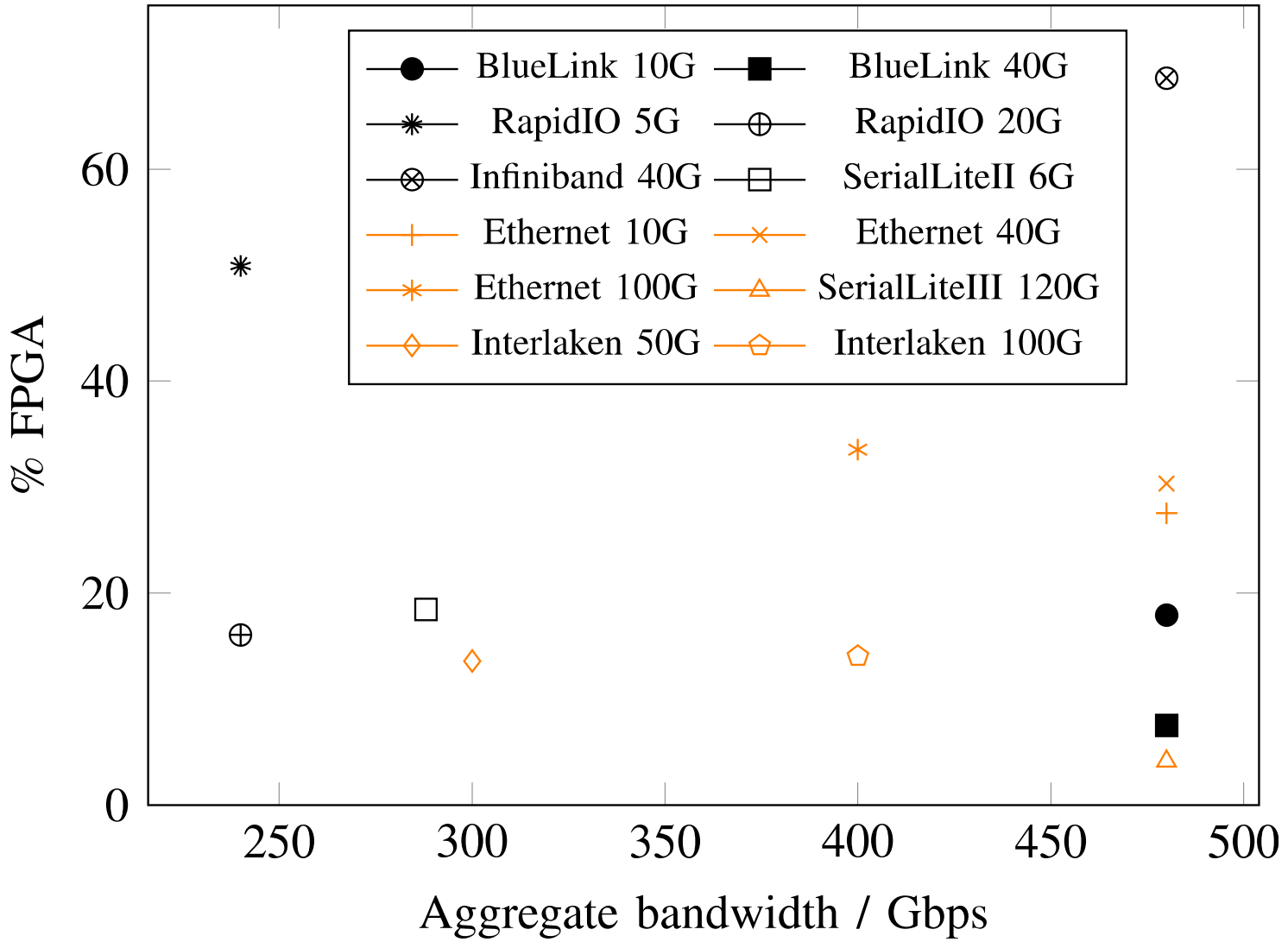


work by James Srinivasan, University of Cambridge

