



Новосибирский государственный университет
Факультет информационных технологий
Кафедра параллельных вычислений

Ускоритель Intel Xeon Phi

Введение. Архитектура

Преподаватель:
Киреев С.Е.

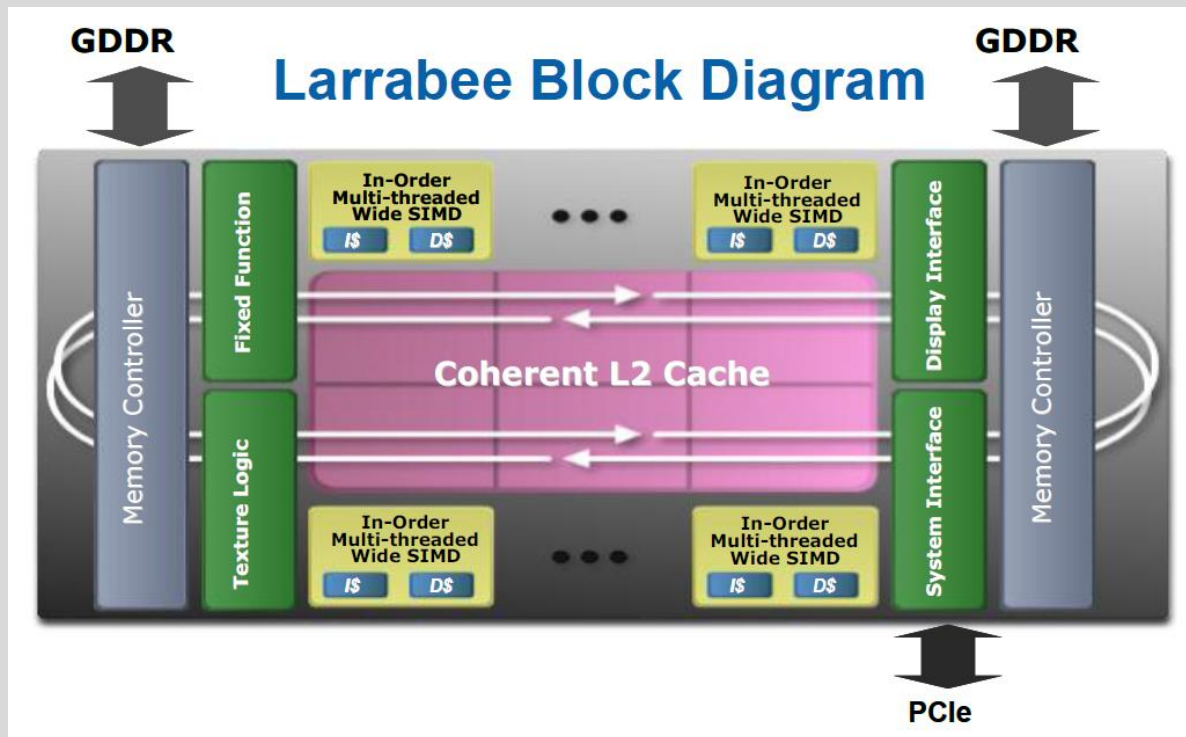


ПРЕДЫСТОРИЯ

Предыстория

Intel Larrabee (2006 – 2011)

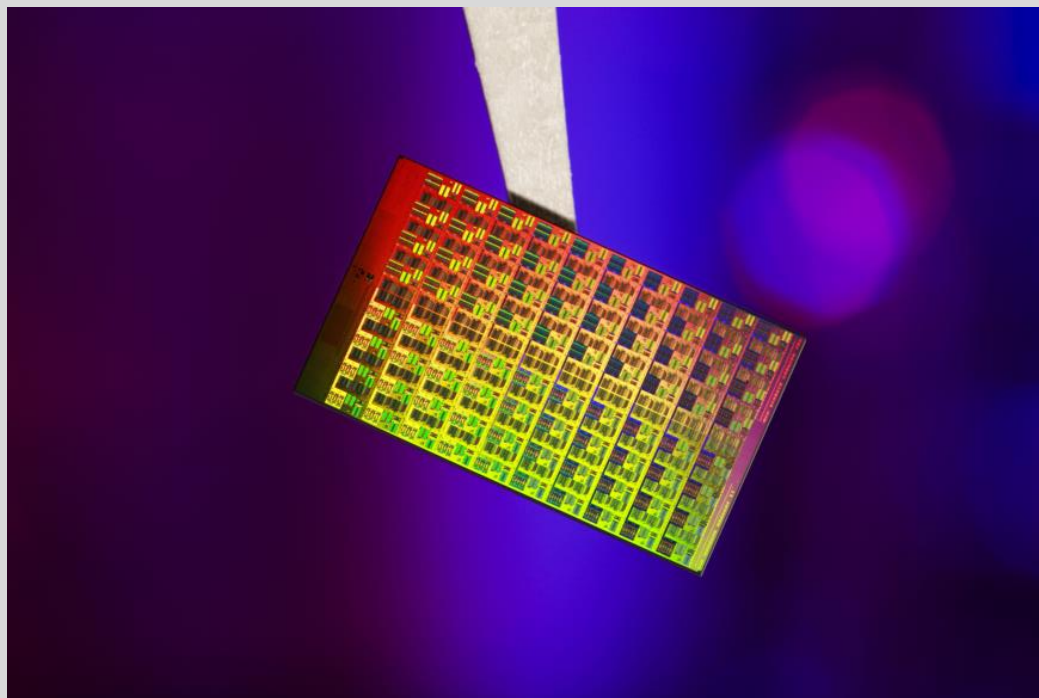
- 16 ядер: на основе архитектуры x86-64
- Гибрид CPU и GPU: включает спец. ядра – стадии графического конвейера
- Сопроцессор – не предназначен для запуска ОС



