SAND2004-4815C

Applications Modeling: Connecting Real World Problems with Supercomputer Design

Моделирование прикладных программ: связь реальных мировых проблем с конструкцией суперкомпьютеров

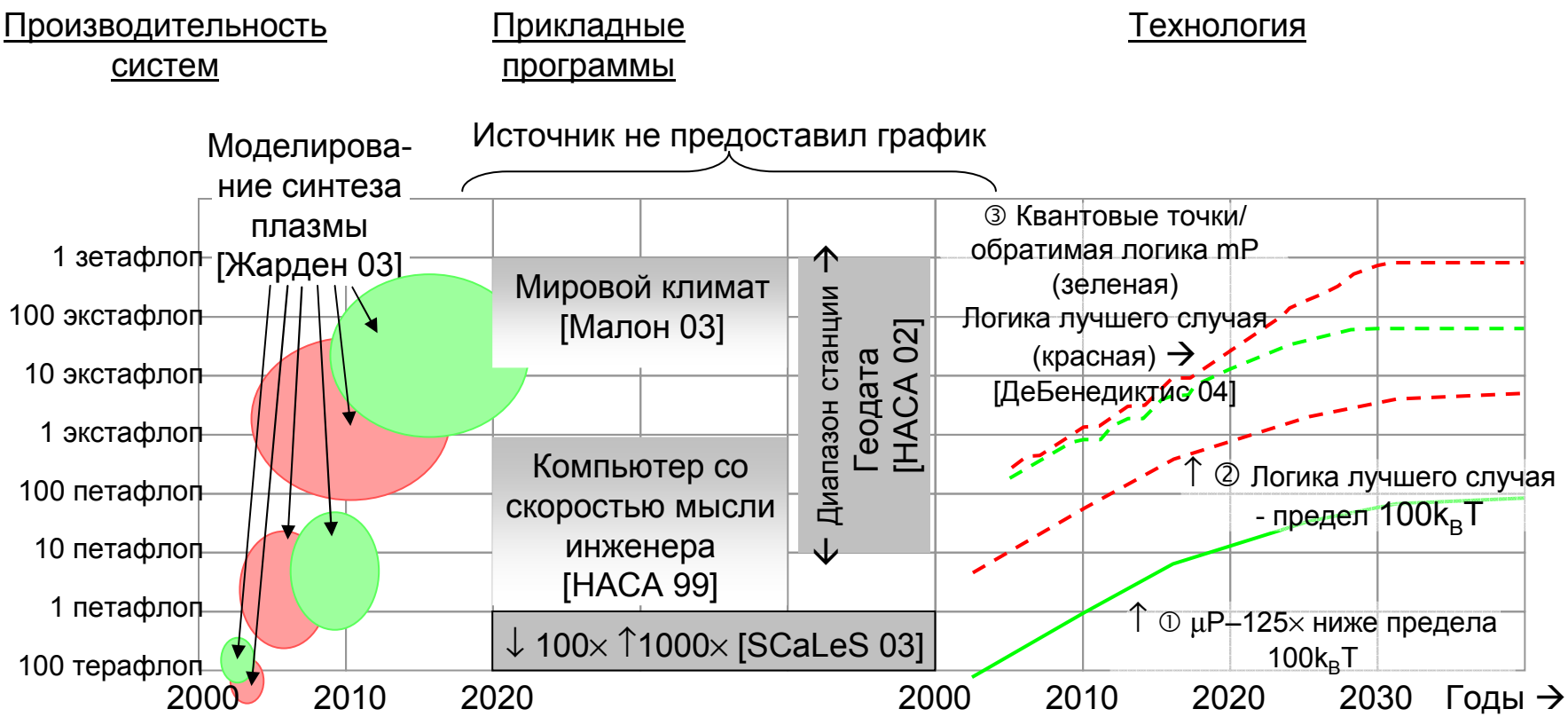
Erik P. DeBenedictis
Sandia National Laboratories, USA
Эрик П. ДеБенедиктис
Сандийские национальные лаборатории, США

Conference on Super Computations and Computer Simulation
Sarov, Russia, 2004
Конференция по суперкомпьютерам и компьютерному моделированию
Саров, Россия, 2004 г.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
Сандия – это многопрофильная лаборатория, управляемая корпорацией «Сандия», дочерней компанией фирмы «Локхид-Мартин», по контракту с Национальной администрацией по ядерной безопасности Министерства энергетики США № DE-AC04-94AL85000

