ETH zürich

Enzian: in stores now

David Cock, and the Enzian team
Systems Group
ETH Zurich Department of Computer Science

www.enzian.systems

DINFK Systems@ETH zürich ENZIAN

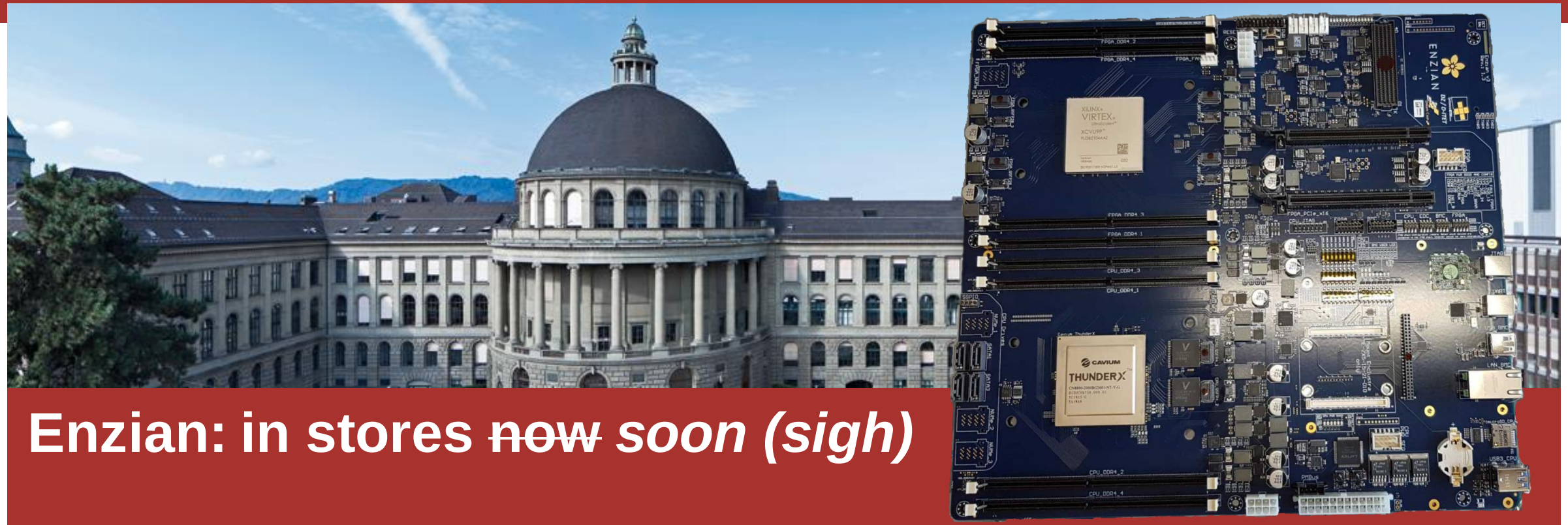
ETH zürich

Enzian: in stores ~~now~~ *soon (sigh)*

David Cock, and the Enzian team
Systems Group
ETH Zurich Department of Computer Science

www.enzian.systems

DINFK Systems@ETH zürich ENZIAN



Enzian: in stores ~~now~~ *soon (sigh)*

David Cock, and the Enzian team
Systems Group
ETH Zurich Department of Computer Science
www.enzian.systems

DINFK Systems@ETH zürich ENZIAN

What Enzian is about:

- ***Relevant*** academic ***systems software research*** in an age of custom hardware.
- A computer ***designed for academic research,*** and what we will use it for.
- How we actually ***built such a machine***.
- How to ***get*** one.

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Acknowledgements

- **Mohsen Owaida**: ECI link bringup
- **Adam Turowski**: Barrelfish, Linux, and FreeBSD OS bringup, BMC software lead
- **Tobias Grosser**: Application use cases
- **Amit Kulkarni**: Verilog hacking
- **Reto Achermann**: Trace processing, software emulation, Barrelfish bringup
- **Zeke Wang**: DDR4 controllers
- **David Sidler**: FPGA network implementation
- **Alexander Hedges**: Simulation environment
- **Nikita Lazarev**: Interconnect protocol specification and modelling
- **Abishek Ramdas**: ECI implementation
- **Dario Korolija**: FPGA Shell design
- **Fabio Maschi**: Verilog hacking
- **Timothy Roscoe, Gustavo Alonso, David Cock**

Systems@ETH zürich

D INFK

CAVIUM

XILINX

Dream CHIP

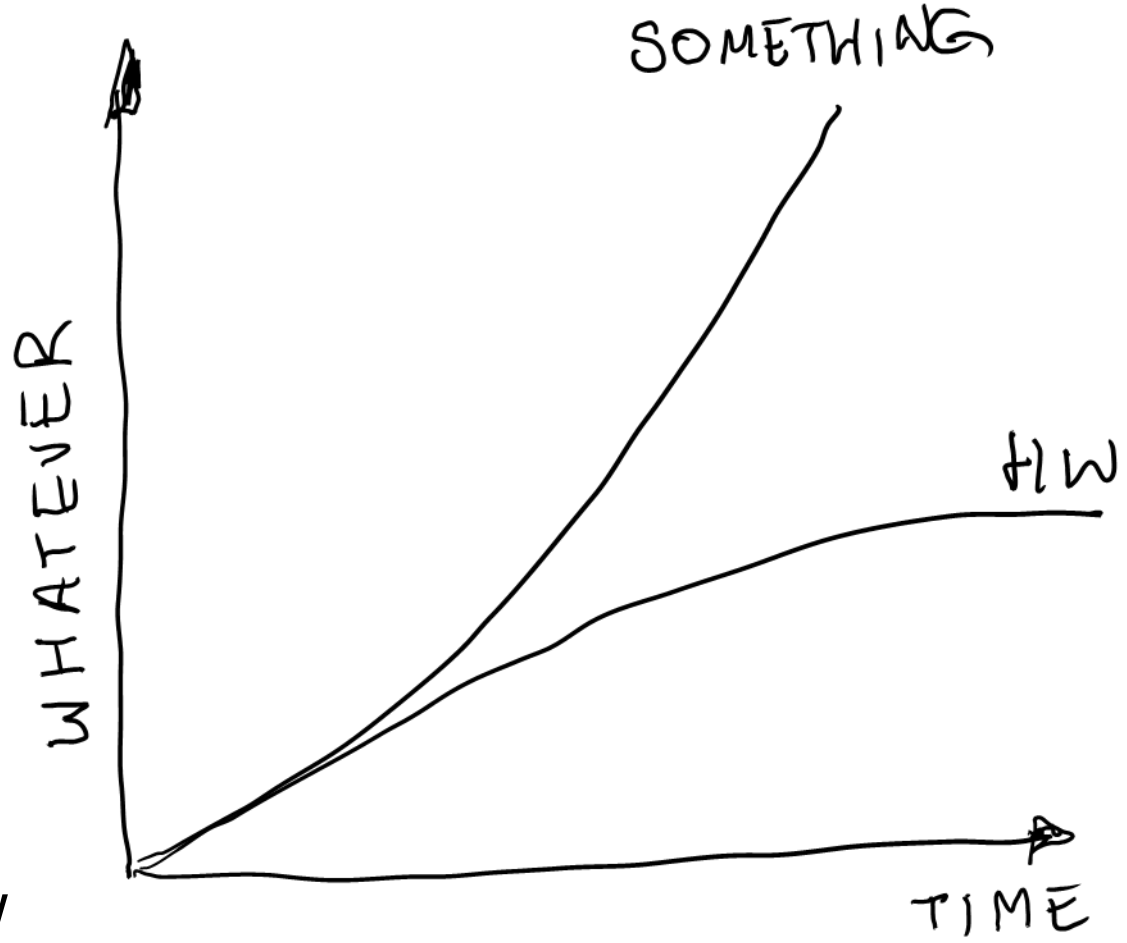
…and everyone else who has helped us so far!

David Cock, Systems Group
Department of Computer Science
ETH Zurich

The usual bla, bla, bla ...

- Moore's Law
- Dennar scaling, physical limits
- Multicore
- GPU, TPU, FPGA
- Data centers and the cloud
- ...

- Corollary: Hardware is changing really fast (because they do not know what to do with the transistors)



(courtesy Gustavo Alonso)

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Reconfigurable and custom hardware is where the action is

Hardware today is complex and diverse

- Custom ASICs
- FPGAs
- On-chip accelerators

- *It's easier to build what you want*
  - High-end CAD systems
  - Simulators and emulators
  - FPGA development environments
  - Rapid fabrication of boards and ASICs

- Big companies ***do***
  - HPE's The Machine
  - Oracle RAPID, SPARC M7
  - Amazon F1
  - Microsoft Catapult
  - Google CloudTPU
  - Baidu FPGA deployments
  - appliances (e.g. PureStorage)

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Reconfigurable and custom hardware is where the action is

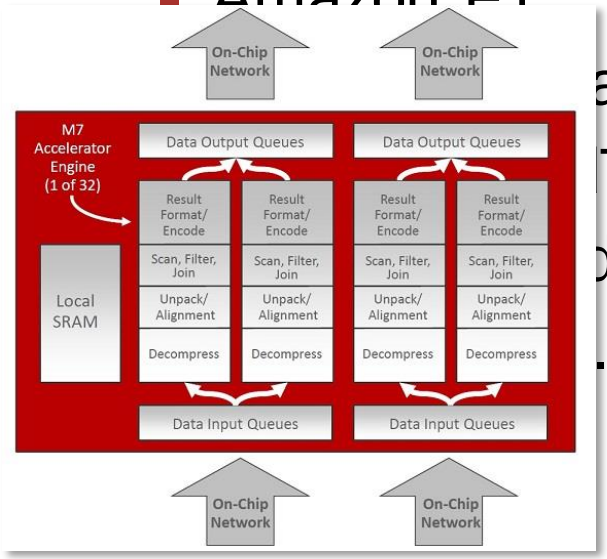
Hardware today is complex and diverse

- Custom ASICs
- FPGAs
- On-chip accelerators

- *It's easier to build what you want*
  - High-end CAD systems
  - Simulators and emulators
  - FPGA development environments
  - Rapid fabrication of boards and ASICs

- Big companies ***do***
  - HPE's The Machine
  - Oracle RAPID, SPARC M7
  - Amazon F1

apult
TPU
deployments
.g. PureStorage)



David Cock, Systems Group
Department of Computer Science
ETH Zurich

Reconfigurable and custom hardware is where the action is

today is complex and diverse

m ASICs

s

ip accelerators

ier to build what you want

nd CAD systems

ators and emulators

- FPGA development environments
- Rapid fabrication of boards and ASICs

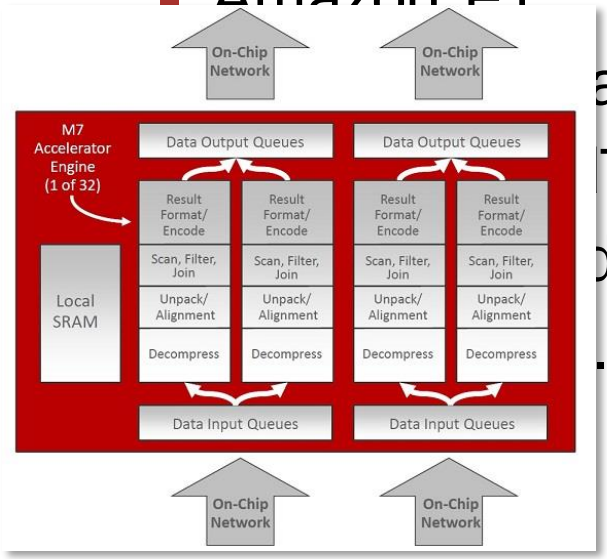
- Big companies ***do***
  - HPE's The Machine
  - Oracle RAPID, SPARC M7
  - Amazon F1

apult

TPU

deployments

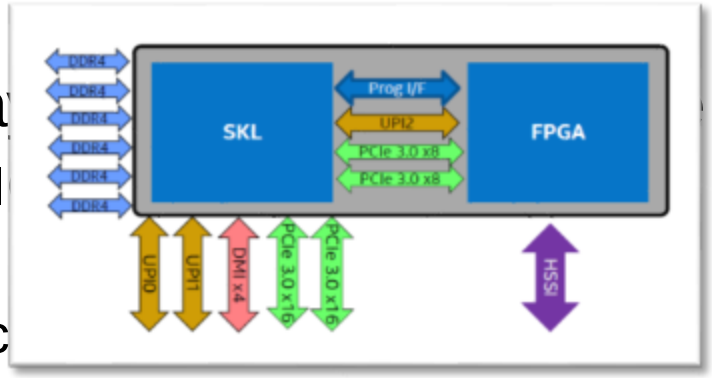
.g. PureStorage)