Предыстория

Intel Polaris (2006) – Teraflops Research Chip

- Матрица из 80 ядер (по 2 FPU на ядро)
- 1.01 TFLOPS (3.16 GHz) – при нормальной температуре
- 1.81 TFLOPS (5.7 GHz) – со спец. охлаждением



Предыстория

Intel Single-chip Cloud Computer (2009)

- Матрица из 48 ядер архитектуры IA



Предыстория

Intel Knights Ferry (2010)

- 1-е поколение архитектуры MIC – Many Integrated Cores



The image shows an Intel Knights Ferry MIC card, a small circuit board with a fan and the Intel logo. It is set against a dark blue background with a grid pattern.

Knights Ferry

- Software development platform
- Growing availability through 2010
- 32 cores, 1.2 GHz
- 128 threads at 4 threads / core
- 8MB shared coherent cache
- 1-2GB GDDR5
- Bundled with Intel HPC tools

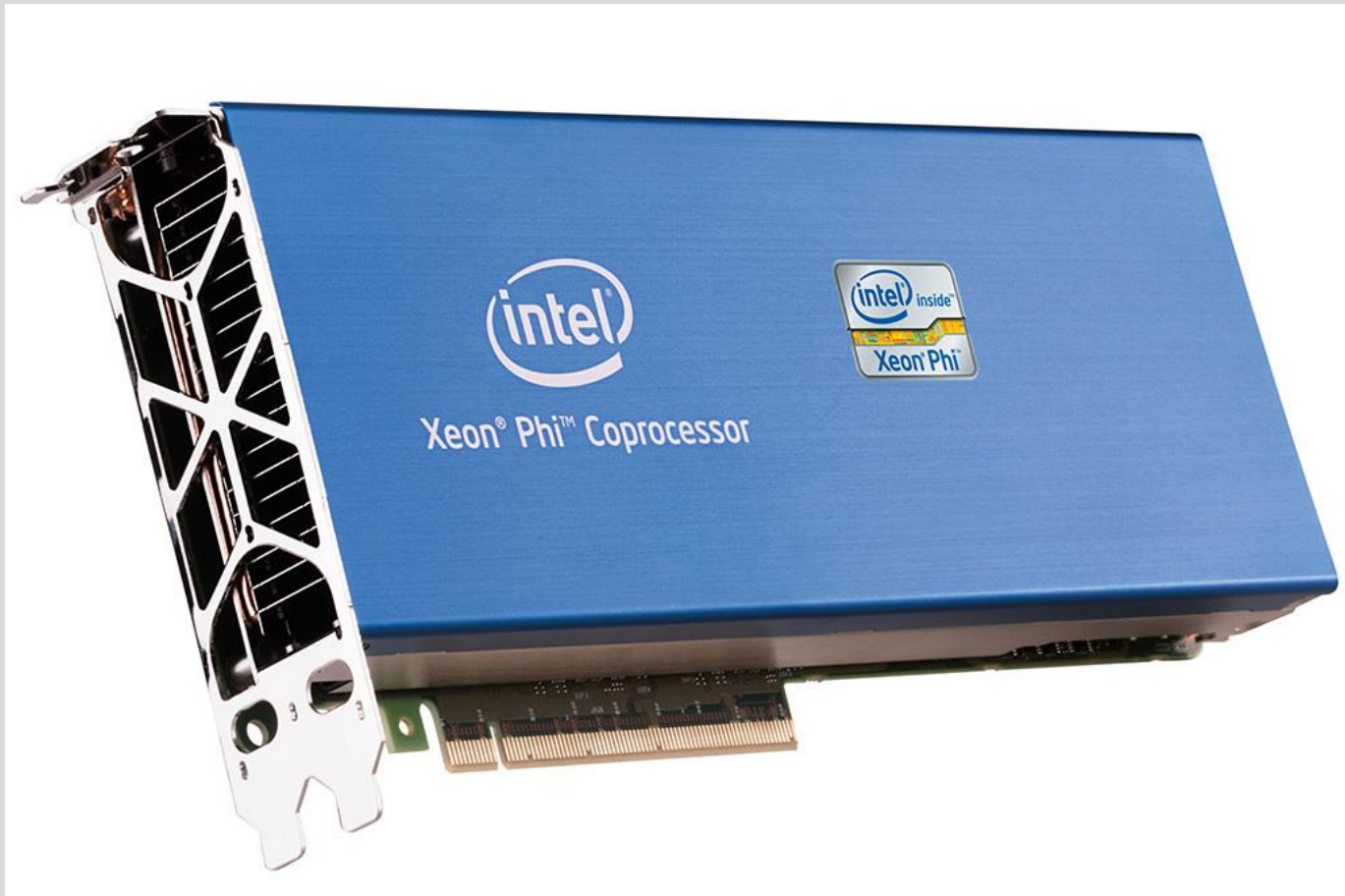
Software development platform for Intel® MIC architecture