Sandia National Laboratories

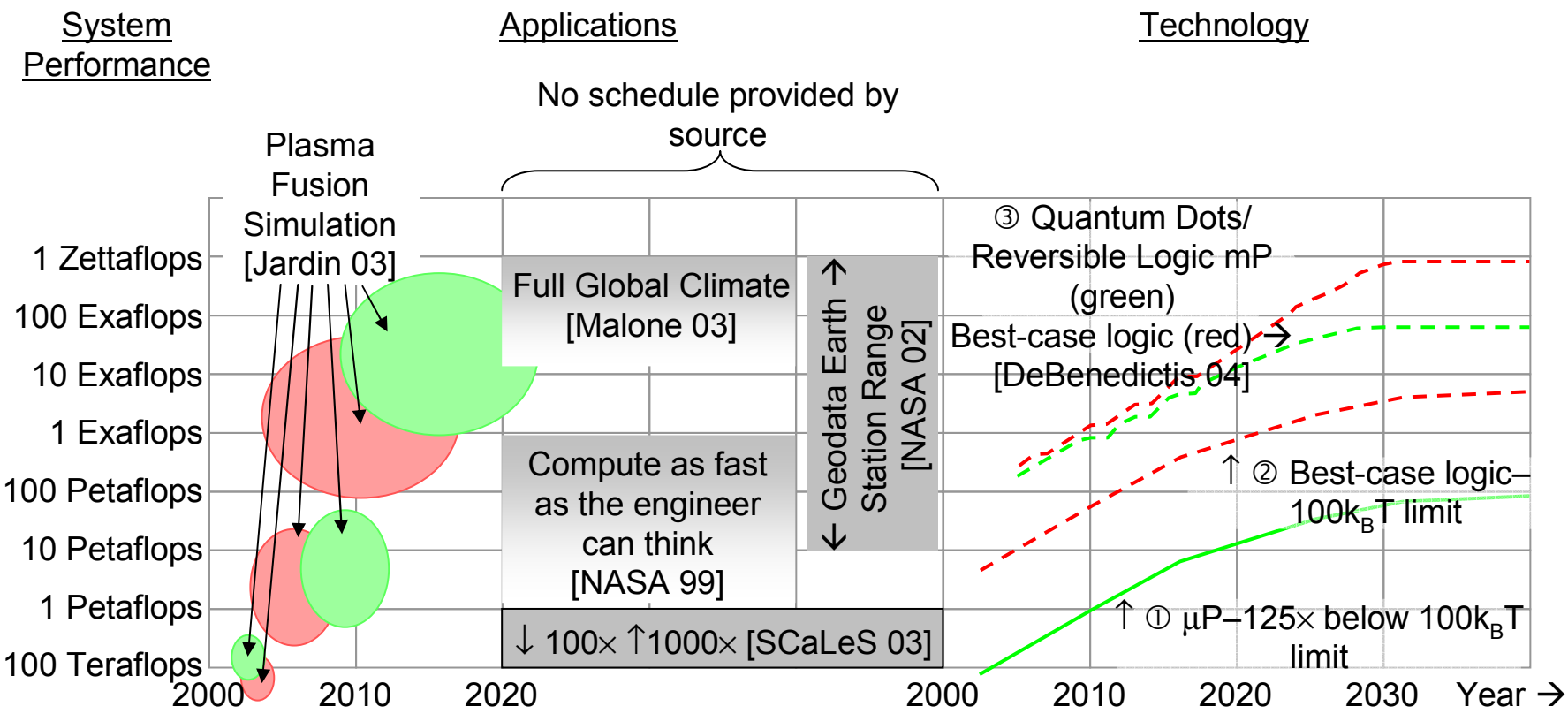
Прикладные программы и компьютерная технология



