



# **ЭКСАФЛОПНЫЕ СУПЕР-ЭВМ КОНТУРЫ АРХИТЕКТУРЫ**

**Степаненко С.А., Южаков В.В.**

# Содержание

**1. Этапы эволюции. Задачи моделирования**

**2. Архитектурные барьеры**

**3. Эксафлопные проекты**

**4. Проблемы и принципы создания эксафлопных супер-ЭВМ**

**5. Ожидаемые результаты**

1

2

3

4

5

# 1. Этапы эволюции. Задачи моделирования

1

2

3

4

5



Konrad Zuse



Z1 (V1 – Versuchsmodell-1)

 $10^{24}$  yotta ?!! $10^{27}$  xona $10^{30}$  weka $10^{33}$  vunda $10^{36}$  uda $10^{39}$  treda $10^{42}$  sorta $10^{45}$  rinta $10^{48}$  quexa $10^{51}$  pepta $10^{54}$  ocha $10^{57}$  nena $10^{60}$  minga $10^{63}$  lumaSISD { механика  
лампы  
транзисторGe+Si  
СБИС,  $\mu$  Пр $10^0$  $10^3$  kilo $10^6$  mega $10^9$  giga $10^{12}$  tera $10^{15}$  peta $10^{18}$  exa $10^{21}$  zetta

1938

1950

1960

1980-1990

1996-2000

2007-2010

2018-2020

2025

MIMD {

? MIMD/SIMD {

?!

# Требуемые значения производительности

Область применения	2011 год	2015 год	2018 год
Атомная энергетика	1 Пфлопс	100 Пфлопс	10-20 Эфлопс
Высокотехнологичные отрасли промышленности: - авиа-, судостроение - автомобилестроение - космическая отрасль	0.3 Пфлопс 0.1 Пфлопс 0.1 Пфлопс	3 Пфлопс 1 Пфлопс 2 Пфлопс	1 Эфлопс 0.5 Эфлопс 1 Эфлопс
Нефтегазовые отрасли	1 Пфлопс	100 Пфлопс	1-10 Эфлопс
Новые материалы на основе нанотехнологий	1 Пфлопс	100 Пфлопс	1-10 Эфлопс
Биотехнологии	1 Пфлопс	10 Пфлопс	1-2 Эфлопс



1

2

3

4

5

## 2. Архитектурные барьеры

## Параллелизм и мощность

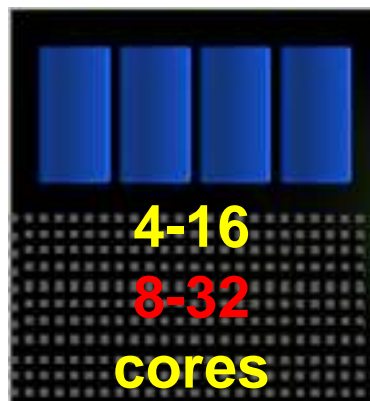
$$\frac{10^{18}}{10^{10}} = 10^8 \text{ шт} - \text{количество ядер ЦП};$$

$$10 \cdot 10^8 = 1000 \text{ МВт на Ефлопс.}$$

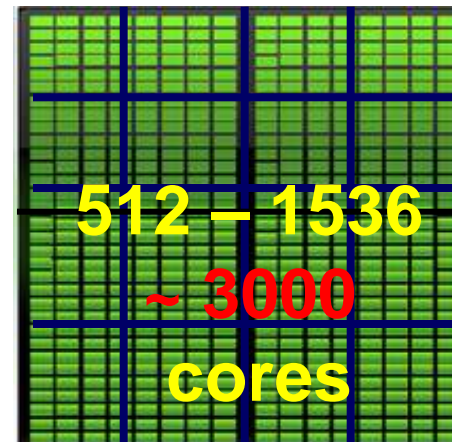
# Гибридная архитектура MIMD-SIMD

● **MIMD** - Multiple Instructions Multiple Data

● **SIMD** – Single Instruction Multiple Data



**MIMD**



**SIMD**

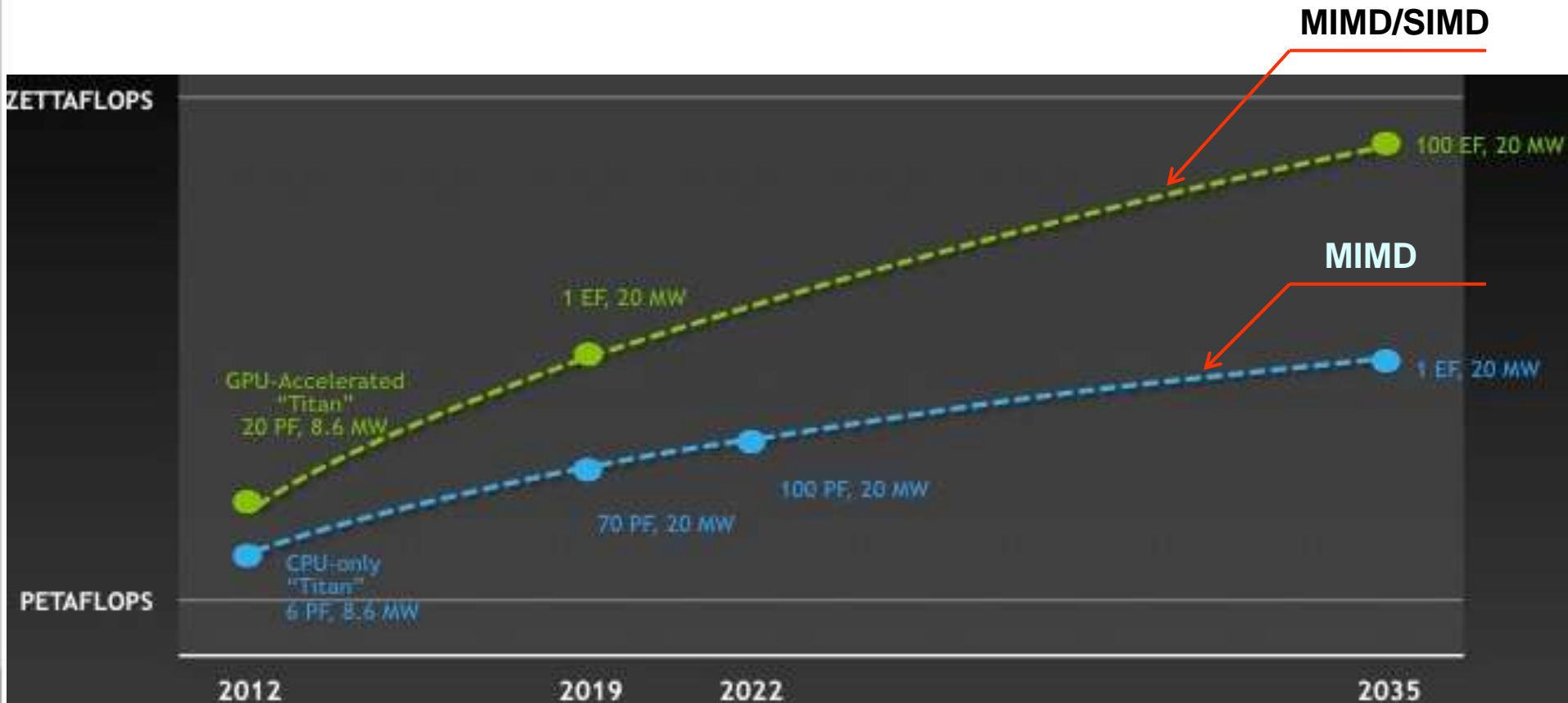


## Удельная производительность

	<b>MIMD</b> $q_m$ Гфлопс/Вт	<b>SIMD</b> $q_s$ Гфлопс/Вт
<b>2010</b>	0,5-1,0	2
<b>2012</b>	1-2	8
<b>2014</b>	2-4	24
<b>2016</b>	4-8	50
<b>2018</b>	10-15	100

Источник [3-5]