Предыстория

Intel Knights Corner (2012) – Xeon Phi

- 2-е поколение архитектуры MIC – Many Integrated Cores



Хеон Phi в мировых рейтингах

Top500
Ноябрь 2014

RANK	SITE	SYSTEM	CORES	RMAX (TFLOP/S)	RPEAK (TFLOP/S)	POWER (KW)
1	National Super Computer Center in Guangzhou China	Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Swiss National Supercomputing Centre (CSCS) Switzerland	Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Cray Inc.	115,984	6,271.0	7,788.9	2,325
7	Texas Advanced Computing Center/Univ. of Texas United States	Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462,462	5,168.1	8,520.1	4,510
8	Forschungszentrum Juelich (FZJ) Germany	JUQUEEN - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	458,752	5,008.9	5,872.0	2,301
9	DOE/NNSA/LLNL United States	Vulcan - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393,216	4,293.3	5,033.2	1,972
10	Government United States	Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40 Cray Inc.	72,800	3,577.0	6,131.8	1,499

Хеон Phi в мировых рейтингах

Green500

Ноябрь 2014

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	5,271.81	GSI Helmholtz Center	L-CSC - ASUS ESC4000 FDR/G2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR, AMD FirePro S9150 Level 1 measurement data available	57.15
2	4,945.63	High Energy Accelerator Research Organization /KEK	Suiren - ExaScaler 32U256SC Cluster, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, PEZY-SC	37.83
3	4,447.58	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104Re-1G Cluster, Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	35.39
4	3,962.73	Cray Inc.	Storm1 - Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40m Level 3 measurement data available	44.54
5	3,631.70	Cambridge University	Wilkes - Dell T620 Cluster, Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62
6	3,543.32	Financial Institution	iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	54.60
...				
18	2,598.21	Exploration & Production - Eni S.p.A.	HPC2 - iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, NVIDIA K20x	1,227.00
19	2,540.02	Tulane University	Cypress - Dell C8220X Cluster, Intel Xeon E5-2680v2 10C 2.8GHz, InfiniBand 4x FDR, Intel Xeon Phi 7120P	87.00
20	2,495.12	Mississippi State University	Shadow - Cray CS300-LC, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, Intel Xeon Phi 5110P	101.93
21	2,401.41	St. Petersburg Polytechnic University	RSC PetaStream - RSC PetaStream, Xeon E5-2697v3 14C 2.6GHz, Infiniband FDR, Intel Xeon Phi 5120D	71.00
22	2,386.42	Government	Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40	1,498.90

Хеон Phi в мировых рейтингах

Graph500

Ноябрь 2014

No.	Rank ▲	Machine	Installation Site	Number of nodes	Number of cores	Problem scale	GTEPS
1	1	DOE/NNSA/LLNL Sequoia (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Lawrence Livermore National Laboratory	98304	1572864	41	23751
2	2	K computer (Fujitsu - Custom supercomputer)	RIKEN Advanced Institute for Computational Science (AICS)	82944	663552	40	19585.2
3	3	DOE/SC/Argonne National Laboratory Mira (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Argonne National Laboratory	49152	786432	40	14982
4	4	JUQUEEN (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Forschungszentrum Juelich (FZJ)	16384	262144	38	5848
5	5	Fermi (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	CINECA	8192	131072	37	2567
6	6	Tianhe-2 (MilkyWay-2) (National University of Defense Technology - MPP)	Changsha, China	8192	196608	36	2061.48
7	7	Turing (IBM - BlueGene/Q, Power BQC 16C 1.60GHz)	CNRS/IDRIS-GENCI	4096	65536	36	1427
8	7	Blue Joule (IBM - BlueGene/Q, Power BQC 16C 1.60 GHz)	Science and Technology Facilities Council - Daresbury Laboratory	4096	65536	36	1427

