A Multicore Processor Designed For PetaFLOPS Computation

Weiwu Hu

Institute of Computing Technology, Chinese Academy of Sciences
Loongson Technologies Corporation Limited
hww@ict.ac.cn

Contents

- **Background**
- **The Scalable Godson-3 Multicore Architecture**
- **The Vector Extension of the CPU Core**
- **HPCs Based on Godson-3**

Godson is the academic name of Loongson™

TOP10 in 2008.11

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband / 2008 IBM	129600	1105.00	1456.70	2483.47
2	Oak Ridge National Laboratory United States	Jaguar - Cray XT5 QC 2.3 GHz / 2008 Cray Inc.	150152	1059.00	1381.40	6950.60
3	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.66 GHz / 2008 SGI	51200	487.01	608.83	2090.00
4	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
5	Argonne National Laboratory United States	Blue Gene/P Solution / 2007 IBM	163840	450.30	557.06	1260.00
6	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband / 2008 Sun Microsystems	62976	433.20	579.38	2000.00
7	NERSC/LBNL United States	Franklin - Cray XT4 QuadCore 2.3 GHz / 2008 Cray Inc.	38642	266.30	355.51	1150.00
8	Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz / 2008 Cray Inc.	30976	205.00	260.20	1580.71
9	NNSA/Sandia National Laboratories United States	Red Storm - Sandia/ Cray Red Storm, XT3/4, 2.4/2.2 GHz dual/quad core / 2008 Cray Inc.	38208	204.20	284.00	2506.00
10	Shanghai Supercomputer Center China	Dawning 5000A - Dawning 5000A, QC Opteron 1.9 Ghz, Infiniband, Windows HPC 2008 / 2008 Dawning	30720	180.60	233.47	

Top10 in 2009.11

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
2	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.50
3	National Institute for Computational Sciences/University of Tennessee United States	Kraken XT5 - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	98928	831.70	1028.85	
4	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution / 2009 IBM	294912	825.50	1002.70	2268.00
5	National SuperComputer Center in Tianjin/NUDT China	Tianhe-1 - NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband / 2009 NUDT	71680	563.10	1206.19	
6	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0 GHz/Nehalem EP 2.93 Ghz / 2009 SGI	56320	544.30	673.26	2348.00
7	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
8	Argonne National Laboratory United States	Blue Gene/P Solution / 2007 IBM	163840	458.61	557.06	1260.00
9	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband / 2008 Sun Microsystems	62976	433.20	579.38	2000.00
10	Sandia National Laboratories / National Renewable Energy Laboratory United States	Red Sky - Sun Blade x6275, Xeon X55xx 2.93 Ghz, Infiniband / 2009 Sun Microsystems	41616	423.90	487.74	

Top10 in 2010.6

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
2	National Supercomputing Centre in Shenzhen (NSCS) China	Nebulae - Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU / 2010 Dawning	120640	1271.00	2984.30	
3	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband / 2009 IBM	122400	1042.00	1375.78	2345.50
4	National Institute for Computational Sciences/University of Tennessee United States	Kraken XT5 - Cray XT5-HE Opteron Six Core 2.6 GHz / 2009 Cray Inc.	98928	831.70	1028.85	
5	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution / 2009 IBM	294912	825.50	1002.70	2268.00
6	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX/8400EX, Xeon HT QC 3.0/Xeon Westmere 2.93 Ghz, Infiniband / 2010 SGI	81920	772.70	973.29	3096.00
7	National SuperComputer Center in Tianjin/NUDT China	Tianhe-1 - NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband / 2009 NUDT	71680	563.10	1206.19	
8	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution / 2007 IBM	212992	478.20	596.38	2329.60
9	Argonne National Laboratory United States	Intrepid - Blue Gene/P Solution / 2007 IBM	163840	458.61	557.06	1260.00
10	Sandia National Laboratories / National Renewable Energy Laboratory United States	Red Sky - Sun Blade x6275, Xeon X55xx 2.93 Ghz, Infiniband / 2010 Sun Microsystems	42440	433.50	497.40	

What Next?

No. 1 ?

Design CPU for HPCs!

CPU Plan of China

- **National Strategic Product**
  - **Supported by the National S&T Major Project for CPU**
- **10th Five Year Plan (2001-2005):**
  - **Startup and key technology research**
  - **Four-Issue OOO Architecture, 1.0GHz**
- **11th Five Year Plan (2006-2010):**
  - **From emulation to innovation, low- to high-end, research to product**
  - **Multi-core CPU with leading performance, CPU company setup**
  - **Desktop, servers, and HPC products based on domestic designed CPU**
- **12th and 13th Five Year Plan (2011-2020):**
  - **Build a new ecosystem to support the IT industry in China**
  - **Start from National Security, education, e-government, ...**

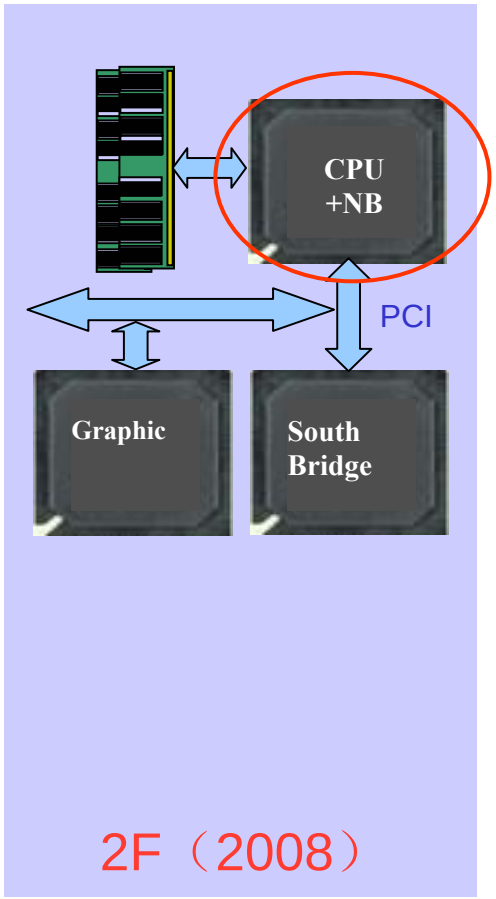
National Sci.&Tech. Major Project

- **Part of the National Mid-Term Plan**
- **16 Major Projects, each fund more than USD 10B from 2006-2020**
  - **CPU and OS**
  - **VLSI process techmology**
  - **Next-generation (4G) wireless network**
  - **High-end digital machine tool**
  - **Advanced nuclear fission power plant**
  - **Water pollution control and treatment**
  - **Large aircraft**
  - **High-resolution earth-observation system**
  - **Manned space flight and lunar exploration**
  - **.....**

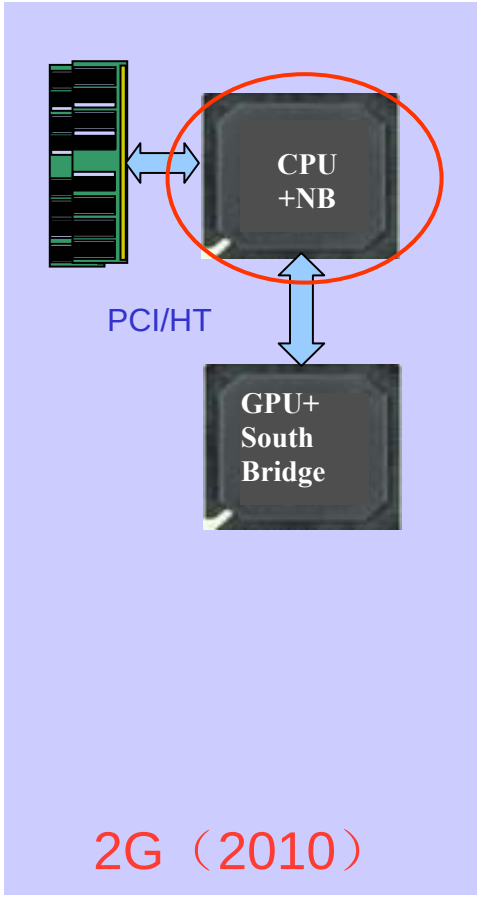
Godson CPU Briefs

- **Research Stage: started in 2001.**
  - **The 32-bit Godson-1 in 2002 is the first CPU in China**
  - **The 64-bit Godson-2B in 2003.10**
  - **The 64-bit Godson-2C in 2004.12**
  - **The 64-bit Godson-2E in 2006.03**
  - **Each Triple the performance of its previous one**
  - **SPEC int2000 and SPEC fp2000 of Godson-2E> 500**
- **Product Stage: started in 2008**
  - **Godson-2F is the product version of Godson-2E**
  - **The 4-core Godson-3A product in 2010**