# Оценки производительности и мощности



### 3. Эксафлопные проекты

# Эксафлопные проекты

1

2

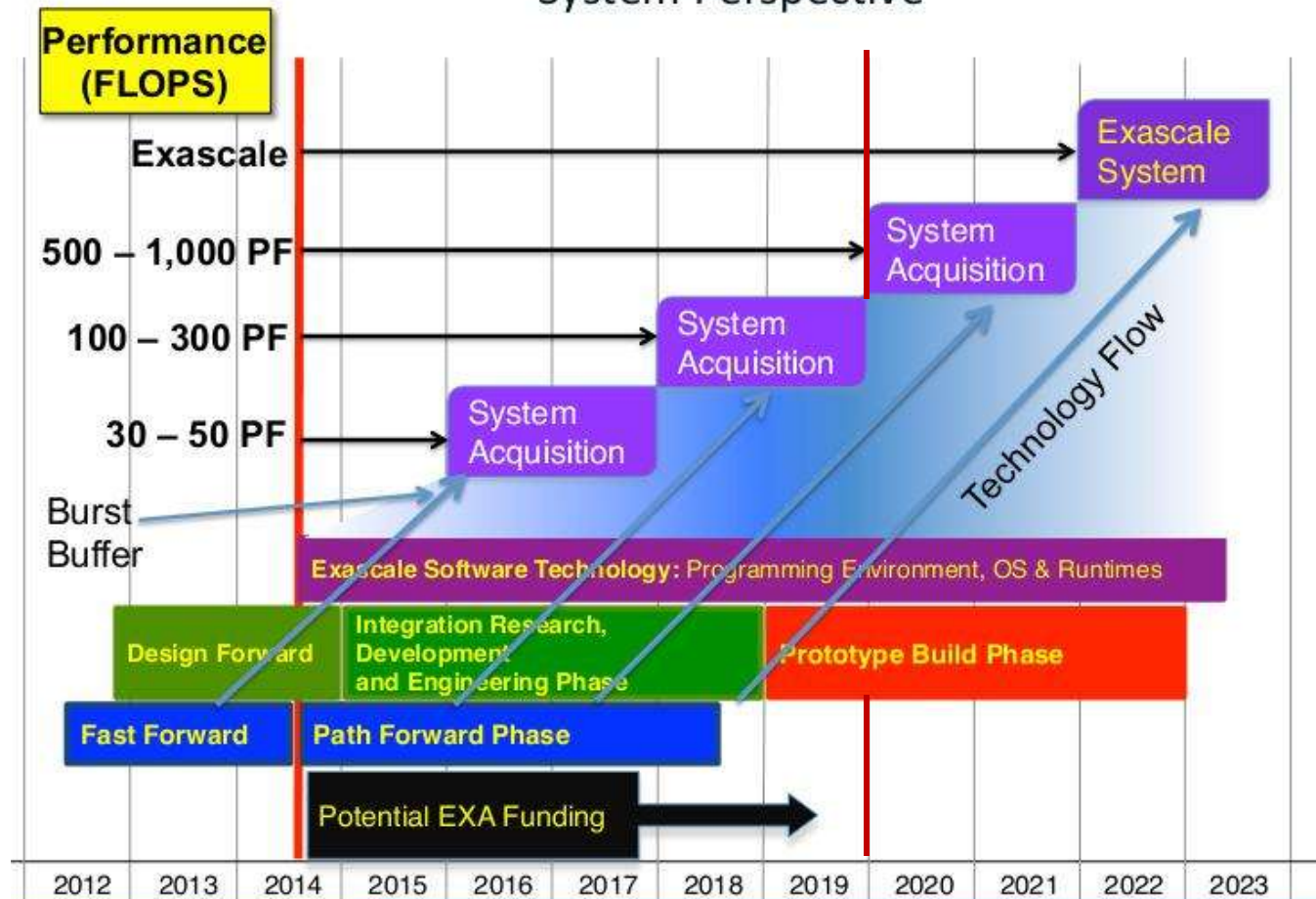
3

4

5

	США	Евросоюз	Китай	Япония
	Массированная финансовая поддержка правительства через DOE (более 100 млн.\$/год); DOD DARPA (десятки млн.\$/год); NSF, ВУЗы десятки млн.\$/год; Всего на исследования без стоимости машин около 150 млн.\$/год	Еврокомиссия, программы PRACE, EESI, ETP4HPC национальные программы, университеты и индустрия. Только ETP – 700 млн.€ до 2020 года	Пятилетние планы по развитию суперкомпьютерной индустрии как на зарубежной, так и собственной элементной базе. Открытая часть финансирования – 60-70 млн.\$ в год	МEXT 1,38 млрд. \$ до 2020 г.
1-10 Пфлопс	20 Пфлопс: Sequoia, Titan 10 Пфлопс: Mira, Blue Waters, Stampede 5 Пфлопс: Vulcan и т.д. 2012-2014 гг.	PizDaint – 8 Пфлопс JUQUEEN – 6 Пфлопс SuperMUC – 3 Пфлопс ... На начало 2014 г. – 14 петафлопных систем	Tianhe-2 – 54 Пфлопс Tianhe-1 – 5 Пфлопс Возможно есть секретные проекты 2012-2014 гг.	K-Computer - 10 Пфлопс TSUBAME2/5 - 6 Пфлопс 2012-2014 гг.
100 Пфлопс	DOE Exascale Initiative 2016-17 г.г. – 30-50 Пфлопс 2018-19 г.г. – 100-300 Пфлопс 2020-21 г.г. – >500 Пфлопс Есть и другие проекты	2015-2020 - несколько экспериментальных суперЭВМ от 10 Пфлопс и выше	2015-2020 г.г. Несколько суперЭВМ 5-10 петафлопной производительности, хотя бы одна 50-100 Пфлопс Бюджет: 615 млн. \$	Системы промежуточной производительности
1000 Пфлопс	2022 г.г. – несколько суперЭВМ с производительностью свыше 1 Эфлопс, есть частные проекты с более ранними сроками	2020 г. – европейский Эксафлопс	2020г. 1~10 Эксафлопс Бюджет: неизвестен	2020 г. 1 Эксафлопс
Многочисленные международные проекты по созданию системного и прикладного ПО для систем эксафлопного класса с участием США, стран Евросоюза, Японии, Китая, России (IESP, G8 exascale, CRESTA и другие)				Совместно с DOE US

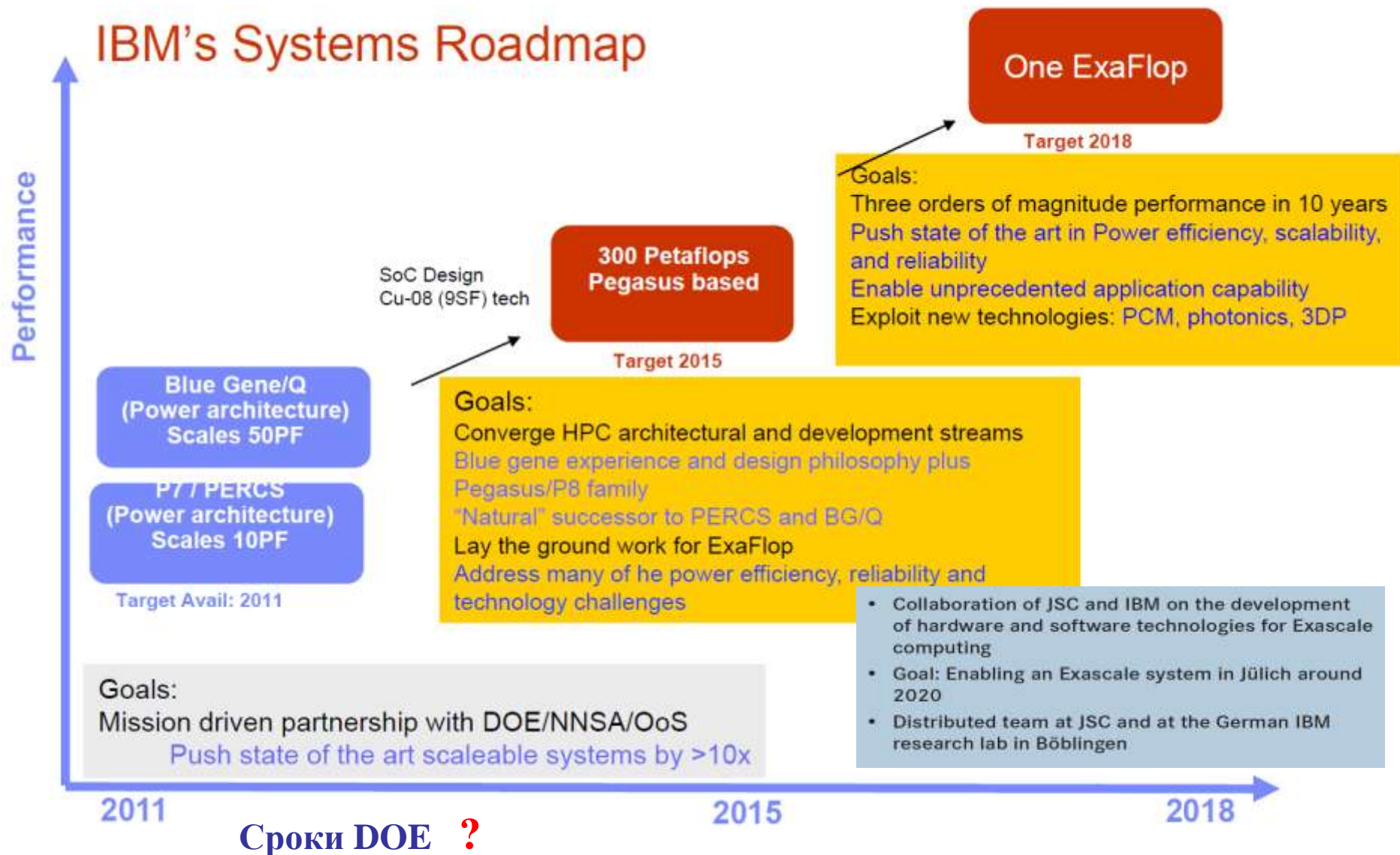
## Exascale Computing Initiative Timeline System Perspective



Источник [14]