Хеон Phi в мировых рейтингах

Top50
СНГ

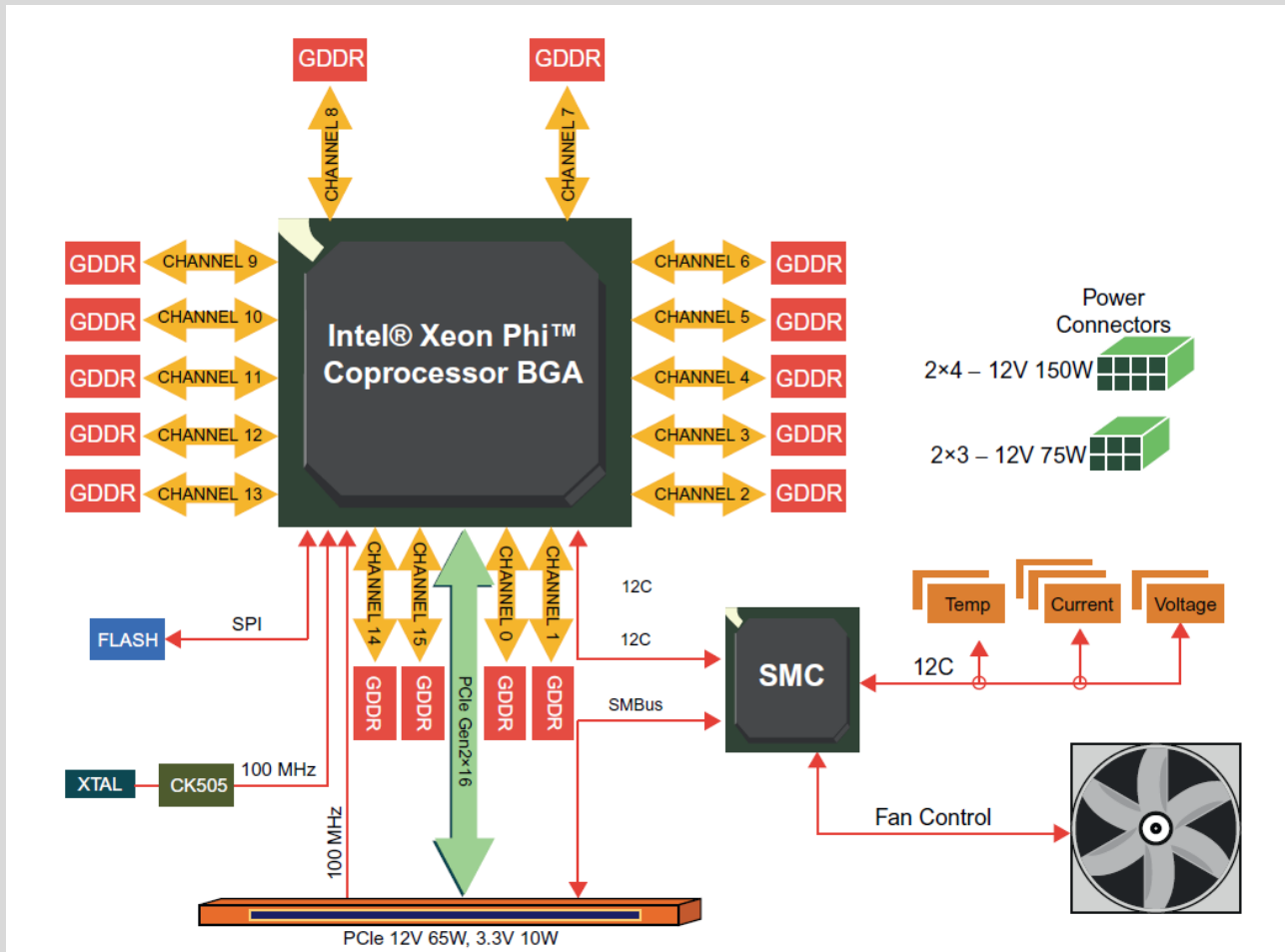
Текущий рейтинг

21-ая редакция от 23.09.2014

N	Место	Кол-во CPU/ядер	Архитектура (тип процессора / сеть)	Производительность (Tflop/s)		Разработчик
				Linpack	Пиковая	
1	Москва Московский государственный университет имени М.В.Ломоносова 2012 г.	12422/82468	узлов: 4160 (2xXeon 5570 2.93 GHz 12 GB RAM) узлов: 777 (2xXeon E5630 [Acc: 2xTesla X2070] 2.53 GHz 12 GB RAM) узлов: 640 (2xXeon 5670 2.93 GHz 24 GB RAM) узлов: 288 (2xXeon E5630 [Acc: 2xTesla X2070] 2.53 GHz 24 GB RAM) узлов: 260 (2xXeon 5570 2.93 GHz 24 GB RAM) узлов: 40 (2xXeon 5670 2.93 GHz 48 GB RAM) узлов: 30 (2xPowerXCell 8i 3.2 GHz 16 GB RAM) узлов: 4 (4xXeon E7650 2.26 GHz 512 GB RAM) сеть: Infiniband QDR/Gigabit Ethernet/Gigabit Ethernet	901.90	1,700.21	T-Платформы
2	Москва Поставщик ИТ-услуг 2014 г.	300/8250	узлов: 150 (2xXeon E5-2660 [Acc: 3x Kepler K20] 2.2 GHz 64 GB RAM) сеть: Gigabit Ethernet/Gigabit Ethernet	408.98	568.74	Hewlett-Packard
3	Москва ФГБУН Межведомственный суперкомпьютерный центр Российской академия наук 2012 г.	416/28704	узлов: 208 (2xXeon E5-2690 [Acc: 2x Xeon Phi 7110X] 2.9 GHz 80 GB RAM) сеть: FDR Infiniband/Gigabit Ethernet/Fast Ethernet	375.70	523.83	Группа компаний РСК
4	Москва Московский государственный университет имени М.В.Ломоносова 2014 г.	256/6400	узлов: 256 (Xeon E5-2680v2 [Acc: Tesla K40M] 2.8 GHz 32.768 GB RAM) сеть: FDR Infiniband/2xGigabit Ethernet/Gigabit Ethernet	319.80	423.40	T-Платформы
5	Челябинск Южно-Уральский государственный университет 2013 г.	768/28032	узлов: 384 (2xXeon X5680 [Acc: Xeon Phi SE10X] 3.33 GHz 24.576 GB RAM) сеть: QDR Infiniband/QDR Infiniband/Gigabit Ethernet	288.20	473.64	Группа компаний РСК

