Golden Age of FPGAs

Kubilay Atasu

September 2, 2015 @ FPL

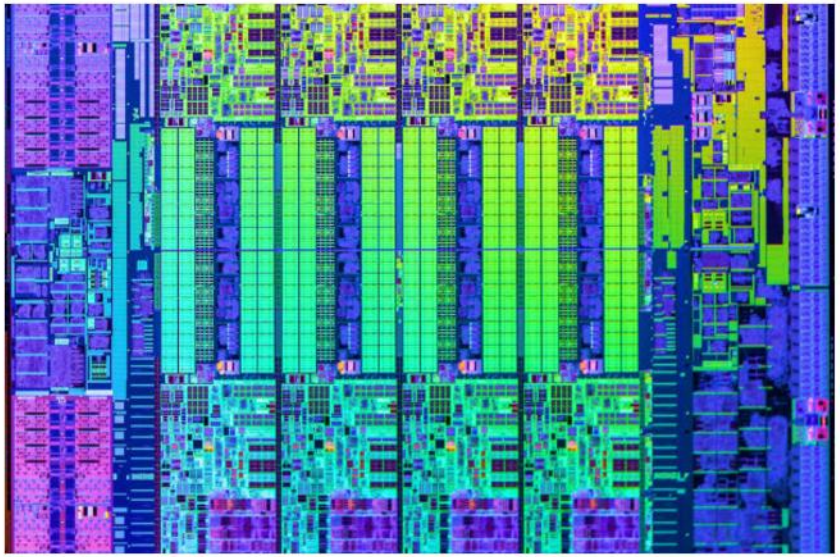
Disclaimer

Any views or opinions expressed in this talk are solely those of the author and do not necessarily represent those of IBM.

End of technology scaling

Intel TOCK BLOCK: 10nm Cannonlake delayed to 2017, bonus 14nm Kaby Lake to '16

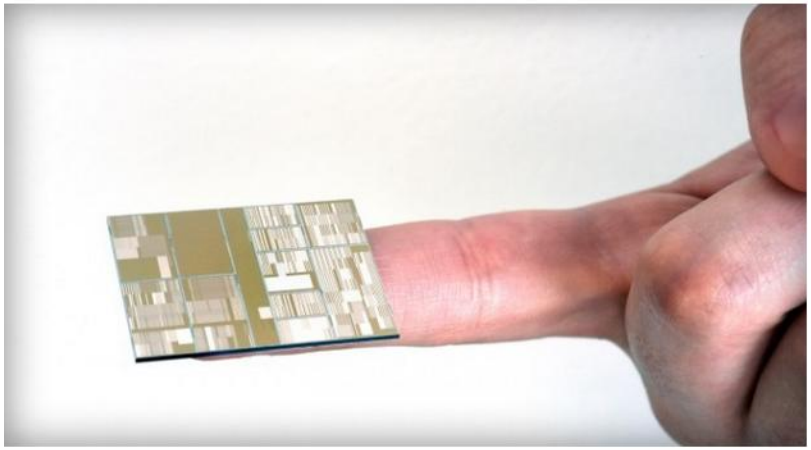
Chipzilla chiefzilla admits tiny gates are really hard