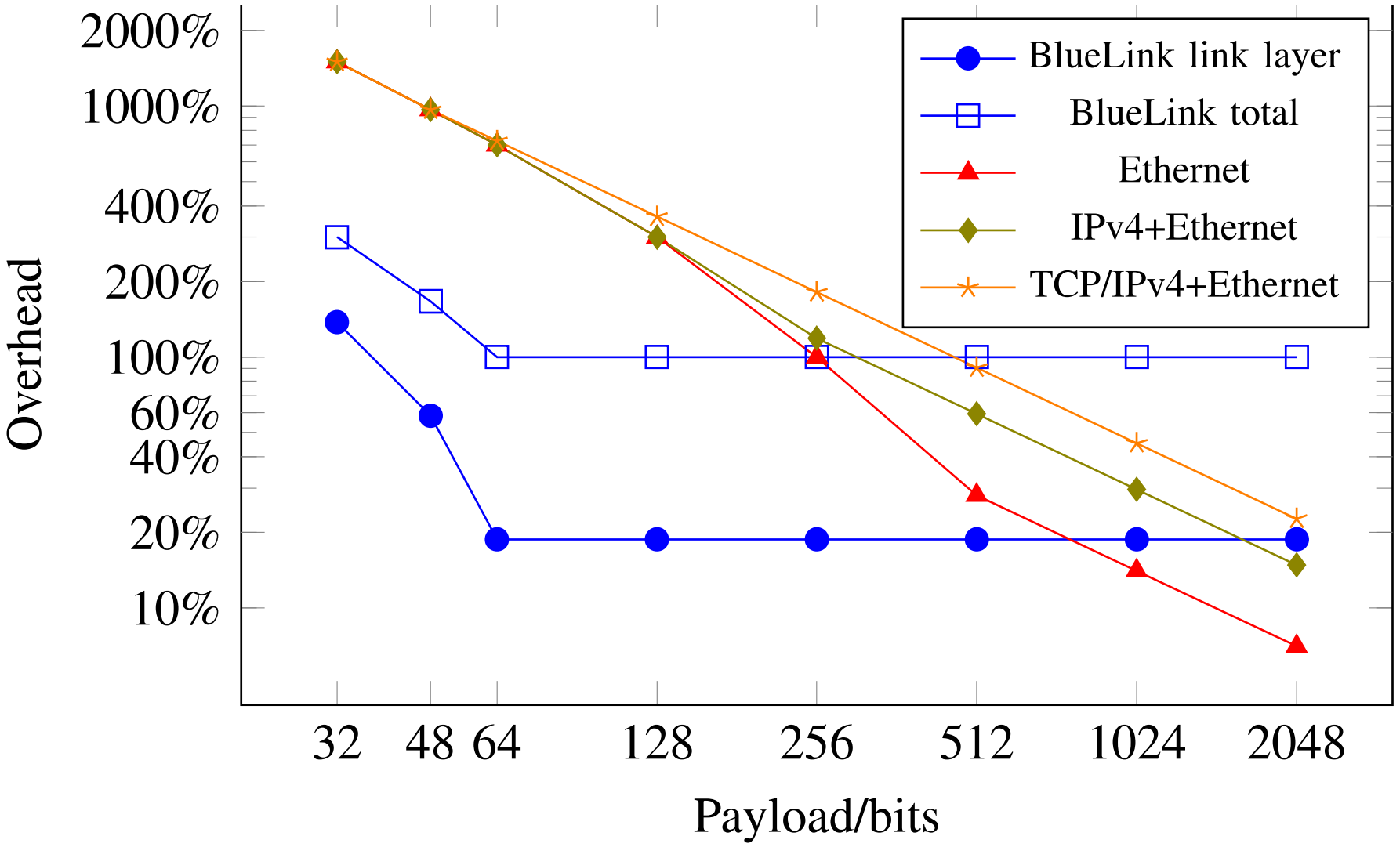
Why focus on custom processors?

What about custom communication?