[Jardin 03] S.C. Jardin, “Plasma Science Contribution to the SCaLeS Report,” Princeton Plasma Physics Laboratory, PPPL-3879 UC-70, available on Internet.
[Malone 03] Robert C. Malone, John B. Drake, Philip W. Jones, Douglas A. Rotman, “High-End Computing in Climate Modeling,” contribution to SCaLeS report.
[NASA 99] R. T. Biedron, P. Mehrotra, M. L. Nelson, F. S. Preston, J. J. Rehder, J. L. Rogers, D. H. Rudy, J. Sobieski, and O. O. Storaasli, “Compute as Fast as the Engineers Can Think!” NASA/TM-1999-209715, available on Internet.
[NASA 02] NASA Goddard Space Flight Center, “Advanced Weather Prediction Technologies: NASA’s Contribution to the Operational Agencies,” available on Internet.
[SCaLeS 03] Workshop on the Science Case for Large-scale Simulation, June 24-25, proceedings on Internet a http://www.pnl.gov/scales/.
[DeBenedictis 04], Erik P. DeBenedictis, “Matching Supercomputing to Progress in Science,” July 2004. Presentation at Lawrence Berkeley National Laboratory, also published as Sandia National Laboratories SAND report SAND2004-3333P. Sandia technical reports are available by going to http://www.sandia.gov and accessing the technical library.

Sandia National Laboratories

Applications and Computer Technology



[Jardin 03] S.C. Jardin, "Plasma Science Contribution to the SCaLeS Report," Princeton Plasma Physics Laboratory, PPPL-3879 UC-70, available on Internet.
[Malone 03] Robert C. Malone, John B. Drake, Philip W. Jones, Douglas A. Rotman, "High-End Computing in Climate Modeling," contribution to SCaLeS report.
[NASA 99] R. T. Biedron, P. Mehrotra, M. L. Nelson, F. S. Preston, J. J. Rehder, J. L. Rogers, D. H. Rudy, J. Sobieski, and O. O. Storaasli, "Compute as Fast as the Engineers Can Think!" NASA/TM-1999-209715, available on Internet.
[NASA 02] NASA Goddard Space Flight Center, "Advanced Weather Prediction Technologies: NASA's Contribution to the Operational Agencies," available on Internet.
[SCaLeS 03] Workshop on the Science Case for Large-scale Simulation, June 24-25, proceedings on Internet a http://www.pnl.gov/scales/.
[DeBenedictis 04], Erik P. DeBenedictis, "Matching Supercomputing to Progress in Science," July 2004. Presentation at Lawrence Berkeley National Laboratory, also published as Sandia National Laboratories SAND report SAND2004-3333P. Sandia technical reports are available by going to http://www.sandia.gov and accessing the technical library.

Sandia National Laboratories

Outline - Программа

• Future Applications	• Будущие прикладные программы
• Limits of Moore's Law	• Пределы закона Мора
• Applications Modeling	• Моделирование прикладных программ
• Implications to Design	• Осложнения с дизайном
• Conclusions	• Выводы

Sandia National Laboratories

Рост скорости вычислений и решение задач мирового климата

	Задача	Масштабирование
1 Зетафлоп	Ансамбли, сценарии 10×	Постыдно параллельное
100 Экзафлоп	Продолжительность 100×	Большее время обсчета
1 Экзафлоп	Новые параметризации 100×	Более сложные физические задачи
10 Петафлоп	Полнота модели 100×	Более сложные физические задачи
100 Терафлоп	Пространст. разрешение 10^4× (10^3×-10^5×)	Разрешение
10 Гигафлоп	Используемые кластеры (100 узлов, эффективность 5%)	

Ref. “High-End Computing in Climate Modeling,” Robert C. Malone, LANL, John B. Drake, ORNL, Philip W. Jones, LANL, and Douglas A. Rotman, LLNL (2004)

Sandia National Laboratories

FLOPS Increases for Global Climate

	Issue	Scaling
1 Zettaflops	Ensembles, scenarios 10×	Embarrassingly Parallel
100 Exaflops	Run length 100×	Longer Running Time
1 Exaflops	New parameterizations 100×	More Complex Physics
10 Petaflops	Model Completeness 100×	More Complex Physics
100 Teraflops	Spatial Resolution $10^4\times$ ($10^3\times$-$10^5\times$)	Resolution
10 Gigaflops	Clusters Now In Use (100 nodes, 5% efficient)	

Ref. "High-End Computing in Climate Modeling," Robert C. Malone, LANL, John B. Drake, ORNL, Philip W. Jones, LANL, and Douglas A. Rotman, LLNL (2004)

Sandia National Laboratories

Outline - Программа

- Future Applications
- Limits of Moore's Law
- Applications Modeling
- Implications to Design
- Conclusions

