Are we going back to a "RISC era" ?

Or moving to some other recognizable era?

28 October 2015

Sverre Jarp, CERN honorary staff

RISC: Reduced Instruction Set Computing

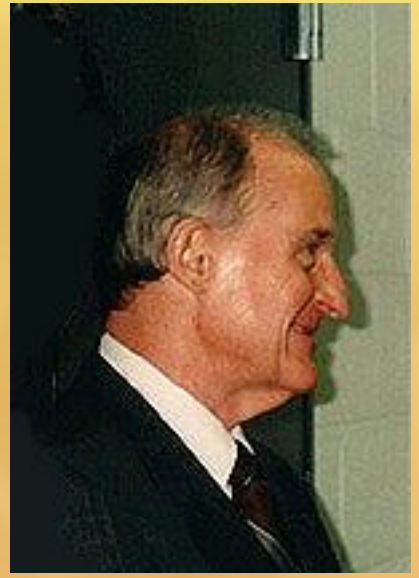
Disclaimer

- Predicting the future
- Going back?
- Personal biases
- Evening lecture

My own background at CERN

- I started at CERN in 1974
- Brand-new computer building (B513)
- Supercomputers:
  - CDC 6600
  - CDC 7600 with CDC6500/6400 front-ends
- Mainframes
  - IBM S/370 systems
  - Initially to get reliable tape storage; running MVS
  - IBM 303x, 309x families and plug-compatible systems
  - VM/CMS
- RISC era
  - DEC, HP, IBM, SGI, SUN desktop/server systems
- PC era
  - Until now

CDC supercomputers



From Wikipedia

- Designed by Seymour Cray
  - Serial number “00000N”
  - “Which instruction would you like next?”
- R-RISC (Really Reduced)
- People were “cheap”, machines expensive
- 7600
  - 27.5 ns, peak 36 Mflops, 60-bit words
  - Fragile
  - Introduced instruction pipelining
  - Smart instruction stack
  - Peripheral Processing Units for I/O
  - In production at CERN for **ten** years

IBM mainframes

- Monolithic systems for “everything”
  - Batch, time sharing, data handling, networking, etc.
- Very expensive – just like the CDC supercomputers
- Started with 370/168 (in 1976)
  - 4 MB of memory !
  - MVS operating system, FORTRAN-77
  - Wylbur/Milten for text editing on start/stop terminals
    - Each user was represented in a 4K page
- Later 303x, 3090 systems and Plug-Compatible Systems
  - Instruction Set enhanced via a Vector Facility
  - Operating system moved to VM/CMS, VM/XA
    - “The poor man’s PC”

Nevertheless, these “tiny” mainframes allowed C.Rubbia (UA-1) to find the W boson in 1982

The RISC era
(Reduced Instruction Set Computing)

- First RISC system was probably the IBM 801
  - John Cocke et al
- Idea pursued by teams at Stanford (J.Hennessy) and Berkeley (D.Patterson)
  - The former led to MIPS, latter to SPARC
- Big breakthrough with Apollo DN/10000
  - 4 CPUs
  - Custom-made RISC core (PRISM)
  - Could attach to the mainframe

Why did we jump on the RISC bandwagon ?

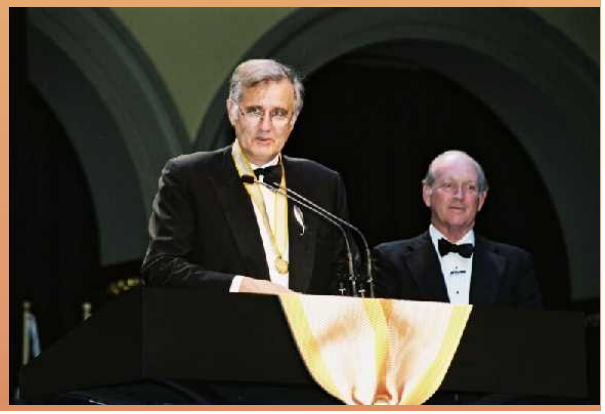
- Price/performance
- Flexibility for incremental additions/ upgrades
- Distributed ownership
- UNIX
- Software development environment on personal workstations
- The mainframe was still around (for a long while) performing more mundane tasks (tape handling, etc.)