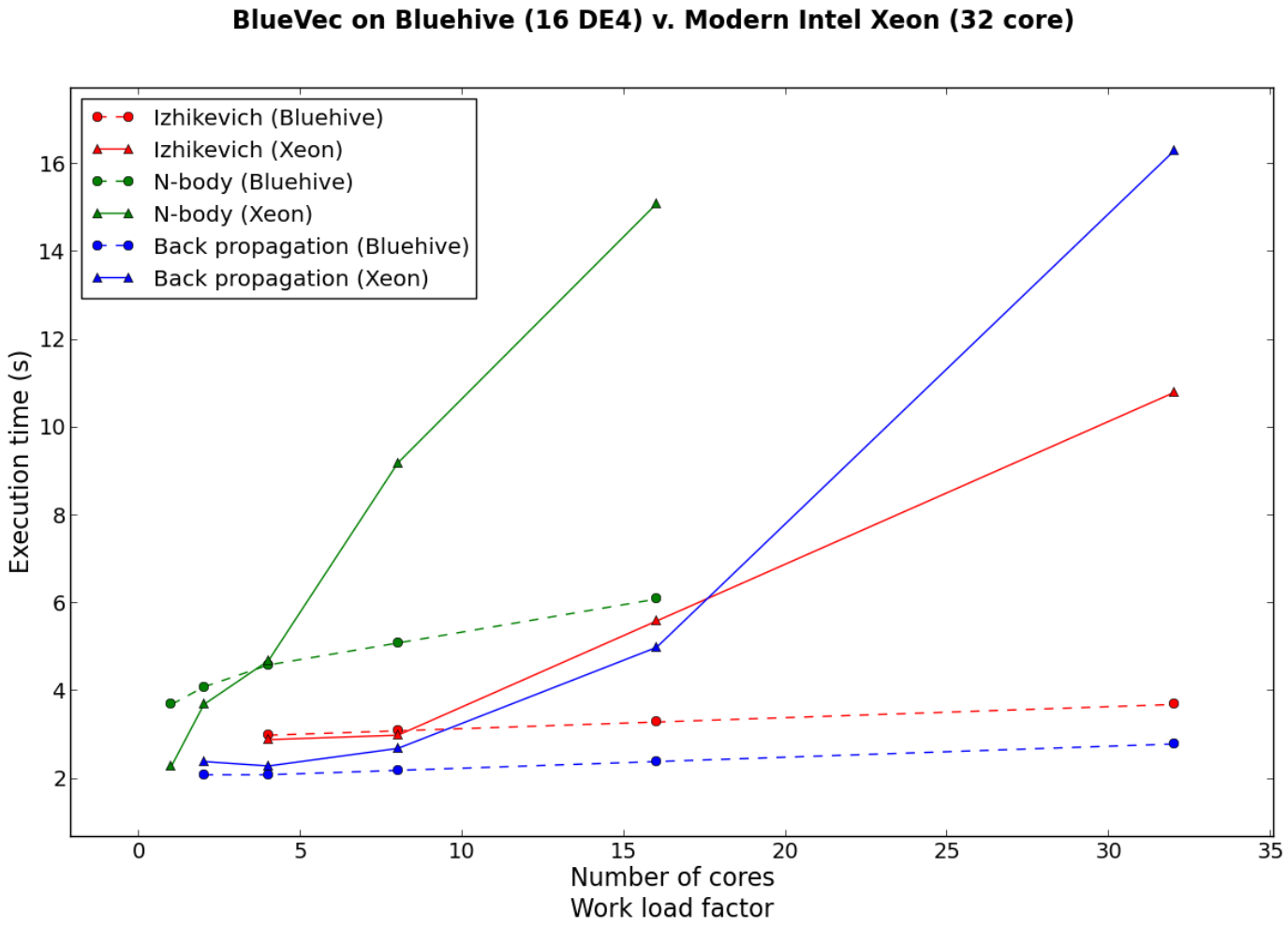
Custom vs. Standard Protocols



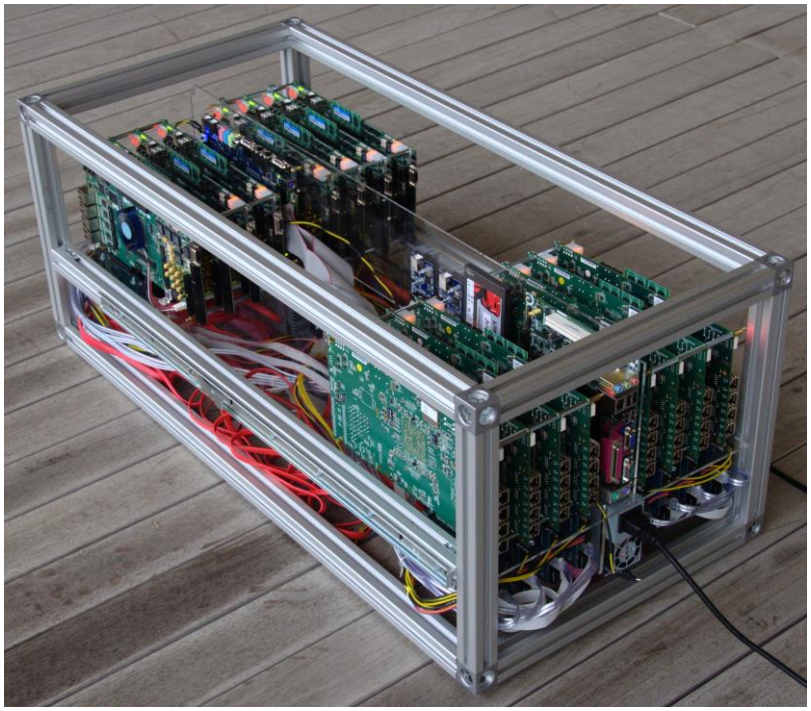
Optimize for (small) packet size



