Mars: A 64-core ARMv8 Processor

Charles Zhang
Phytium Technology Co., Ltd

Statements

The following slides are presented to introduce the general features of one of our products, instead of any commitment about it. It is for information purposes only, and may not be incorporated into any contract. It is not suggested to make purchasing decisions accordingly. The development, release, and timing of any features or functionality described here remains at the sole discretion of Phytium.

A Brief Introduction of Phytium

- China corporation, founded in 2012
  - Guangzhou
  - Tianjin
- Vision
  - Leading edge CPU and ASIC provider in China
- Market focuses on chips for
  - Internet & Cloud Computing infrastructure
  - Traditional workload mainframe servers

China is a Fast-growing Server Market

Company	1Q15 Revenue	1Q15 Market Share (%)	1Q14 Revenue	1Q14 Market Share (%)	1Q15-1Q14 Growth (%)
Inspur	332,613,480	21	227,328,256	17	46
Dell	322,063,140	20	246,281,271	19	31
Lenovo	295,914,571	18	80,084,826	6	270
HP	217,487,450	14	167,775,923	13	30
Huawei	197,490,419	12	189,963,266	14	4
Sugon	140,377.091	9	70,705,366	5	99
Others	104,566,737	6	329,549,621	25	-68
Total	1,610,512,888	100.0	1,311,688,529	100.0	23

China

Company	1Q15 Revenue	1Q15 Market Share (%)	1Q14 Revenue	1Q14 Market Share (%)	1Q15-1Q14 Growth (%)
HP	3,191,694,948	23.8	2,890,992,229	25.5	10.4
Dell	2,296,473,026	17.1	2,006,639,006	17.7	14.4
IBM	1,887,939,141	14.1	2,244,631,789	19.8	-15.9
Lenovo	970,254,659	7.2	127,973,470	1.1	658.2
Cisco	890,179,930	6.6	616,620,000	5.4	44.4
Others	4,157,871,704	31.0	3,469,383,444	30.6	19.8
Total	13,394,413,409	100.0	11,356,239,939	100.0	17.9

WW

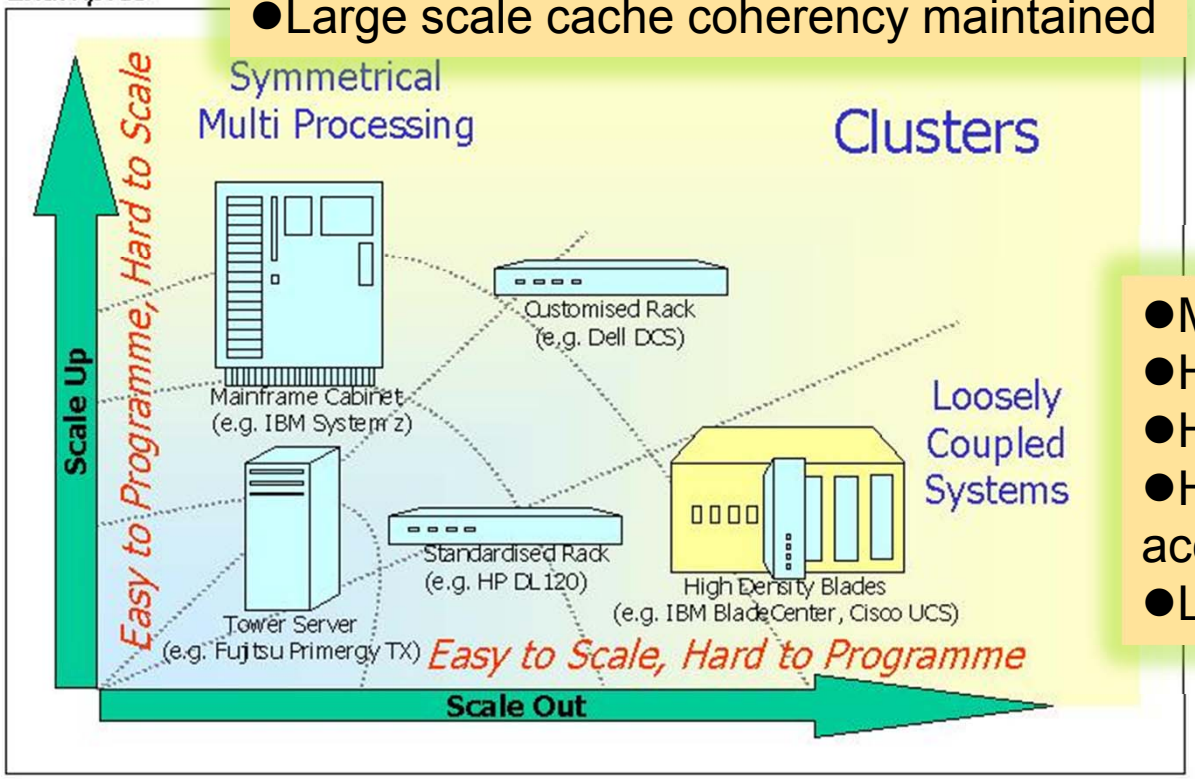
Source: Gartner (May 2015)

What is Mars for?

- High performance
- High volume of memory
- High bandwidth memory access
- High bandwidth I/O access
- Large scale cache coherency maintained