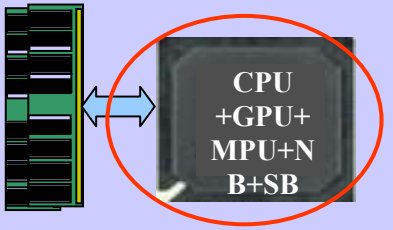
Low end roadmap: From CPU to SOC



2F（2008）

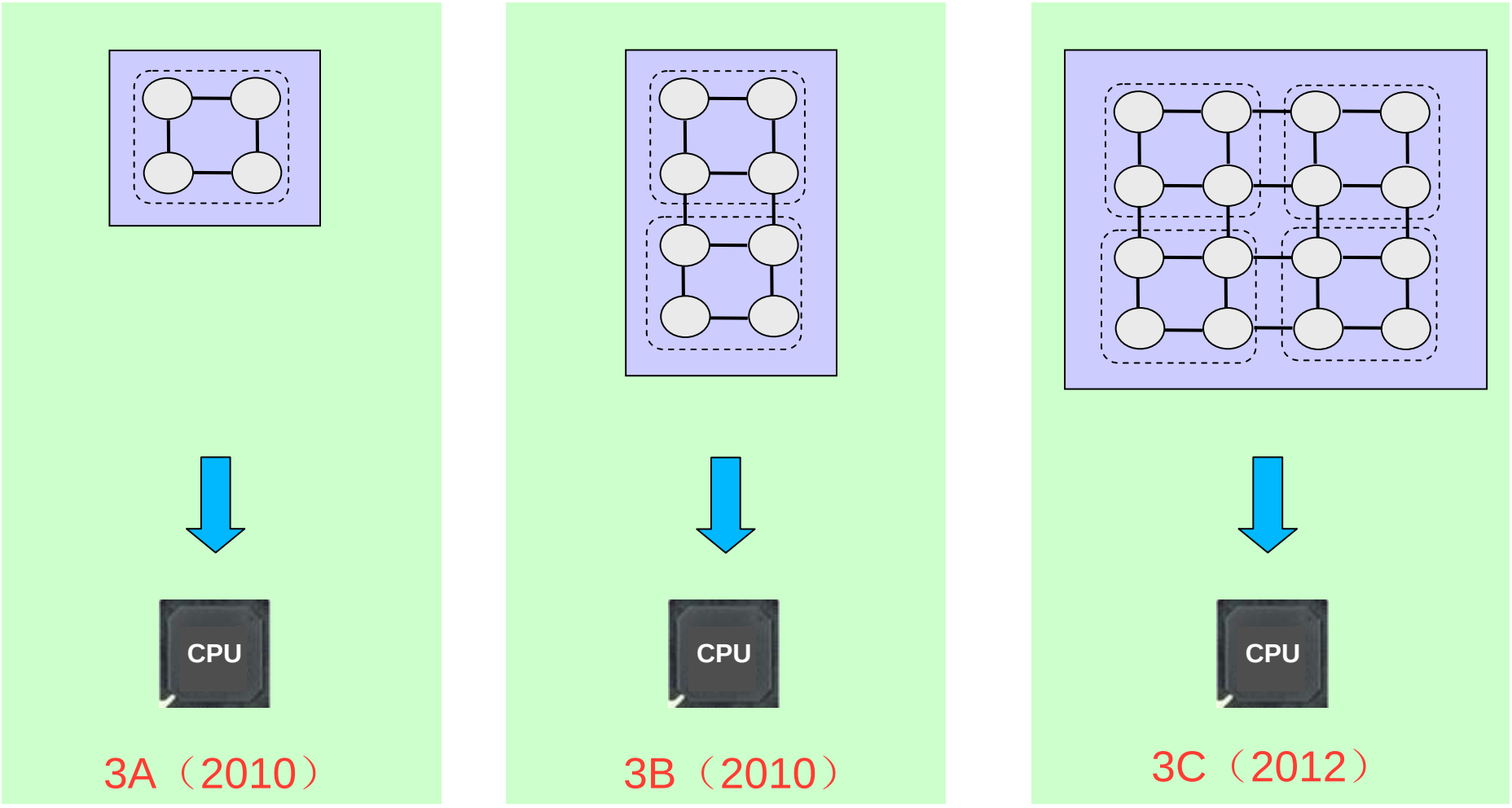


2G（2010）



2H（2011）

High end roadmap: More cores on a chip

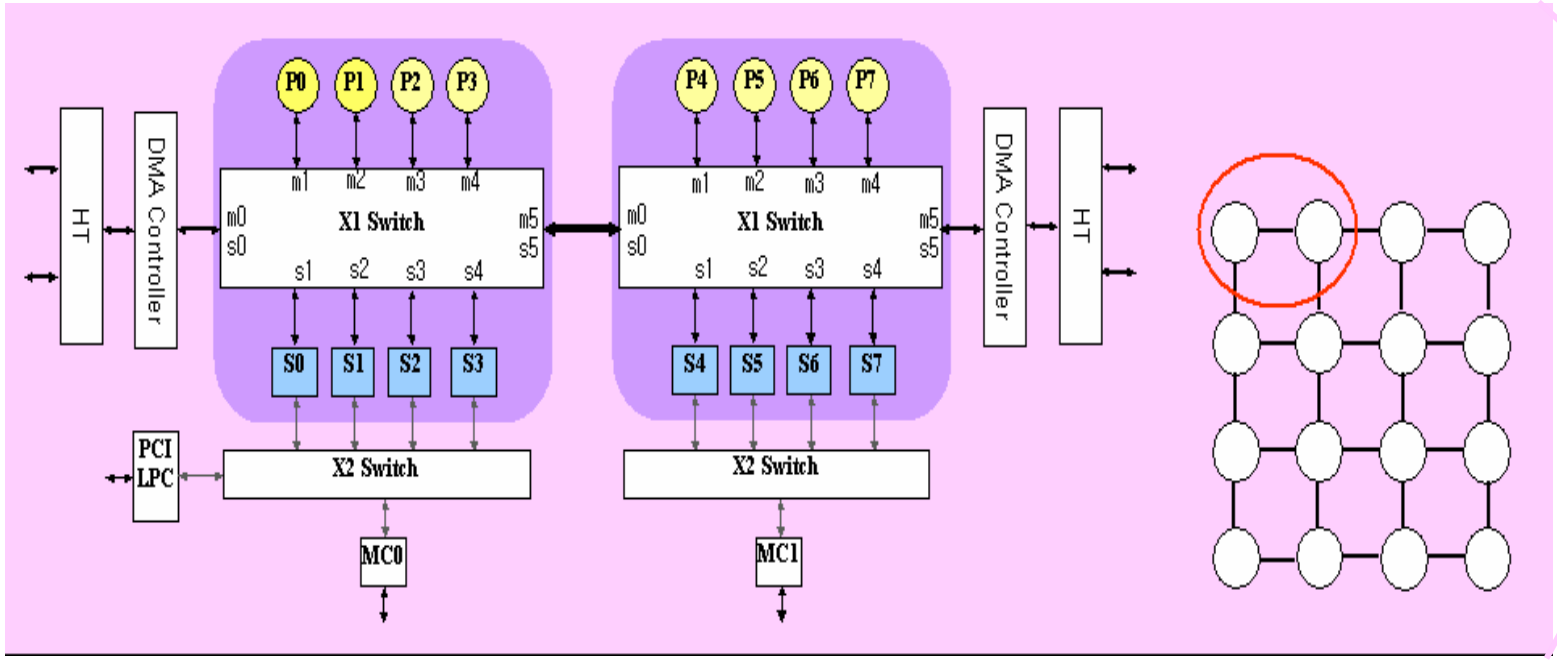


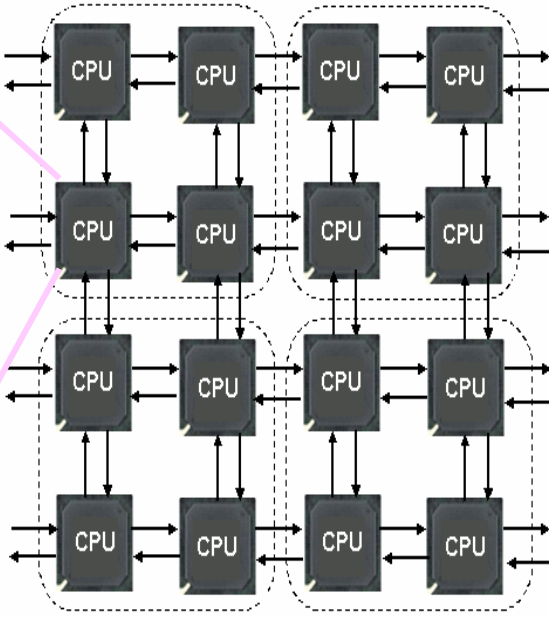
Contents

- **Background**
- **The Scalable Godson-3 Multicore Architecture**
- **The Vector Extension of the CPU Core**
- **HPCs Based on Godson-3**

Godson-3 Scalable Architecture

- **Scalable interconnection network:**
  - **Crossbar (intra-node) + Mesh (inter-node)**
- **Shared L2 Cache, on-chip memory controller**
  - **Directory-based cache coherence protocol for intra and inter chip CC**