- Будущие прикладные программы
- Пределы закона Мора
- Моделирование прикладных программ
- Осложнения с дизайном
- Выводы

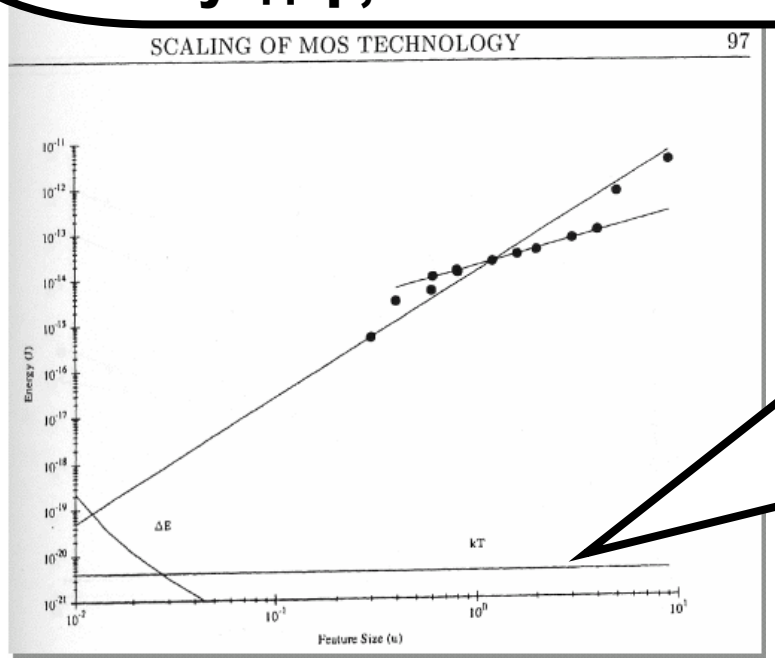
Sandia National Laboratories

Предел теплового шума

Эта логическая необратимость связана с физической необратимостью и требует минимальной генерации тепла на машинный цикл, обычно порядка kT на каждую необратимую функцию.
– Р. Лаундер, 1961 г.

R. Landauer

Irreversibility and Heat Generation in the Computing Process

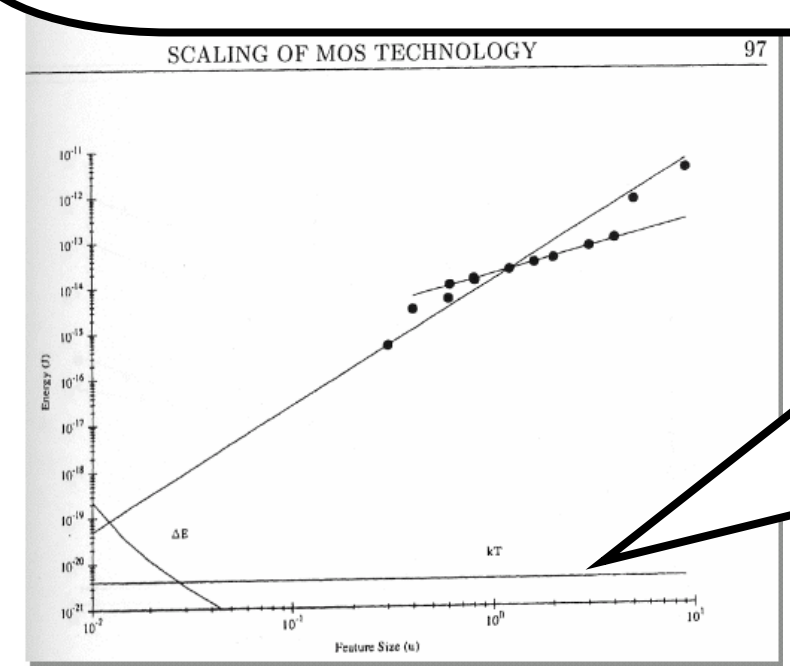


«Линия-помощник» kT проведена из фокусной точки наблюдателя, поскольку при написании она была не важна
– Карвер Мид, Масштабирование технологии МОП, 1994

Sandia National Laboratories

Thermal Noise Limit

This logical irreversibility is associated with physical irreversibility and requires a minimal heat generation, per machine cycle, typically of the order of kT for each irreversible function.
– R. Landauer 1961



kT "helper line," drawn out of the reader's focus because it wasn't important at the time of writing
– Carver Mead, Scaling of MOS Technology, 1994