Mars

Figure 3 – Sca
Examples



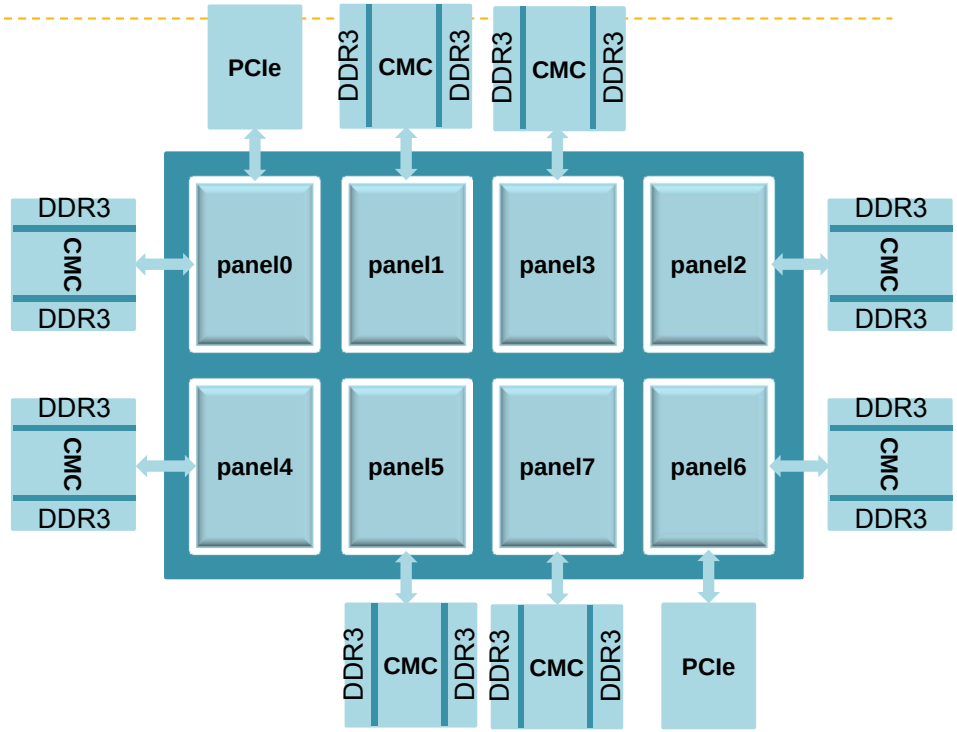
Source: ITCandor, April 2010

Earth

- Moderate performance
- High power efficiency
- High density computing
- High bandwidth memory access
- Low cost

Mars Overview

- Architecture Features
  - 64 Xiaomi cores, ARMv8 compatible
  - Hardware-maintained global cache coherency
  - Panel-based data affinity architecture
  - Mesh topology on chip network
  - 32MB L2 cache
  - 8 Cache & Memory Chips (CMC)
    - 128MB L3 cache
    - 16 DDR3-1600 channels
  - Two 16-lane PCIE3.0 i/f
  - ECC and parity protection on all caches, tags and TLBs

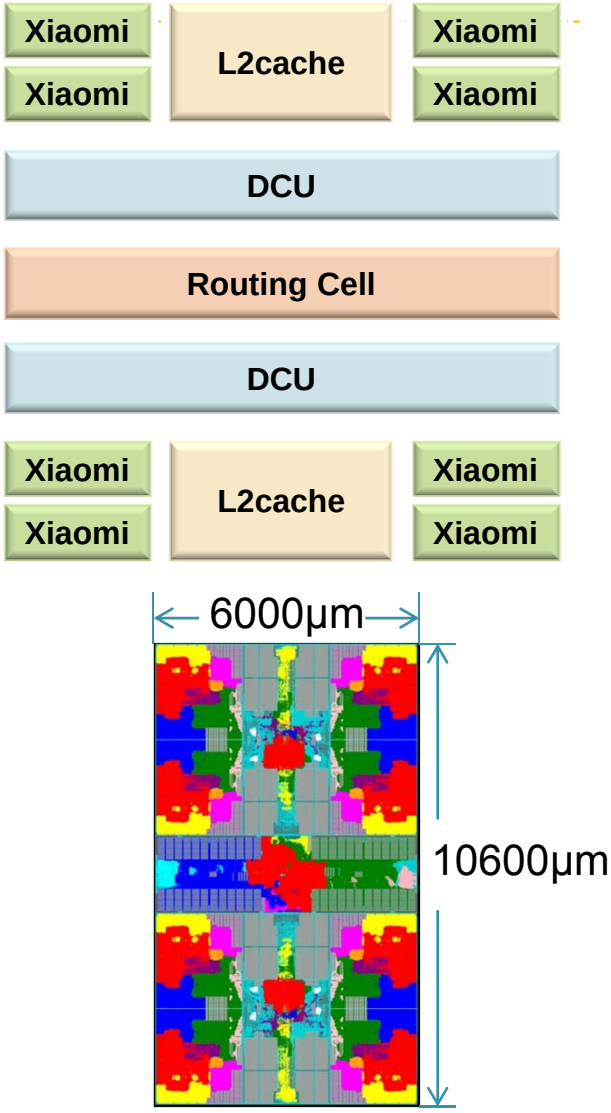


Physical
- ~180M instances
- 2.0GHz@28nm
- 120W

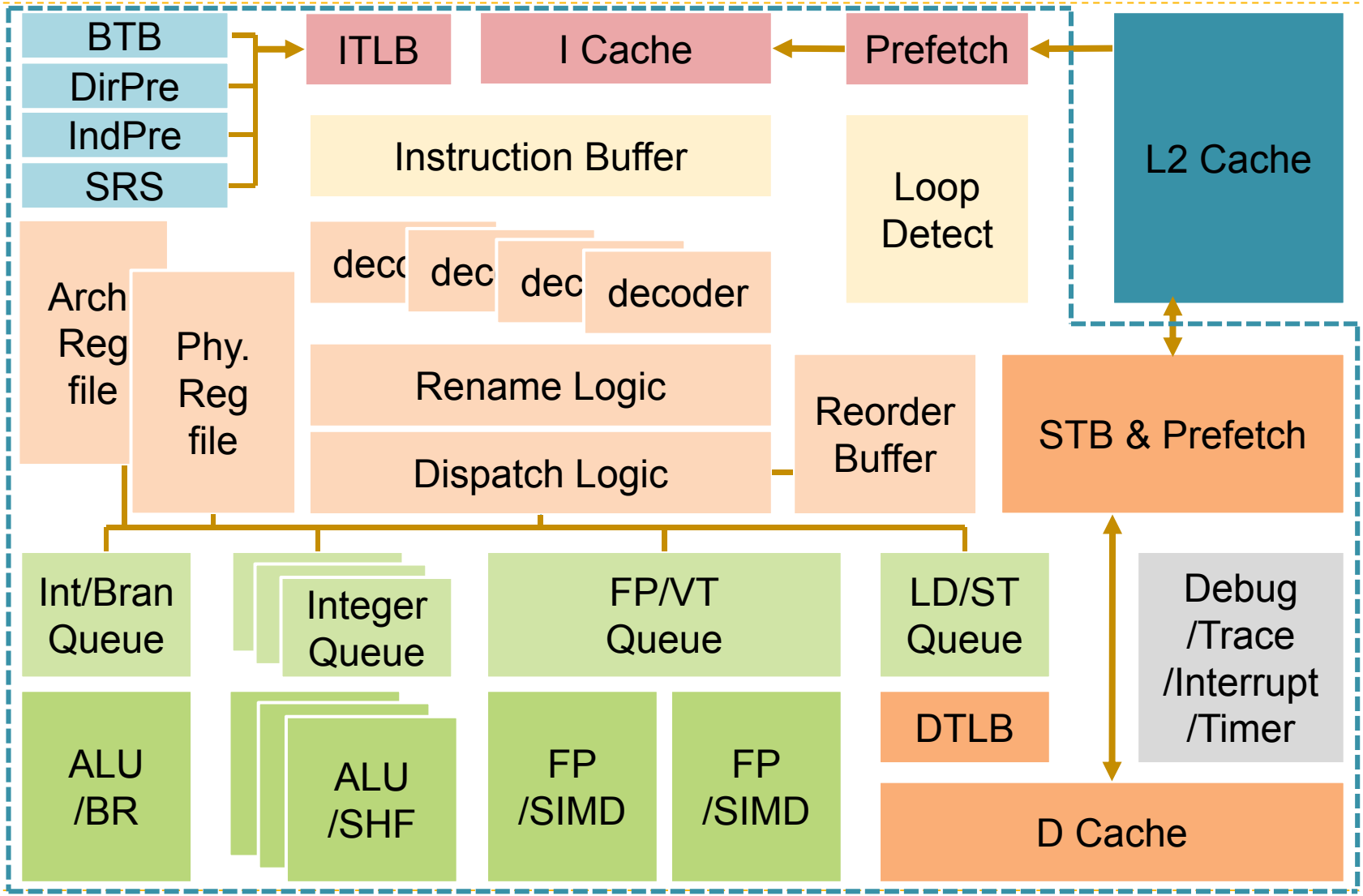
Performance
- Peak：512GFLOPS
- Mem BW：204GB/s
- I/O BW： 32GB/s