Was everything “good” ?

- Of course, not!
- Lots of “waste” due to replication of effort
- Every RISC vendor offered a different UNIX and a different compiler
  - System maintenance was expensive:
    - “Six support staff - one primary and one back-up per architecture”
  - Porting and verification took a huge toll on resources
    - The results often differed: Endianness, FMA, optimisation levels, etc.
  - People had to become porting heroes

SHIFT architecture

- "Scalable Heterogeneous Integrated FaciliTy"
- Allowed seamless integration of the RISC-based systems
  - CPU servers
  - Disk servers
  - Tape servers
- Won a 21st Century Achievement Award from the Computerworld Honors Program
- Software based on tcp/ip:
  - Remote File I/O (RFIO)
  - Disk Pool Manager (DPM)

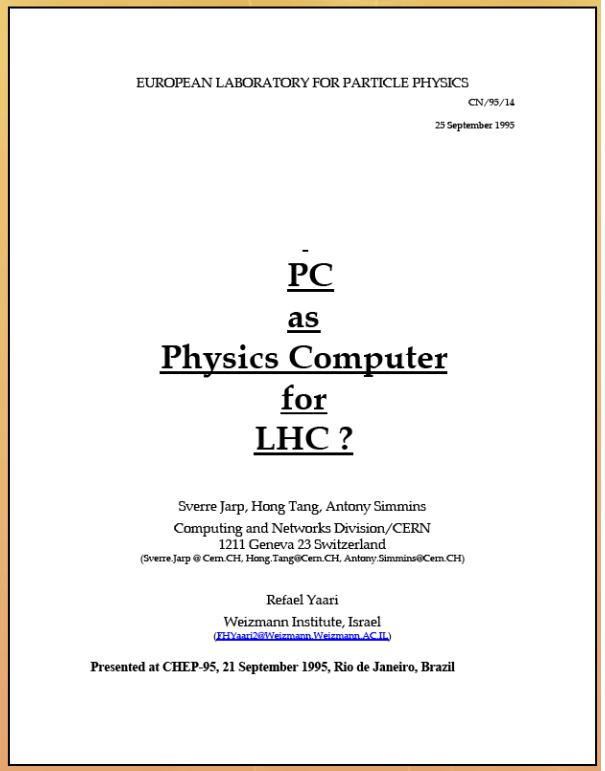


HP and PA-RISC

- HP's workstations were preferred by many developers
- Servers were small, but price/performance was good
  - We bought, for instance, many servers for simulations
  - The "snake farms"
- HP acquired Apollo
  - But, it took almost a decade to get a 4-CPU server back in the market,
    - The Kittyhawk

RD47: Our seminal CHEP-95 paper

- Paper presented in Rio
  - Computing in High Energy Physics
- Using PCs
  - Pentium and Pentium Pro
- Not faster than the RISC systems
  - But much better price/performance
- In our community,
  - this started the move to PCs
  - Plus: Linux



EUROPEAN LABORATORY FOR PARTICLE PHYSICS

CN/95/14

25 September 1995

PC
as
Physics Computer
for
LHC ?