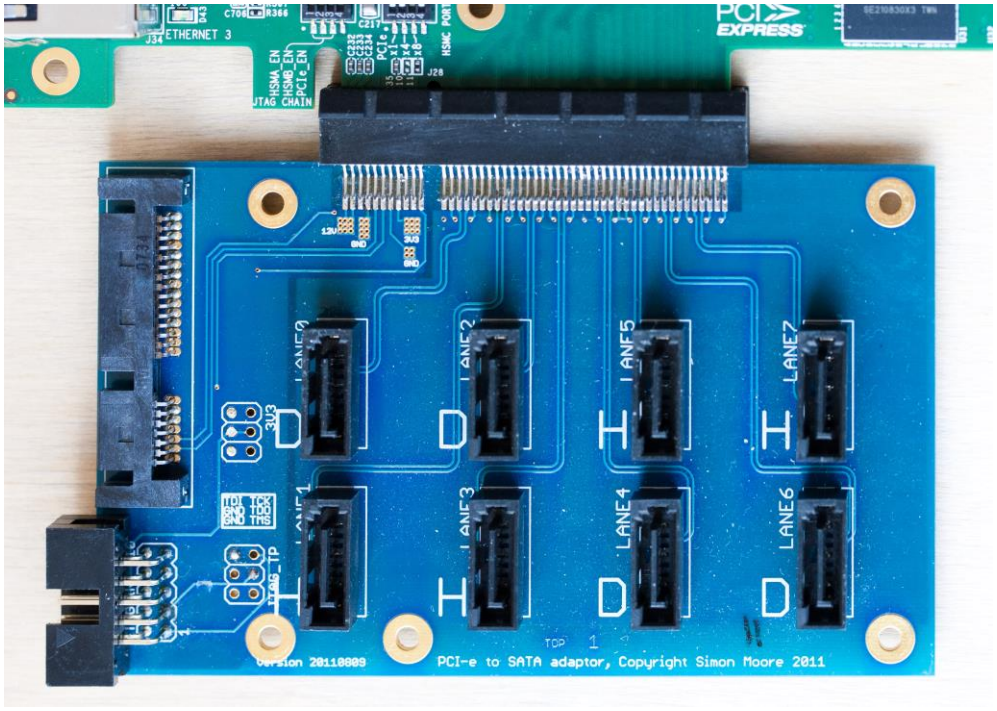
Example – cache coherence on Xeon vs. custom communication on FPGA



Bluehive (2012)