МИКРОАРХИТЕКТУРА

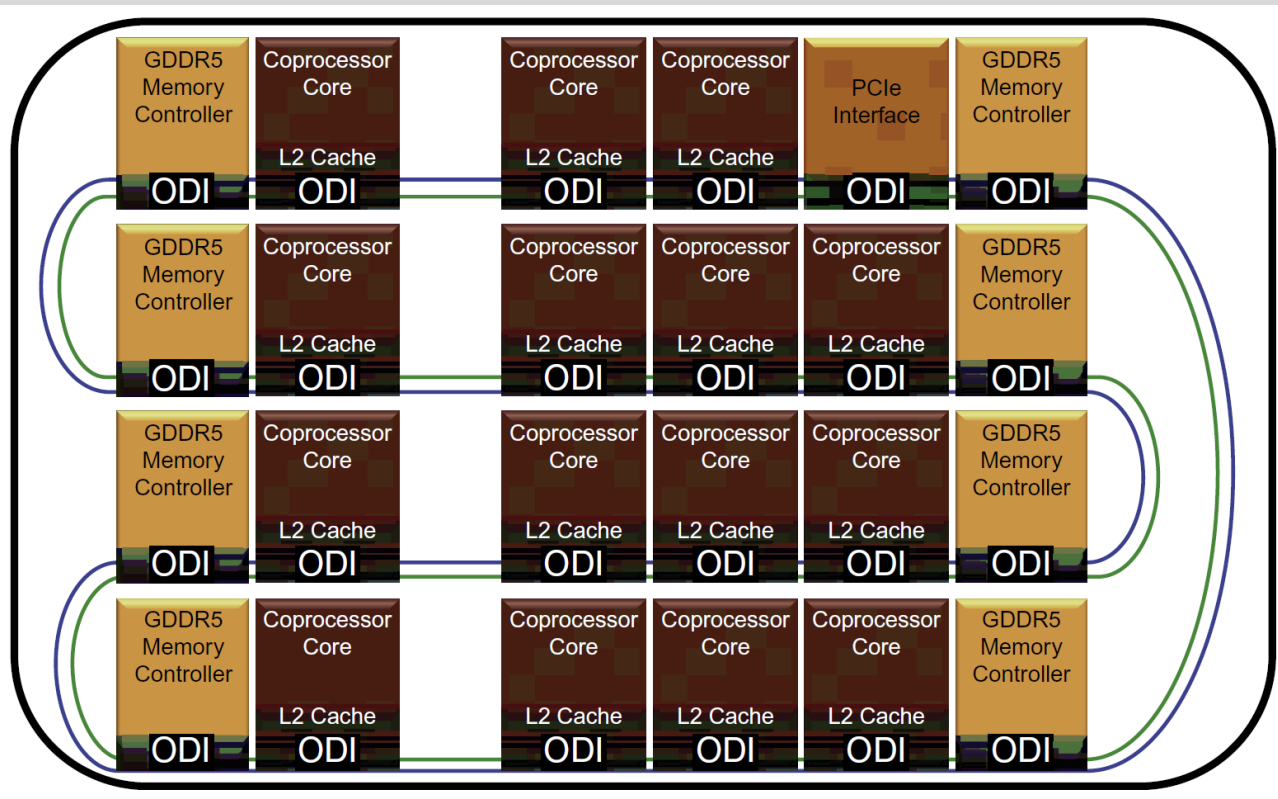
Схема карты сопроцессора



Компоненты

- Процессор
- Память
- FLASH-память с ОС Linux
- PCIe
- ...

Схема процессора



Компоненты

- Ядра (57/60/61)
- Контроллеры памяти
- Интерфейс PCIe
- Двухнаправленная кольцевая шина (on-die interconnect)