# IBM's Systems Roadmap



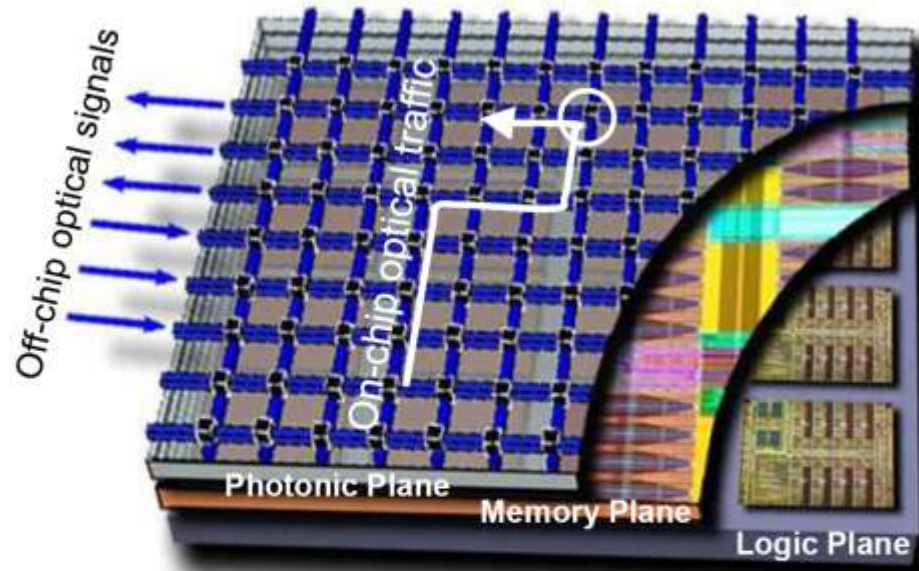
## IBM делает упор на смену базовых технологий

### Vision for 2020: Optically connected 3-D Supercomputer Chip

2020

1mW/Gb/s

\$0.025/Gb/s



- 36 “Cell” 3-D chip
- Silicon photonics layer integrated with high performance logic and memory layers
- Layers separately optimized for performance and yield

Logic plane

~300 cores, ~5TF  
(36 “supercores”)

Memory plane

~30GB eDRAM

Photonic plane

**On-Chip Optical Network**>20 Tbps (bidirectional) optical on-chip  
(between supercores)

&gt;20 Tbps optical off-chip

Photonic layer not only connects the multiple cores, but also routes the traffic

System level study:  
IBM, Columbia, Cornell, UCSB



## An NVIDIA ExaScale Machine in 2017

### • 2017 GPU Node – 300W

- 2,400 throughput cores (7,200 FPU), 16 CPUs – single chip
- 40TFLOPS (SP) 13TFLOPS (DP)
- Deep, explicit on-chip storage hierarchy
- Fast communication and synchronization

### • Node Memory

- 128GB DRAM, 2TB/s bandwidth
- 512GB Phase-change/Flash for checkpoint and scratch

### • Cabinet – 100kW

- 384 Nodes – 15,7PFLOPS (SP), 50TB DRAM
- Dragonfly network – 1TB/s per node bandwidth

### • System – 10MW

- 128 Cabinets – 2EFLOPS (SP), 6,4PB DRAM
- Dragonfly network with active optical links

### • RAS

- ECC on all memory and links
- Option to pair cores for self-checking (or use application-level checking)
- Fast local checkpoint

Источник [6]

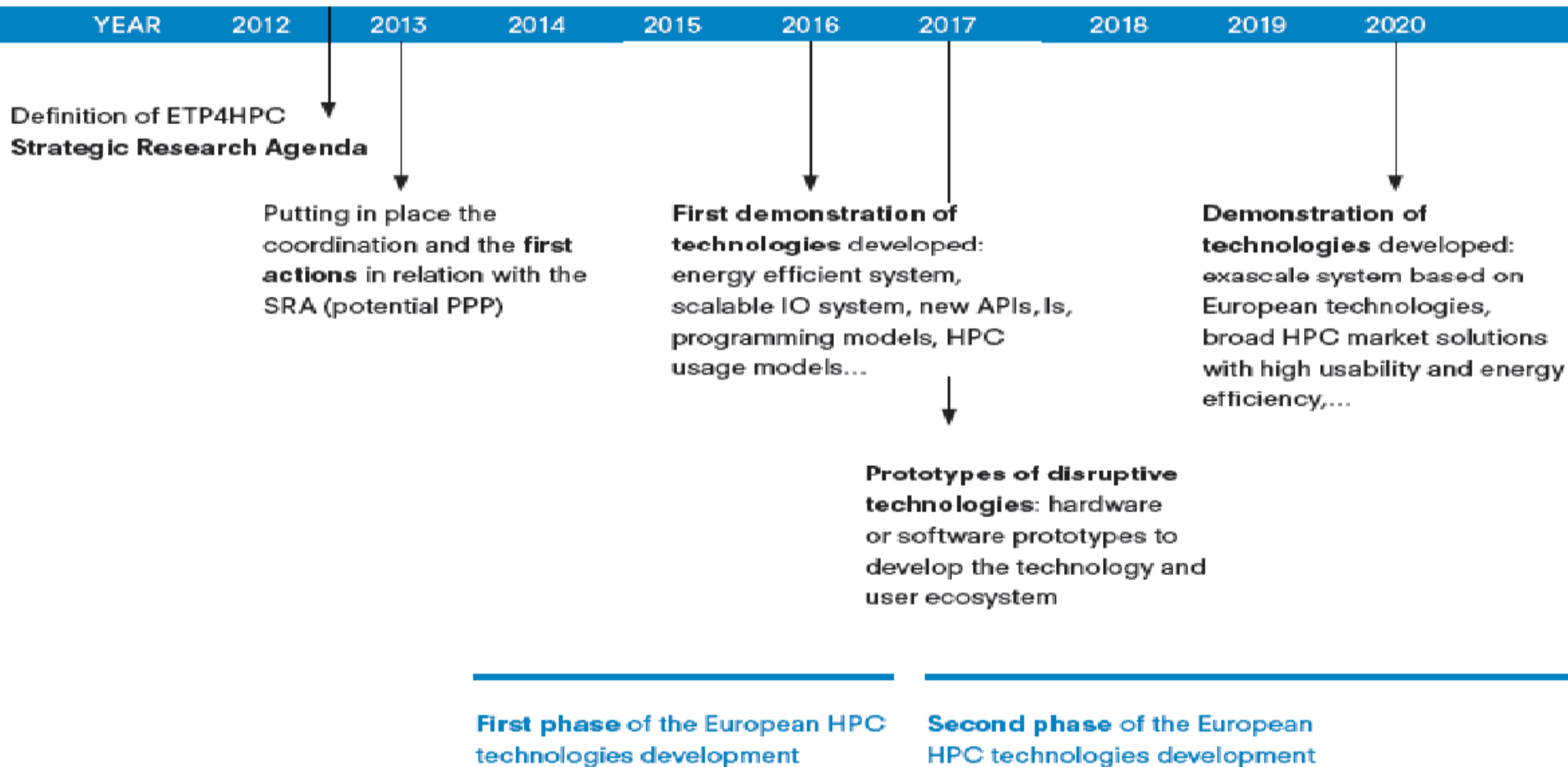
## Cray-ORNL



Titan Cray
21900 (2700/19200) Тфлопс
18688 MIMD/SIMD процессоров
$\approx 3 \cdot 10^5 / 2.8 \cdot 10^7$ ядер
8,6 мВт
200 стоек (45 кВт)
Процессор 16/1536 ядер $\pi = 141/1027$ Гфлопс