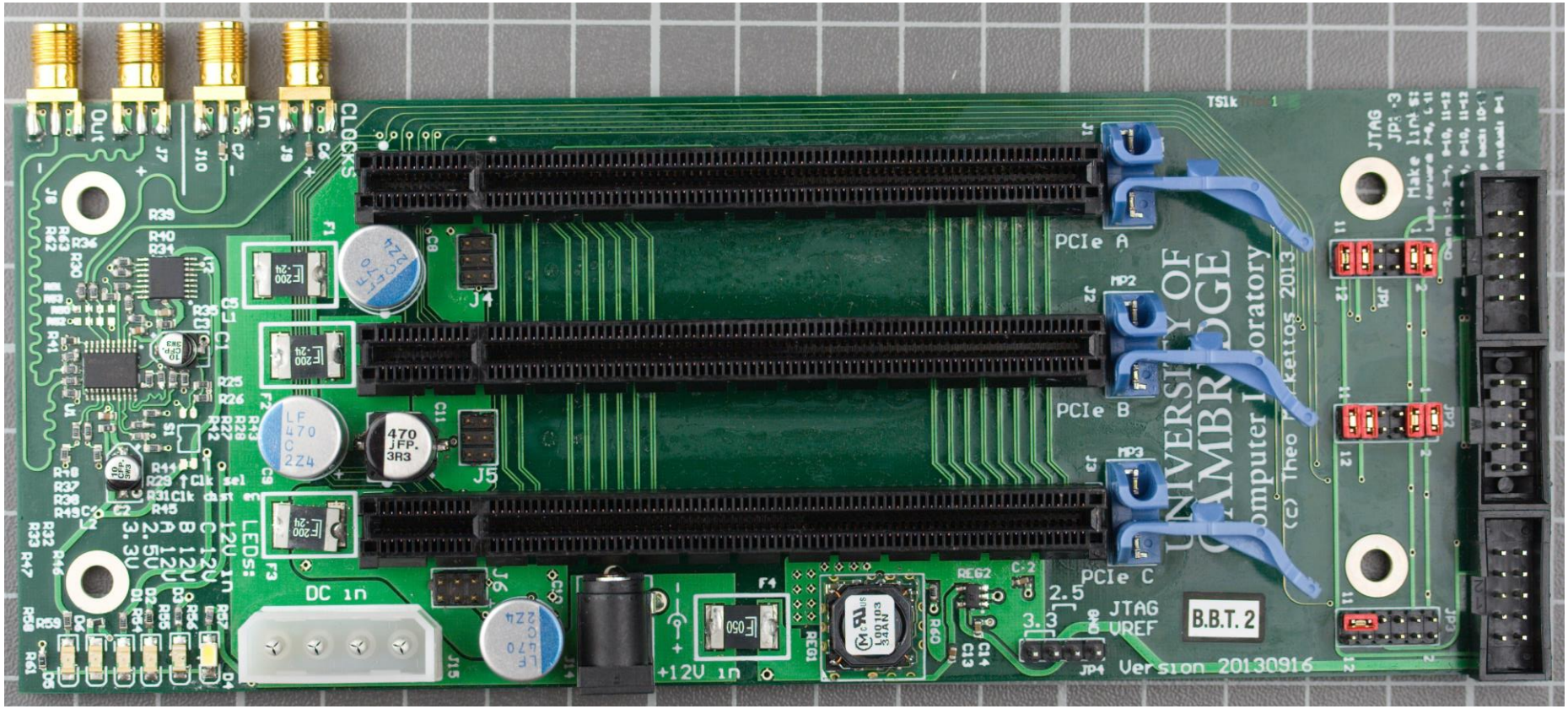


16 x DE4 Stratix IV FPGA boards
Connected by SATA cables



Our PCIe to SATA break-out board

Our 3-slot PCIe Passive Backplane



Where Next?

- POETS (Partially Ordered Event Triggered Systems) Programme Grant
  - Explore massively-parallel highly-connected FPGA-based machines for scientific compute
- Commercialise the CHERI secure processor and software system

Conclusions

- Communication costs are becoming more important than computation costs
  - Corollary:
    optimise communication, do more compute!
- Great time to be doing computer architecture work on FPGAs, and many opportunities to collaborate

Simon Moore
Computer Architecture Group
Computer Laboratory, University of Cambridge
http://www.cl.cam.ac.uk/research/comparch/