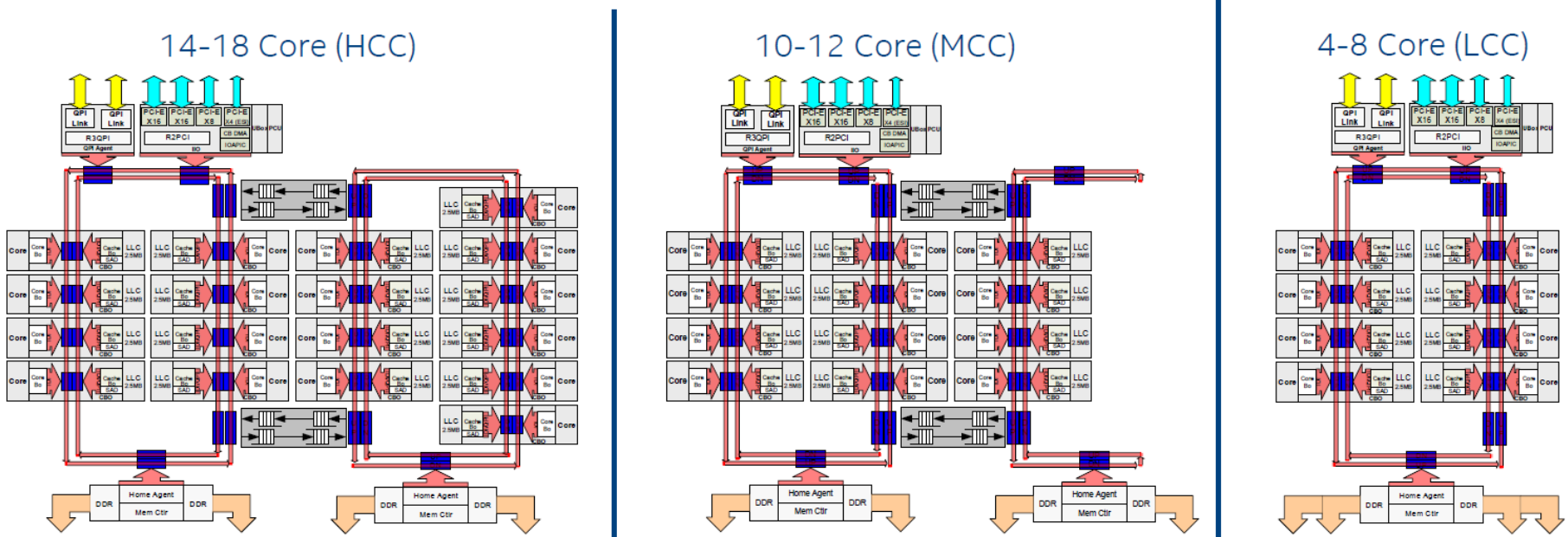
Частота

- 1.1 – 1.238 ГГц

Схема процессора

- Сравнение с современными процессорами общего назначения

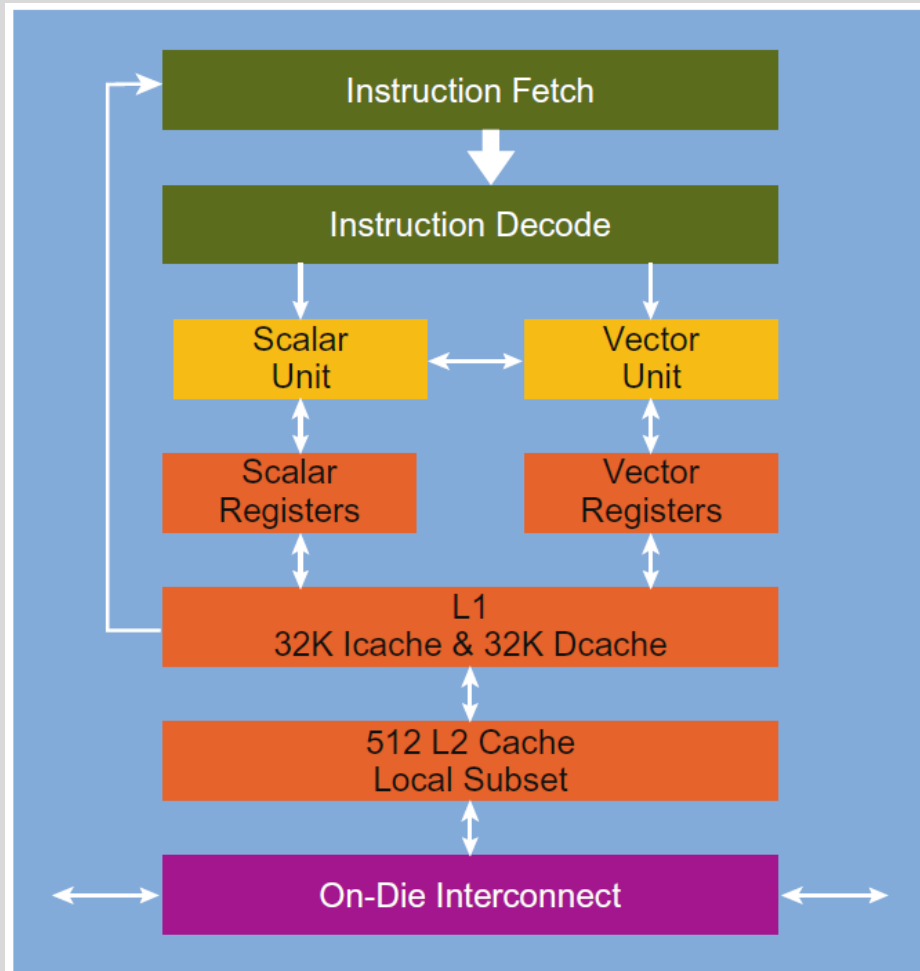
Haswell EP Die Configurations



Not representative of actual die-sizes, orientation and layouts – for informational use only.

Chop	Columns	Home Agents	Cores	Power (W)	Transistors (B)	Die Area (mm ²)
HCC	4	2	14-18	110-145	5.69	662
MCC	3	2	6-12	65-160	3.84	492
LCC	2	1	4-8	55-140	2.60	354