IBM gets green light to sell off chips biz to GlobalFoundries

US regulators furrow brows, frown, nod

IBM announces 7nm breakthrough, builds first test chips on new process



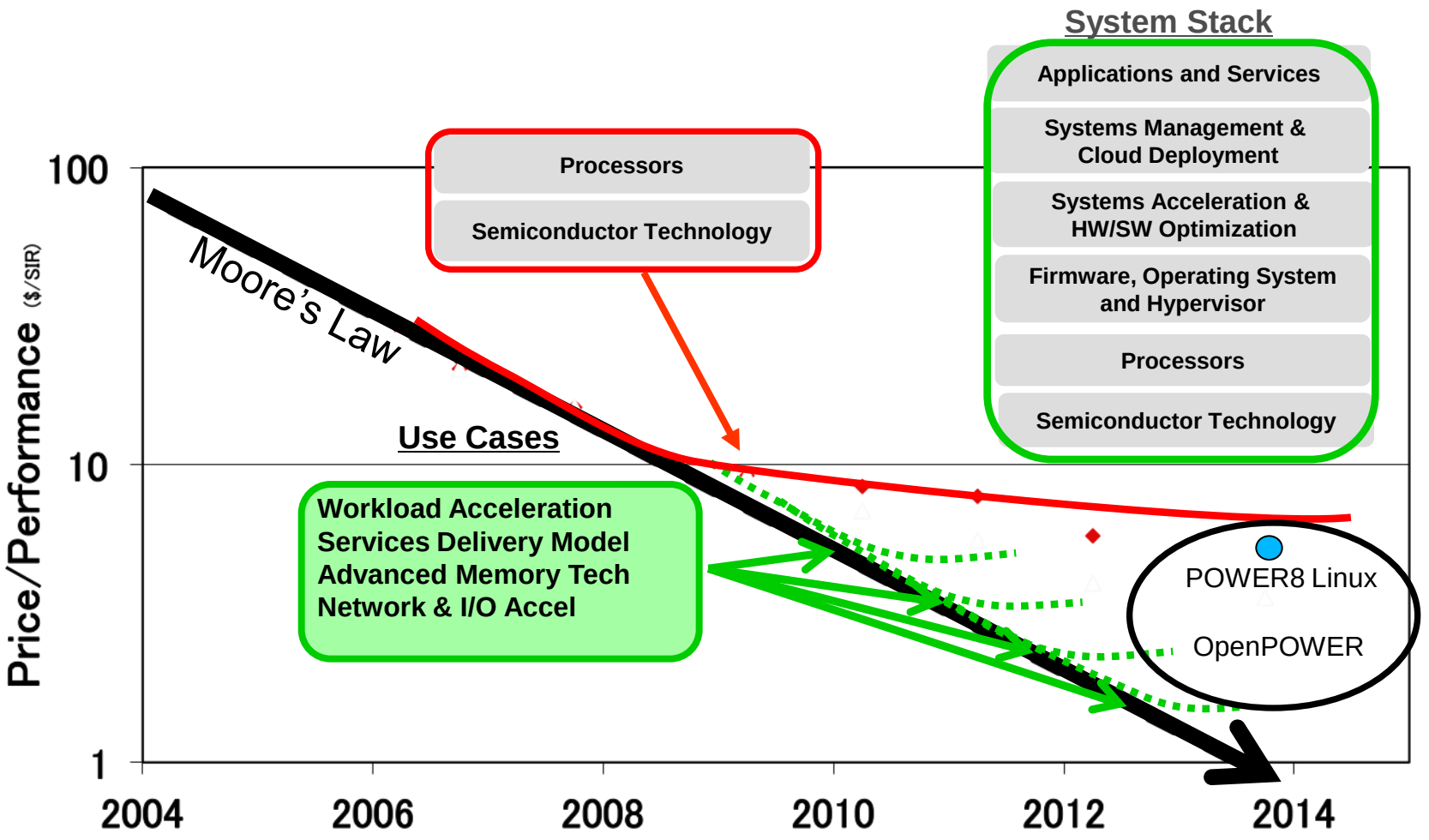
Challenges:

- leakage current
- EUV lithography
- low yield
- high cost

Sources: theregister.co.uk, extremetech.com

Industry trends create new opportunities

Microprocessors alone no longer drive sufficient Cost/Performance improvements

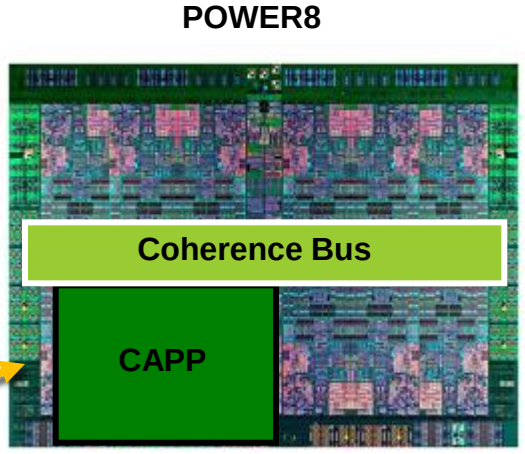


System stack innovations are required to drive Cost/Performance

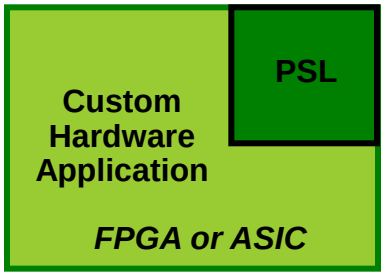
Coherent Accelerator Processor Interface (CAPI)

Virtual Addressing
- Accelerator can work with same memory addresses that the processors use
- Pointers de-referenced same as the host application
- Removes OS & device driver overhead

Hardware Managed Cache Coherence
- Enables the accelerator to participate in “Locks” as a normal thread
- Lower Latency over IO communication model