David Cock, Systems Group
Department of Computer Science
ETH Zurich

Reconfigurable and custom hardware is where the action is

toda
m ASI
s
ip acc

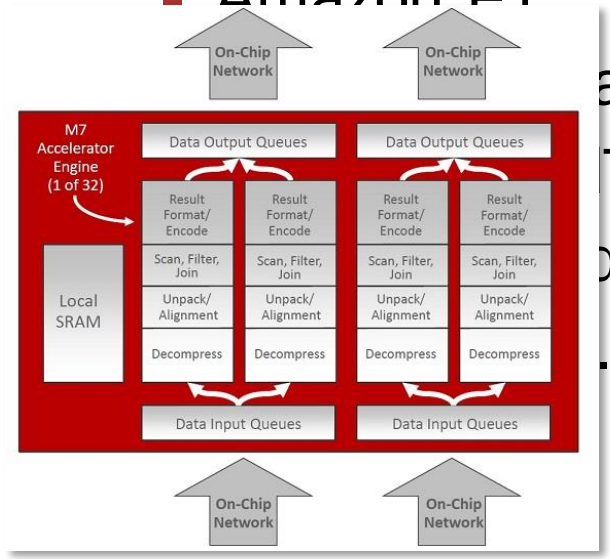


ier to build what you want

- end CAD systems
- ators and emulators
- FPGA development environments
- Rapid fabrication of boards and ASICs

- Big companies ***do***
  - HPE's The Machine
  - Oracle RAPID, SPARC M7
  - Amazon F1

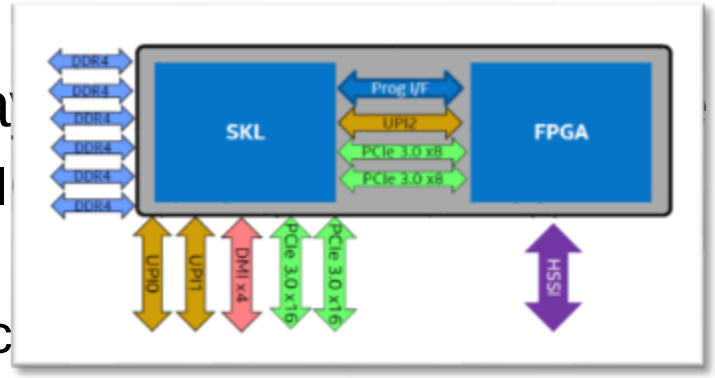
apult
TPU
deployments
.g. PureStorage)



David Cock, Systems Group
Department of Computer Science
ETH Zurich

Reconfigurable and custom hardware is where the action is

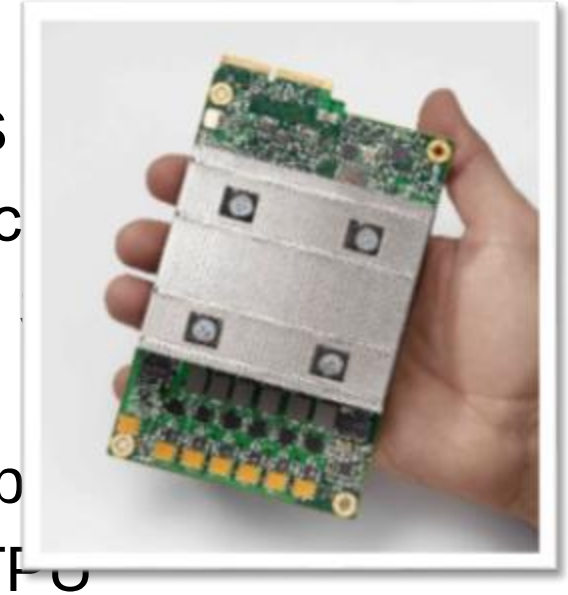
toda
m ASI
s
ip acc



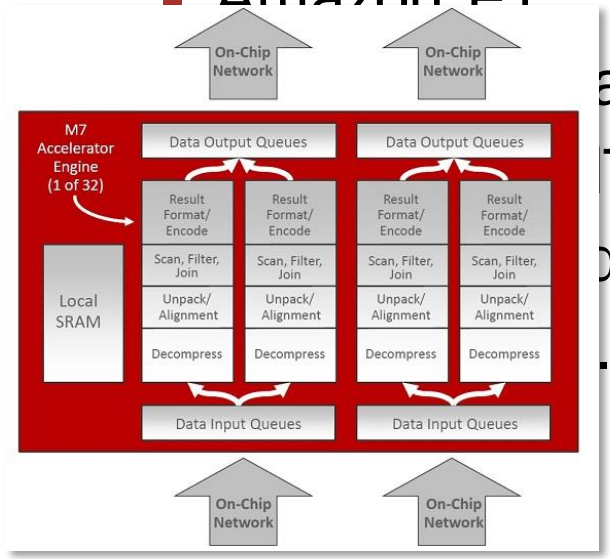
ier to build what you want

- nd CAD systems
- ators and emulators
- FPGA development environments
- Rapid fabrication of boards and ASICs

- Big companies
  - HPE's The Mac
  - Oracle RAPID,
  - Amazon F1



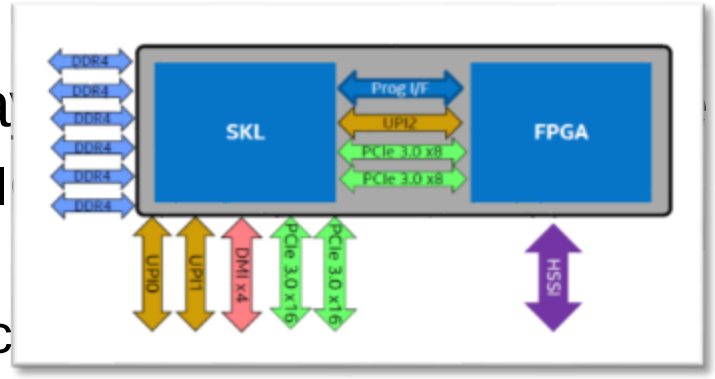
ap
TPU
deployments
.g. PureStorage)



David Cock, Systems Group
Department of Computer Science
ETH Zurich

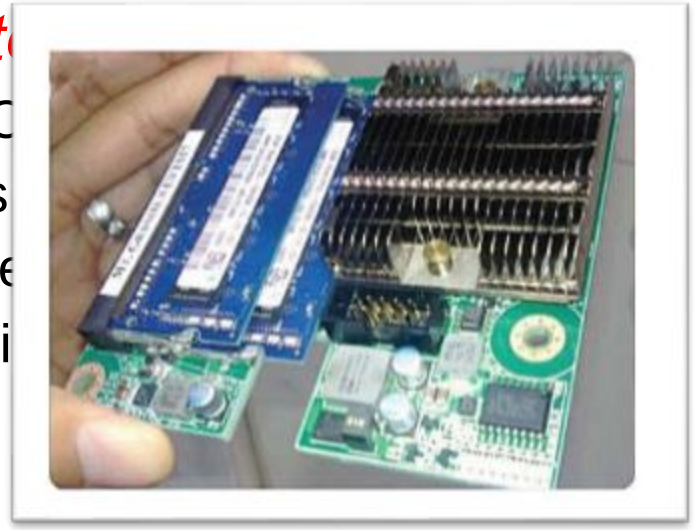
Reconfigurable and custom hardware is where the action is

toda
m ASI
s
ip acc

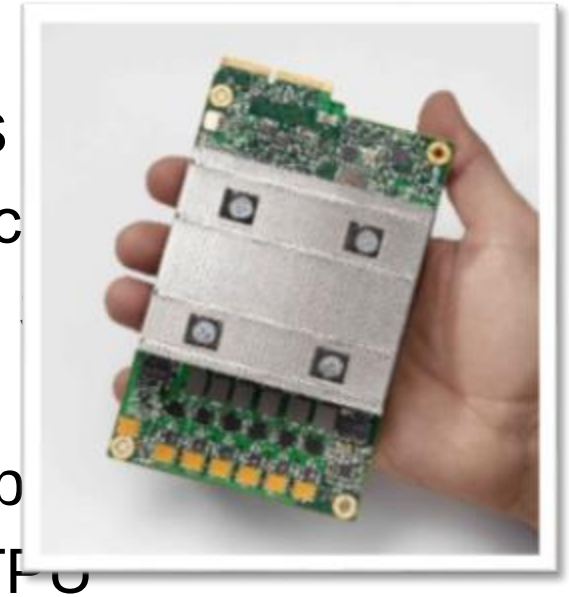


ier t
nd C
ators

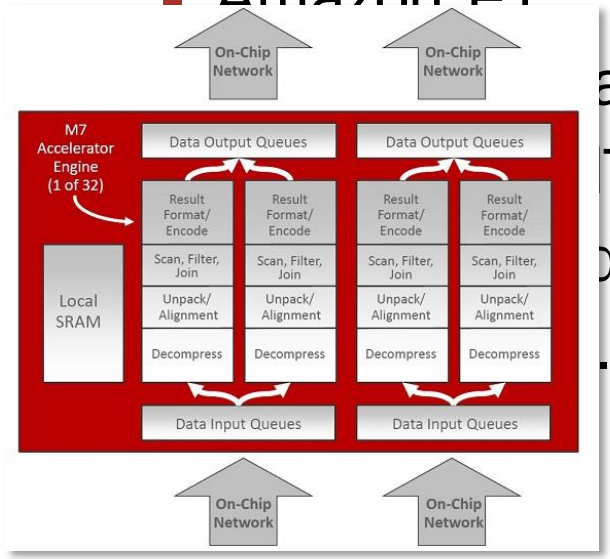
- FPGA deve
- Rapid fabri



- Big companies
  - HPE's The Mac
  - Oracle RAPID,
  - Amazon F1

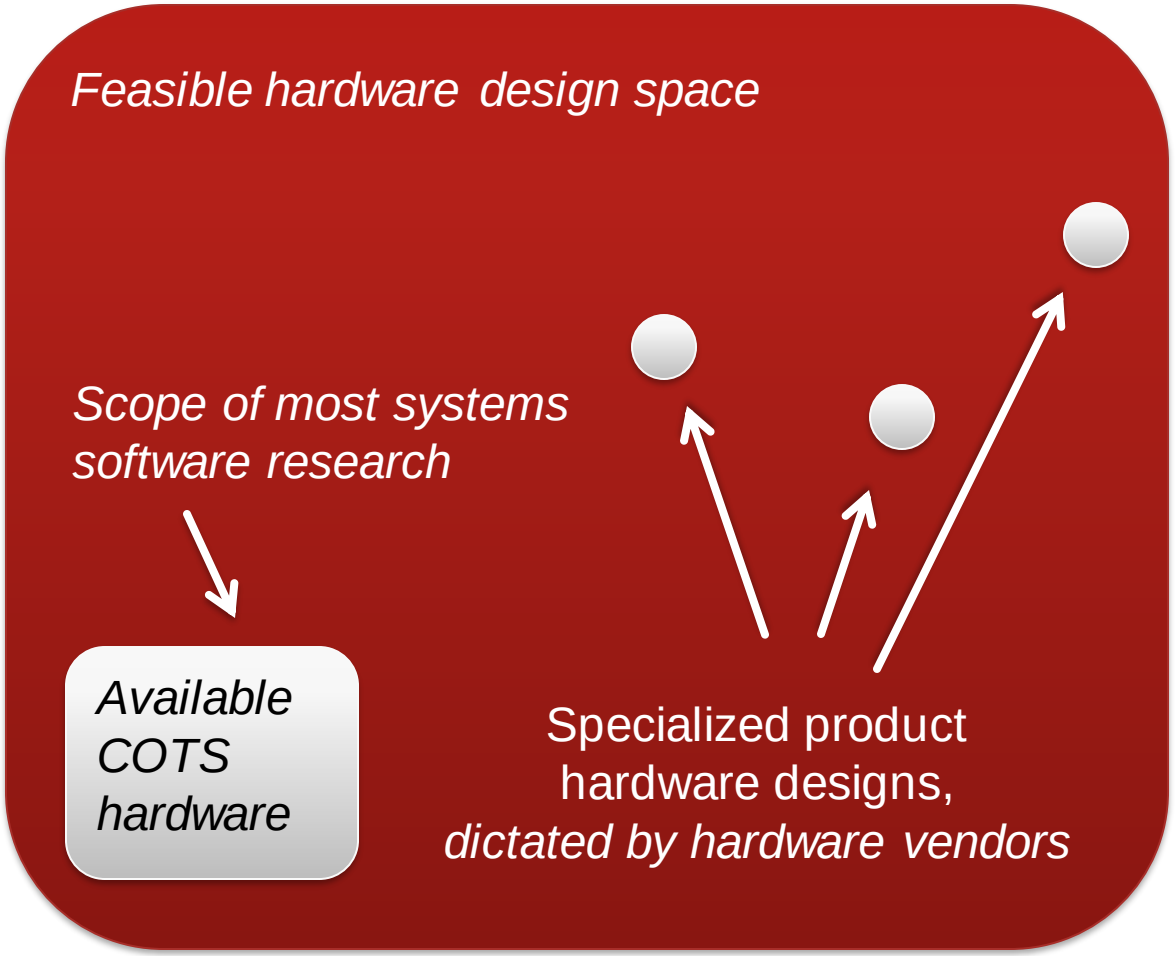


ap
TPU
deployments
.g. PureStorage)