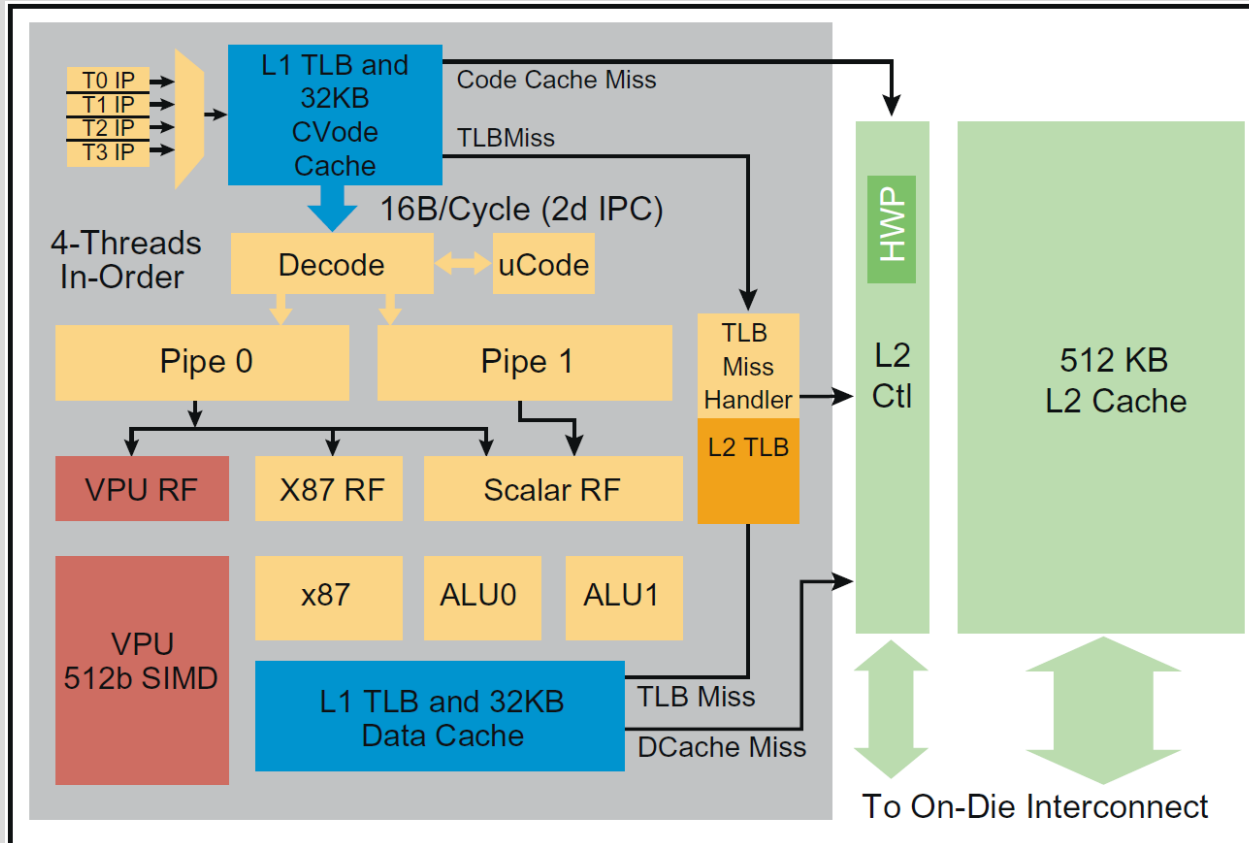
Схема ядра процессора



Особенности

- Выполнение до 2-х команд за такт
- Для каждого конвейера свои команды
- Выполнение команд по порядку

Схема ядра процессора



Особенности

- Выполнение до 2-х команд за такт
- Для каждого конвейера свои команды
- Выполнение команд по порядку
- 4 потока на ядро
- Декодер может брать команды из одного потока минимум через такт
- Потoki переключаются циклически

Кэш-память:

L1 data: 32 KB

L1 code: 32 KB

L2: 256 KB

Иерархия памяти

Parameter	L1	L2
Size	32 KB + 32 KB	512 KB
Associativity	8-way	8-way
Line Size	64 bytes	64 bytes
Banks	8	8
Access Time	1 cycle	11 cycles
Policy	pseudo LRU	pseudo LRU

- Организация кэшей инклюзивная
- Аппаратная предвыборка в кэш L2 – 16 потоков (в прямом или обратном направлении)
- Быстрый доступ к кэшам соседних ядер (реализован распределённый каталог)

АРХИТЕКТУРА

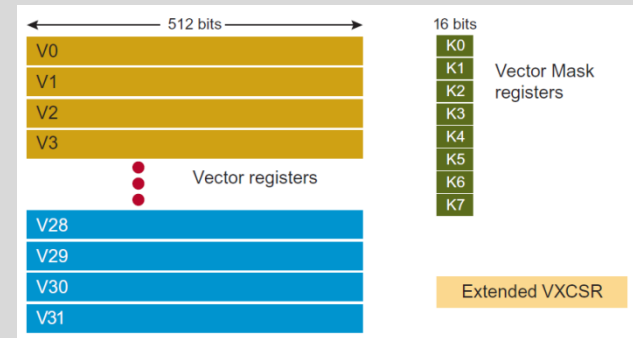
Особенности архитектуры

- Базовая архитектура: x86-64
- Новое векторное расширение: 512-битное
- Другие векторные расширения на поддерживаются
 - MMX, SSE, AVX, ...

Особенности архитектуры

Векторное расширение архитектуры

- Регистры
 - 32 векторных 512-битных регистра
 - 8 16-битных регистров-масок
- Команды
 - Трёхоперандные команды
 - Поддерживаемые типы данных: (u)int32, (u)int64, float, double
 - Выполнение команд по маске
 - Базовые операции: mul, div, add, ...
 - Умножение со сложением/вычитанием: fma, fms
 - Допускаются различные перестановки операндов
 - Операции рассылки/сборки вектора: scatter/gather
 - Быстрое преобразование float/double
 - Спецоперации для float: exp2, log2, rec, recsqrt