4-core Godson-3A

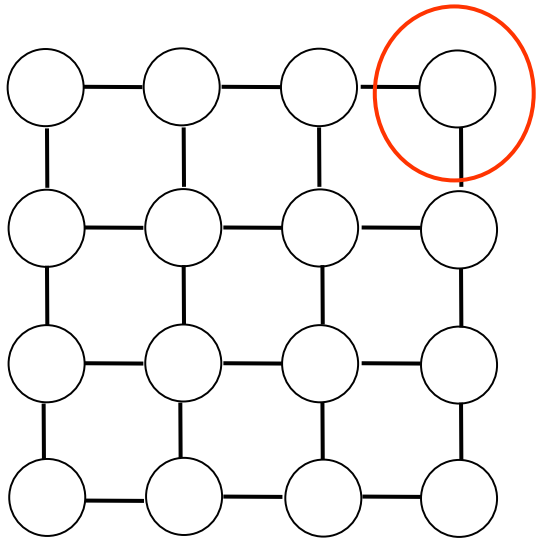
4 four-issue 64-bit Core

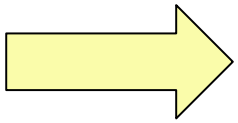
1.0GHz@65nm

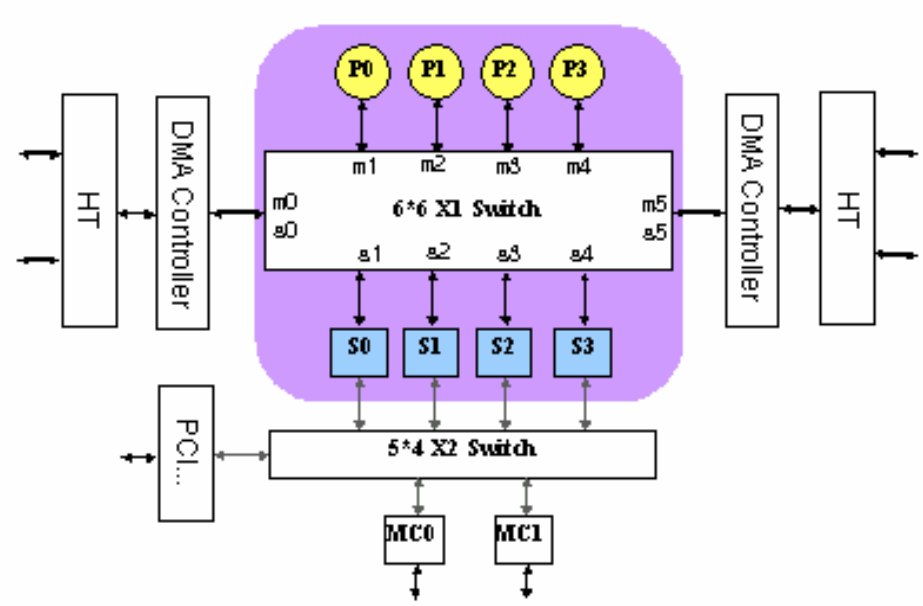
16GFLOPS@15W

2 DDR3, 2 HT Controllers

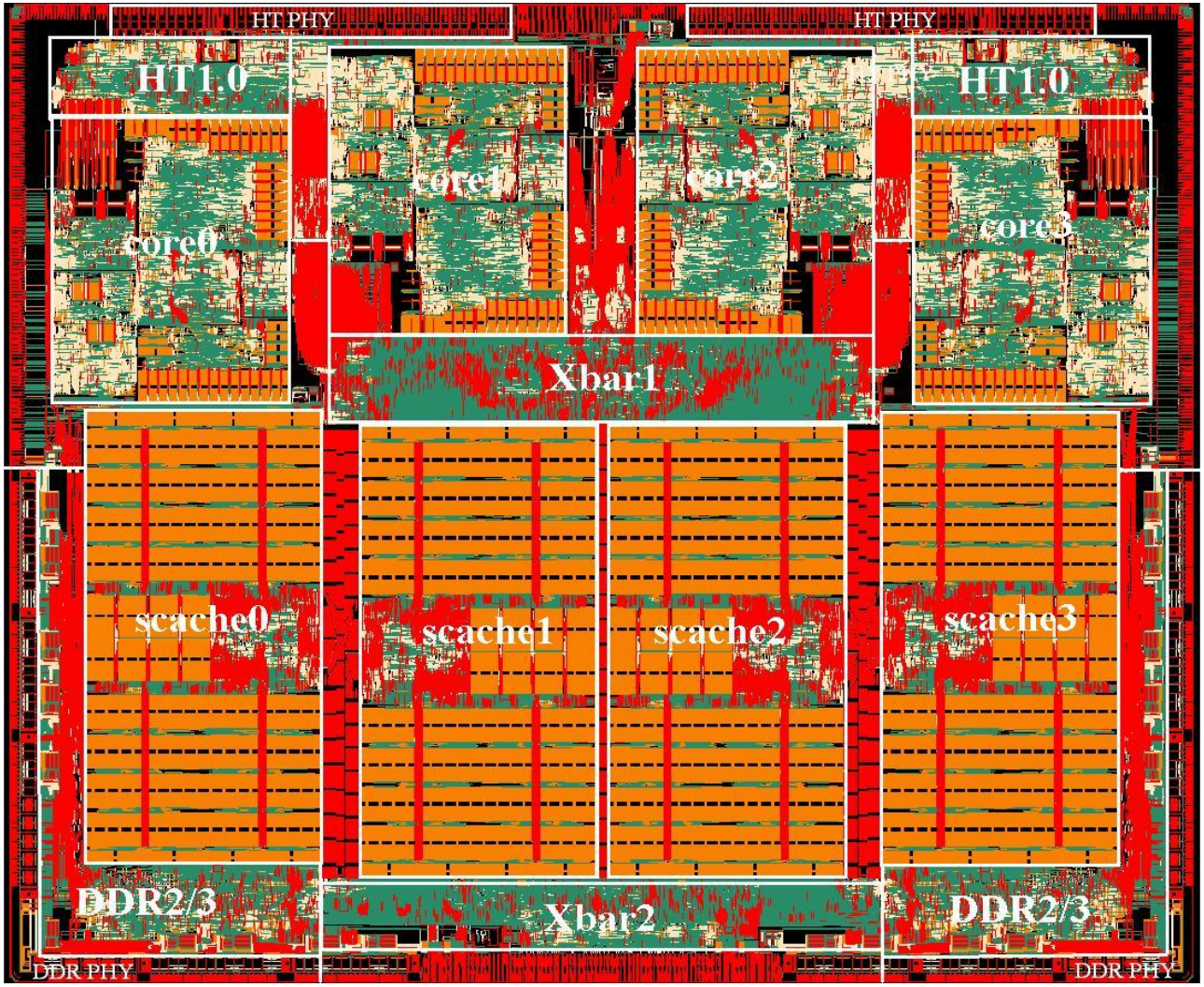
In product stage







Layout of Godson-3A



8-core Godson-3B

8 four-issue 64-bit core

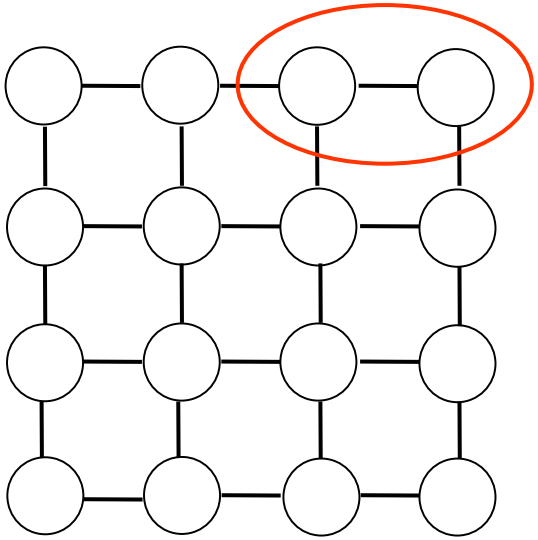
2*256-bit Vector Ext. per core

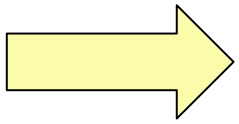
1.0GHz@65nm

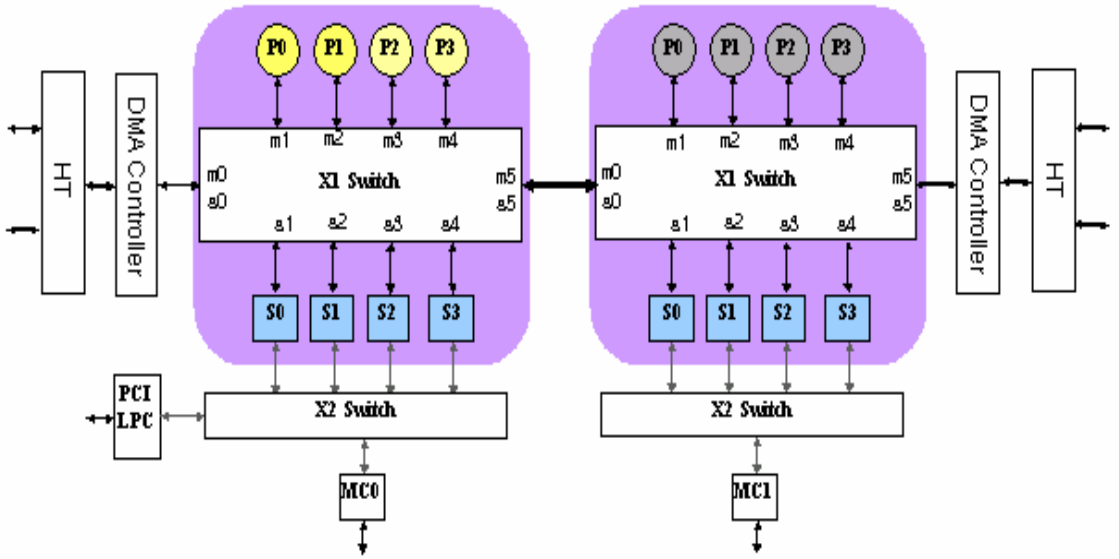
128GFLOPS@40W

2 DDR3, 2 HT Controllers

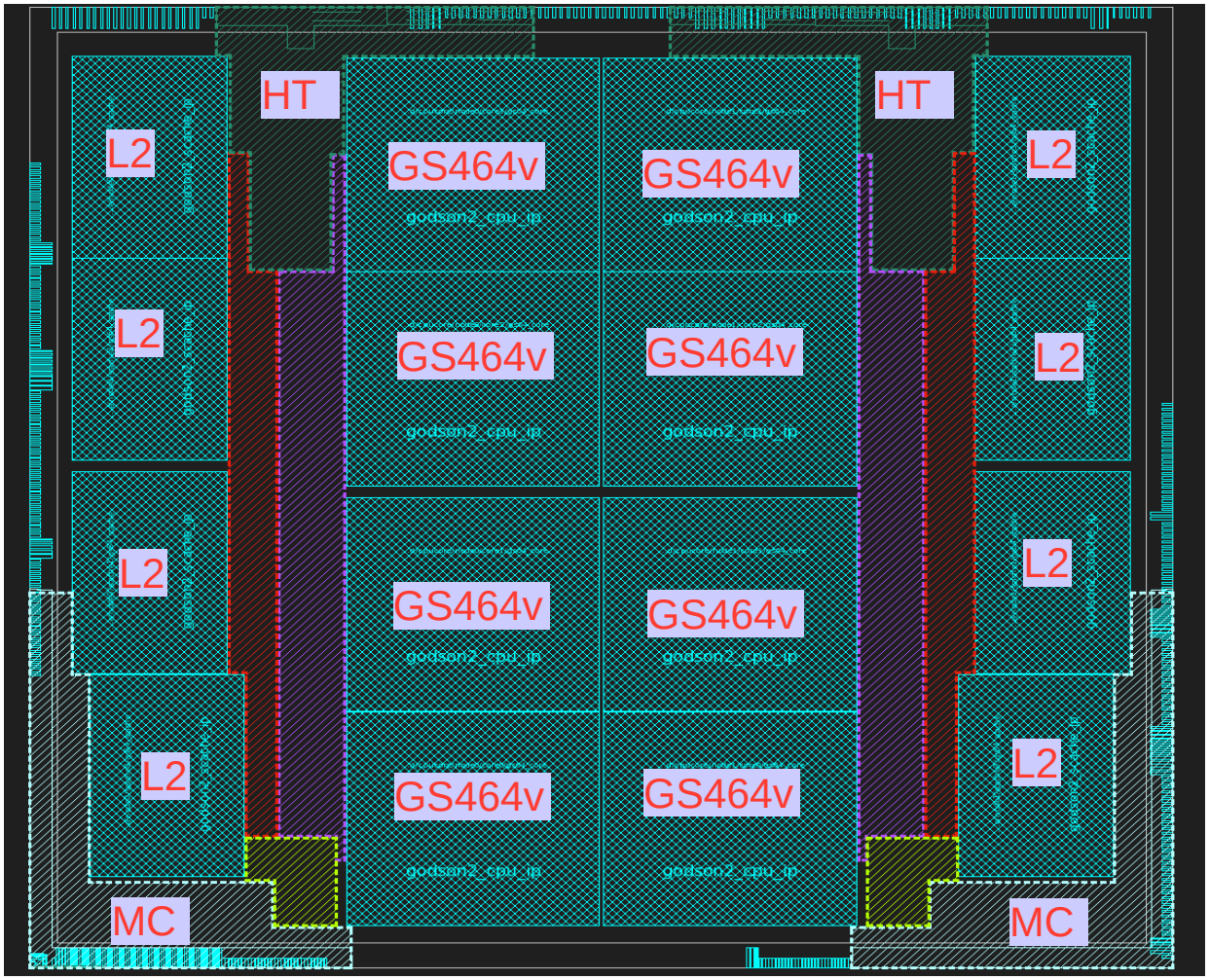
Taped out 2010.5







Layout of Godson-3B



16-core Godson-3C

16 four-issue 64-bit Core

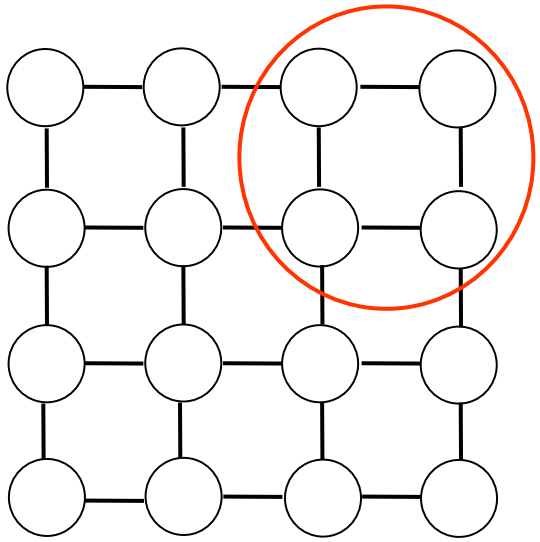
2*256-bit Vector Ext. per core

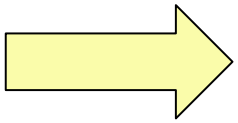