David Cock, Systems Group
Department of Computer Science
ETH Zurich

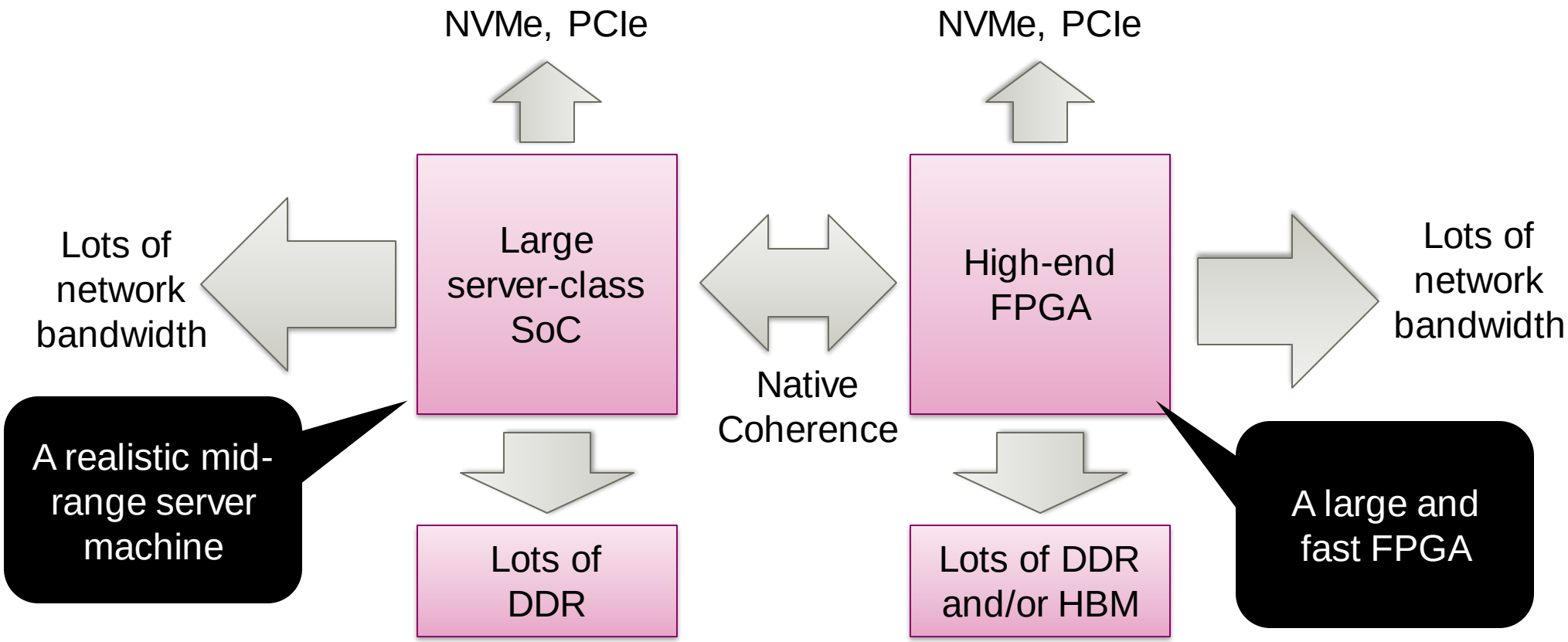
What if we had…



- A research platform for system software
  - No unrealistic commodity platforms
  - Leapfrog cost-optimized specialized designs
- Optimized for exploration
  - Not for unit cost or performance/watt
  - Overengineered rel. products
  - Flexibile and reconfigurable
- Viable rackscale building block
  - Able to function as a variety of different components at scale

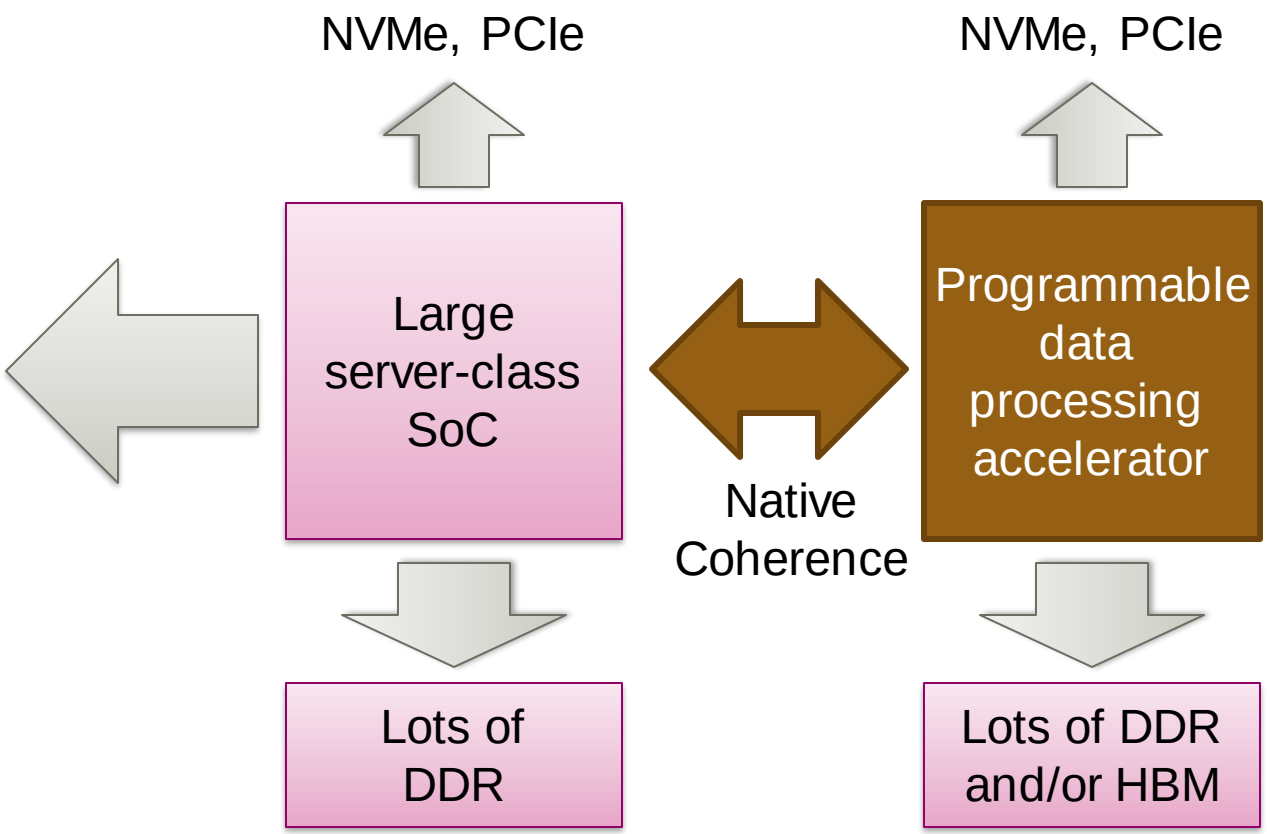
David Cock, Systems Group
Department of Computer Science
ETH Zurich