Sandia National Laboratories

Semiconductor Roadmap
Карта развития полупроводников

YEAR OF PRODUCTION	2010	2013	2016
DRAM ½ PITCH (nm)	45	32	22
MPU / ASIC ½ PITCH (nm)	50	35	25
MPU PRINTED GATE LENGTH (nm)	25	18	13
MPU PHYSICAL GATE LENGTH (nm)	18	13	9
Physical gate length high-performance (HP) (nm) [1]	18	13	9
Equivalent physical oxide thickness for high-performance T_{ox} (EOT)(nm) [2]	0.5-0.8	0.4-0.6	0.4-0.5
Gate depletion and quantum effects electrical thickness adjustment factor (nm) [3]	0.5	0.5	0.5
T_{ox} electrical equivalent (nm) [4]	1.2	1.0	0.9
Nominal power supply voltage (V_{dd}) (V) [5]	0.6	0.5	0.4
Nominal high-performance NMOS sub threshold leakage current, I_{sdleak} (at 25°C) (μA/μm) [6]	3	7	10
Nominal high-performance NMOS saturation drive current, I_{dd} (at V_{dd} at 25°C) (μA/μm) [7]	1200	1500	1500
Required percent current-drive "mobility/transconductance improvement" [8]	30%	70%	100%
Parasitic source/drain resistance (Rsd) (ohm-um) [9]	110	90	80
Parasitic source/drain resistance (Rsd) pe	25%	30%	35%
Parasitic capacitance percent of ideal gat	31%	36%	42%
High-performance NMOS device τ (C_{gate} * V_{dd} / I_{dd}-NMOS)(ps) [12]	0.39	0.22	0.15
Relative device performance [13]	4.5	7.2	10.7
Energy per (W/L_{gate}=3) device switching transition (C_{gate}*(3*L_{gate})*V^2) (fJ/Device) [14]	0.015	0.007	0.002
Static power dissipation per (W/Lgate=3) device (Watts/Device) [15]	9.7E-08	1.4E-07	1.1E-07

1000 k_BT/транзистор

White—Manufacturable Solutions Exist, and Are Being Optimized

Yellow—Manufacturable Solutions are Known

Red—Manufacturable Solutions are NOT Known

Sandia National Laboratories

Пределы научных суперкомпьютеров

Логика лучшего случая	Архитектура микропроцессора	Физический фактор	Источник авторитета
2×10^{24} логич. оп/с		Предел надежности 750KW/(80k_BT)	Ведущие физики (T=60°C температура стыка)
		Сниж. 20,000 преобр. логич. оп. в плав. запят.	Техника с плав. запятой (прецизия 64 бит)
Мнение экспертов	100 Экзафлоп 800 Петафлоп ← 125:1 →	Сниж. на произв. допуск (4×)	Оценка
Оценка	25 Экзафлоп 200 Петафлоп	Неопределенность (6×)	Разрыв в диаграмме
	4 Экзафлоп 32 Петафлоп	Улучшенные устр-ва (4×)	Оценка
	1 Экзафлоп 8 Петафлоп	Планир. Улечшение ITRS на 22 нм (100×)	Комитет экспертов ITRS
	80 Терафлоп	Снижение напряже-ния питания (2×)	Комитет экспертов ITRS
	40 Терафлоп	«Red Storm»	Контракт

Положение: суперкомпьютер размером и ценой «Ред сторм»: бюджет US$100M; энергия сети 2 МВт; 750 кВт на акттивные компоненты