1.5GHz@28nm

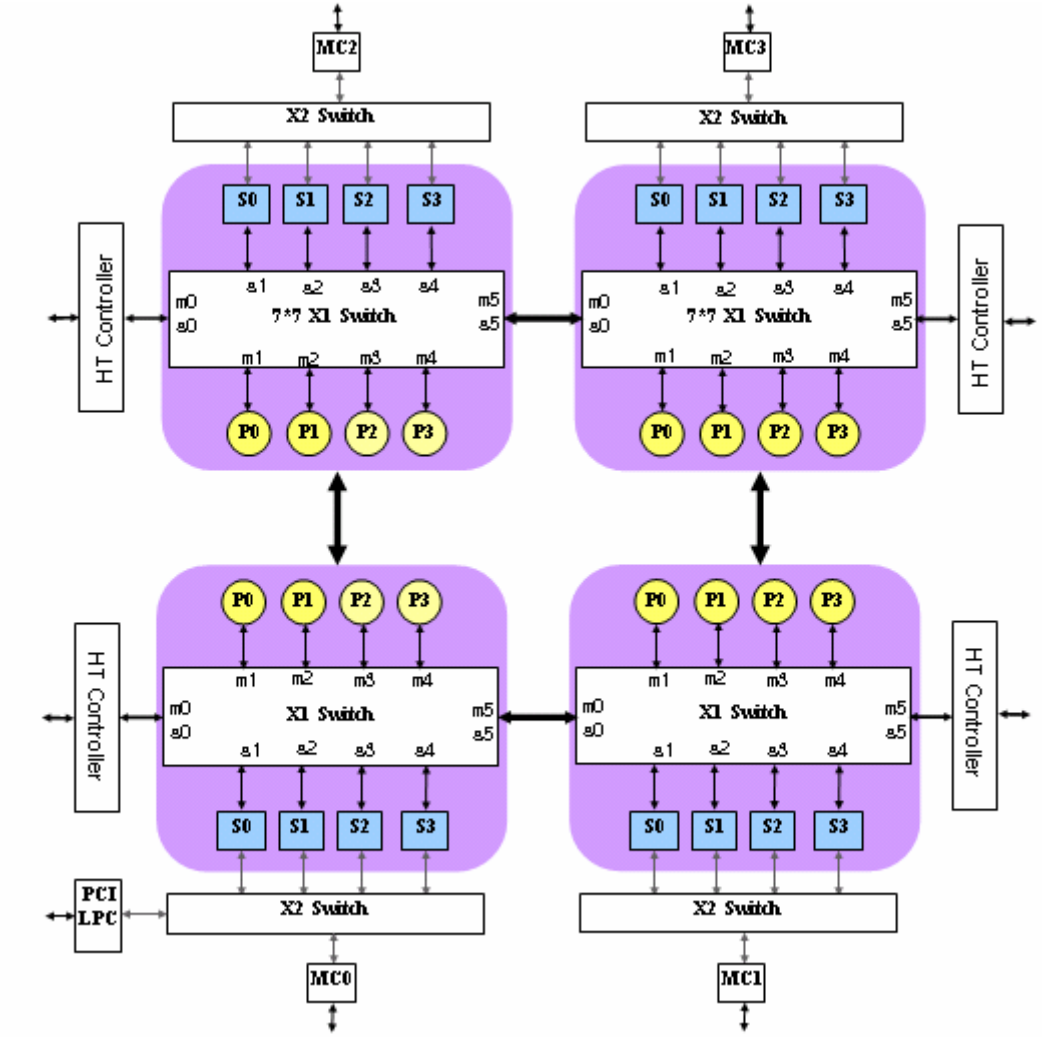
384GFLOPS@15W

4 DDR3, 4 HT Controllers

To be taped out 2011







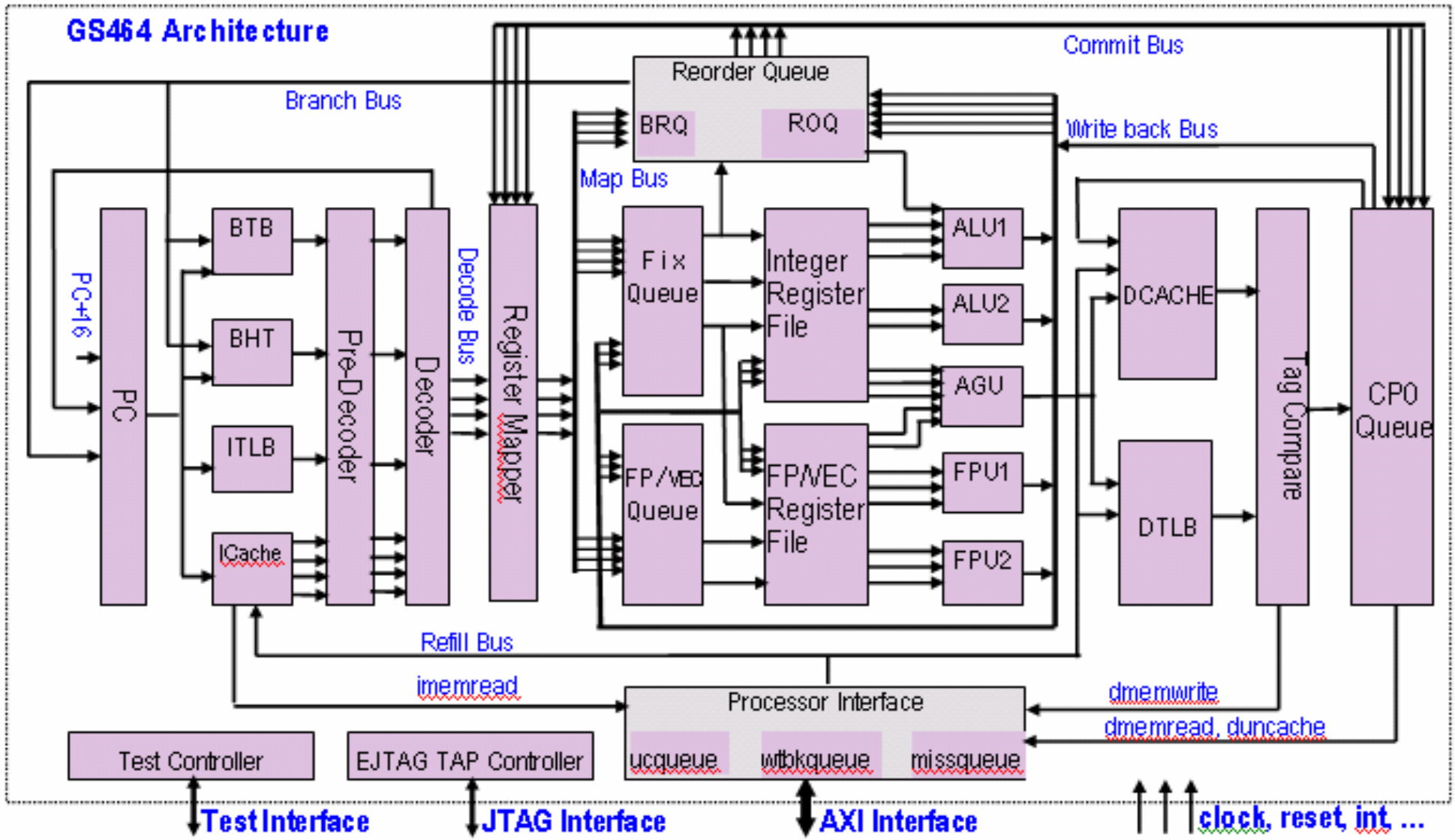
Contents

- **Background**
- **The Scalable Godson-3 Multicore Architecture**
- **The Vector Extension of the CPU Core**
- **HPCs Based on Godson-3**

The State-of-Art GS464 CPU Core

- **MIPS64 compatible, 200+ instructions for X86 emulation**
- **Four-issue 64-bit superscalar OOO pipeline**
- **Two fix, two FP, one memory units**
- **64KB icache and 64KB dcache, 4-way**
- **64-entry TLB, 16-entry ITLB**
- **Directory-based cache-coherence**
- **Parity check for icache, ECC for dcache**

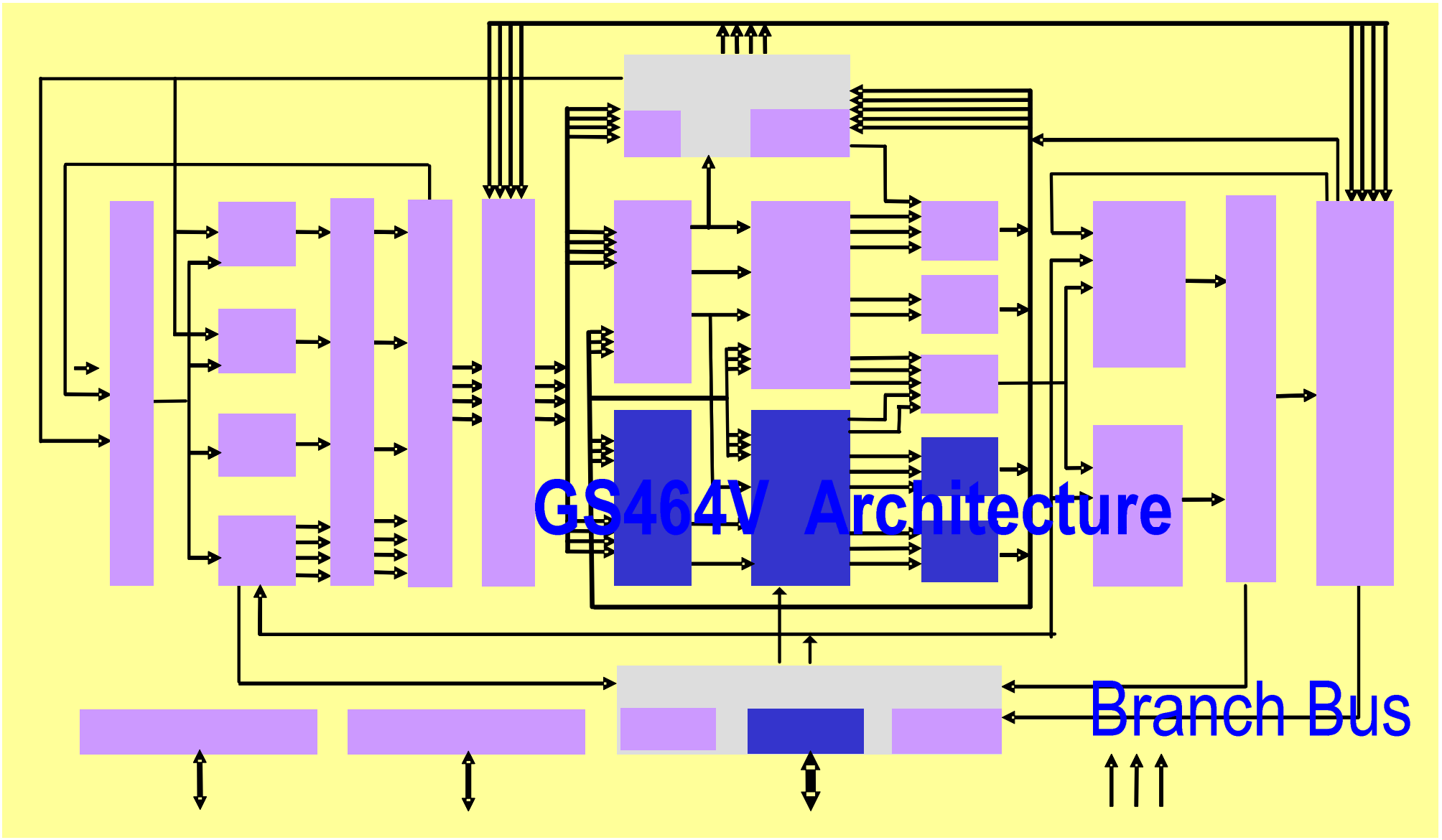
GS464 Architecture



GS464V Architecture Features

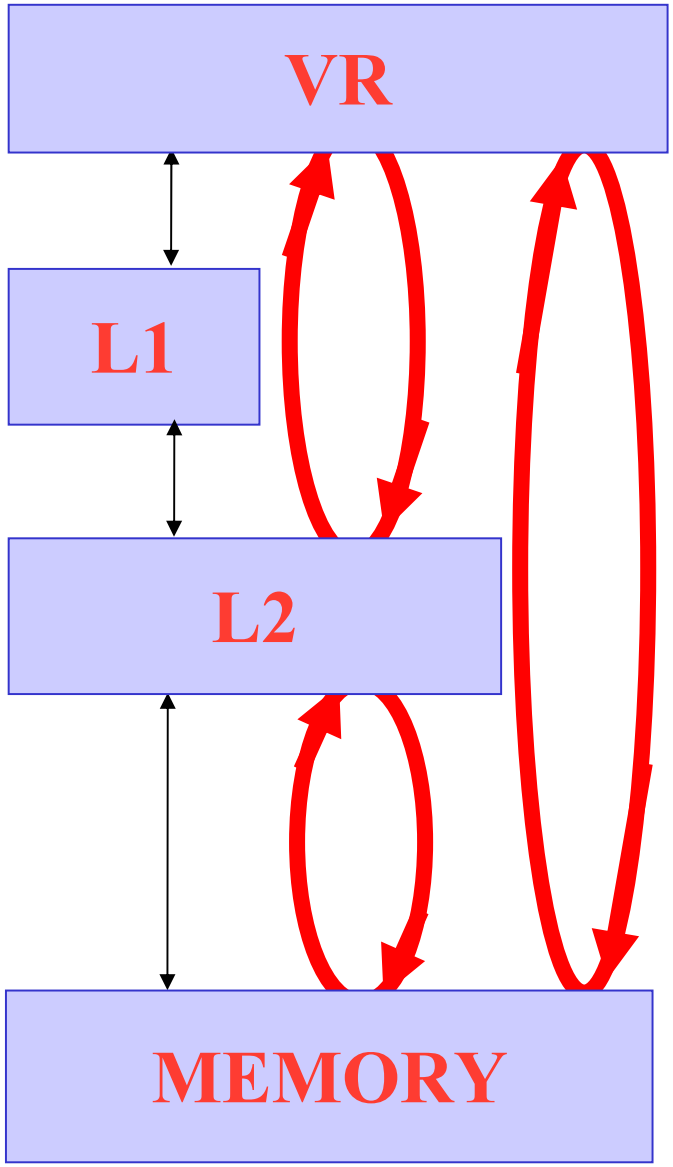
- **Keeps all GS464 features**
- **Extend each 64-bit FP unit to 256-bit SIMD vector unit**
  - **Two vector units**
  - **Each core has eight 64-bit MACs**
  - **Compatible with MIPS64 FP instruction**
- **128-entry 256-bit register file**
- **300+ SIMD instructions (Linpack, FFT, media……)**

Microarchitecture of GS464V



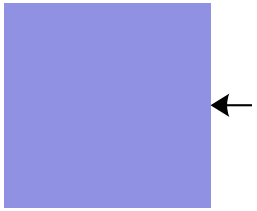
Feeding the Starving Vector Unit

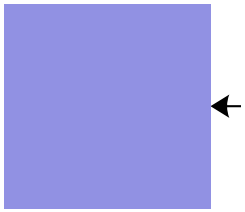
- **With limited bandwidth, how to provide enough data in required format to feed the vector unit ?**
  - **Use the bandwidth more efficiently**
- **Special data link for vector unit**
  - **Data moves in parallel with computation**
  - **Reorganizing data in the way from memory/cache to VR, as required by matrix computation, digital signal processing, media processing, etc.**