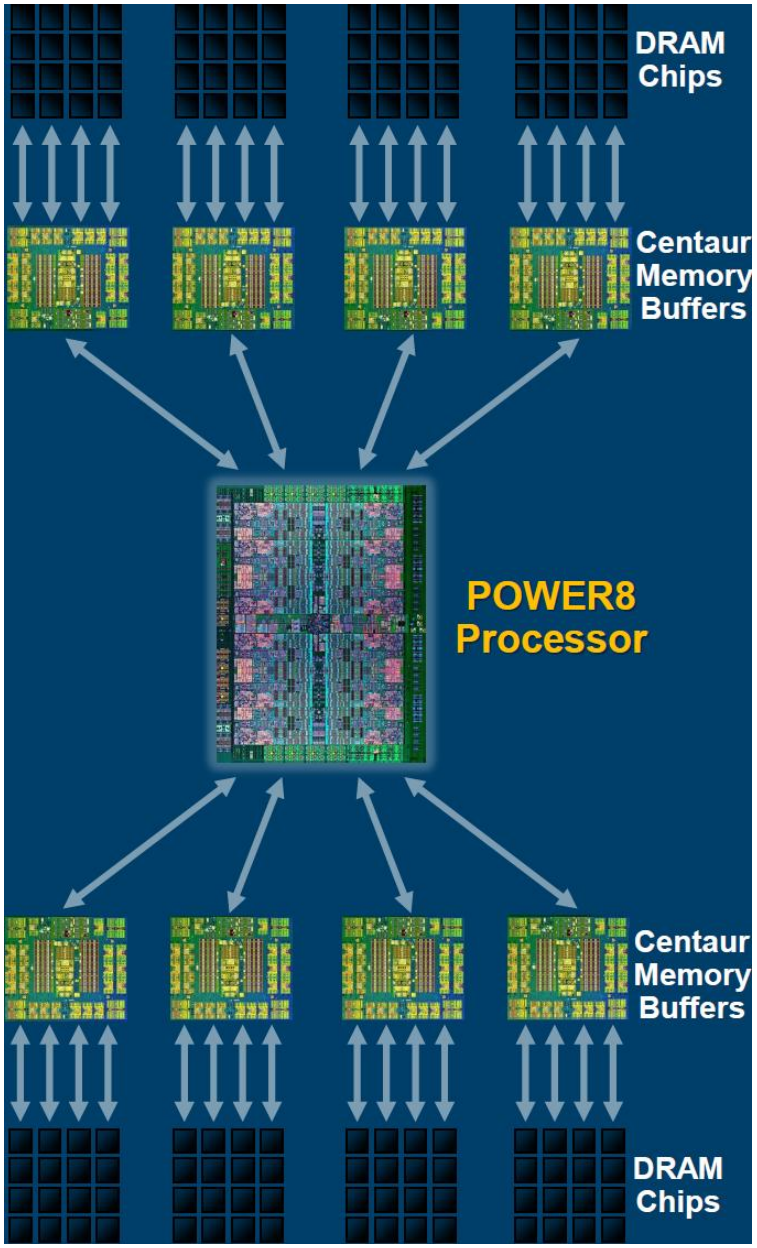
PCIe Gen 3
Transport for encapsulated messages



Processor Service Layer (PSL)
- Present robust, durable interfaces to applications
- Offload complexity / content from CAPP

Customizable Hardware Application Accelerator
- Specific system SW, middleware, or user application
- Written to durable interface provided by PSL