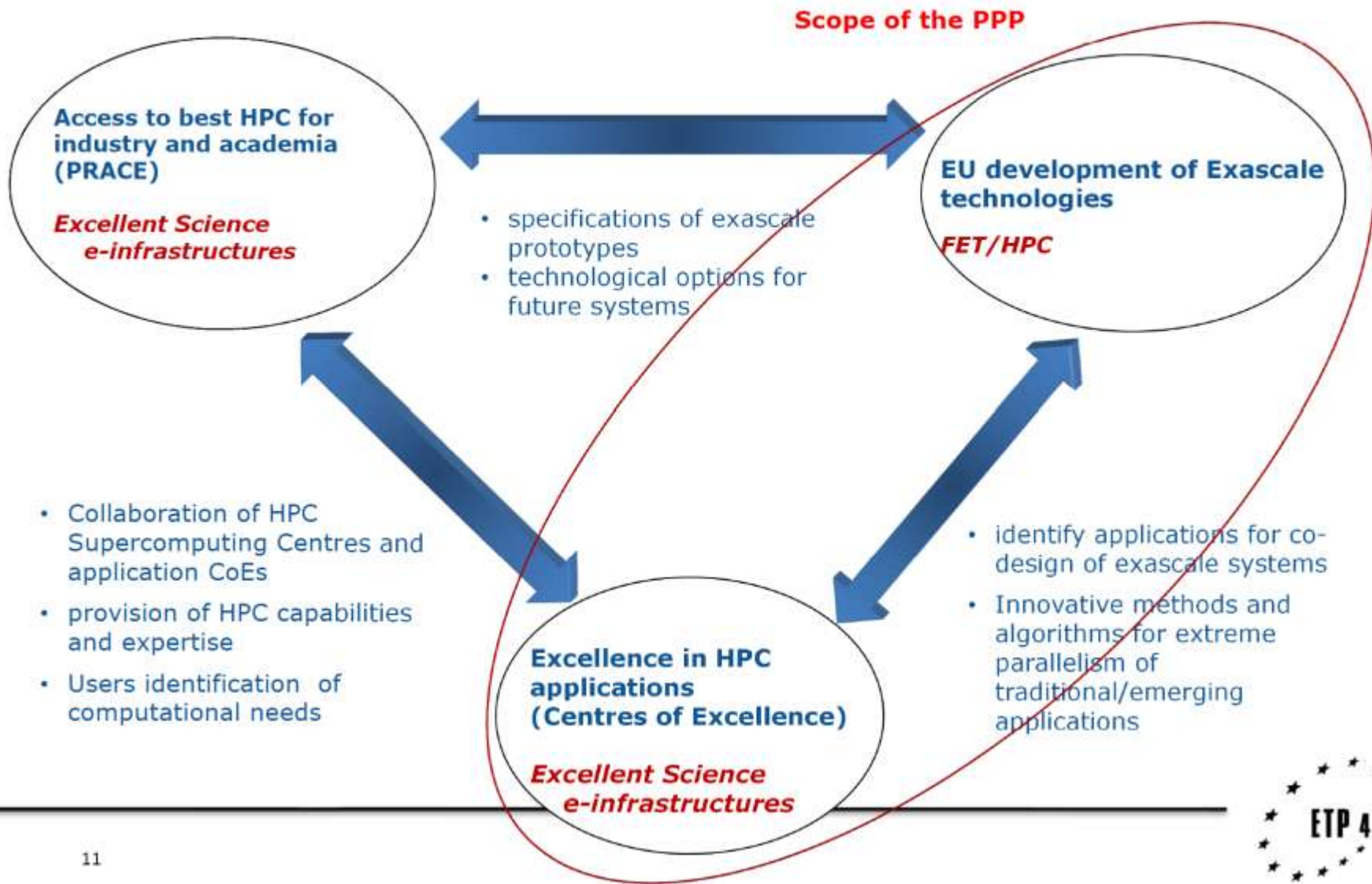
Источник [7]

## EU ExaScale ETP4HPC



Европейская технологическая платформа для  
высокопроизводительных вычислений – 2014-2020 **700 млн.€**

## EU ExaScale HORIZON 2020





## Japan's Future Exascale System

Basic design for GP & Accelerator components and thorough review the system specification

- Architecture: general-purpose plus accelerator components
- Target performance of 1 Exaflops, 100 times as powerful as K
- Power consumption of 30-40MW (cf. K computer: 12.7 MW)



С апреля 2014 года японским министерством MEXT финансируется **FLAGSHIP 2020 Project**

## Continuous Development of HPCI (High Performance Computing Infrastructure)

Peak Performance [FLOPS]  
FLOPS: Floating point Operations Per Second

Continuous Development of Top-class Supercomputers

Raise the Level of Low-end Supercomputers

Exascale Supercomputer

K-Computer

Earth Simulator

CP-PACS  
NWT

Universities  
National Institutes

Industry  
Laboratories

Universities  
National Institutes

Industry  
Laboratories

Universities  
National Institutes

Industry  
Laboratories

Industry  
Laboratories

Industry  
Laboratories

Industry  
Laboratories

Industry  
Laboratories

Industry  
Laboratories

Industry  
Laboratories

Industry  
Laboratories

В сентябре 2014 г. исполнителем проекта выбрана фирма **FUJITSU**



\*At Joint High Level Committee Meeting on Science and Technology cooperation (April 30, 2013)

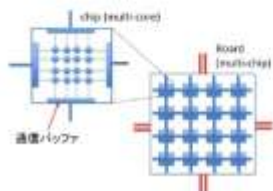
- Participating Organizations: National Lab. (DOE, US), RIKEN (Japan)
- Cooperative Area: System Software
- Major Contents: Description for Forms of Cooperation and Management of Project etc.

Источник [15]

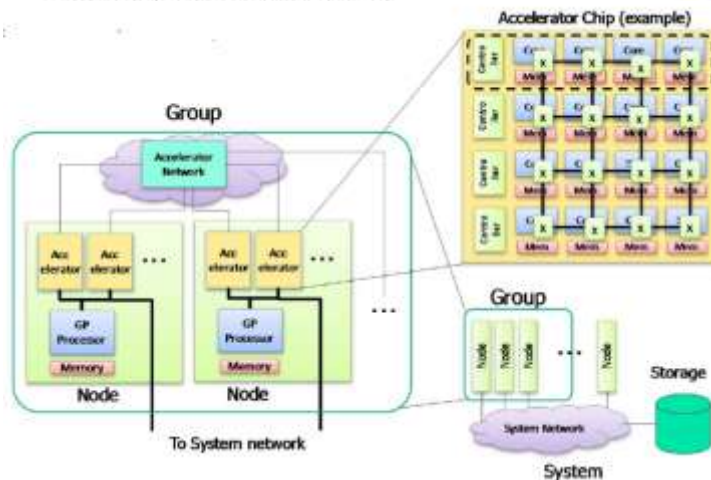
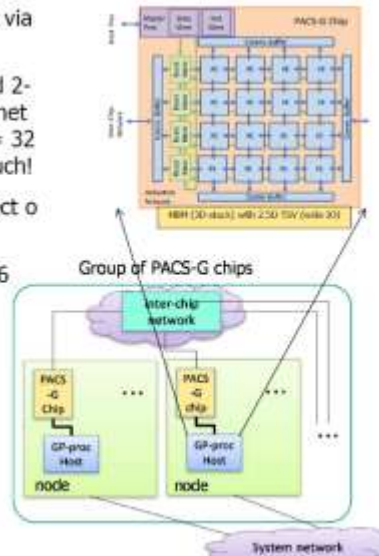
# Japan's Future Exascale System

## PACS-G: a straw man architecture

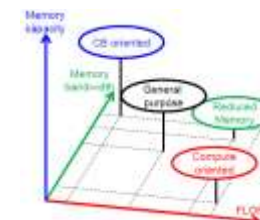
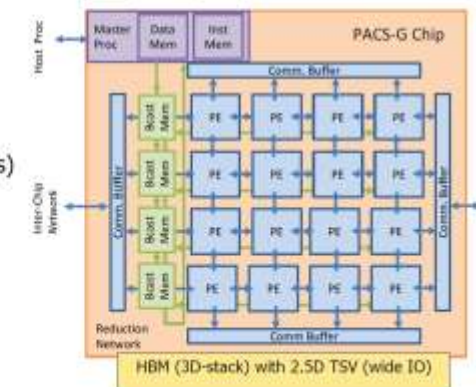
- A group of 1024~2048 chips are connected via accelerator network (inter-chip network)
- 25 ~ 50Gbps/link for inter-chip: If we extend 2-D mesh network to the (2D-mesh) external network in a group, we need 200~400GB/s (= 32 ch. x 25~50Gbps x 2(bi-direction)) $\Rightarrow$ too much!
- For 50Gbps data transfer, we may need direct optical interconnect from chip.
- I/O Interface to Host: PCI Express Gen 4 x16 (not enough!!!)



Interconnect between chips (2D mesh)



- SIMD architecture, for compute oriented apps (N-body, MD), and stencil apps.
- 4096 cores (64x64), 2FMA@1GHz, 4GFlops x 4096 = 16TFlops/chip
- 2D mesh (+ broadcast/reduction) on-chip network for stencil apps.
- We expect 10nm technology available in the range of year 2018-2020, Chip-dai size: 20mm x 20mm
- Mainly working on on-chip memory (size 512 MB/chip, 128KB/core), and,
- with module (global) memory by HBM (3D-stack/wide IO DRAM memory via 2.5D TSV), bandwidth 1TB/s, size 16-32GB/chip (block access only, no random access)
- No external memory (DIM/DDR)
- 250 W/chip expected
- 64K chips for 1 EFLOPS (at peak)



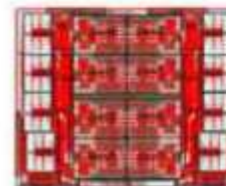


## China Exascale Plans

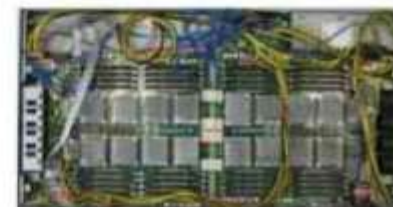
- **12th 5-year Plan (2011-15)**
  - Several petascale HPCs
    - E..g. Chinese Academy of Sciences – 10PF (CNY 700mil)
    - Godson Processor R&D?
  - At least one 50-100 PFLOPS
  - Budget: CNY 4 Billions
    - MOST: 60%, Local Government: 40%
- **13th 5-year Plan (2016-20)**
  - 1~10 ExaFLOPS HPC
  - Budget: unknown yet



China National Grid (CNGrid)



Godson-3B



## Общие аспекты эксафлопных проектов:

- Централизация, финансирование государственными институтами. «Частные» проекты – часть государственных.
- Модернизация промышленности (проектные нормы элементной базы 10-20 нм, оптические компоненты).
- Разработка принципиально новых архитектуры и средств математического моделирования (системное ПО, прикладное ПО, физические модели).