Panel Architecture

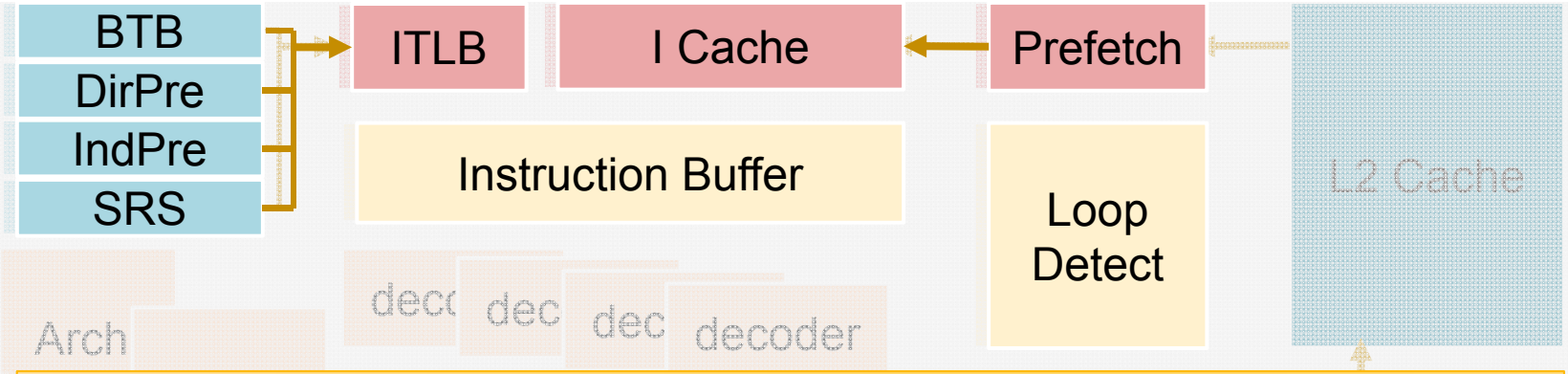
- Eight Xiaomi Cores
  - Compatible design with ARMv8 arch license
  - Both AArch32 and AArch64 modes
  - EL0~EL3 supported
  - ASIMD-128 supported
  - Adv. hybrid Branch Prediction
  - 4 fetch/4 decode/4 dispatch Out-of-Order superscalar pipeline
- Cache Hierarchy
  - Separated L1 ICache and L1 Dcache
  - Shared L2 cache, totally 4MB
- Directory-based cache coherency maintenance
  - Directory Control Unit (DCU)
- Routing Cell