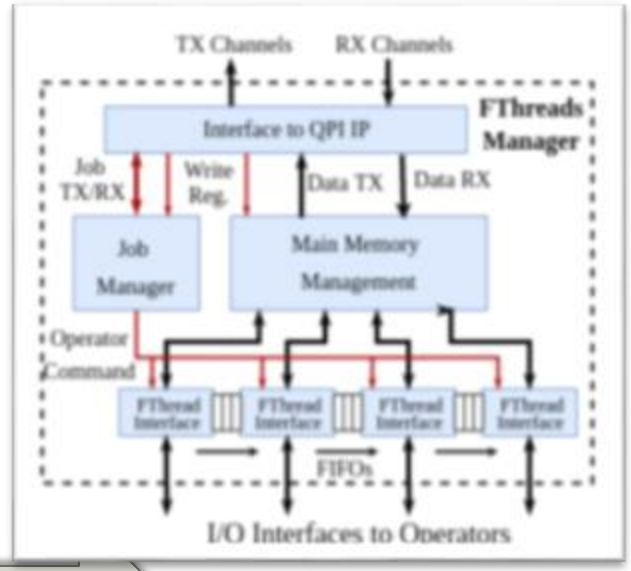
Sketch: the basic building block

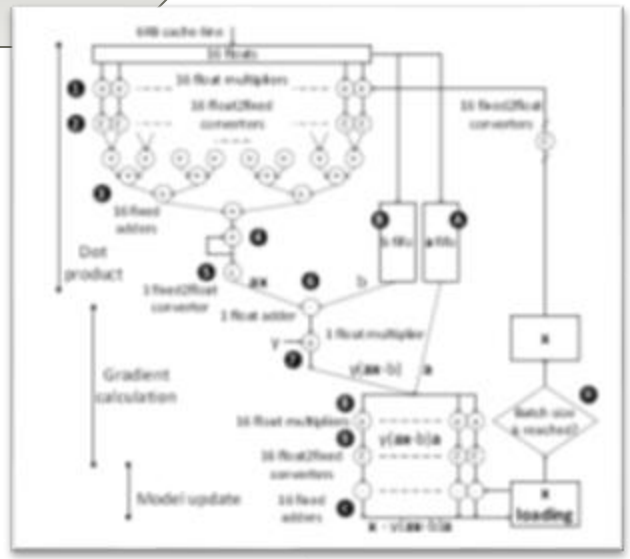


David Cock, Systems Group
Department of Computer Science
ETH Zurich

Applications
Accelerator design

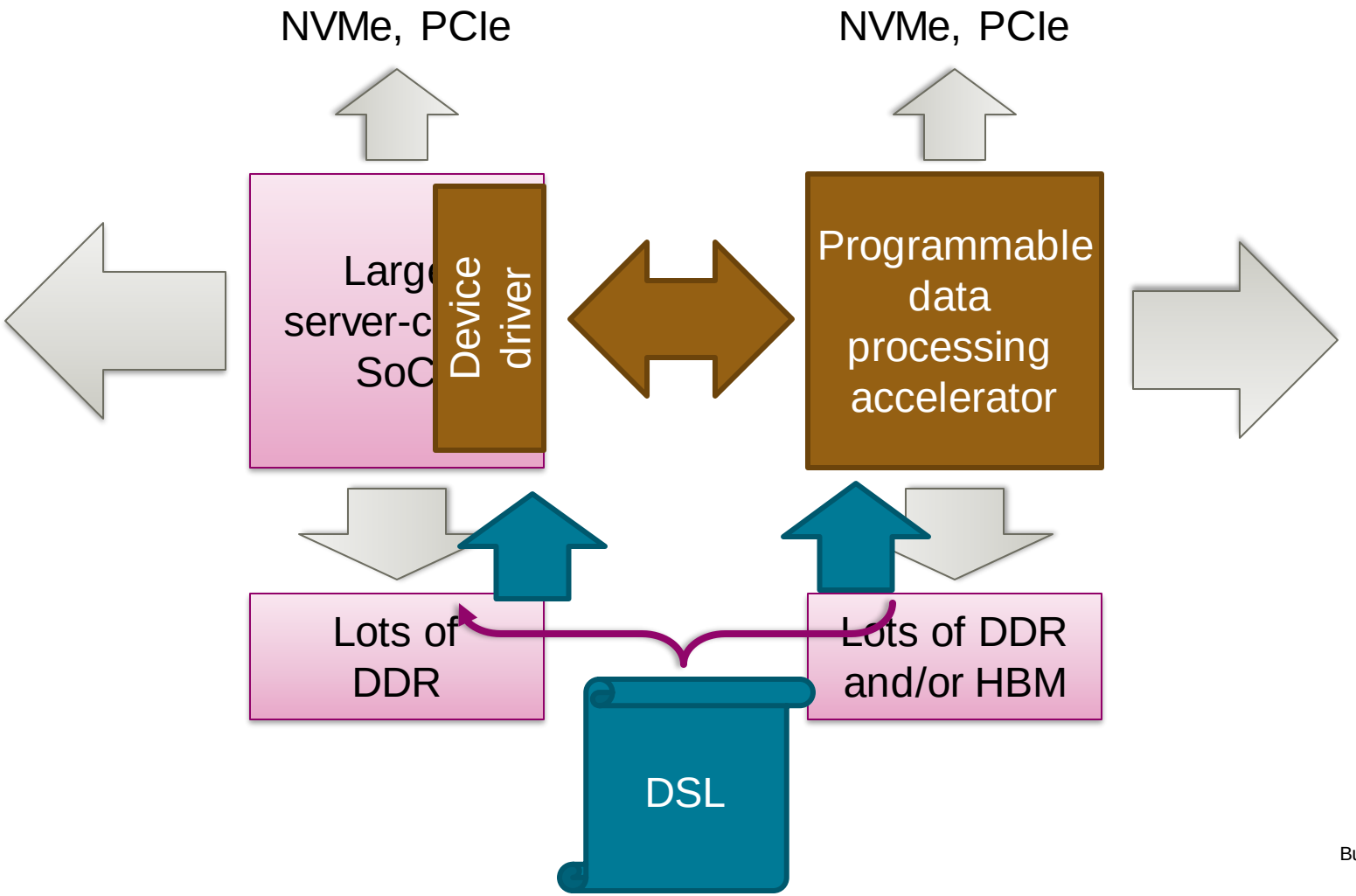






David Cock, Systems Group
Department of Computer Science
ETH Zurich

Applications
Hardware/software co-synthesis



David Cock, Systems Group
Department of Computer Science
ETH Zurich

Applications
Smart storage devices

NVMe, PCIe

Large server-class SoC

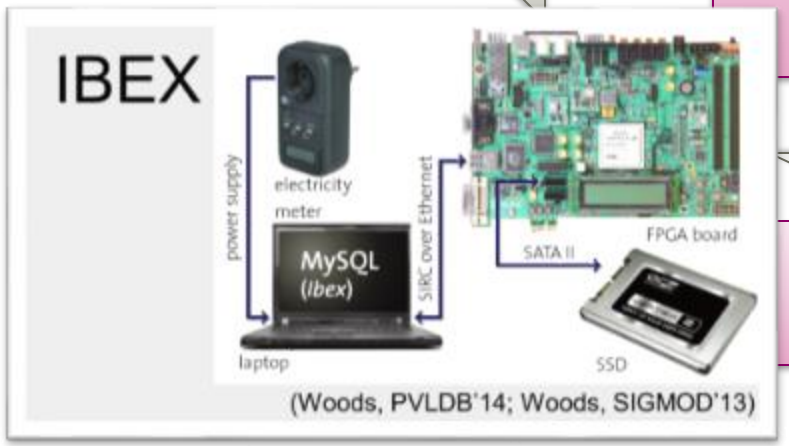
Lots of DDR

Native Coherence

NVMe, PCIe

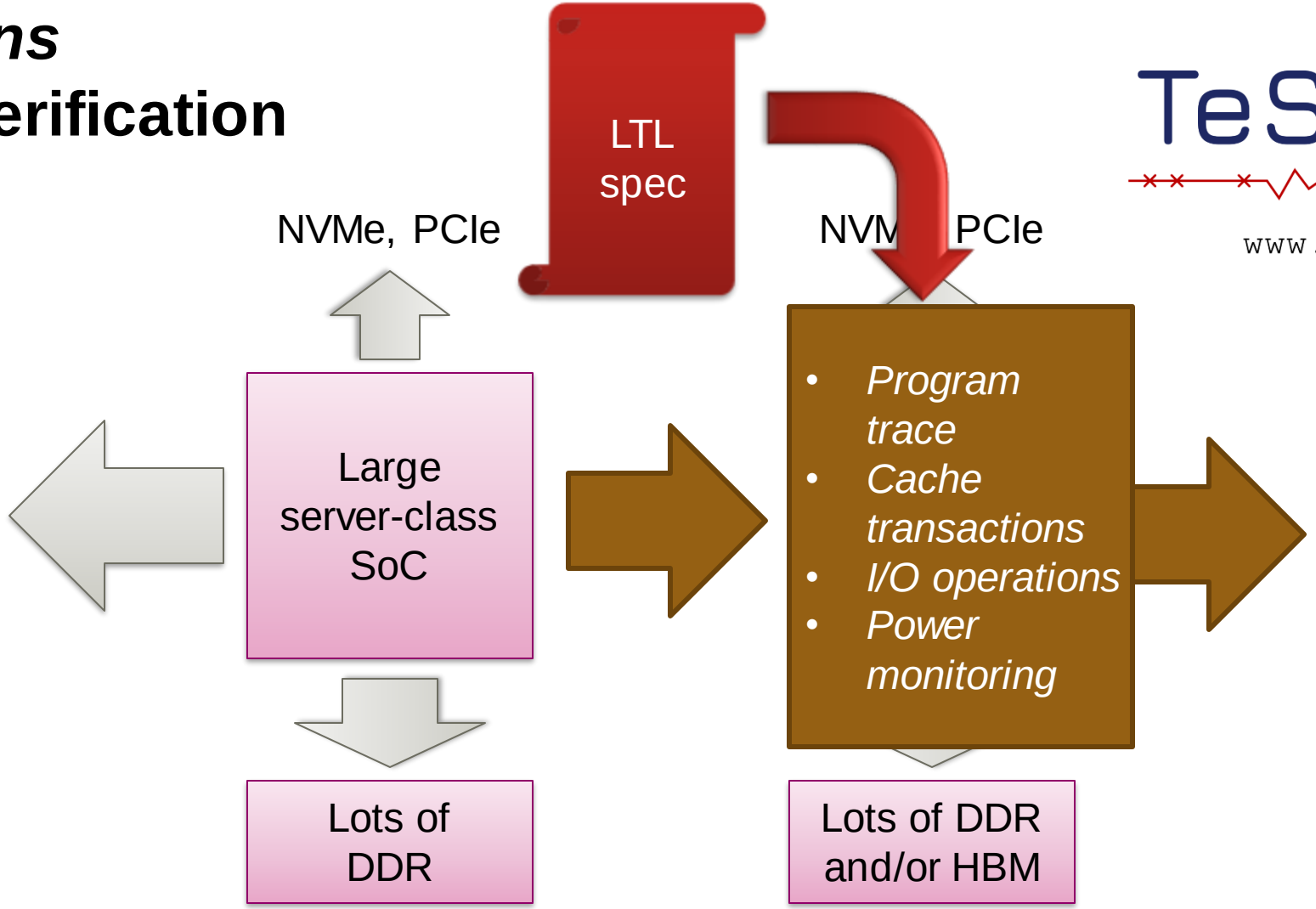
- *Scans*
- *Indexing*
- *Prefetching*
- *etc.*

Lots of DDR and/or HBM



David Cock, Systems Group
Department of Computer Science
ETH Zurich

Applications
Runtime verification



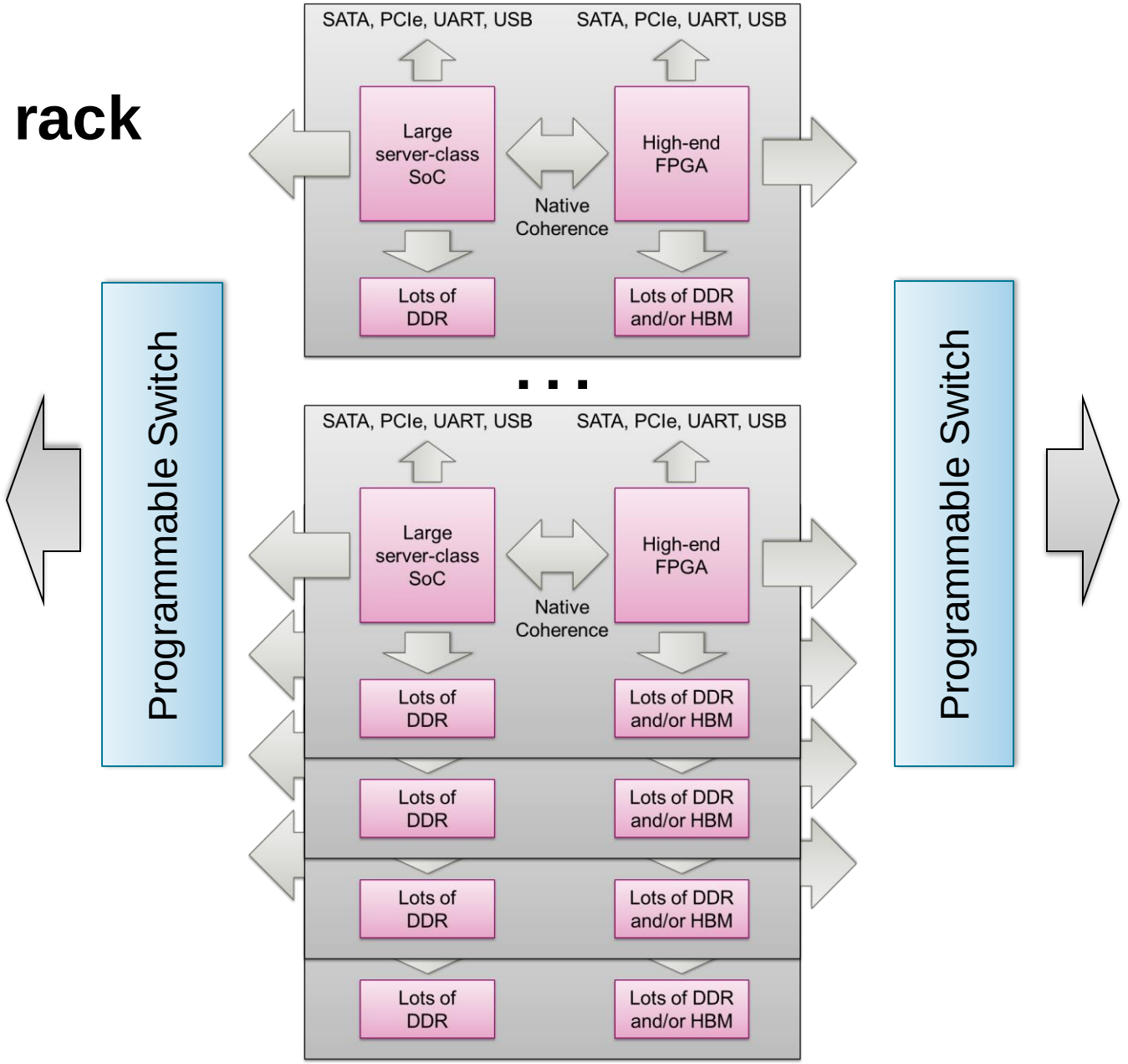
TeSSLa
www.tessla.io

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Applications
Runtime verification