## 4. Проблемы и принципы создания эксафлопных супер-ЭВМ



1

2

3

4

5

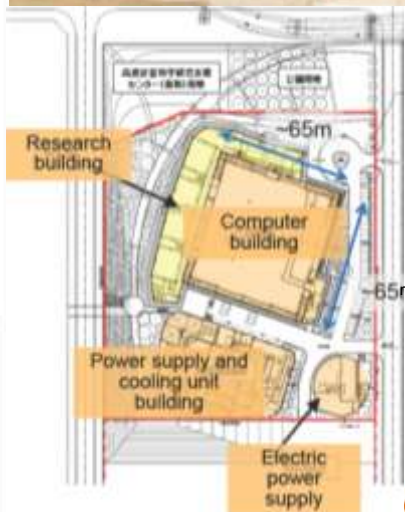
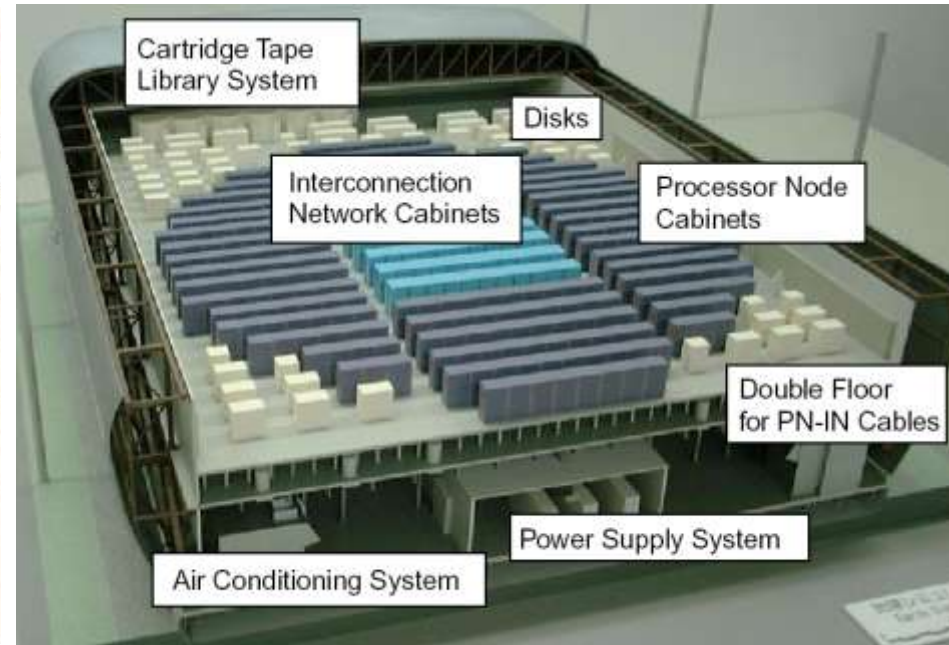
# Уровни параллелизма



	1	2	3	4
$q$ (шт. ядер)	$(100-200)/(1000-2000)$	$(10^3 \cdot 2 \times 10^3)/(10^4 \cdot 2 \times 10^4)$	$10^5/10^6$	$10^7/10^8$
$\pi$ (Тфлпс)	$(1-2)/(10-20)$	$(16-32)/(160-256)$	$(1,6-3,2) \times 10^3/(1,6-2,5) \times 10^4$	$(1,6-3,2) \times 10^5/(1,6-2,5) \times 10^6$
$V$ (ТБ/с)	$\sim 1=0,060$ (21-18)	$0,060 \times 16 \times 16 = 15$	$0,06 \times 8 \times 16 \times 100 = 768$	
$V_k$ (ТБ/с)	$0,060 = (0,005 \times 12)$	$0,060 \times 16 = 0,960$	$0,06 \times 16 \times 100 = 96$	
$l$ (м)	0,1	2,0	20	
$C$ , шт/ $L$ , км	$4 \times 10^6/432$	$1,5 \times 10^6/3 \times 10^3$	$3 \times 10^6/62 \times 10^3$	

3D

# Инженерная инфраструктура



Производительность систем, Пфлопс	Цена энергии MIMD/SIMD, млн руб.	Цена энергии MIMD, млн руб.
10 2014 год	65 97,2	162
100 2017 год	145 162	810
1000 2020 год	~ 648	3241





# Проблемы создания эксафлопных технологий

Проблемы	Средства решения	Планируемый результат
ЭФФЕКТИВНОЕ ЗАДЕЙСТВОВАНИЕ $10^8$ - $10^9$ ядер (рост степени параллелизма в 1000-10000 раз)	<ul style="list-style-type: none"> <li>- парадигмы параллелизма;</li> <li>- создание численных методов;</li> <li>- <u>архитектурное масштабирование эффективности</u>;</li> <li>- выполнение баланса между скоростью передачи информации и производительностью системы.</li> </ul>	Эффективность использования суперЭВМ не менее 50%
При стандартных подходах: ЭНЕРГОПОТРЕБЛЕНИЕ 100МВт КОНСТРУКТИВНЫЕ ПАРАМЕТРЫ ~2000-3000 стоек	<ul style="list-style-type: none"> <li>- <u>применение принципиально новых гибридных архитектур со специализированными арифметическими ускорителями</u>;</li> <li>- уменьшение удельного энергопотребления в 50-100 раз;</li> <li>- проектные нормы элементной базы 15-20 нм;</li> <li>- повышение коэффициента теплообмена в 5 раз</li> </ul>	Энергопотребление суперЭВМ: <b>10-20 МВт;</b> Конструктивные параметры суперЭВМ: <b>100-200 стоек.</b>
НАДЕЖНОСТЬ наработка на отказ О (мин)	<ul style="list-style-type: none"> <li>- <u>рекурсивное топологическое резервирование</u>;</li> <li>- проектные нормы элементной базы 15-20 нм; рост плотности компоновки элементов в 10000 раз.</li> </ul>	Время безотказной работы среды суперЭВМ более 24 часов



# Top 10 Challenges to Exascale

- ? V 1. **Energy efficiency:**
  - Creating more energy efficient circuit, power, and cooling technologies.
- V 2. **Interconnect technology:**
  - Increasing the performance and energy efficiency of data movement.
- ? 3. **Memory Technology:**
  - Integrating advanced memory technologies to improve both capacity and bandwidth.
- V 4. **Scalable System Software:**
  - Developing scalable system software that is power and resilience aware.
- ? V 5. **Programming systems:**
  - Inventing new programming environments that express massive parallelism, data locality, and resilience
- ? V 6. **Data management:**
  - Creating data management software that can handle the volume, velocity and diversity of data that is anticipated.
- V 7. **Exascale Algorithms:**
  - Reformulating science problems and refactoring their solution algorithms for exascale systems.
- V 8. **Algorithms for discovery, design, and decision:**
  - Facilitating mathematical optimization and uncertainty quantification for exascale discovery, design, and decision making.
- V 9. **Resilience and correctness:**
  - Ensuring correct scientific computation in face of faults, reproducibility, and algorithm verification challenges.
- V 10. **Scientific productivity:**
  - Increasing the productivity of computational scientists with new software engineering tools and environments.
  - Unless researcher productivity increases, the time to solution may be dominated by application development, not computation.