Xiaomi Core

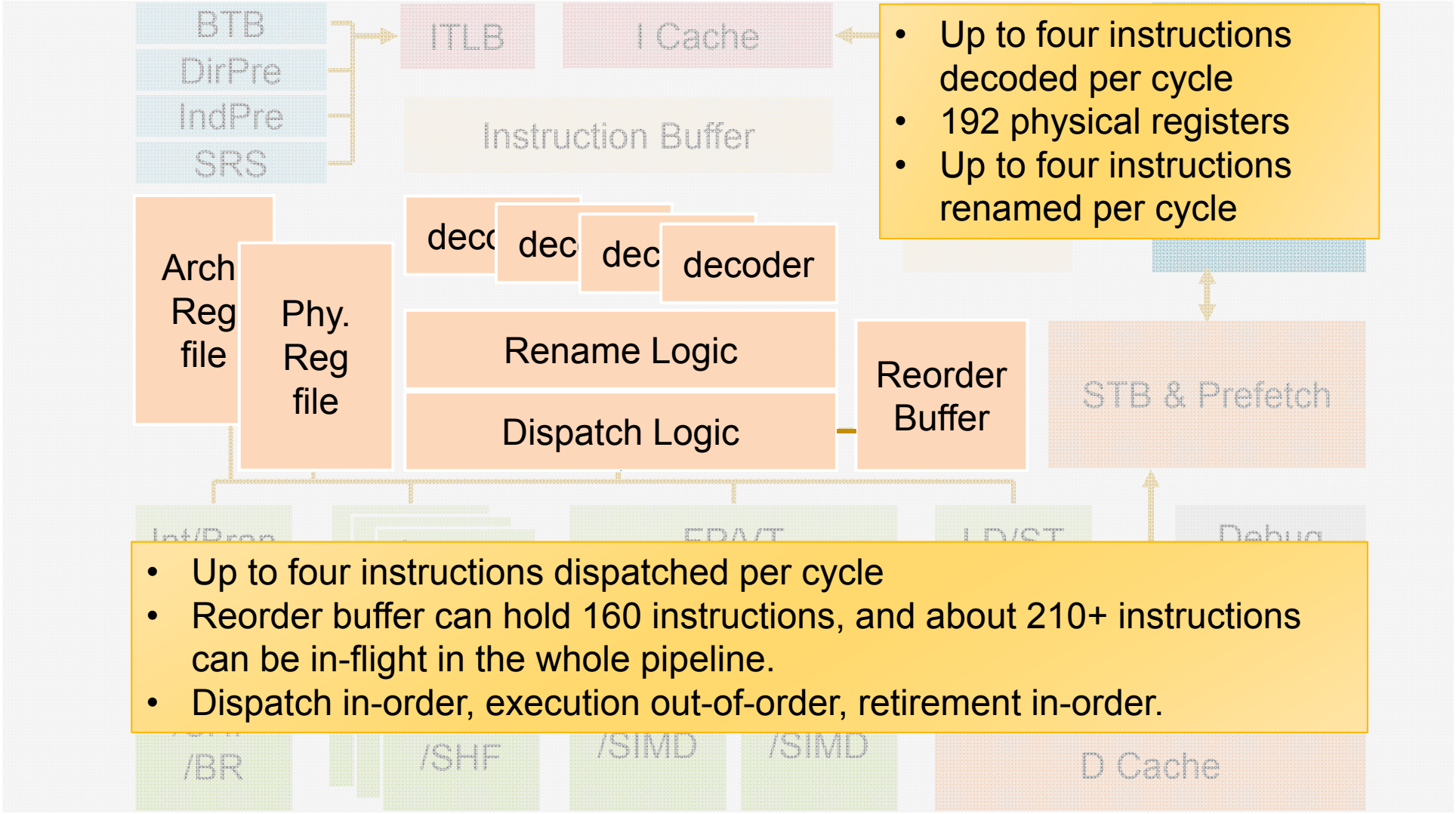

BTB
DirPre
IndPre
SRS
ITLB
I Cache
Prefetch
L2 Cache
Instruction Buffer
Loop Detect
Arch Reg file
Phy. Reg file
decc
dec
dec
decoder
Rename Logic
Dispatch Logic
Reorder Buffer
STB & Prefetch
Int/Bran Queue
Integer Queue
FP/VT Queue
LD/ST Queue
Debug /Trace /Interrupt /Timer
ALU /BR
ALU /SHF
FP /SIMD
FP /SIMD
DTLB
D Cache

Xiaomi Core Front End



- 32KB L1 instr. Cache
- Next line prefetch
- Hybrid Branch Predictor
  - 2048-entry BTB
  - Direction predict with TAGE predictor
  - 512-entry indirect predictor
  - 48-entry Speculative Return Stack
- Four instructions fetched per cycle
- 32-entry instruction buffer
- Loop detect and Instr. Cache bypass

Xiaomi Core Decode, Rename & Dispatch

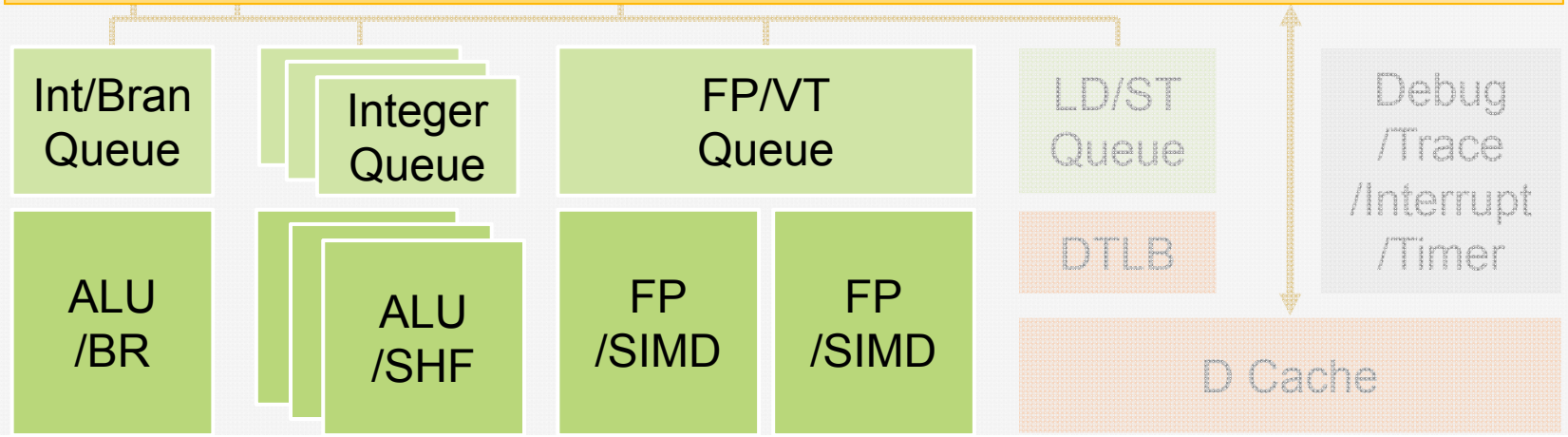


- Up to four instructions decoded per cycle
- 192 physical registers
- Up to four instructions renamed per cycle

- Up to four instructions dispatched per cycle
- Reorder buffer can hold 160 instructions, and about 210+ instructions can be in-flight in the whole pipeline.
- Dispatch in-order, execution out-of-order, retirement in-order.