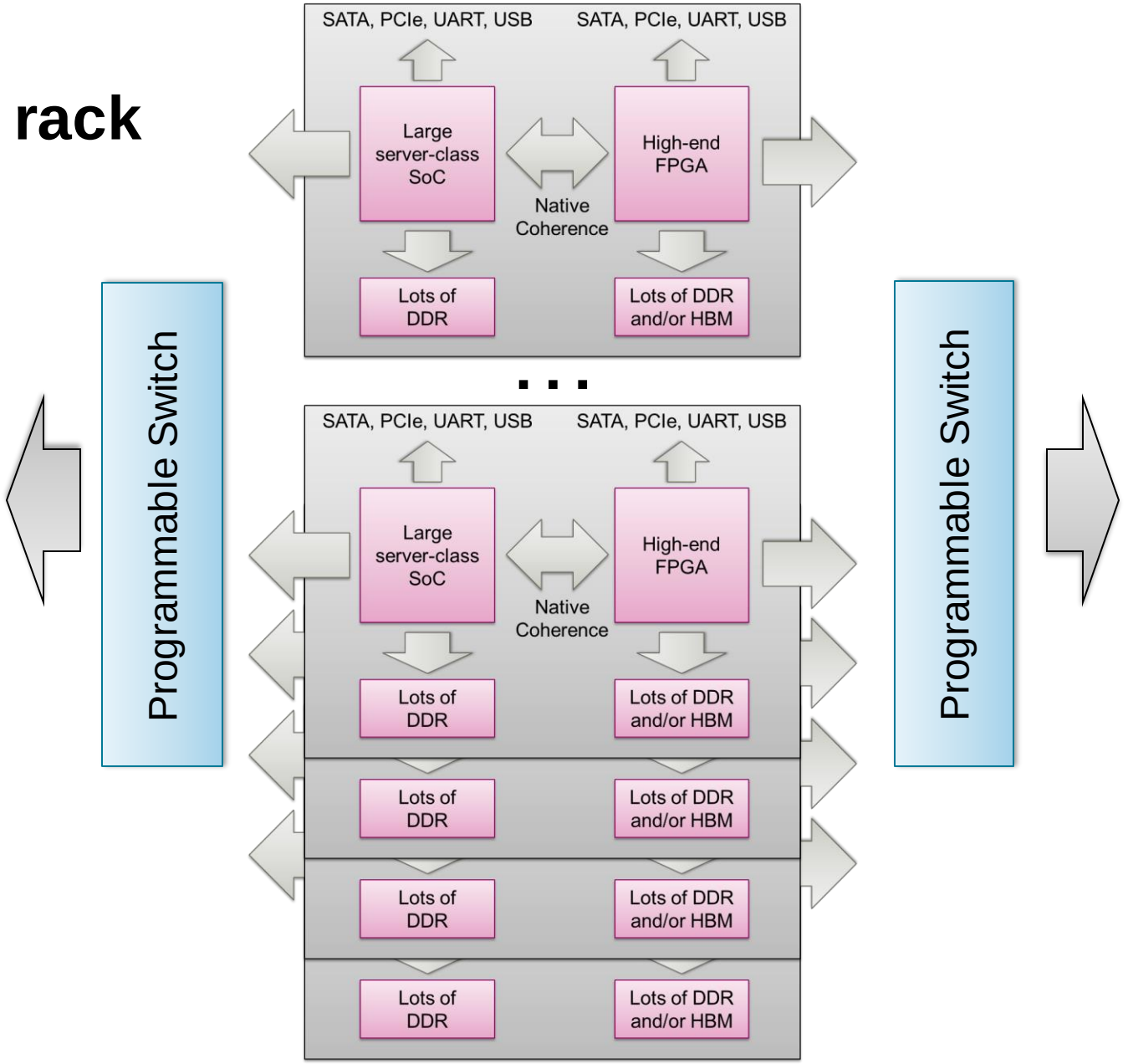
- TeSSLa runtime monitoring framework merged with our runtime verification system by a great student (Pirmin Schmid).
- Port to Enzian ongoing (trace via ECI).
- Joint project with Lübeck.

But imagine:
50 of these in a rack



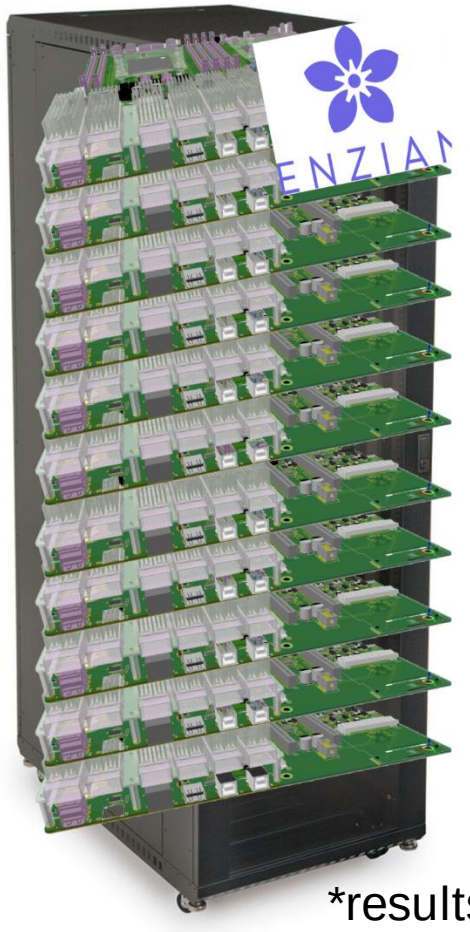
David Cock, Systems Group
Department of Computer Science
ETH Zurich