Sandia National Laboratories

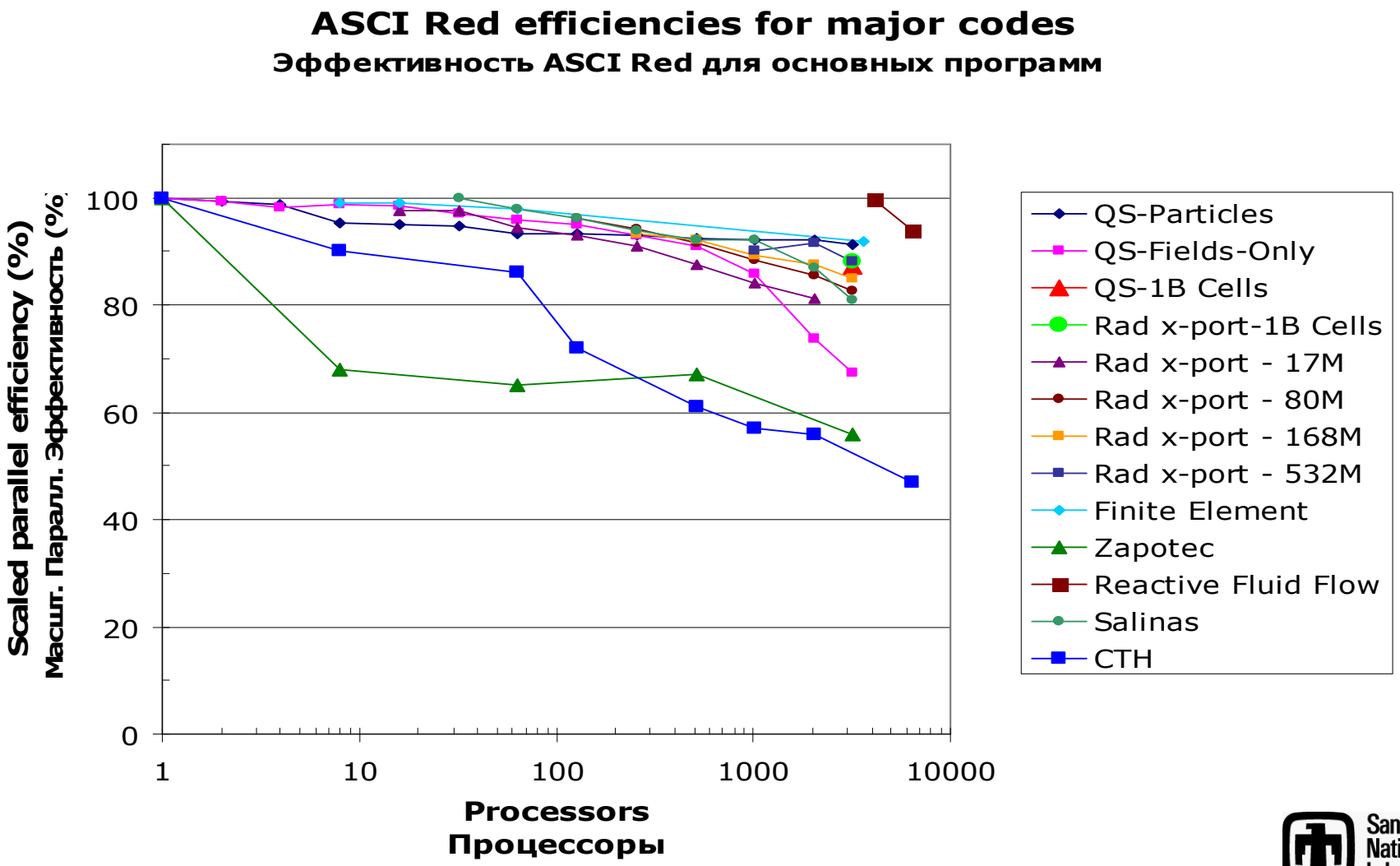
Outline - Программа

- Future Applications
- Limits of Moore's Law
- Applications Modeling
- Implications to Design
- Conclusions

- Будущие прикладные программы
- Пределы закона Мора
- Моделирование прикладных программ
- Осложнения с дизайном
- Выводы

Sandia National Laboratories

Benchmarks
Достижения



Sandia National Laboratories

Measurement Tools
(Измерительные инструменты)

- Measurements may be impacted by system stability or timing noise
- It may be necessary to measure and compensate for timing stability
- На измерения может повлиять стабильность системы или помехи синхронизации
- Возможно, придется измерять и компенсировать стабильность синхронизации