V – реализуемо в России  
? – не реализуемо в России

1

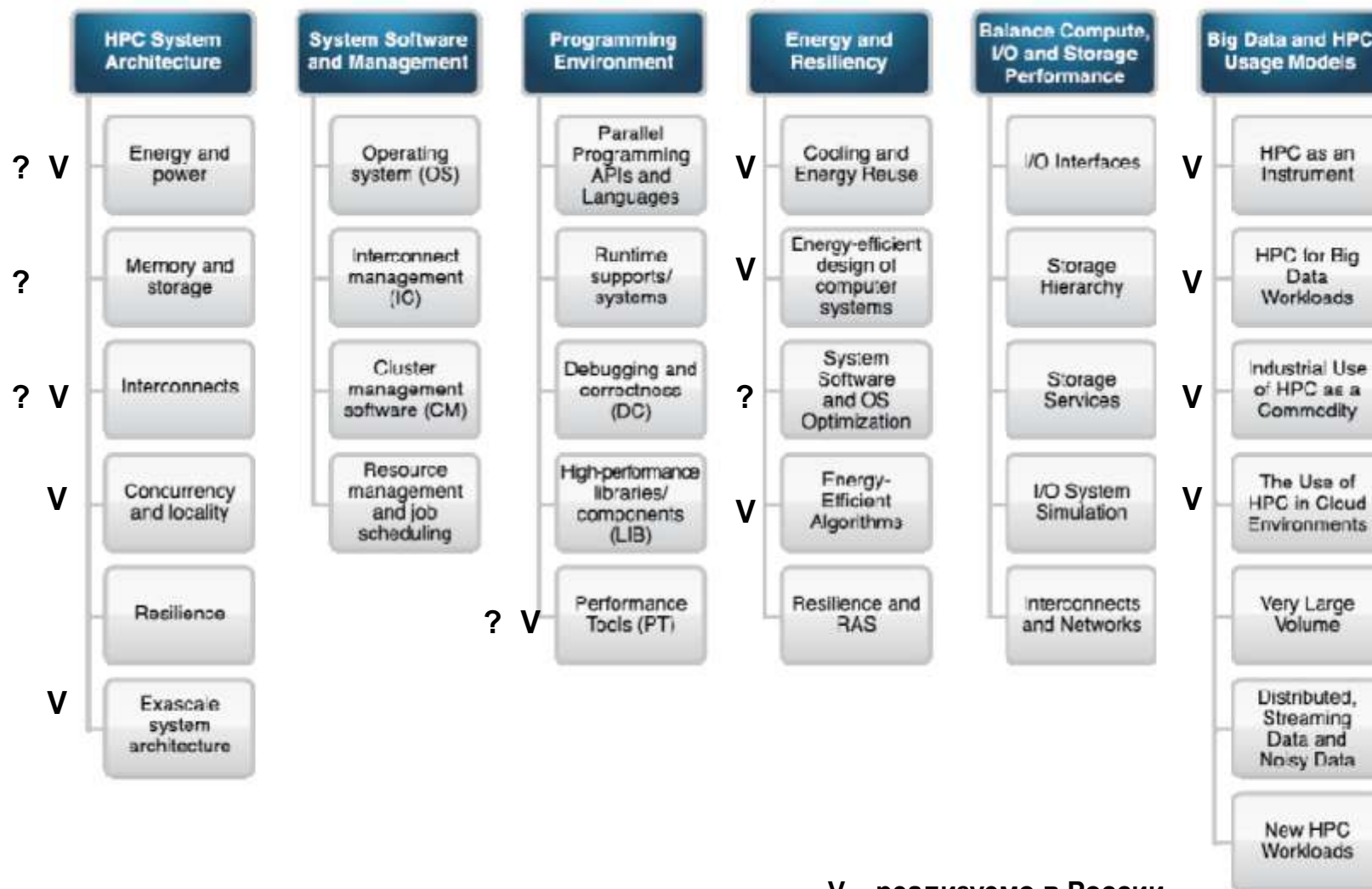
2

3

4

5

## ETP4HPC приоритетные исследования



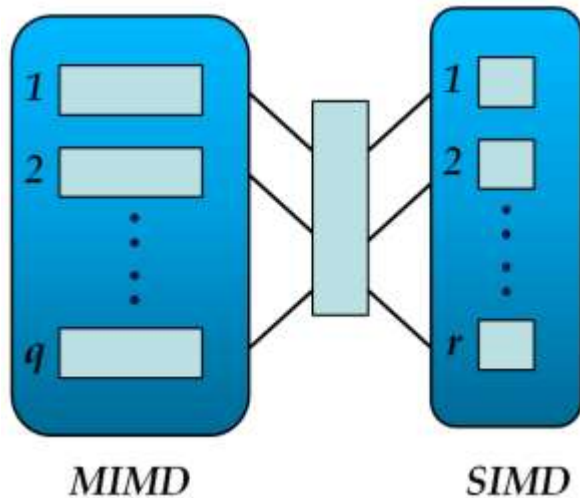
V – реализуемо в России

? – не реализуемо в России



# Архитектурное масштабирование эффективности Гибридные реконфигурируемые структуры

## Уровень 1



Параметры процесса  $\varphi, \rho$

Эксперимент

Размерность 55x55x55

$$T_1 = 22,96c$$

$$T_{1,1} = 9,87c$$

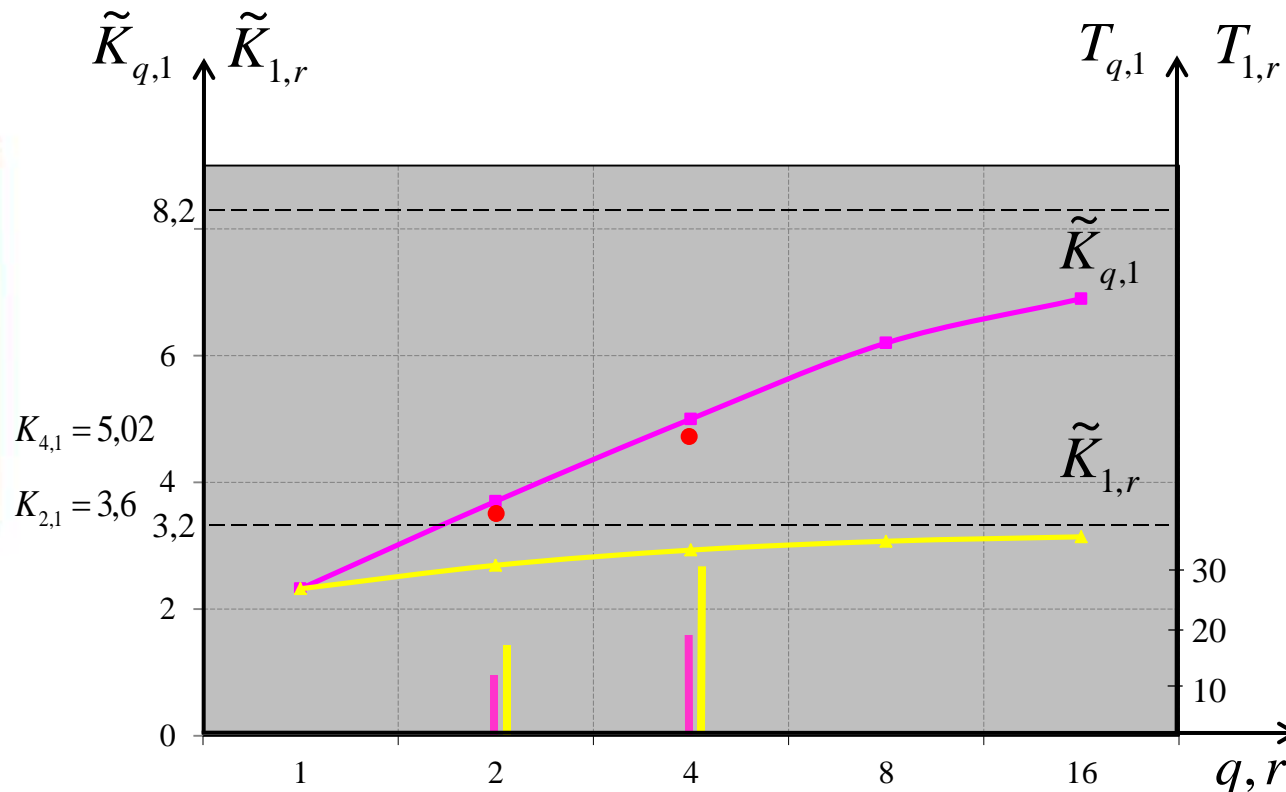
$$\varphi = \frac{T_M}{T_1} = 0,31$$

$$T_M = 7,07c$$

$$T_S = 2,8c$$

$$K_{1,1} = 2,3$$

$$\rho = \frac{(1-\varphi)T_1}{T_S} = 5,67$$



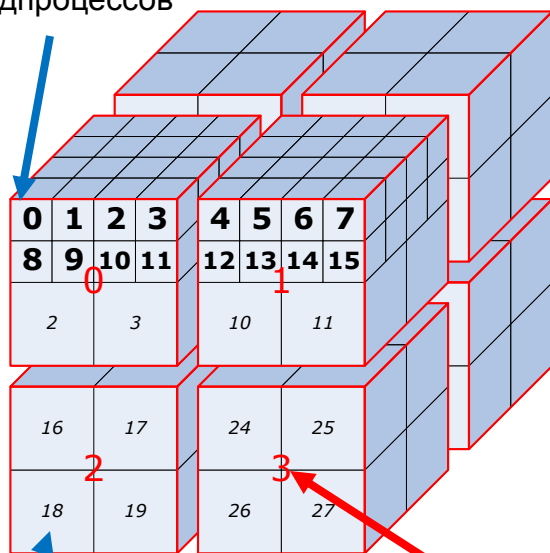
Источник [8]

# Архитектурное масштабирование эффективности Минимизация длительностей обменов Декомпозиция и распределение процессов

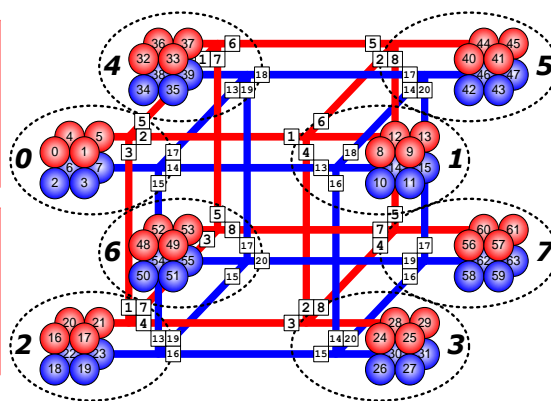
## Уровни 2, 3, 4

Значения производительности  
на NPB 3.2 LU

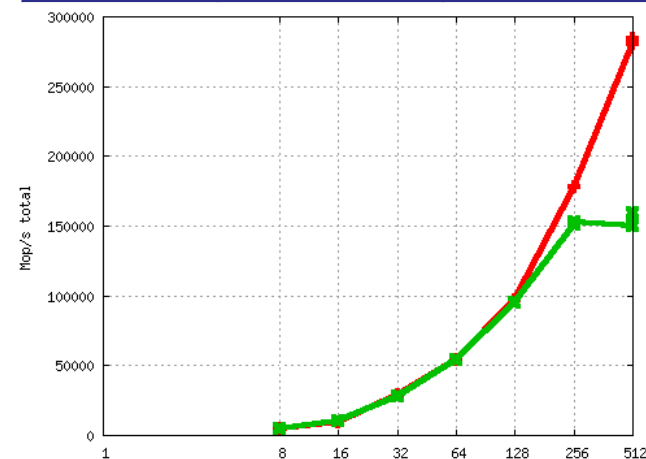
Номера  
подпроцессов



Номера ПЭ



	Класс C	
Количество ядер	Вариант 1 Mop/s	Вариант 2 Mop/s
128	97072	95398
256	177953	152613
512	283926	164283