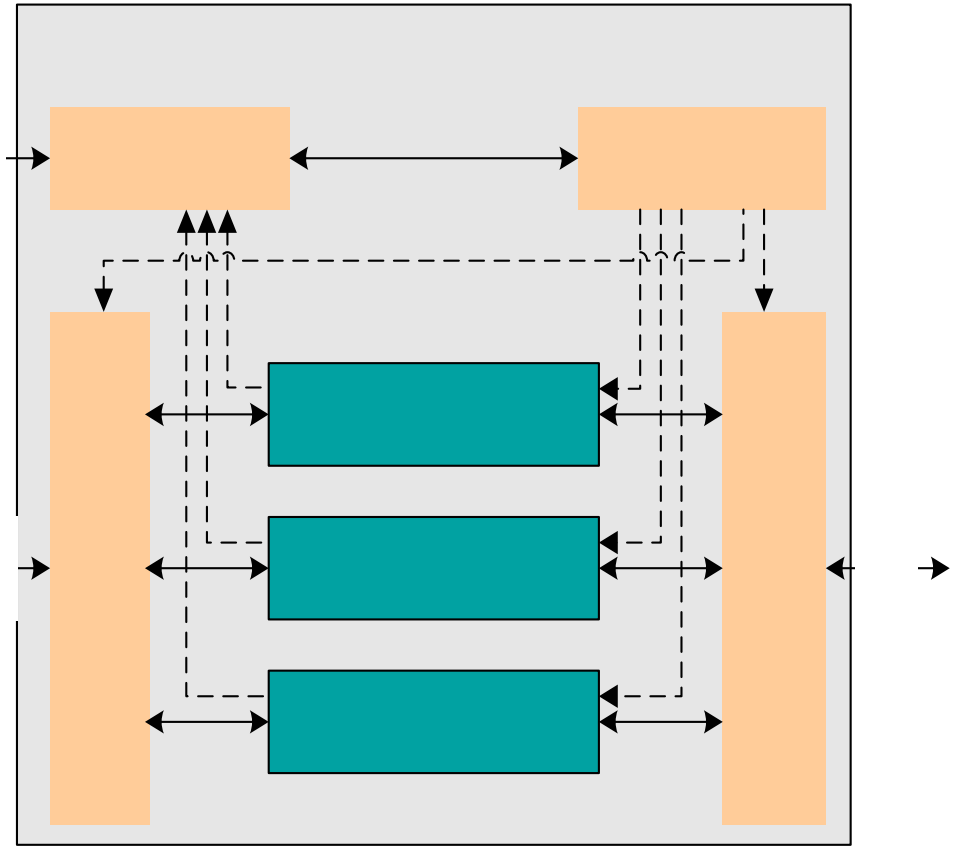


Godson Super Link (GSL)

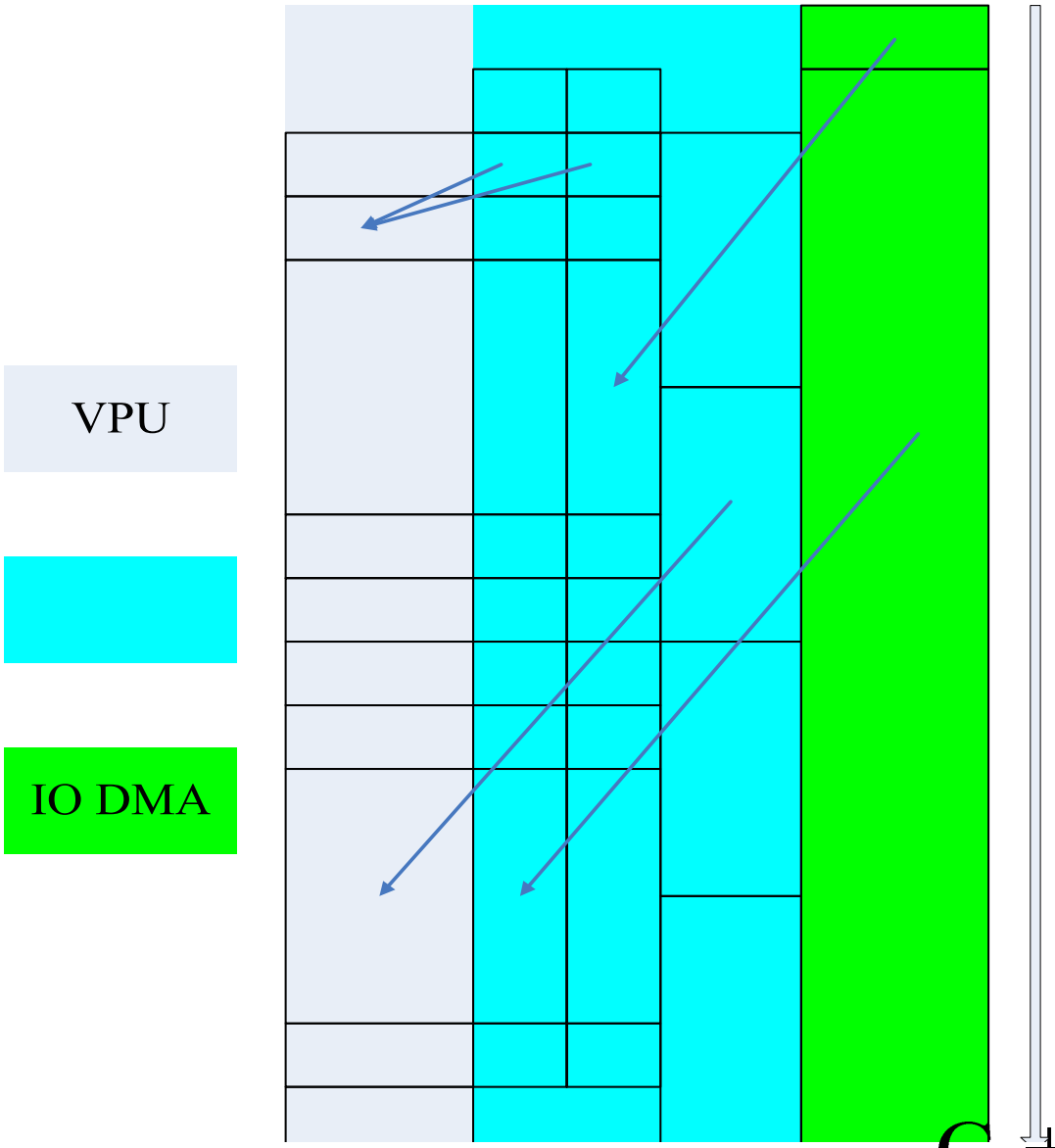
- **Direct & reconfigurable data transfer btw. cache/memory and VR**
  - **Matrix transposing**
  - **Bit revert**
  - **Entropy decoding**
  - **……**
- **Memory access coprocessor**
  - **Three Godson Super-Link**
  - **Flow control among GSLs**
  - **Synchronize with GS464V**







Parallel data movement and computation: Linpack as an example



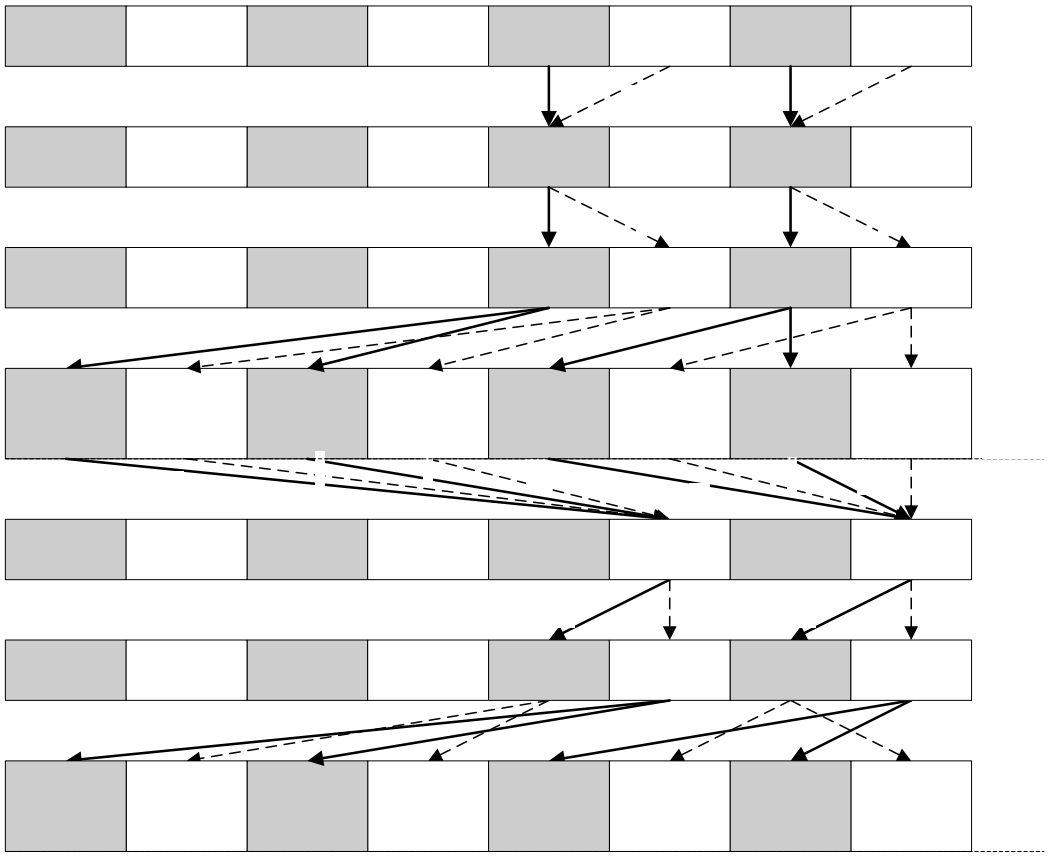
- **Computation and shuffle in parallel**
- **Computation and data transferring in parallel**
  - **VR↔L2**
  - **VR↔mem**