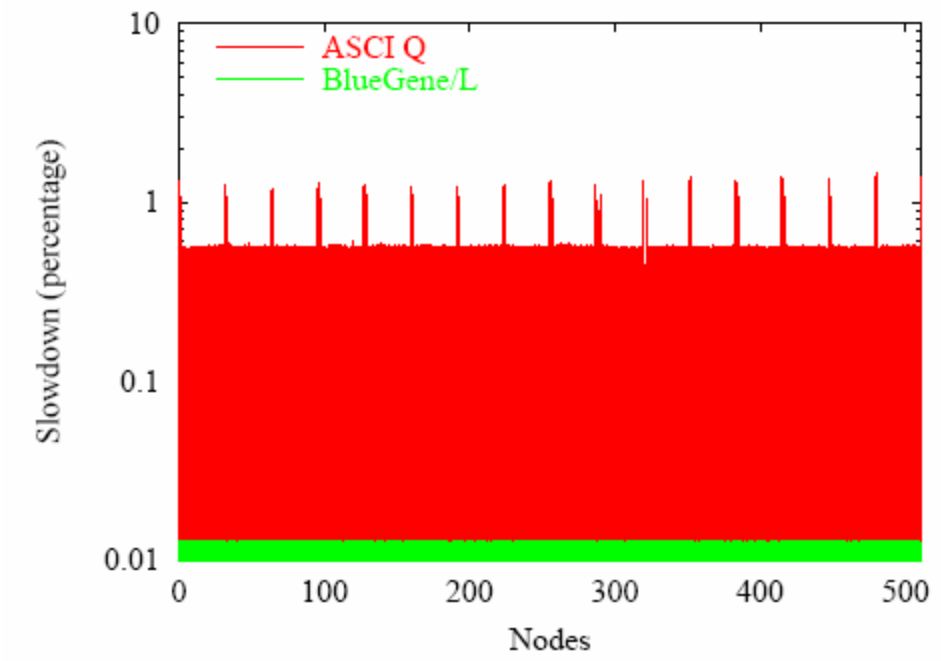


Figure 2. Slowdown due to computational noise on a per node basis.