ВЫВОДЫ

Общий взгляд

Минусы

- Низкая частота
- Простая архитектура ядра

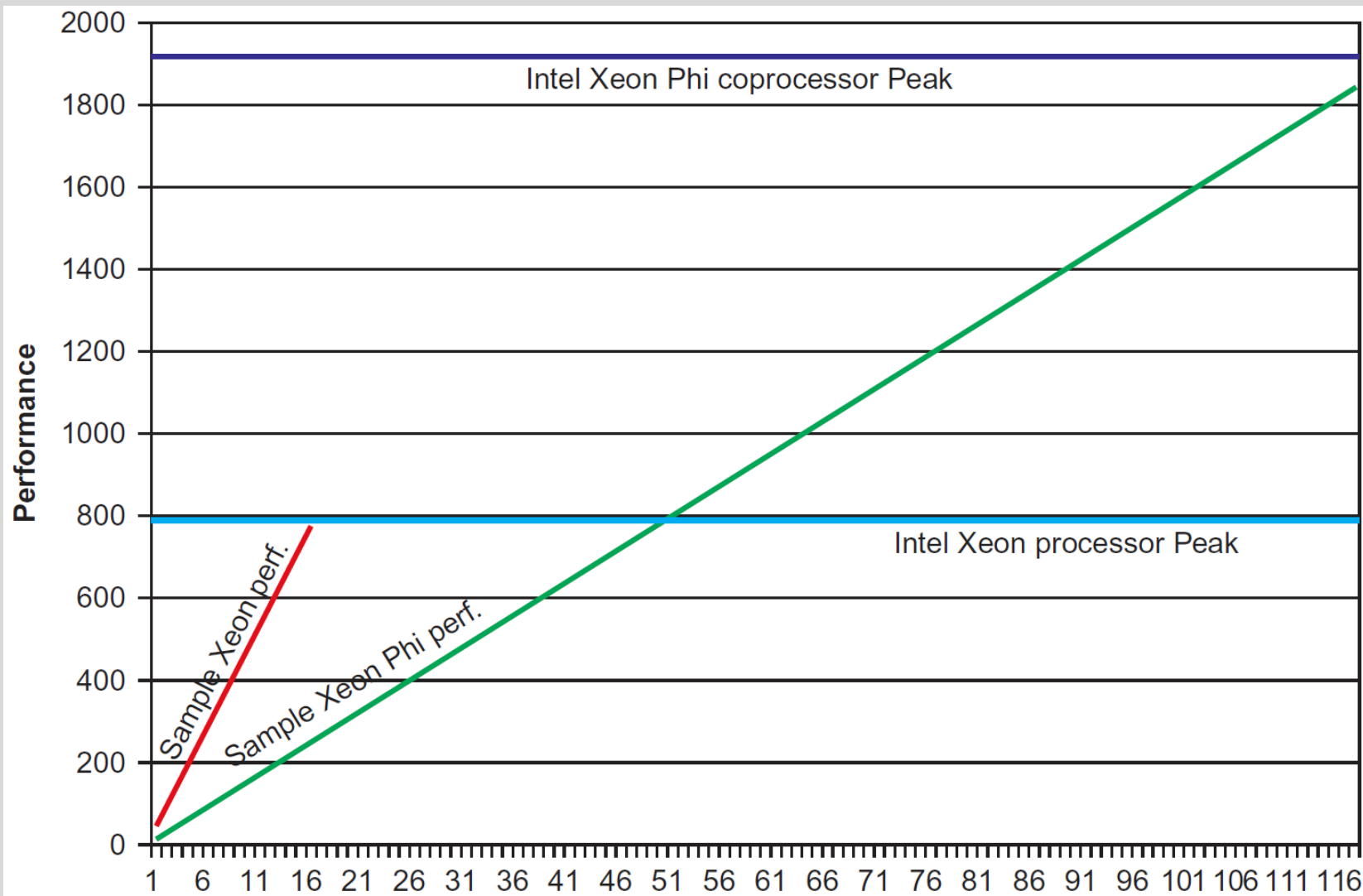
Следствие: Низкая производительность на одном потоке

Плюсы

- Большие вектора
- Много ядер
- Быстрый доступ к памяти

Следствие: высокий потенциал производительности при векторизации и распараллеливании

Сравнение ожидаемой производительности процессоров Xeon и Xeon Phi



КЛАСТЕР

Вычислительный комплекс МВС-10П

МСЦ РАН



МВС-10П. Пиковая производительность - 523.8 ТФЛОПС.
Производительность на тесте LINPACK - 375.7 ТФЛОПС. В состав кластера входит 207 вычислительных узлов. Каждый вычислительный узел имеет в своем составе 2 процессора Xeon E5-2690, 64 ГБ оперативной памяти, два сопроцессора Intel Xeon Phi 7110X.

МВС-10П является энергоэффективным суперкомпьютером, что подтверждено результатами рейтинга Green500 (30 место, 1949 МФЛОПС/Ватт).

Все вычислительные узлы объединяются в 3 сети:

- коммуникационную и транспортную сеть на базе FDR Infiniband;
- сеть мониторинга и управления на базе Gigabit Ethernet;
- управления заданиями на базе Gigabit Ethernet.

На комплексе установлены две системы управления прохождением задач: СУПЗ и SLURM. По умолчанию – SLURM.

Xeon Phi 7110X

- 61 ядро, 244 потока
- Частота: 1094.792 МГц
- Память: 8 Гбайт

Хеон Phi 7110X

пиковые характеристики

- Производительность (на float):
 - Поток
 - $1094.792 \text{ MHz} \times 2 \text{ FLOP} \times 16 \text{ vector size} / 2 \text{ clks} = \mathbf{17.52 \text{ GFLOPS}}$
 - Ядро
 - $1094.792 \text{ MHz} \times 2 \text{ FLOP} \times 16 \text{ vector size} = \mathbf{35.03 \text{ GFLOPS}}$
 - Процессор
 - $1094.792 \text{ MHz} \times 2 \text{ FLOP} \times 16 \text{ vector size} \times 61 \text{ cores} = \mathbf{2137.03 \text{ GFLOPS}}$
- Пропускная способность памяти:
 - Процессор:
 - $5.5 \text{ GT/s} \times 4 \text{ B} \times 16 \text{ channels} = \mathbf{352 \text{ GB/s}}$

A detailed, colorful micrograph of an Intel Xeon Phi processor die. The die is rectangular and densely packed with intricate circuitry. The colors range from deep reds and purples to bright yellows and greens, highlighting different functional blocks and interconnects. The central area is dominated by a grid of repeating vertical structures, likely representing the many cores of the processor. The edges show more complex, less regular patterns, possibly indicating peripheral logic or memory controllers.

intel[®] *inside*[™]

Xeon Phi[™]