- Вариант 1 — размещение процессов согласно структуре связей системы
- Вариант 2 — последовательное размещение процессов

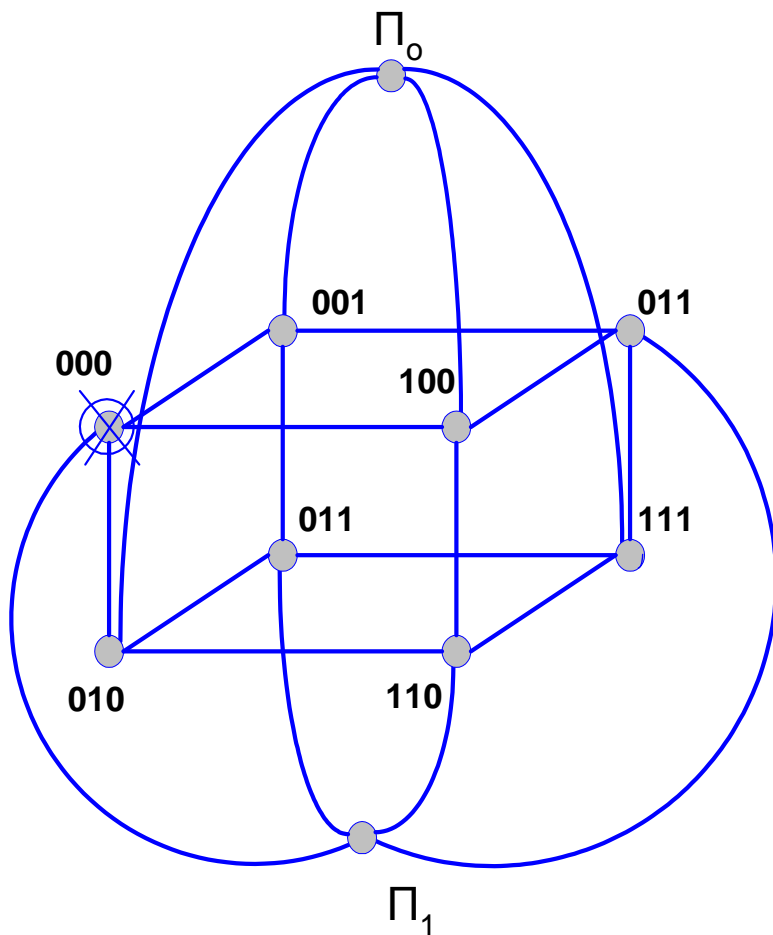
Результат – ускорение в 1,73 раза на 512 ядрах

Источник [9]

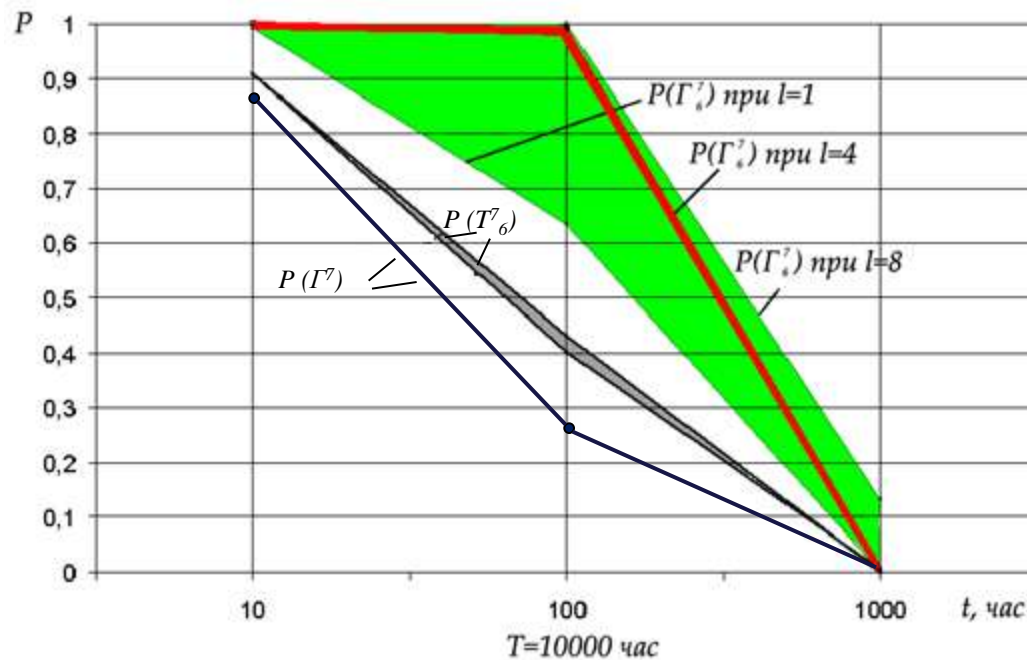
# Топологическое резервирование

## Вероятность выполнения процесса

Уровни 1, 2, 3



Резервирование  $\Gamma^3$



# МАСШТАБИРУЕМАЯ ЭКСТРАПОЛЯЦИОННАЯ ПЛАТФОРМА ЭКСАФЛОПНОЙ СУПЕР-ЭВМ

- динамическая адаптируемость архитектуры к особенностям исполняемой программы;
- комплекс программно-аппаратных средств отказоустойчивого масштабирования эффективности;
- средства предсказания параметров супер-ЭВМ эксафлопсной производительности, полученной масштабированием этой платформы применительно к эталонному набору прикладных программ.



1

2

3

4

5

## 5. Ожидаемые результаты

# Значения параметров супер-ЭВМ

Параметр	2014 г.	2017 г.*	2020 г.*
Пиковая производительность MIMD/SIMD, Пфлопс	10 (1,0+9,0)	100 (10+90)	1000 (100+900)
Количество MIMD/SIMD ядер, шт	$10^5/4,5 \cdot 10^6$	$10^6/10^7$	$10^7/10^8$
Потребляемая мощность, КВт	1625 (500+1125)	4300 (2500+1800)	19000 (10000+9000)
Емкость оперативной памяти, Пбайт	0,5-1,0	2-4	30-40
Производительность MIMD/SIMD процессора, Гфлопс	(256-512)/ (1500-2000)	(500-1000)/ (4000-8000)	(1000-2000)/ (10000-16000)
Производительность вычислительного модуля MIMD/SIMD, Тфлопс	(1-2)/(6-8)	(4-8)/(32-64)	(16-32)/ (160-256)
Пропускная способность вычислительного модуля, Гбайт/с	100-250	500-2000	2500-10000
Количество MIMD/SIMD процессоров, шт	4500-6000	11250-22500	50000-90000
Количество вычислительных модулей, шт	1000-1500	1400-3000	3500-5600
Внешняя память параллельной ФС, ПБайт	10	100	1000
Скорость обмена с внешней памятью, ТБайт/с	0,5-1	5-10	50-100

\* – обязательное владение технологической, элементной и инструментальной базой!!!

# ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СУПЕР-ЭВМ

## 1. Базовое системное программное обеспечение:

- Операционная система (ОС)
- Управление межсоединениями (интерконнектом)
- ПО управление кластером
- Управление ресурсами и диспетчеризация заданий

## 2. Среда программирования:

- Языки и прикладные программные интерфейсы для параллельного программирования
- Средства динамической поддержки/ системы исполнения
- Отладка и корректность (DC)
- Высокопроизводительные библиотеки/компоненты (LIB)
- Средства представления (PT)

## 3. Эффективность ввода/вывода:

- Интерфейсы ввода/вывода
- Иерархия памяти
- Сервис памяти
- Имитация системы ввода/вывода
- Межсоединения и сети

## 4. Прикладное программное обеспечение и физические модели.

# ЗАКЛЮЧЕНИЕ

1

2

3

4

5

Для создания супер-ЭВМ эксафлопной производительности необходимы:

- технологическая реализация различных уровней параллелизма
- применение гибридных архитектур, реализующих различные дисциплины вычислений, обеспечивающих реконфигурацию компонент в соответствии с особенностями исполняемого процесса и отказоустойчивое масштабирование эффективности.
- разработка инструментальных системных программных средств, позволяющих реализовывать эффективные (производительность, энергия) программы.
- создание прикладного программного обеспечения для гибридных архитектур (принципиально отличного от существующего).
- координация работ в рамках единого мультидисциплинарного государственного проекта.