Intelligent Memory Systems

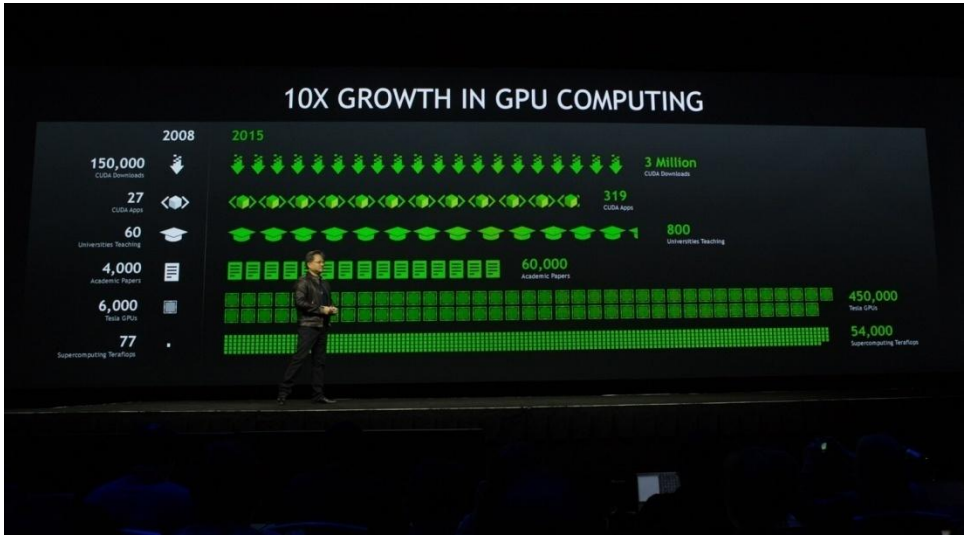


FPGA-based memory controller enables exploitation of new memory technologies (on display @ OpenPOWER Summit).

Source: http://openpowersummit2015.tumblr.com/factsheet

Prediction 1

FPL will have more than 4000 attendees before 2040!



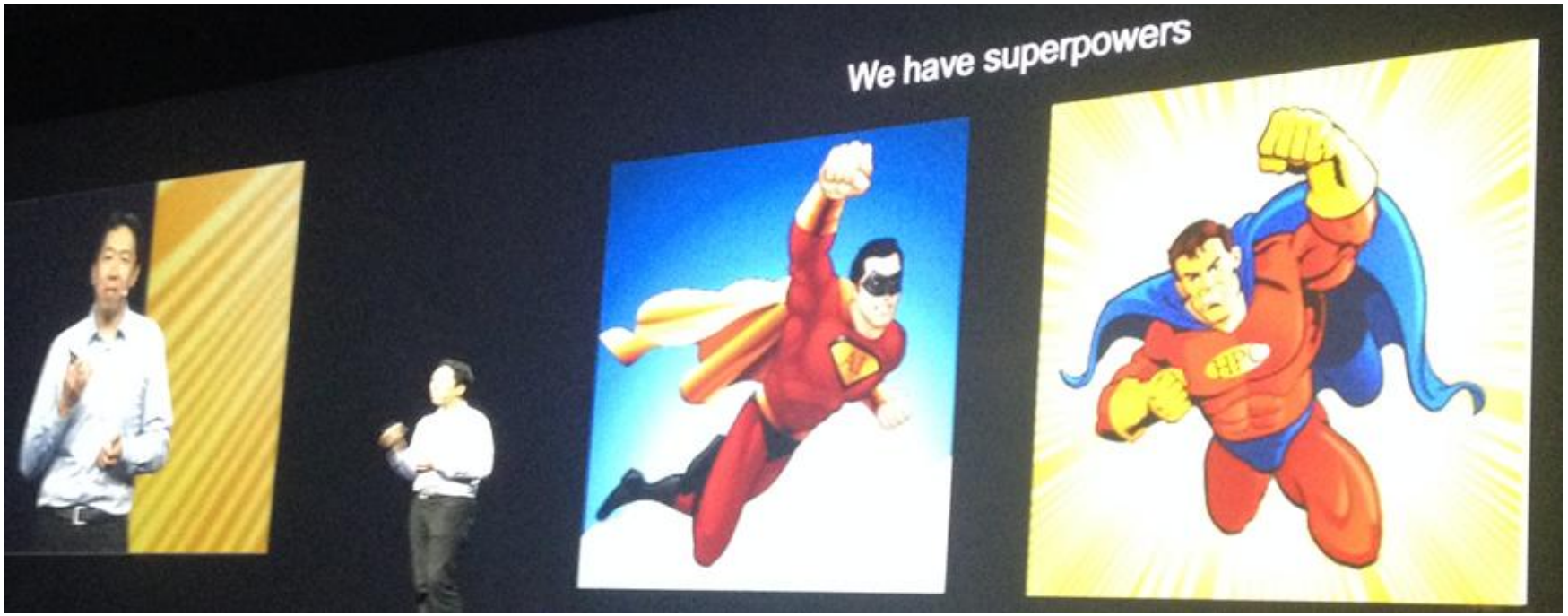
The 2015 GPU Technology Conference @ San Jose Convention Center

Not just scientists! Mostly practitioners and application developers…

Pictures taken from: http://www.gputechconf.com/

Prediction 2

Somebody will say: “Big Data is our fuel and FPGAs our engines”.



Andrew Ng at the 2015 (Sixth) GPU Technology Conference @ San Jose

But, we have to share our superpowers with application developers!

- An open source ecosystem: applications, libraries, standards…

Pictures taken from: http://www.gputechconf.com/

Prediction 3

There will be a standard high-level language for programming FPGAs

- Open source libraries and tools will develop around this language.