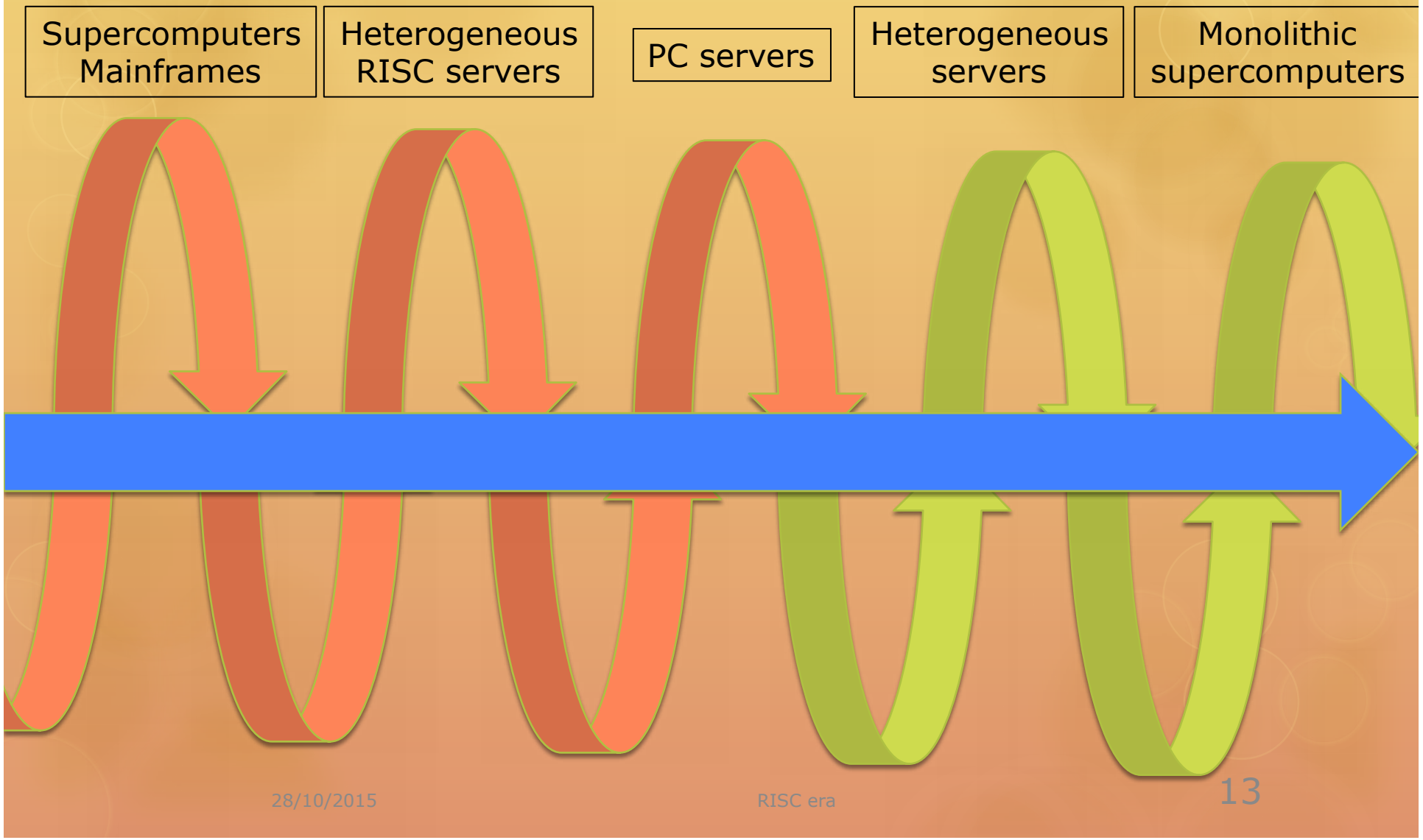
Sverre Jarp, Hong Tang, Antony Simmins
Computing and Networks Division/CERN
1211 Geneva 23 Switzerland
(Sverre.Jarp @ Cern.CH, Hong.Tang@Cern.CH, Antony.Simmins@Cern.CH)

Refael Yaari
Weizmann Institute, Israel
(FHYaari2@Weizmann.Weizmann.AC.IL)

Presented at CHEP-95, 21 September 1995, Rio de Janeiro, Brazil

The first PC farms

Are we moving forward in spirals ?



Who is pushing for heterogeneity?

IBM & the OpenPower Foundation

- Lots of members
  - Platinum, Gold, Silver levels
- Some key members: IBM, Tyan, Nvidia, Mellanox, Google, Micron
- Offerings pitted against Xeon servers
  - Power 8, up to 4 GHz, new CAPI bus
  - Lots of execution units; Wide superscalar design
  - Up to 12 cores/chip, 8-way SMT
  - 4 cache levels (32K+64K, 512K, up to 96 MB, 128 MB)

IBM and supercomputing

- Systems based on Power 8 being offered today
- Two systems planned for 2017
  - Summit (ORNL) and Sierra (LLNL)
  - Based on Power 9 chips
  - Volta GPUs from Nvidia (NVLINK interconnect)
  - Mellanox EDR networking

ARM Ltd.