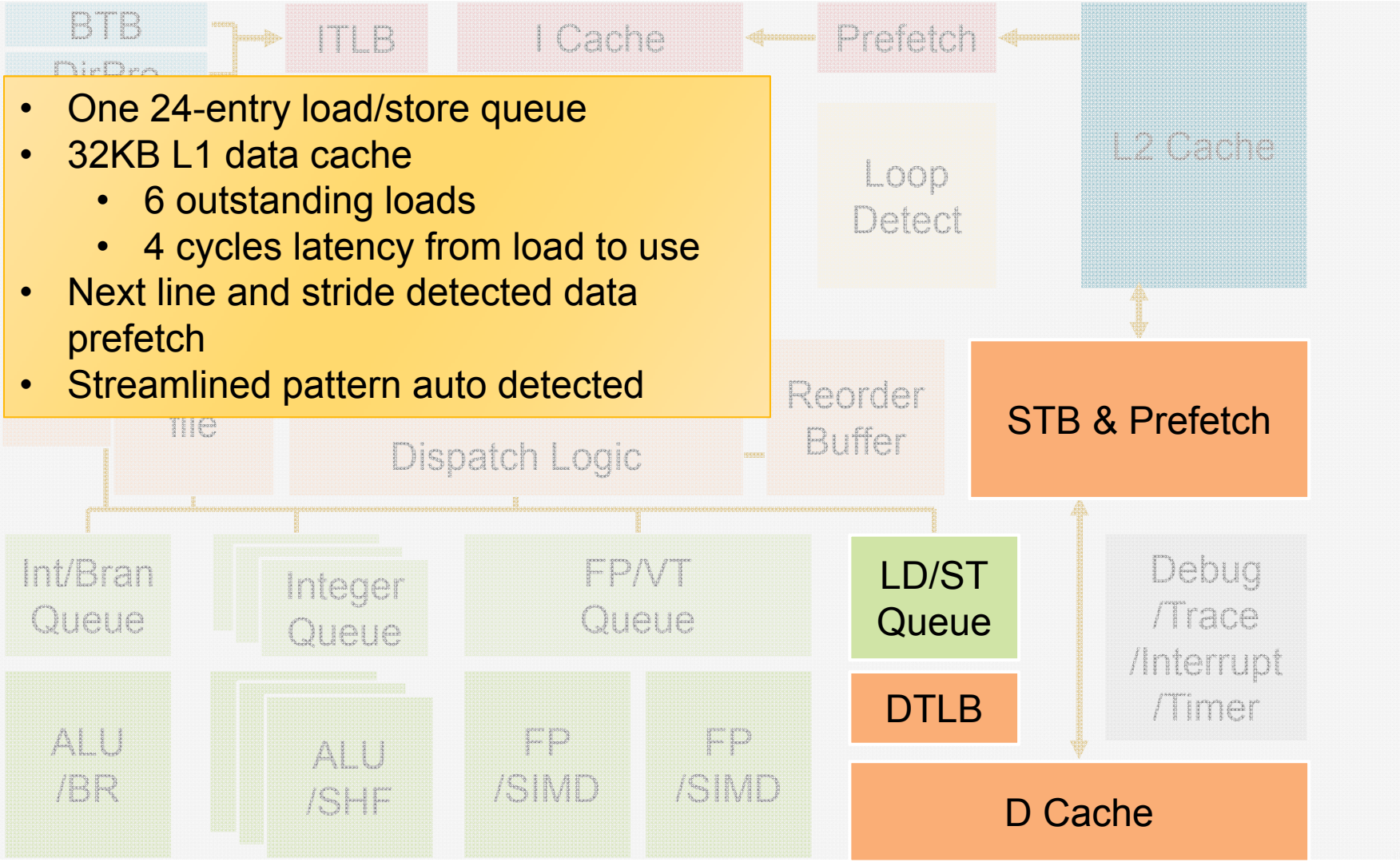
Xiaomi Core Function Units

- Four separated 16-entry integer queues
  - One integer unit can process both multi-cycle integer instructions and branch instructions
  - The other three integer units can only process singe-cycle integer instructions
- One shared16-entry floating point and ASIMD queue
  - Two FP/ASIMD units equipped, which can be combined into one lockstep ASIMD unit.
  - FMA supported in both units.
  - FMUL: 3cycles, FADD: 3cycles, FMA: 6cycles



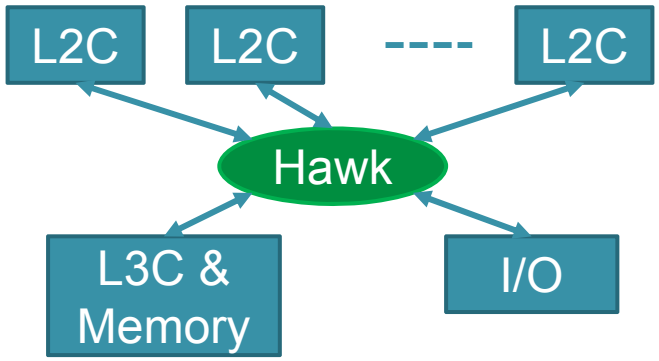
Xiaomi Core Function Units

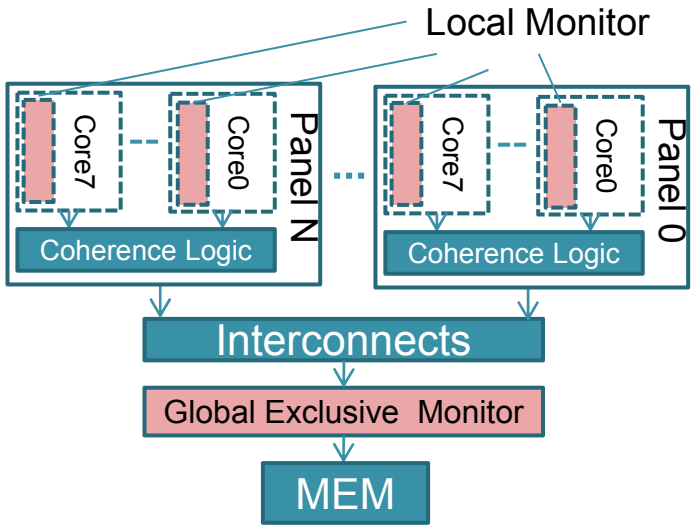
- One 24-entry load/store queue
- 32KB L1 data cache
  - 6 outstanding loads
  - 4 cycles latency from load to use
- Next line and stride detected data prefetch
- Streamlined pattern auto detected



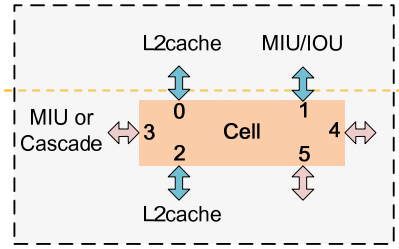
Cache coherence protocol

- Hawk cache coherence protocol
  - Distributed directory-based global cache coherency
  - MOESI-like packet-based coherence protocol
  - A home node DCU(directory control unit) supports
    - Affinitive pairing of L2Cs and CMCs
    - “Infinite” capacity for non-conflicting Reads & Writes
    - Optimized transaction flow for exclusive atomic accesses
    - Reduced latency by cacheline forwarding