OpenCL **C/C++** **SystemC**

Java **Matlab** **OpenMP**

Prediction 4

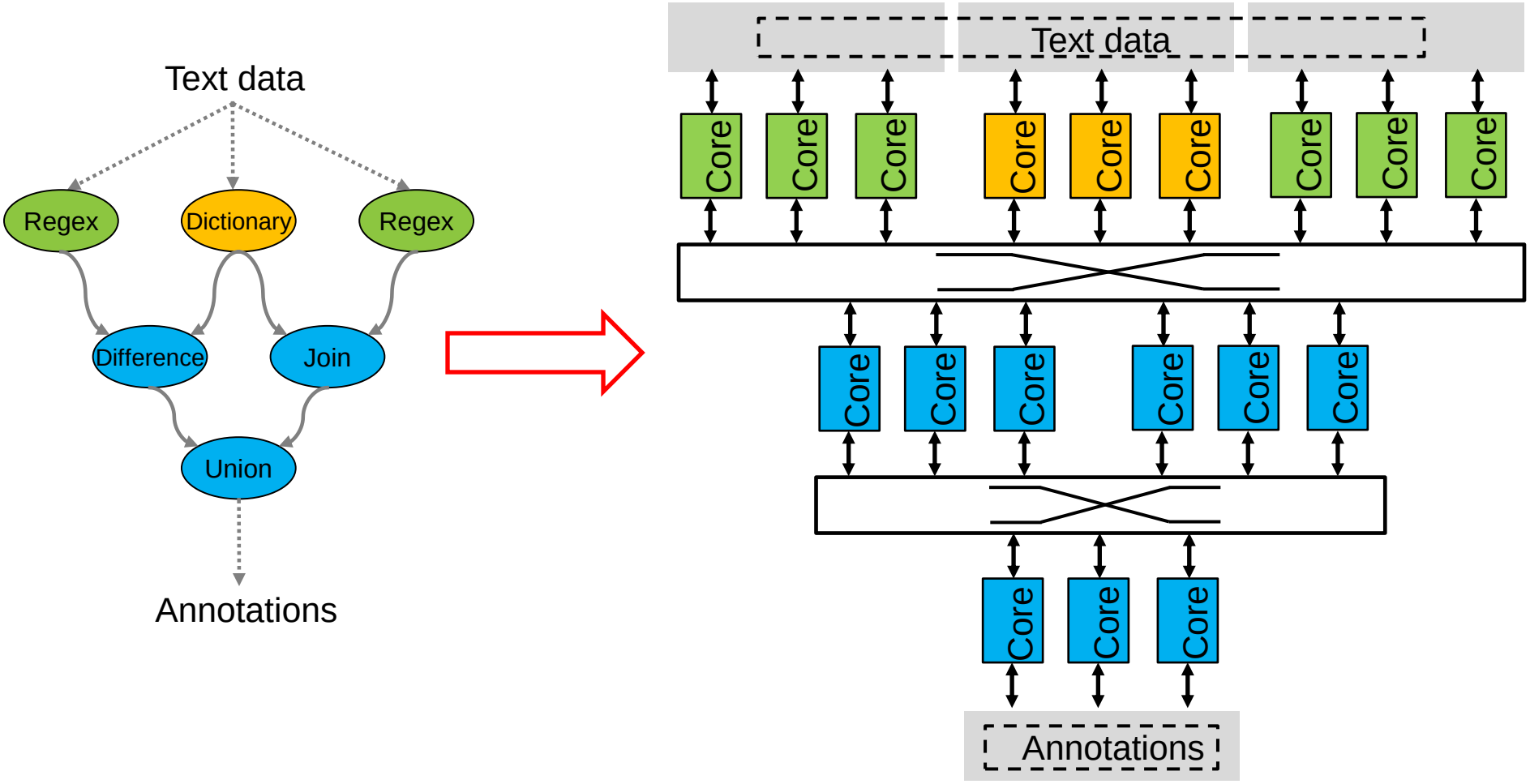
We will observe further convergence between FPGAs and GPUs:

- FPGAs already support hard floating point cores.
- GPUs already support custom precision arithmetic.

What next?

- Small FPGAs in GPUs to support irregular applications?
- Small GPUs in FPGAs optimized for matrix operations?
- FPGAs that can become GPUs (overlay architectures)?

FPGA vendors will support domain-specific overlay architectures.



Conclusions

The Golden Age of FPGAs is about to start!

It will be fun to be a part of it!