- A72: Latest 64-bit high-end design
- Much better than A57
  - Which is/was a “transitional” design
  - Still centred on ARM32 support ?
- 3-way superscalar, OOO execution
- Now, some good DP latencies:
  - FMUL=3, FADD=3, FCVT=2
- Not so good
  - FDIV=18, FSQRT=32

ARM Architecture licensees

- AppliedMicro
- Broadcom
- Cavium
- Apple
- Huawei
- Nvidia
- AMD
- Samsung
- Marvell
- Microsoft
- Qualcomm
- Intel
- Faraday

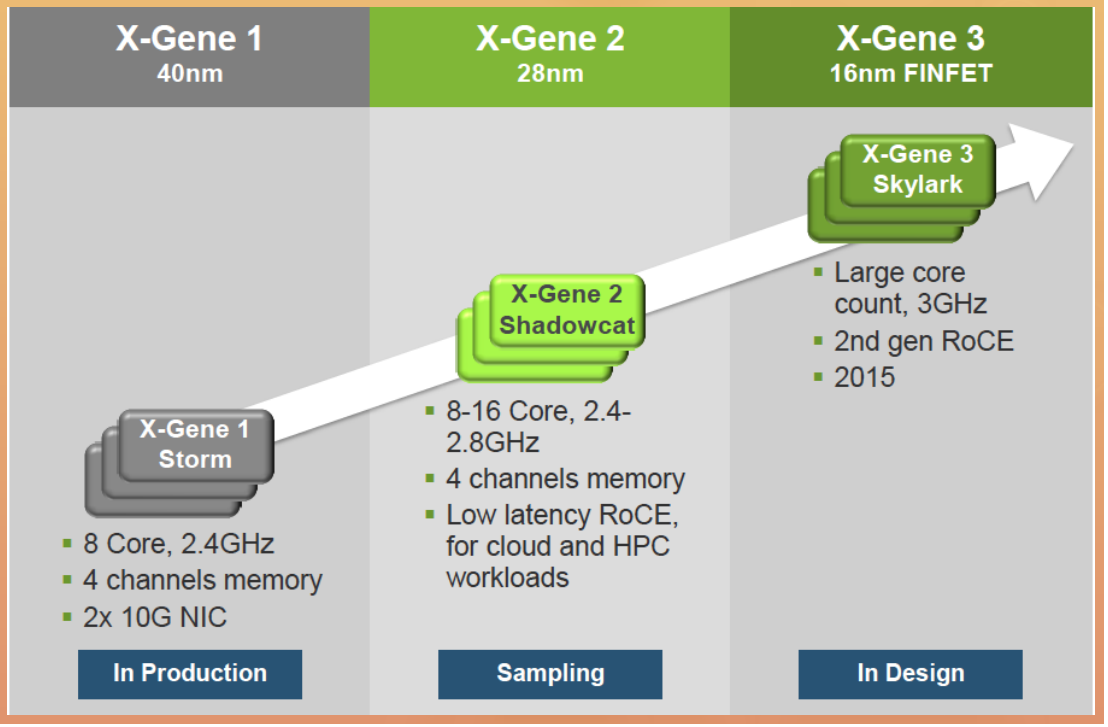
Apple

- Intel x86(-64) chips since 2006
- Acquired P.A. Semi in 2008
  - PWRficient chips
  - Daniel Dobberpuhl [originally DEC]
- ARM64 designs since 2011
- Amazingly aggressive time-wise
  - New designs even before ARM themselves
- As measured by Anandtech the superscalar design is complex
- Most resent design (A9, A9X) saw a doubling of memory bandwidth
  - 80% more CPU performance
- Will Apple stop at the iPad Pro ?