В мире создается принципиально новая дисциплина гибридных вычислений, предназначенная для достижения эксафлопной производительности. В ее основе взаимное соответствие архитектуры и прикладного программного обеспечения, реализуемое специализированными аппаратными и программными компонентами. Эти компоненты не переносимы, их бессмысленно заимствовать вследствие различия программ.

Необходимы разработка масштабируемых гибридных вычислительных систем применительно к собственным программам для решения определенных классов задач и создание на этой базе перспективных средств математического моделирования.



Konrad Zuse



Z1 (V1 – Versuchsmodell-1)

SISD { механика  
лампы  
транзистор  
СБИС,  $\mu$  Пр

MIMD {  
? MIMD/SIMD {

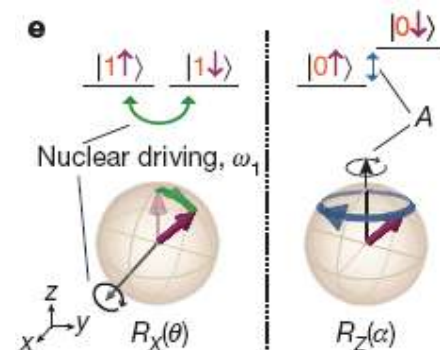
?!

$10^0$   
 $10^3$  kilo  
 $10^6$  mega  
 $10^9$  giga  
 $10^{12}$  tera  
 $10^{15}$  peta  
 $10^{18}$  exa  
 $10^{21}$  zetta

$10^{24}$  yotta  
 $10^{27}$  xona  
 $10^{30}$  weka  
 $10^{33}$  vunda  
 $10^{36}$  uda  
 $10^{39}$  treda  
 $10^{42}$  sorta  
 $10^{45}$  rinta  
 $10^{48}$  quexa  
 $10^{51}$  pepta  
 $10^{54}$  ocha  
 $10^{57}$  nena  
 $10^{60}$  minga  
 $10^{63}$  luma

?!MIMD SIMD  
FPGA

?! Квантовый  
компьютер [11]



# Список литературы

1. Концепция по развитию технологии высокопроизводительных вычислений на базе суперЭВМ эксафлопного класса на 2012-2020 гг. [Электронный ресурс]. Режим доступа: <http://www.rosatom.ru/wps/wcm/connect/rosatom/rosatomsite/aboutcorporation/nauka/>
2. Rick Stevens and Andy White. A DOE Laboratory plan for providing exascale applications and technologies for critical DOE mission needs [Электронный ресурс]. Режим доступа: [http://computing.ornl.gov/workshops/SCIDAC2010/r\\_stevens.pdf](http://computing.ornl.gov/workshops/SCIDAC2010/r_stevens.pdf)
3. Impact of Architecture and Technology for Extreme Scale on Software and Algorithm Design. Jack Dongarra. Cross-cutting Technologies for Computing at the Exascale. February 2-5, 2010 – Washington, D.C.
4. SC'09 Exascale Panel. Steve Scott. Cray Chief Technology Officer.Exhibitor Forum, SC'09.
5. David Barkai (Intel Corporation). HPC in the multi-core and heterogeneous era [Электронный ресурс]. Режим доступа: [http://www.ecmwf.int/newsevents/meetings/workshops/2010/high\\_performance\\_computing\\_14th/presentations/barkai.pdf](http://www.ecmwf.int/newsevents/meetings/workshops/2010/high_performance_computing_14th/presentations/barkai.pdf)
6. SC11 Keynote by Nvidia CEO Jen-Hsun Huang [Электронный ресурс] Режим доступа: <http://blogs.nvidia.com/2011/11/exascale-an-innovator%E2%80%99s-dilemma/>
7. Oak Ridge changes Jaguar's spots from CPUs to GPUs [Электронный ресурс]. Режим доступа: [http://www.theregister.co.uk/2011/10/11/oak\\_ridge\\_cray\\_nvidia\\_titan/](http://www.theregister.co.uk/2011/10/11/oak_ridge_cray_nvidia_titan/)
8. Степаненко С.А. «Способ определения структуры гибридной вычислительной системы». Патент Российской Федерации № 2436151 на изобретение. Приоритет от 01.11.2010.
9. Крючков И.А., Степаненко С.А., Рыбкин А.С. Реализация статической маршрутизации и оптимального размещения вычислительных процессов в мультипроцессорных средах. «Молодежь в науке». Сборник докладов шестой научно-технической конференции. Саров, 2008 г. с.172-176.
10. Степаненко С.А. Топологическое резервирование мультипроцессорных сред выделенными элементами. Труды РФЯЦ-ВНИИЭФ №10, 2005 г. с. 50-60.
11. T. van der Sar, Z.H.Wang, M.S.Blok, H.Bernien, T.H.Taminiau, D.M.Toyli, D. A. Lidar, D.D.Awschalom, R.Hanson & V.V.Dobrovitski. Decoherence-protected quantum gates for a hybrid solid-state spin register. Nature 484. 5 April 2012, p. 82-86. [Электронный ресурс]. Режим доступа: [www.nature.com/nature](http://www.nature.com/nature).
12. Andrew Russell. Intel HPC Technologies. November 2013.
13. Anshel Sag. Intel's New Knig Landing Xeon Phi Combines Omni Sc Fabric with HMC. June 23 2014.
14. Gary Grider. DOE Fast Forward Storage and IO Project. Los Alamos National Laboratory. 04/2013.
15. Japan's Future Exascale System [Электронный ресурс] Режим доступа: <http://www.hpcwire.com/2014/03/18/details-emerge-japans-future-exascale-system/>

# ДОПОЛНИТЕЛЬНЫЕ СЛАЙДЫ



## 10 ПРИОРИТЕТНЫХ ЗАДАЧ ПЕРЕХОДА К ЭКСАФЛОПНЫМ СИСТЕМАМ

1 Эффективное использование энергии:

- создание более эффективных технологий разработки схем, энерготехнологий и технологий охлаждения.

2 Технология интерконнекта:

- повышение производительности и энергоэффективности перемещения данных.

3 Технология памяти:

- интеграция прогрессивных технологий памяти с целью повышения емкости и пропускной способности памяти.

4 Программное обеспечение масштабируемых систем:

- разработка ПО масштабируемой системы, которое позволяет учитывать энергопотребление и обеспечивает отказоустойчивость.

5 Системы программирования:

- изобретение новых сред программирования, которые обеспечивают массивный параллелизм, локальность данных и отказоустойчивость.

6 Управление данными:

- создание ПО управления данными, которое может позволить обрабатывать предполагаемый объем информации, скорость передачи и разнообразие информации.

7 Эксафлопные алгоритмы:

- переформулировка научных задач и реорганизация (реструктуризация) алгоритмов их решения для выполнения расчетов на эксафлопных системах.

## 10 ПРИОРИТЕТНЫХ ЗАДАЧ ПЕРЕХОДА К ЭКСАФЛОПНЫМ СИСТЕМАМ

8 Алгоритмы открытия (изобретения), проектирования и принятия решения:

- облегчение математической оптимизации и количественного расчета неопределенностей для получения результатов поиска, проектирования и принятия решения.

9 Отказоустойчивость (способность к восстановлению) и корректность:

- обеспечение корректного научного расчета в присутствии сбоев, воспроизводительность или сходимость результатов и задачи верификации алгоритмов.

10 Научная продуктивность:

- повышение продуктивности работы специалистов по научным вычислениям с помощью новых программных средств и сред для разработки ПО.
- если продуктивность исследователя не повышается, то время, требуемое для решения, будет определяться разработкой приложений, но не расчетом.

# ЕТР4HPC приоритетные исследования

Архитектура HPC систем	ПО систем и управление	Среда программирования	Энергия и отказо-устойчивость	Сбалансированные вычисления, эффективность ввода/вывода и памяти	Модели использования больших объемов данных и HPC
Энергия и мощность	Операционная система (ОС)	Языки и прикладные программные интерфейсы для параллельного программирования	Охлаждение и повторное использование энергии	Интерфейсы ввода/вывода	HPC как инструмент
ОЗУ и система хранения данных	Управление межсоединениями (интерконнектом)	Средства динамической поддержки/ системы исполнения	Проектирование вычислительных систем с эффективными энергозатратами	Иерархия памяти	HPC для рабочих нагрузок с большим объемом данных
Межсоединения	ПО управление кластером	Отладка и корректность (DC)	Системное ПО и оптимизация ОС	Сервис памяти	Промышленное использование HPC как товара широко потребления
Параллелизм и локальность	Управление ресурсами и диспетчеризация заданий	Высокопроизводительные библиотеки/компоненты (LIB)	Алгоритмы, обеспечивающие эффективные энергозатраты	Имитация системы ввода/вывода	Использование HPC в облачных средах
Отказоустойчивость (возможность восстановления)		Средства представления (PT)	Отказоустойчивость и RAS	Межсоединения и сети	Очень большой объем данных
Архитектура эксафлопных систем					Распределенные данные, потоковые данные и зашумленные данные
					Новые рабочие нагрузки HPC