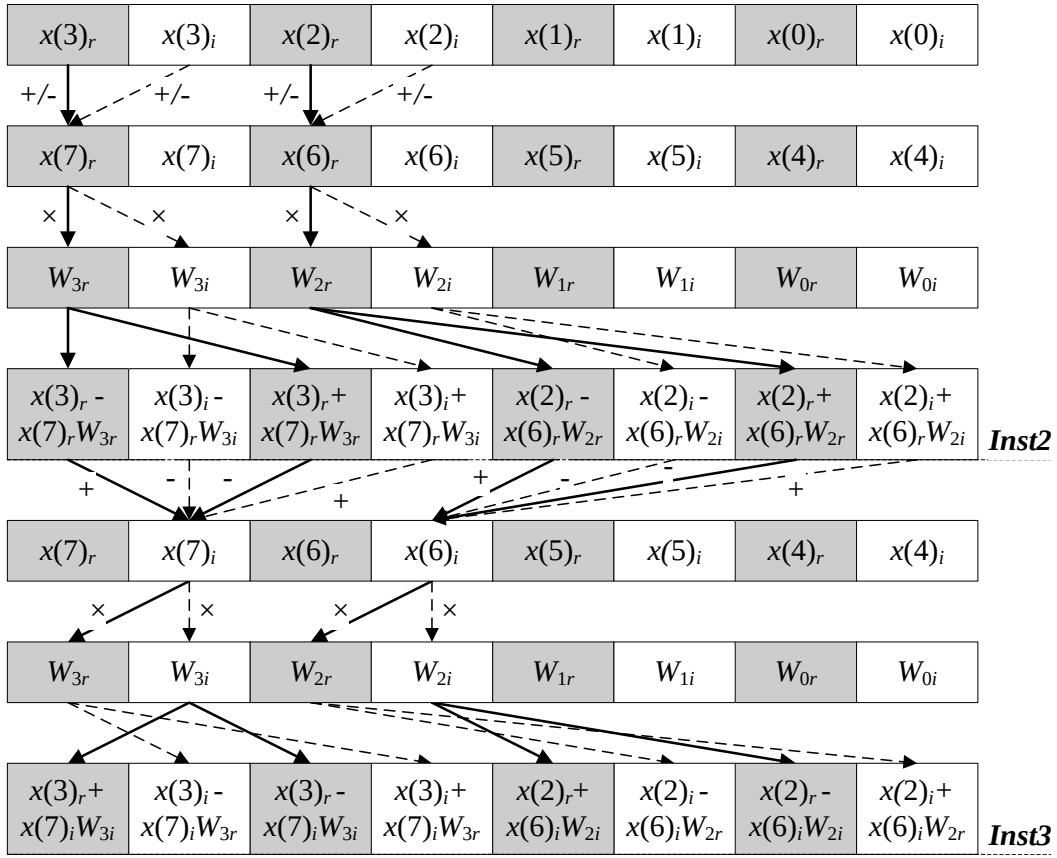
$C_0 += A_0 * B_0$

A_0 L2->vr B_0 L2->vr

A_1 L2->vr B_1 L2->vr

Computation and Shuffling in one Instruction: Vector Instruction for FFT

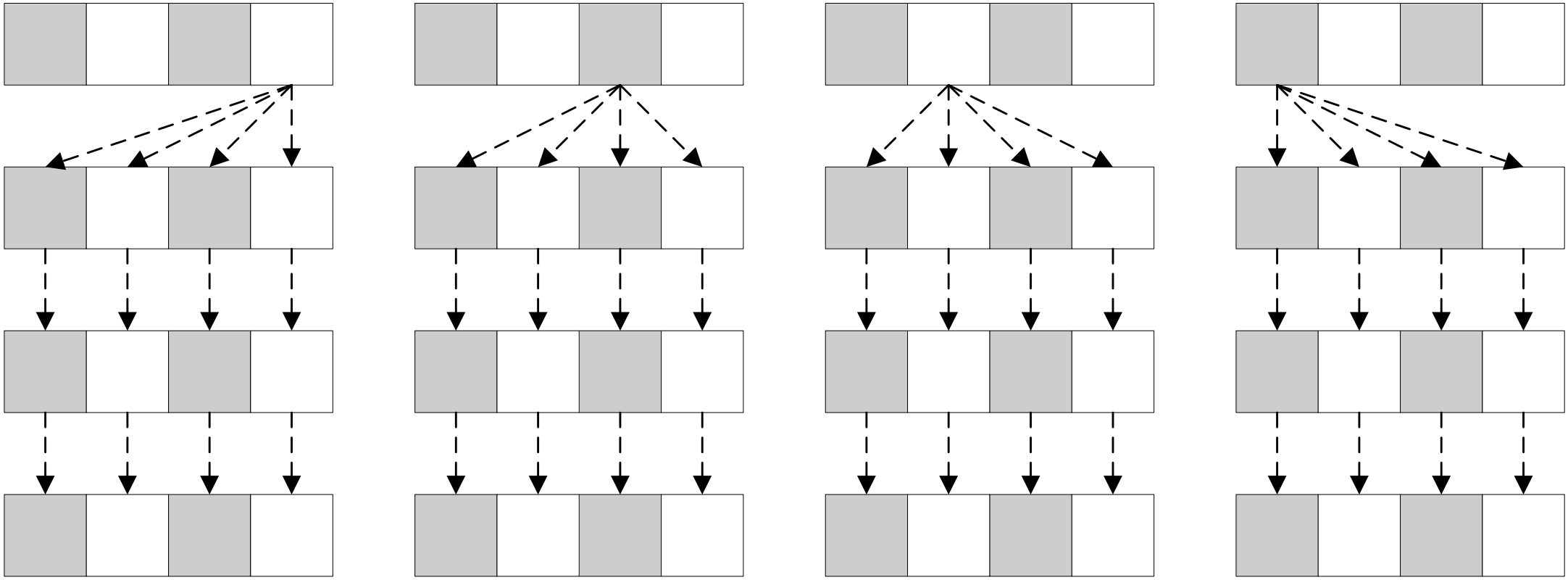




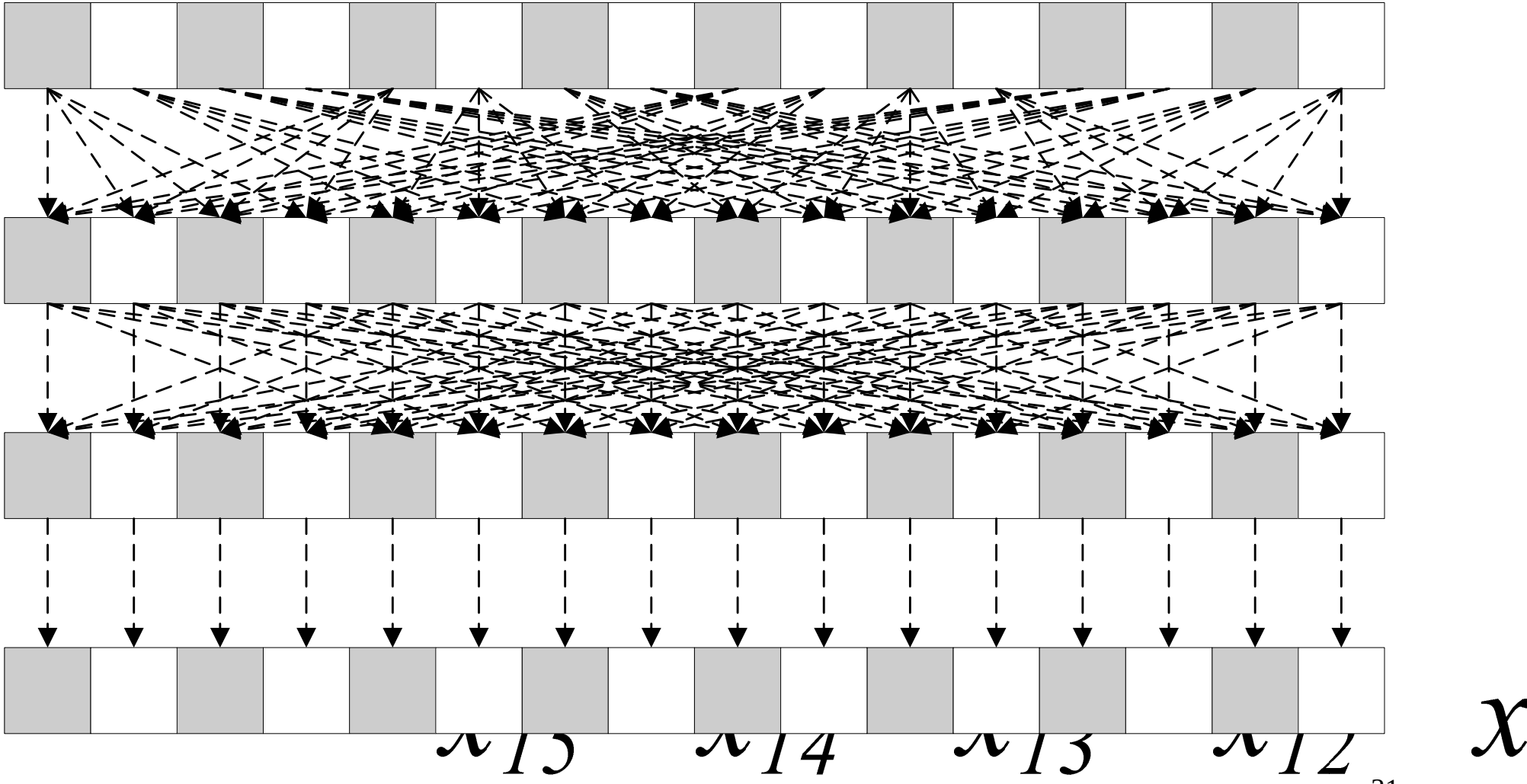
$x(3)_r$ $x(3)_i$ $x(2)_r$ $x(2)_i$ $x(1)_r$ $x($