Applied Micro

- Early out (2011)
- X-Gene 1 available since some time
  - 40 nm, 8 cores, 2.4 GHz
  - Not even Atom-level performance
- X-Gene 2
  - 28 nm, 16 cores, 2.8 GHz
  - 4-way superscalar
- X-Gene 3
  - 16nm FINFET
  - 32 cores

From HotChips 2014



QUALCOMM

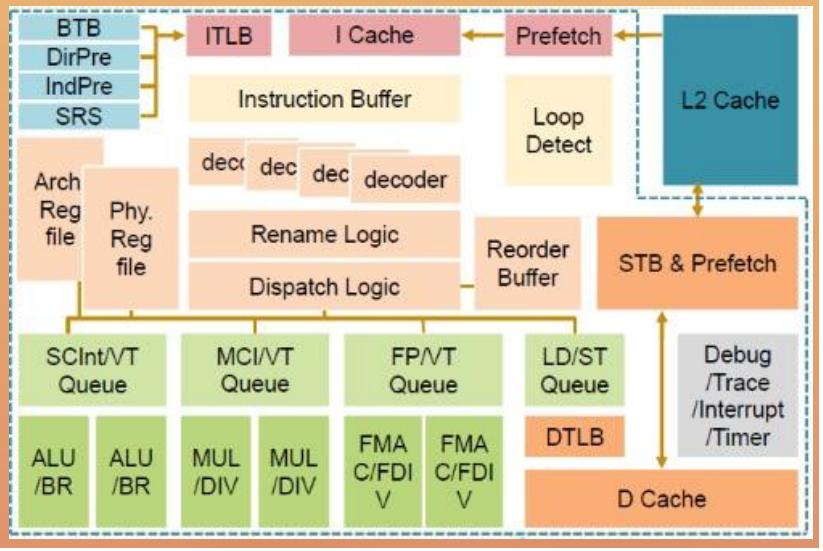
- Best known for its mobile chips
  - Snapdragon, mainly ARM32
- Recently, they announced an alliance around a “server” ARM64 chip
  - Partners: Mellanox, Xilinx
- Prototype based on single 24-core chip
- Working with US and Chinese “hyperscalers"
- Production systems “only” in 2017

Cavium

- ThunderX
  - 28 nm
- Two SoCs launched
  - 8-16 cores
  - 24-48 cores
  - 16 MB L2
- Currently, only provider of ARM64 SMP chips

Phytium: a Chinese ARM chip

- Presented at Hot Chips this year
- Phytium Technology, Ltd (2012)
- Chip with 64 custom cores in 28 nm
  - 2 GHz, 4 flops/cycle → 512 Gflops peak
  - 8-core panels in which four cores share a 4-MB L2
  - 128 MB L3 cache
  - Multithreading
- But, when?



But, does ARM need to care about servers?

- Maybe it is enough to concentrate on the "lower end"
  - Billions of smart phones
    - Still NN-core 64-bit chips with sufficient memory
  - Trillions of IoT (Internet-of-Things) devices
- At least when these devices have power
  - Most of the computing in the world might be done here?

A possible 2020 statement: "The CMS collaboration does more simulation than ever, because every collaborator lends the processors in his/her automobile over night for scientific computing ! "

Back to heterogeneity: What is different, now ?

- OS: Linux for everybody
- Common compilers
  - **GNU** compiler suite. Fortran and C/C++ [5.2]
    - http://gcc.gnu.org/
  - **LLVM** (C/C++) compiler framework [3.7]
    - Originated from U. of Illinois; Strongly supported by Apple
      - http://www.llvm.org/
- Myriads of open source software packages:
  - OpenStack, MySql, Hadoop, Lustre, etc.

LLVM and its design

- Three parts:

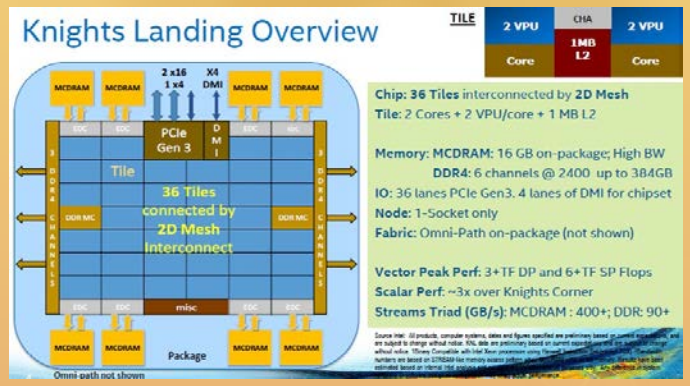
- Advantage:
  - Add Middle part for performing hardware-independent optimisation
- As a result the compiler becomes more attractive for new hardware!
  - Easier to port a “slim” back-end
- Lots of back-end versions exist