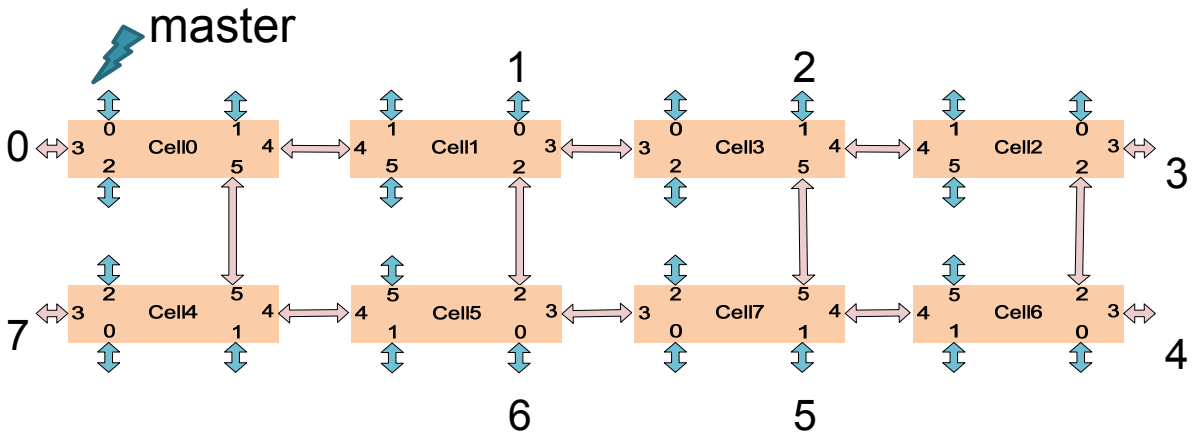


Network on Chip



2D Concentrated Mesh Architecture

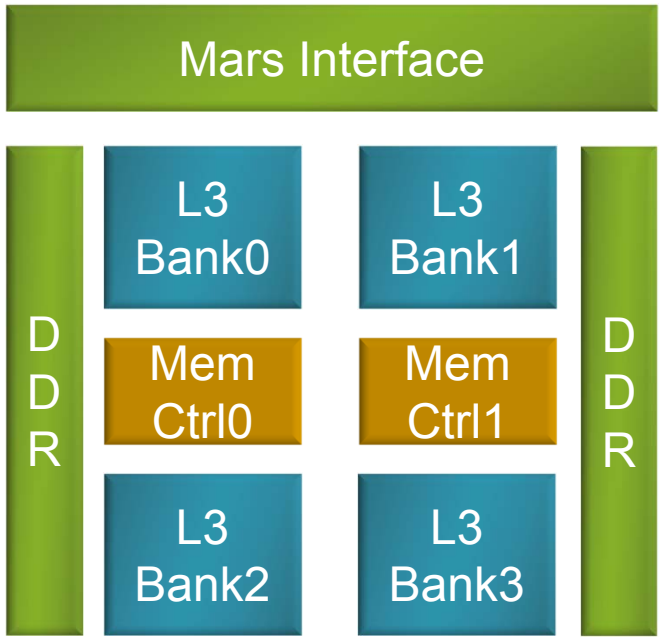
- Cell based switch with 6 bidirectional ports
- Uniform package format for each port, a port can be configured to be connected with a device or cascade cell
- 4 physical channels for CC and 1 channel for debug, DOR Y-X routing
- Low latency: 3 cycles for each hop
- High bandwidth: 384GB/s each cell



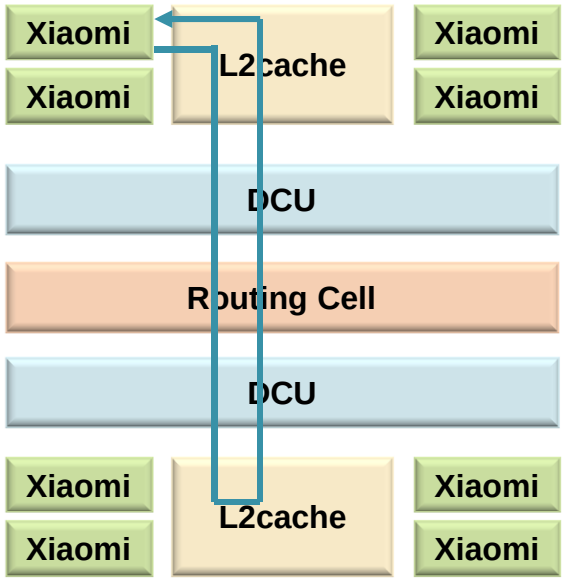
Dest.	Lat. (cycles)
0	3
1	6
2	9
3	12
4	15
5	12
6	9
7	6
Avg.	9

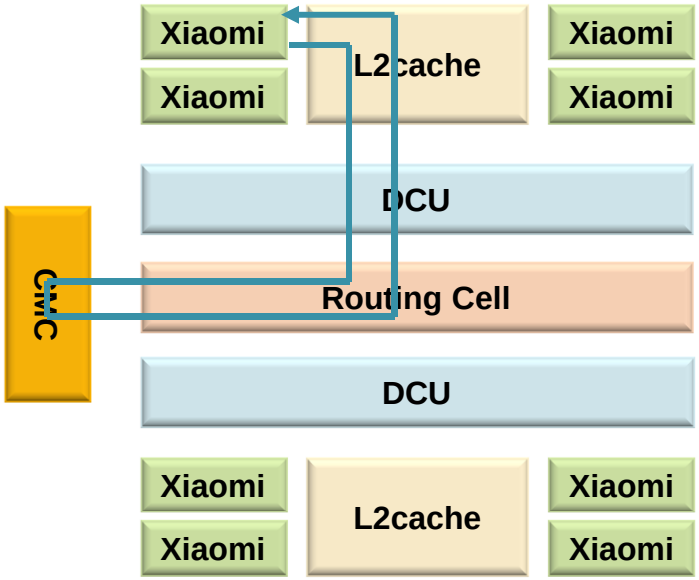
Cache & Memory Chip

- L3 cache
  - 16MB Data Array
  - 2MB Data ECC
- DDR bandwidth
  - 2 x DDR3-800：25.6GB/s
- Proprietary interface between Mars & CMC
  - Parallel interface
    - Needs more pins, but lower latency than serdes
  - Separate write/cmd and read data channel
    - Effective read channel bandwidth：12.8GB/s
    - Effective write/cmd channel bandwidth：6.4GB/s



Latency of affinitive access





Memory access	latency(ns)
Local L1 cache hit	~2
Local L2 cache hit	~8
Affinitive L2 cache hit	~20
Affinitive L3 cache hit	~36
Affinitive DDR access	~70