+/- +/-

Computation and Shuffling in one Instruction:
Vector Instruction for Linpack



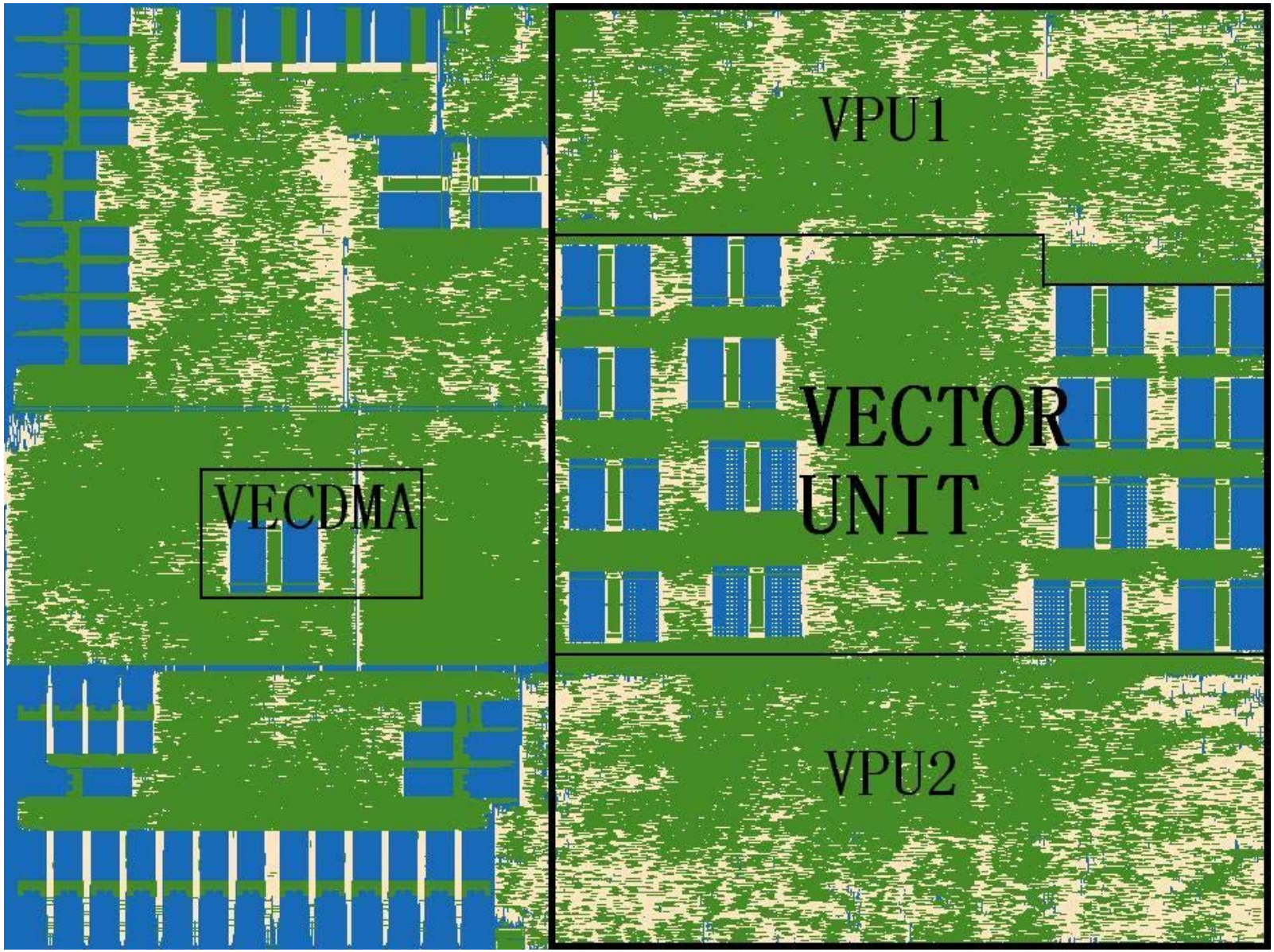
Computation and Shuffling in one Instruction: Vector Instruction for Media



Simulation Results

- **With RTL simulation, GS464V achieves excellent performance**
- **Eight-core godson-3B (64 MACs, 2 DDR3) achieves**
  - **>93% of peak performance for matrix multiplication**
  - **>87% of peak performance for 1024 point complex FFT (0.37us for 1024 point floating point FFT at 1GHz)**
- **1080p high definition H.264 decoding with single core at 1GHz**
  - **>200 frames per second**

GDSII of GS464V



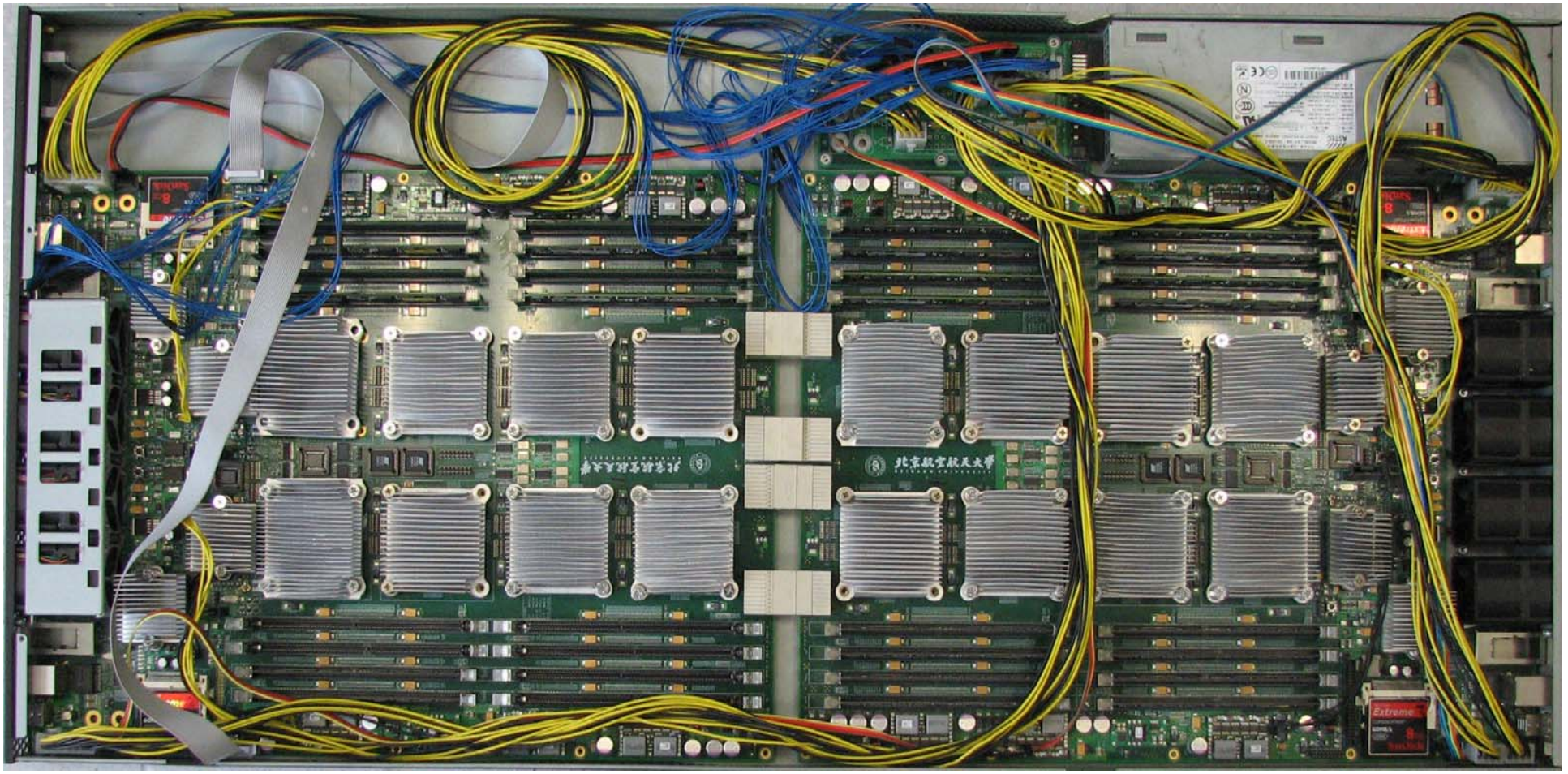
Contents

- **Background**
- **The Scalable Godson-3 Multicore Architecture**
- **The Vector Extension of the CPU Core**
- **HPCs Based on Godson-3**

HPCs based on Godson-3

- **Personal TeraFLOPS HPC**
  - **Desktop HPC**
- **PetaFLOPS in 2011 based on Godson-3B**
  - **10,000 Godson-3B Chips**
- **10 PetaFLOPS in 2013/2014 based on Godson-3C**
  - **30,000 Godson-3C Chips**

1U16P Board for HPC



- **1U2T with Godson-3B**
- **1U6T with Godson-3C**

Conclusion

- **Scalable Godson-3 Multi-core Architecture**
- **Aggressive vector extension to achieve high performance**
- **Godson Super Link helps efficient usage of limited bandwidth**
  - **Computation and data movement in parallel**
  - **Reorganize data in data movement**

Thanks