But imagine:
50 of these in a rack

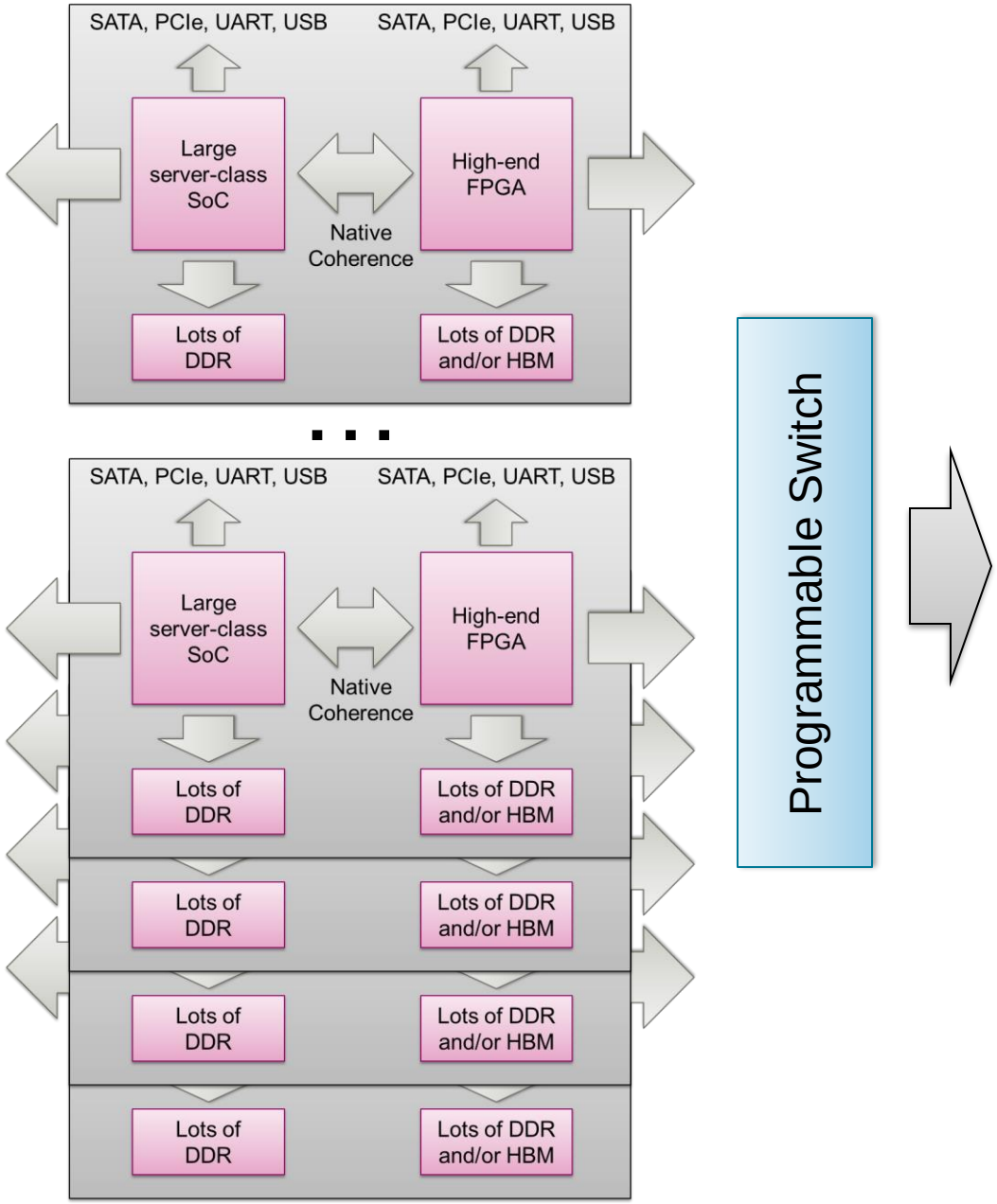


David Cock, Systems Group
Department of Computer Science
ETH Zurich

But imagine:
50 of these in a rack

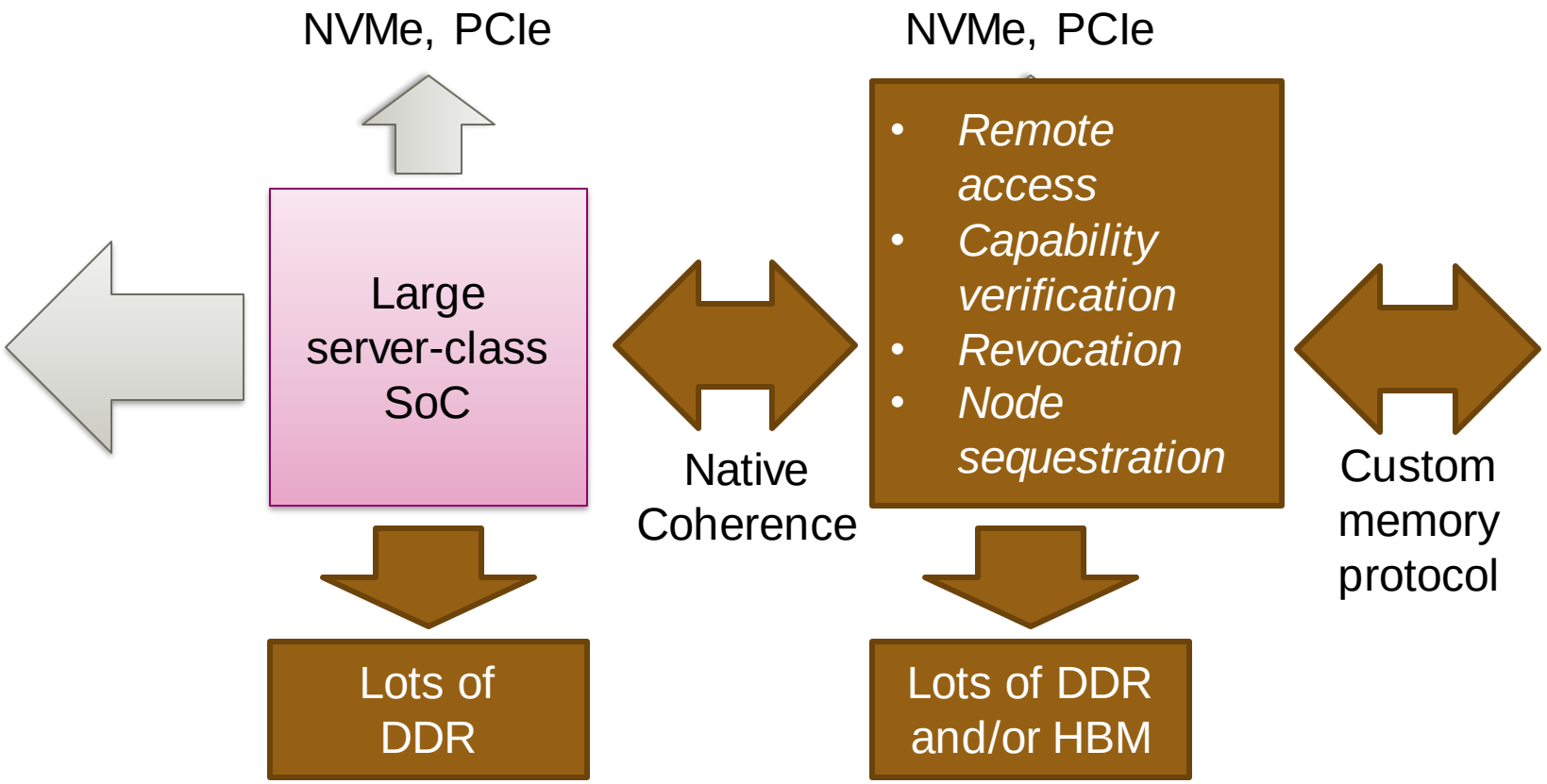


*results may vary

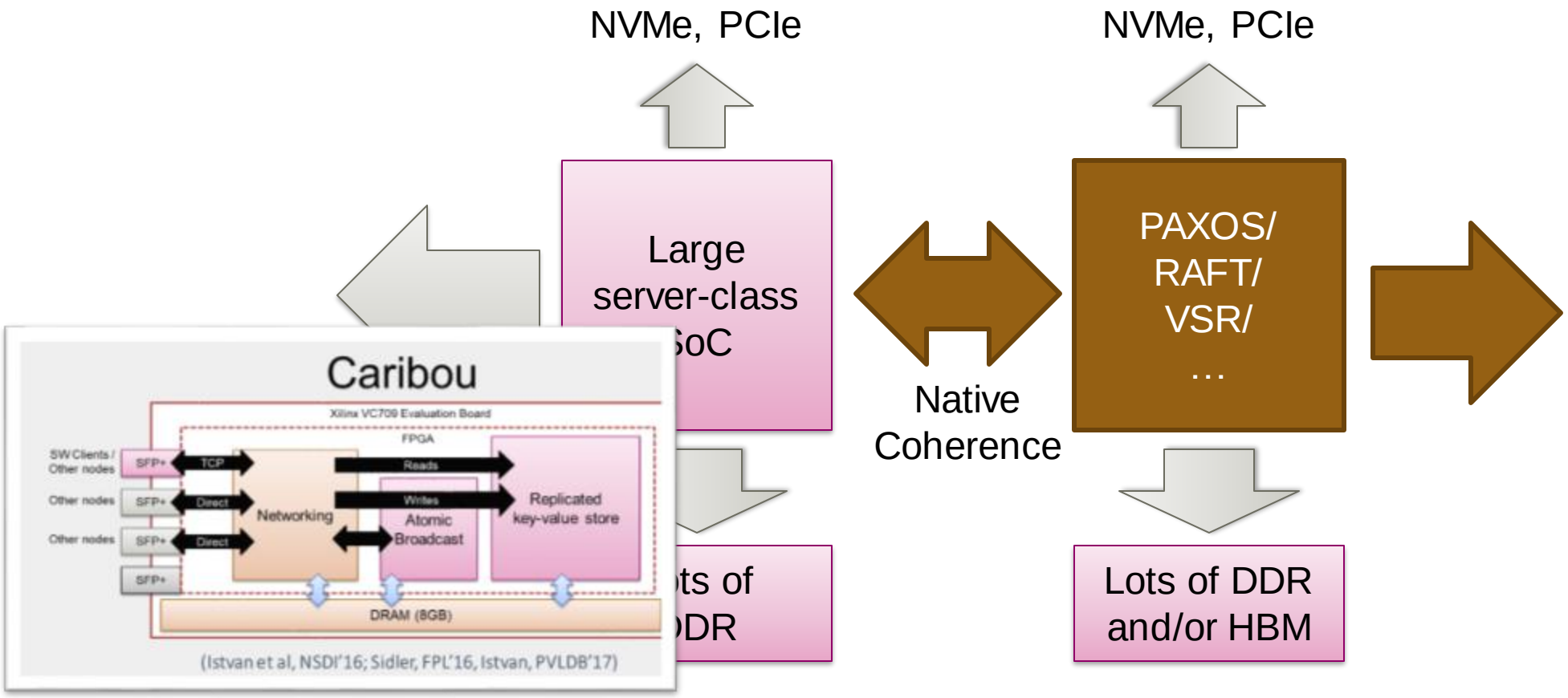


David Cock, Systems Group
Department of Computer Science
ETH Zurich

Applications
Remote memory, and a cure for RDMA



Applications
Offloading consensus



Obviously, this is an impossible dream

- No company will build this for us
  - We asked many: they all said either “No!”, or “Why? No!”
- Universities obviously can’t build real computers any more
  - Sheer complexity of 18-layer 5Ghz board design
  - Cost of design resources

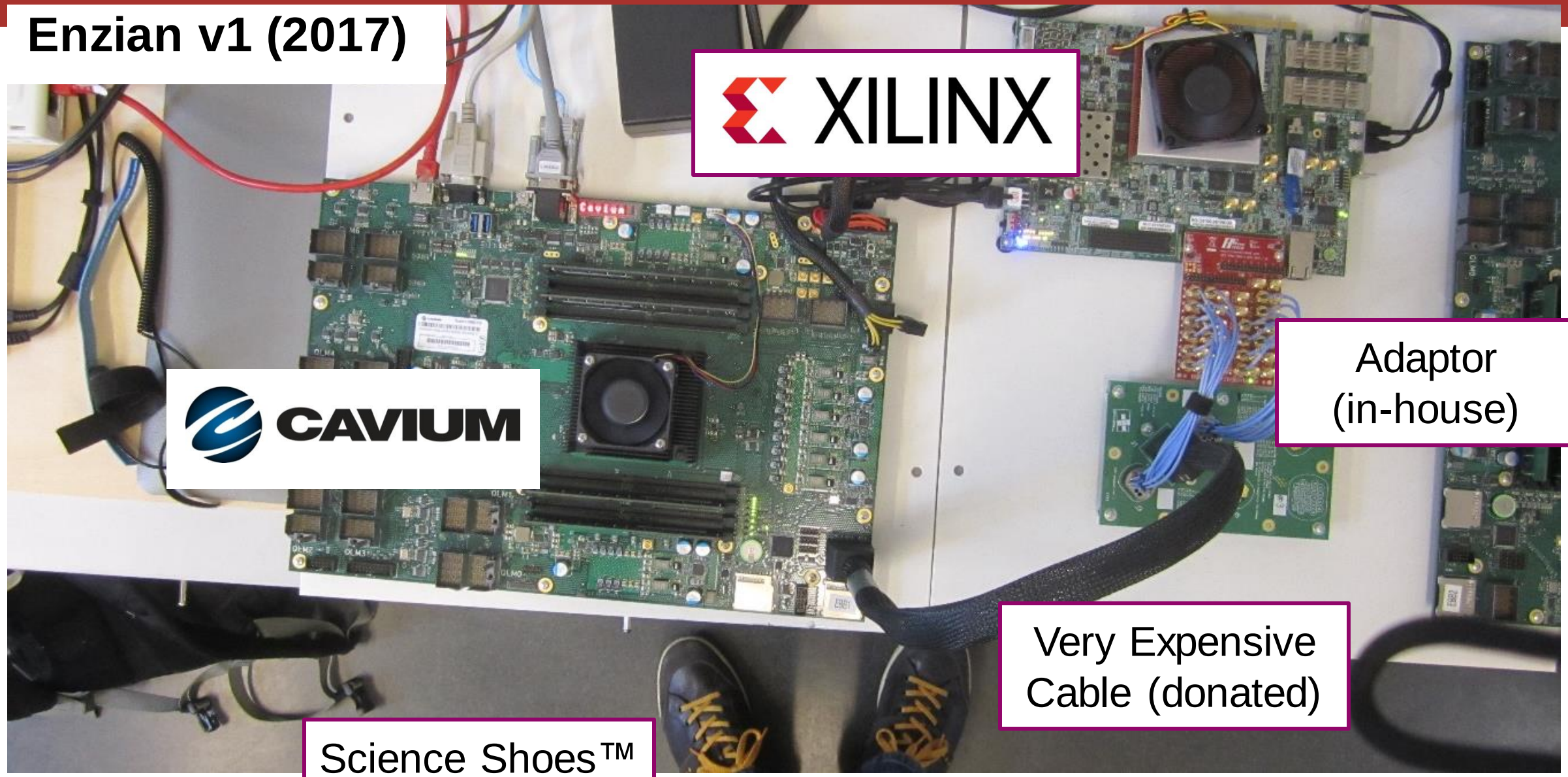
Obviously, this is an impossible dream

- No company will build this for us
  - We asked many: they all said either “No!”, or “Why? No!”
- Universities obviously can’t build real computers any more
  - Sheer complexity of 18-layer 5Ghz board design
  - Cost of design resources

Or not…

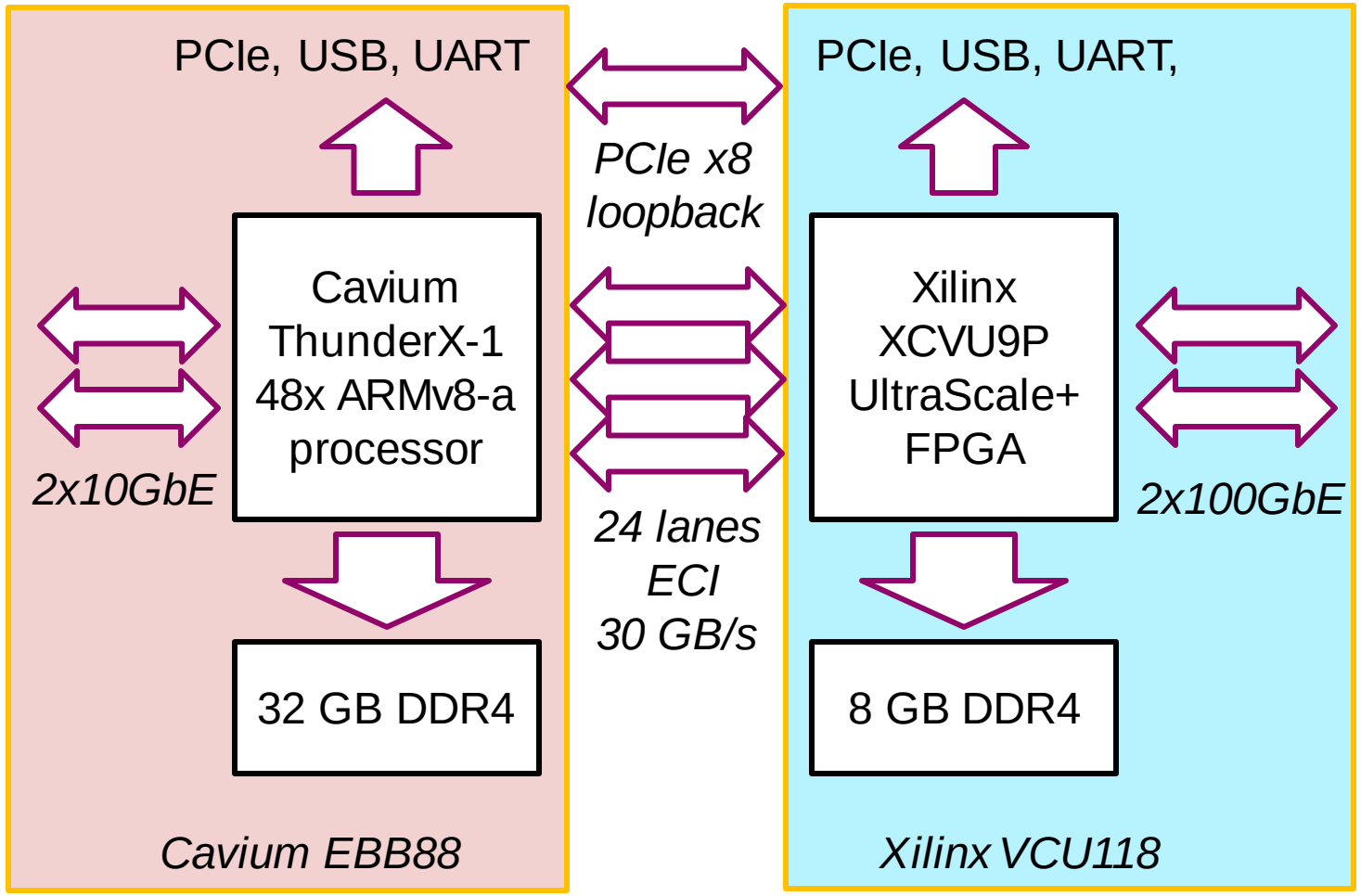
- Friends in the EE department who can help us build small boards
- Outsource final board design to contractors
- Get technical help from friends in industry
- High-end CAD tools available to EU universities at deep discounts