- Panel : 2.0GHz
- NoC: 2.0GHz
- CMC: 1.5GHz

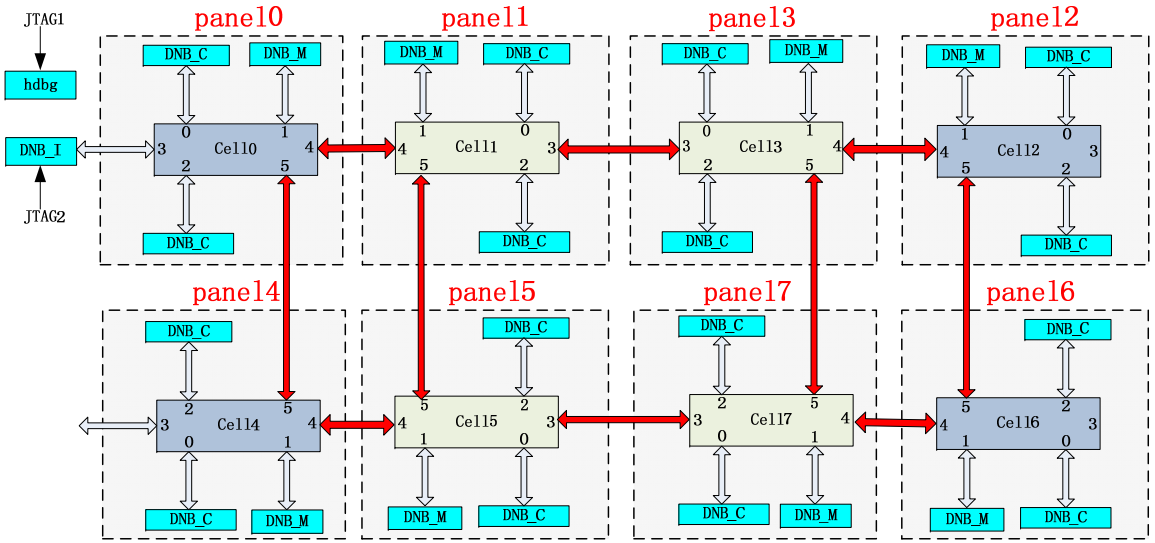
* PCB latency not considered

Memory Tune (mTune)

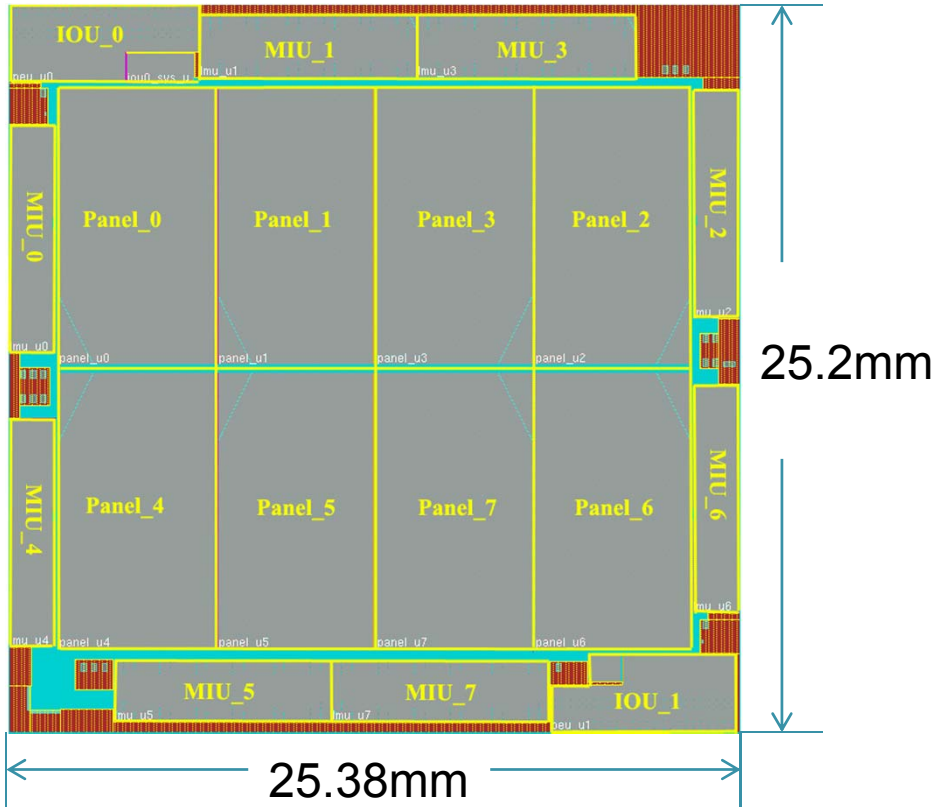
- Rich Data Collection
  - Number of cache hits/misses for L1/L2/L3
  - Workload of cache pipelines
  - Busyness of the NoC
  - ECC corrections of the memory system
- Support Multiple Metrics
  - Average Miss rate/Hit rate
  - Minimal/Maximal/Average Access Latency
  - Bandwidth Analysis
  - Concurrent Average Memory Access Time (CAMAT)
- Support MPI/OpenMP Applications
  - Thread behavior analysis
  - Inter-process behavior analysis

Scalable Debug System

- ARMv8 CoreSight Compatible debug system
- Scalable dedicated debug network across 64 cores
- Distributed debug components
- Configurable events broadcast scope
- Timestamp broadcasts with single signal to simplify implementation