Intel's grip on the market

- Moore's law
- Two decades since the Pentium Pro
- Aggressive new Xeon designs
  - Tick-Tock [.., Haswell, Broadwell, Skylake, …]
- Xeon Phi developed in parallel
  - Although, only combined Tick-Tock
- Not afraid of "deprecating" what they recently invented
  - SSE-128 and AVX-256 in favour of AVX-512
- Huge investments in software
  - Biggest compiler team in the world?
- Entire eco-system accompanies each chip generation

Xeon Phi co-processors

- MIC: Many Integrated Cores
- Grew out of “Larrabee” graphics accelerator
  - Knights Ferry – software dev. vehicle
  - Knights Corner (22 nm – 61 cores)
  - Knights Landing (14 nm – 72 cores)
  - Future versions



- Rather revolutionary vector-based instruction set
  - Evolved from LNI to AVX-512
  - 32 * 512-bit vector registers, 8 * mask registers
- High-speed memory
  - Combination of eDRAM and DDR4

When Moore's law grinds to a halt

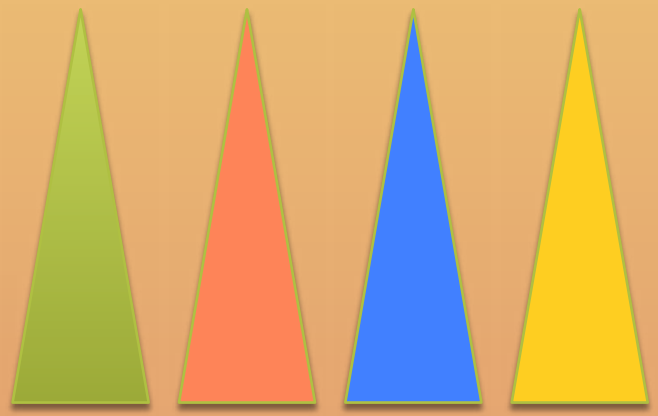
- Increasing competition around smarter and better integrated designs
- Closer integration at all levels
- Hardware/software co-design
- CPUs, GPUs (accelerators), FPGAs, etc.
- Complex memory hierarchies
- Very large non-volatile memories

Intel: A mainframe/supercomputer company (as it was in the past) ?

- The company has most of the ingredients for building an entire system:
- CPUs
- Co-processors
- Memories (in collaboration with Micron and probably others)
- Non-volatile memories (3D XPoint, SSDs)
- Interconnect fabric (OmniPath)
- Software

Scenario-1: the Mega-centers matter

- Billions of edge systems dependent on heterogeneous Megacenters

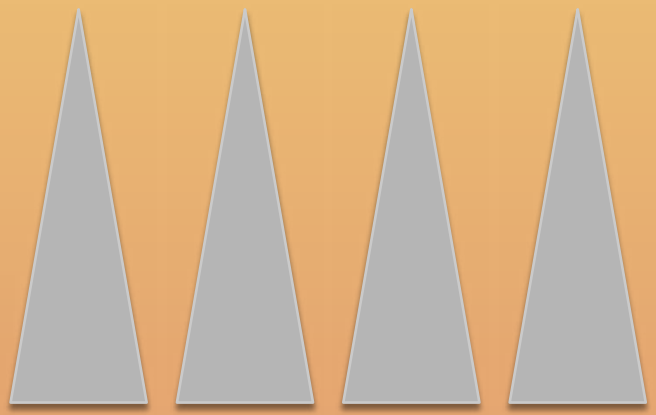


Scenario-2: the Mega-centers matter

- Billions of edge systems dependent on mainly homogeneous Megacenters

Scenario-3: Only the Edge matters

- Big servers are basically anonymous “cloud storage servers”
  - CPU architecture is more or less irrelevant



Thanks!