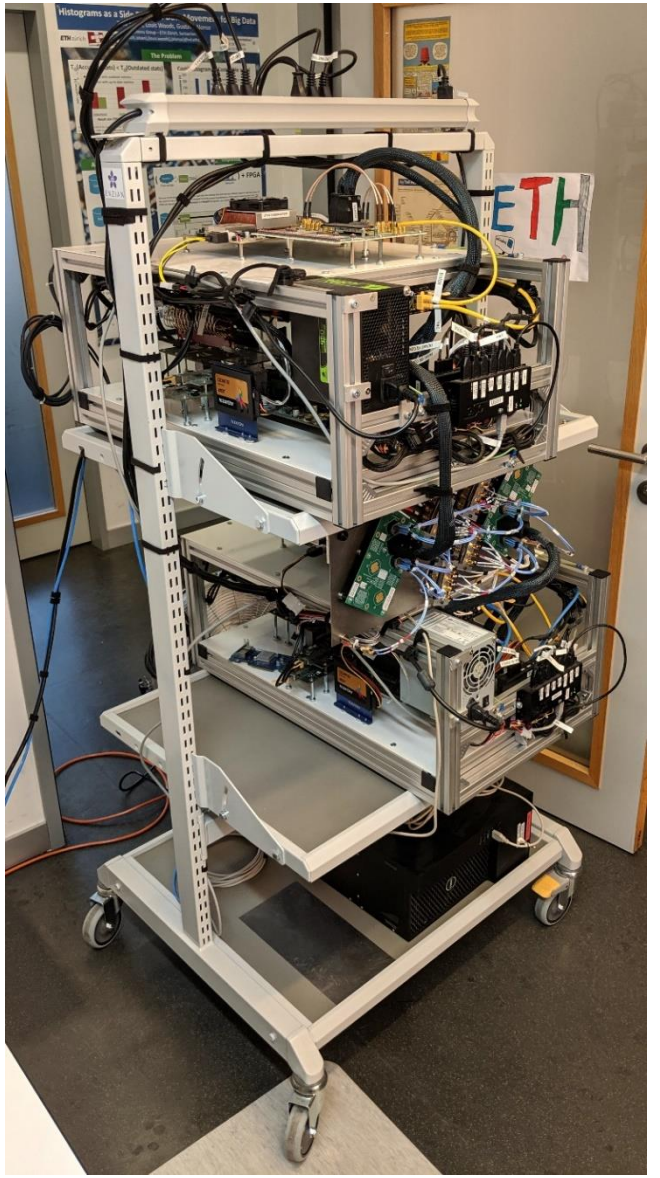
Enzian v1 (2017)



David Cock, Systems Group
Department of Computer Science
ETH Zurich

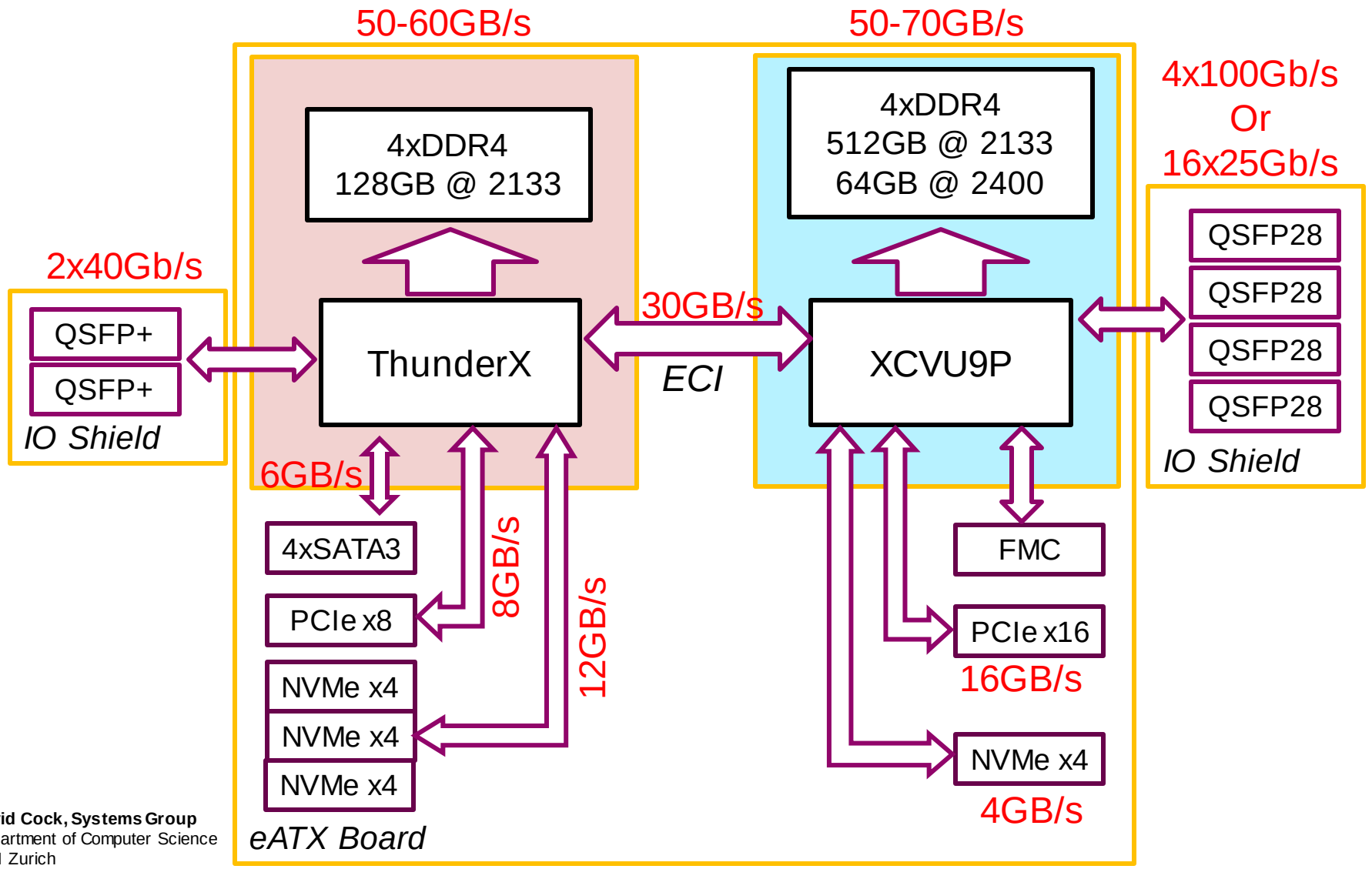
Enzian v2 (2018)

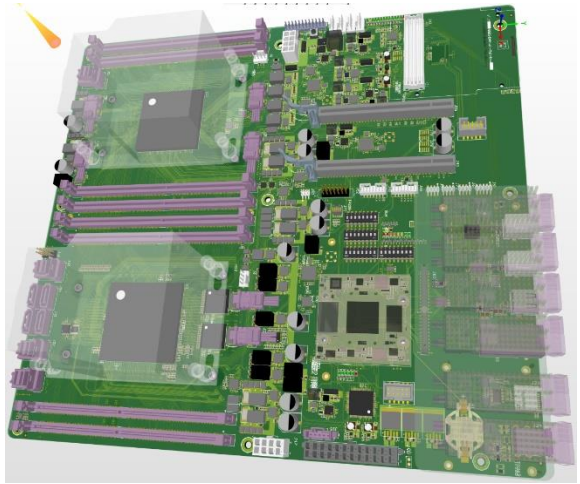




David Cock, Systems Group
Department of Computer Science
ETH Zurich

Enzian v3 (2019)