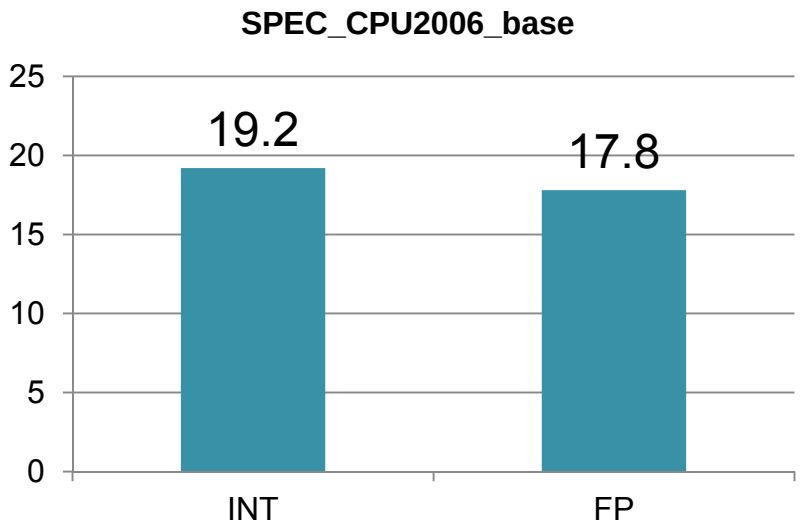
Physical Design



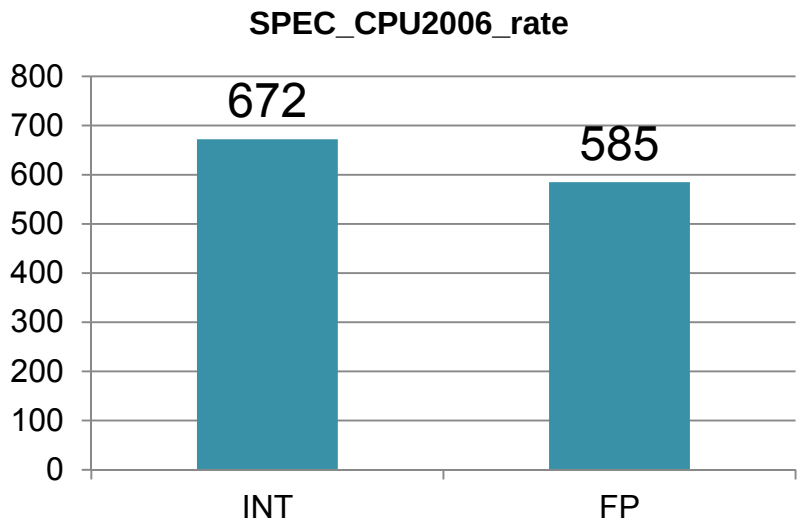
- 28nm process
- 0.9v core/1.8v IO
- 10 metal layers
- ~180M instances
- 2.0GHz
- 120W
- 640mm^2 die size
- FCBGA
- ~3000 pins

Performance Evaluation

- SpecCPU2006



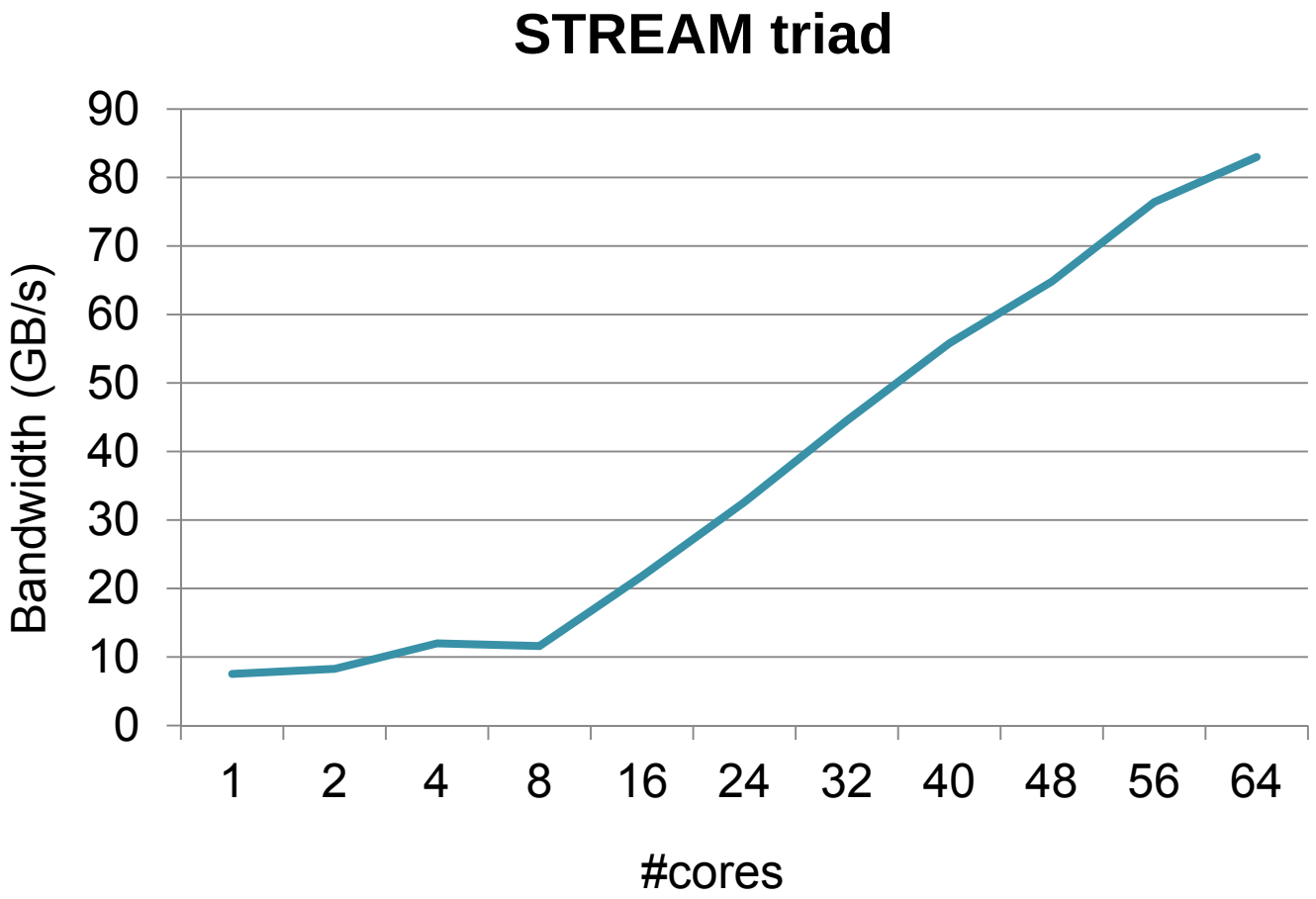
Single copy of SPEC CPU benchmark



64 copies of SPEC CPU benchmark

Performance Evaluation

- STREAM



Next Generation Scale-up CPU

- More powerful core
  - Aggressive Branch Predictor
  - Multithreading
  - More aggressive ILP exploitation
  - Wider SIMD
- More RAS features
- Higher bandwidth memory access
- Higher power efficiency

Mars: A 64-core ARMv8 Processor

Charles Zhang
Phytium Technology Co., Ltd