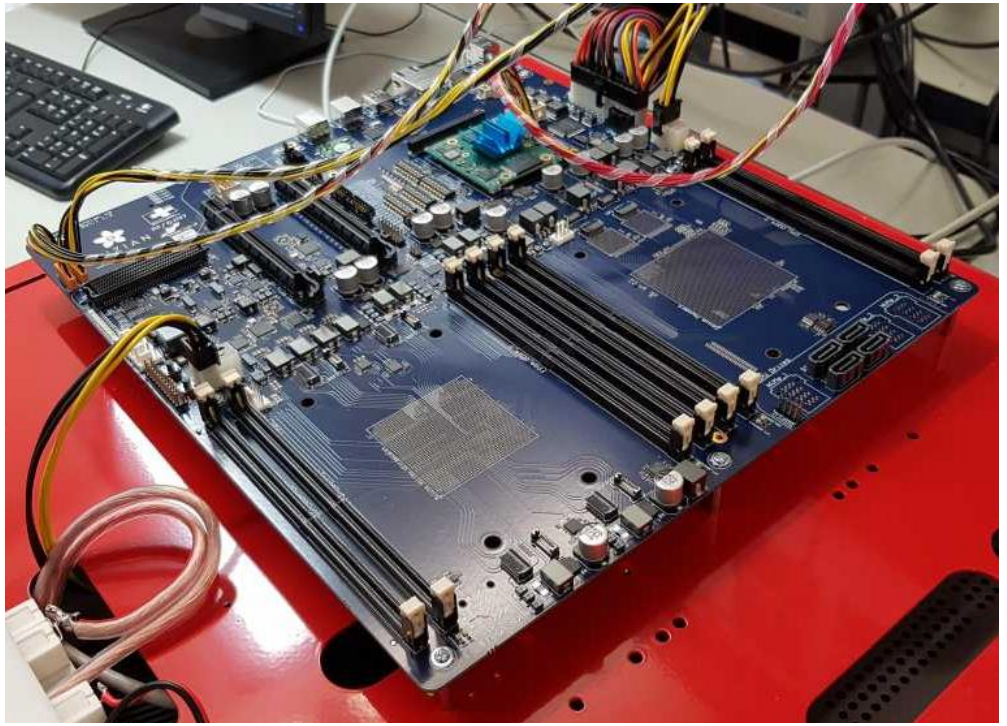


ENZIAN

David Cock, Systems Group
Department of Computer Science
ETH Zurich

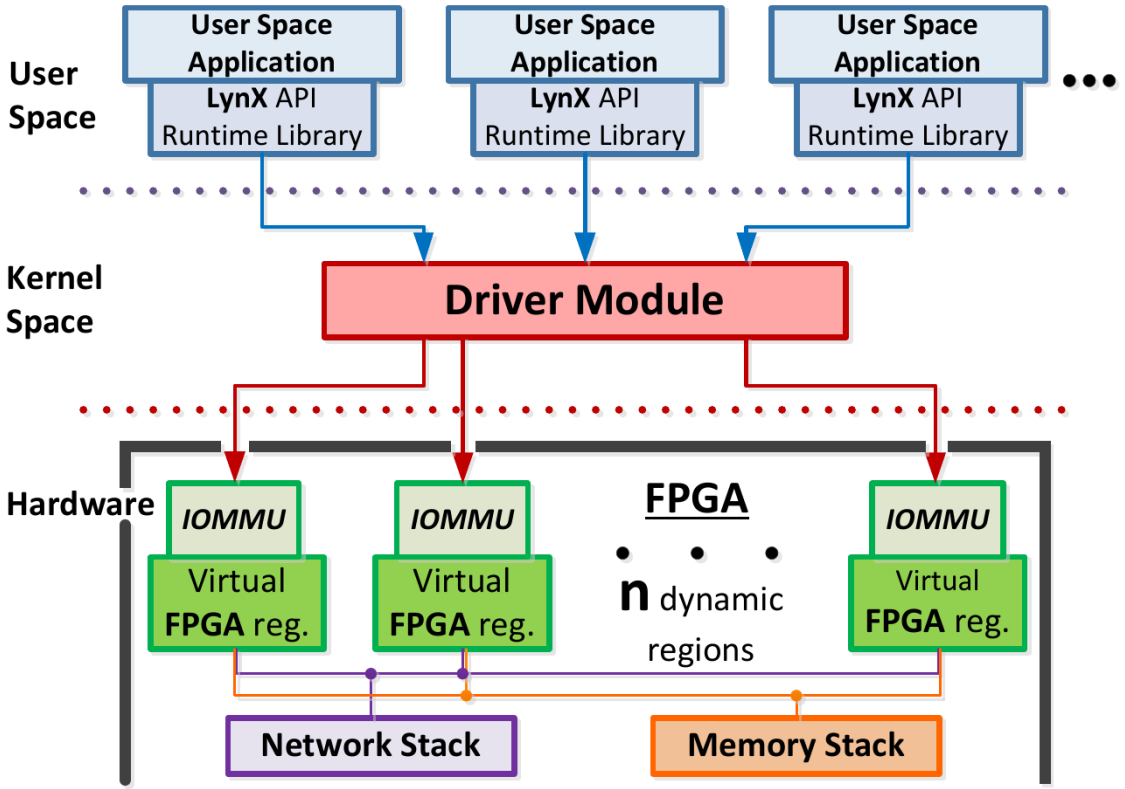
The First Live* Specimen

*TBD



David Cock, Systems Group
Department of Computer Science
ETH Zurich

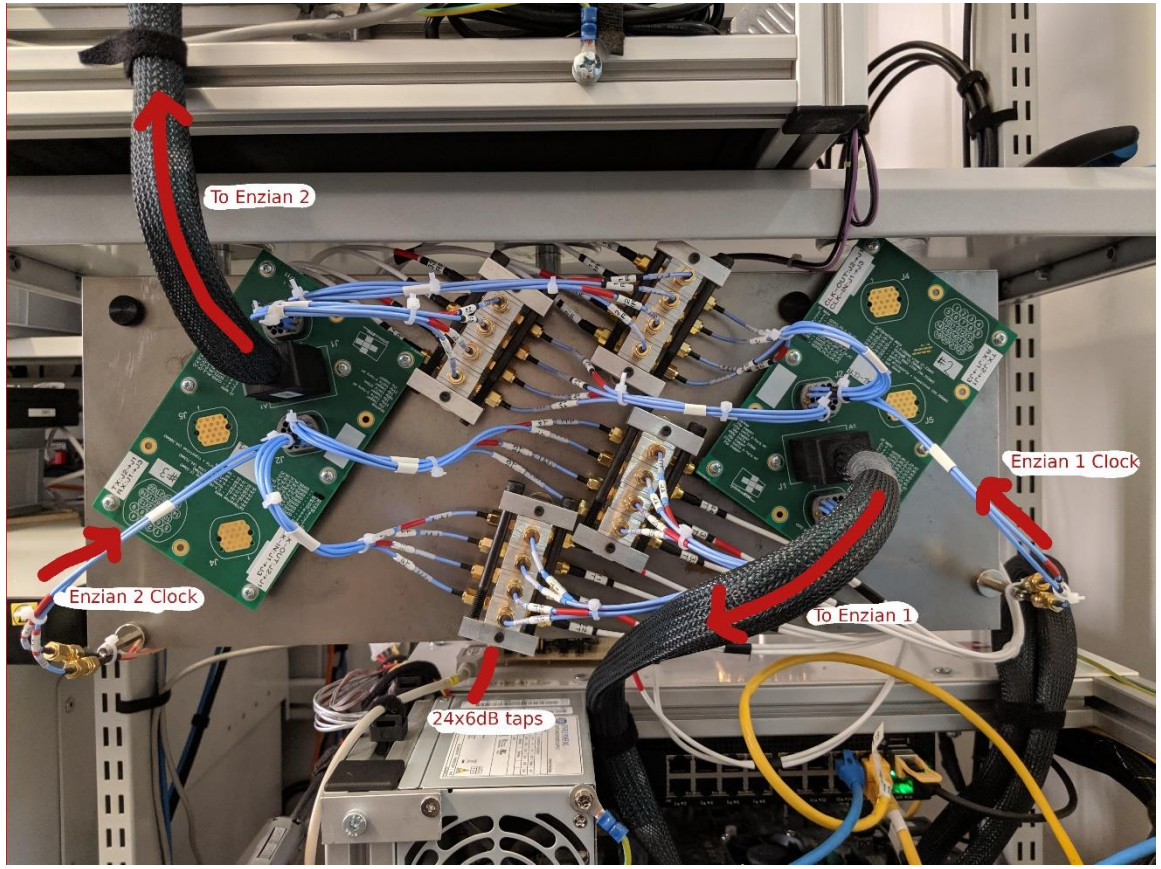
LynX: *A* Shell for Enzian



- With your Enzian, you will get free training wheels.
- You won't *have* to reinvent everything straight away.
- Seriously: these guys (Dario Korolija, Gustavo Alonso, ...) know what they're doing.

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Lessons
Debugging a Fast Interconnect

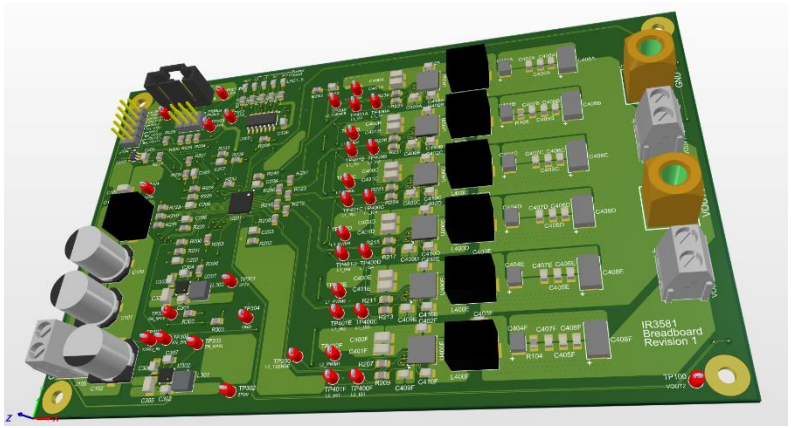


- Turns out, sometimes engineers don't quite get around to writing everything down.
- "The only thing that'll every talk to our chip is our chip."
- …
- *sigh*

Important note: the designers have been *super* helpful.

David Cock, Systems Group
Department of Computer Science
ETH Zurich

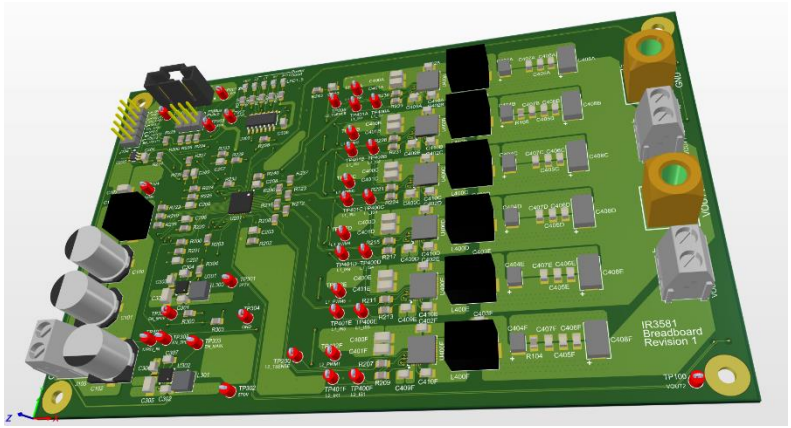
Lessons
Firmware Complexity



- Firmware is **far** more complex than it looked
- One regulator alone has hundreds of registers
- Hard-to-access, poorly-documented tools

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Lessons
Firmware Complexity

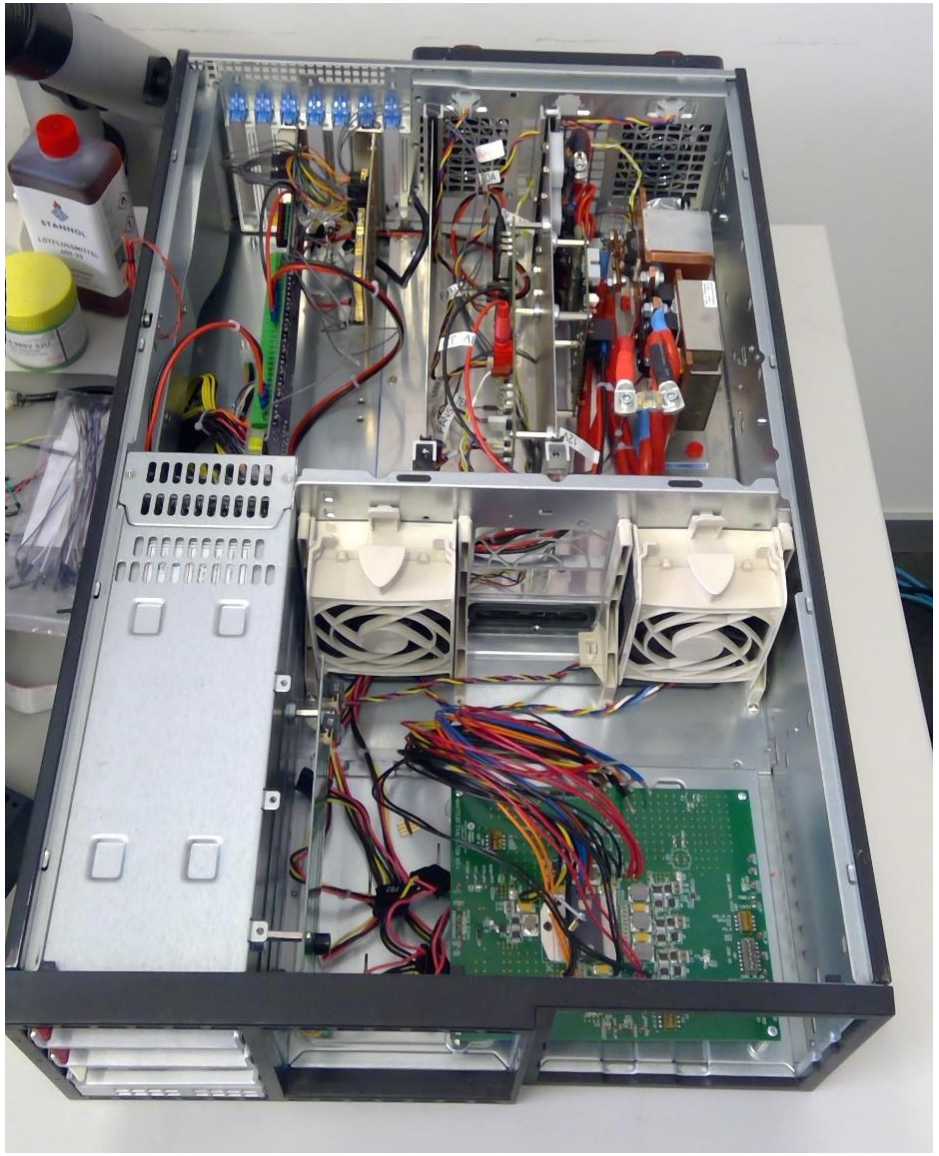


- Firmware is **far** more complex than it looked
- One regulator alone has hundreds of registers
- Hard-to-access, poorly-documented tools

David Cock, Systems Group
Department of Computer Science
ETH Zurich

Lessons
War stories

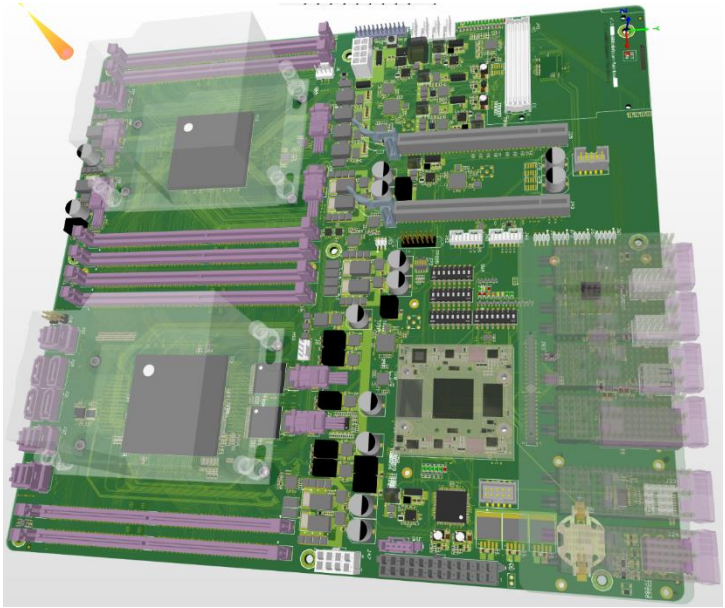
- Hardware is really complex – who knew?
  - 18+-layer board, 5GHz…
  - Weekly phone calls with the board designers
- Procuring chips
  - Complex bureaucracies, POs, quotes,
  - Lots of NDAs just to get documentation
- Friends have been super helpful
  - Donations, technical help from Xilinx and Cavium
  - Dream Chip, our board contractors
  - Colleagues in Electrical Engineering at ETH
- Delays, delays, delays…



David Cock, Systems Group
Department of Computer Science
ETH Zurich

Summary

Systems research needs its own hardware!



The mission:

- We need **overengineered research platforms**
- Our research should deliver **options** and **techniques**

The machine:

- Server-class **SoC**, large **FPGA**, plenty **RAM** & **b/w**
- **Balanced**, **coherent** system

Current status:

- Test boards exist (respin after testing).
- Larger volumes after that (if it works)
- Open source design as much as possible

Who wants one? Who wants 50? Come to our demo.

ENZIAN

www.enzian.systems

David Cock, Systems Group
Department of Computer Science
ETH Zurich