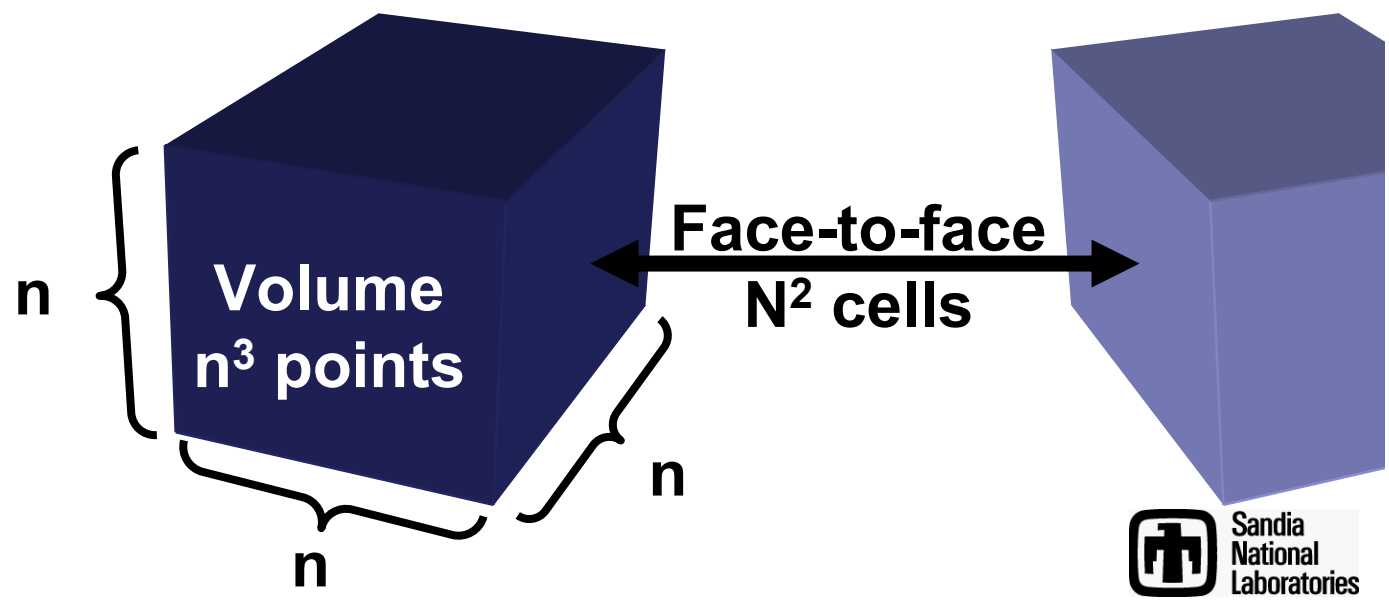
Sandia National Laboratories

Develop Model - Разработка модели

- Simple case: finite difference equation
- Each node holds n×n×n grid points

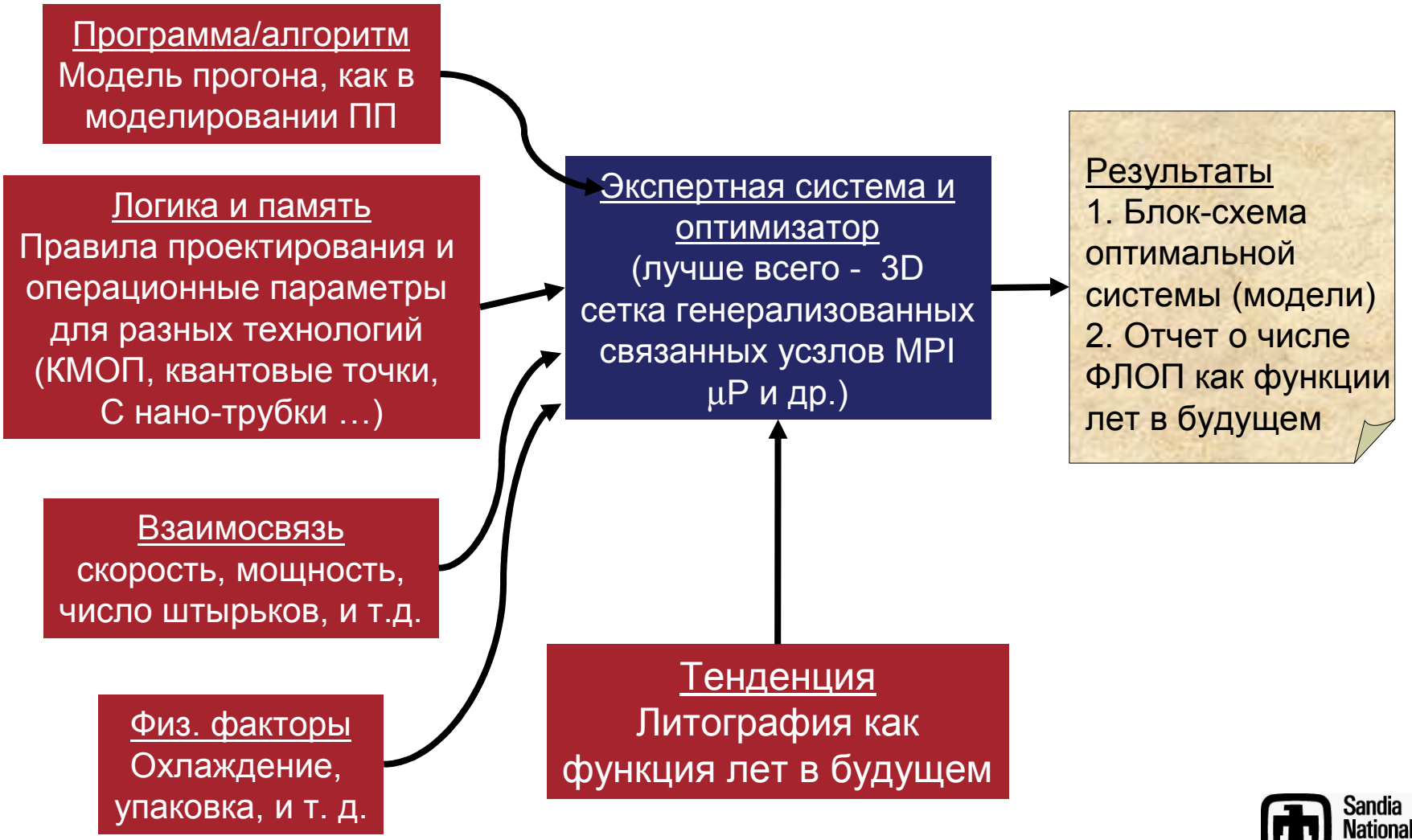
- Простой случай: конечное дифференциальное уравнение
- Каждый узел содержит n×n×n точек матрицы

$$T_{step} = 6\, n^2\, C_{bytes}\, T_{byte} + n^3\, F_{grind}/\text{floprate}$$



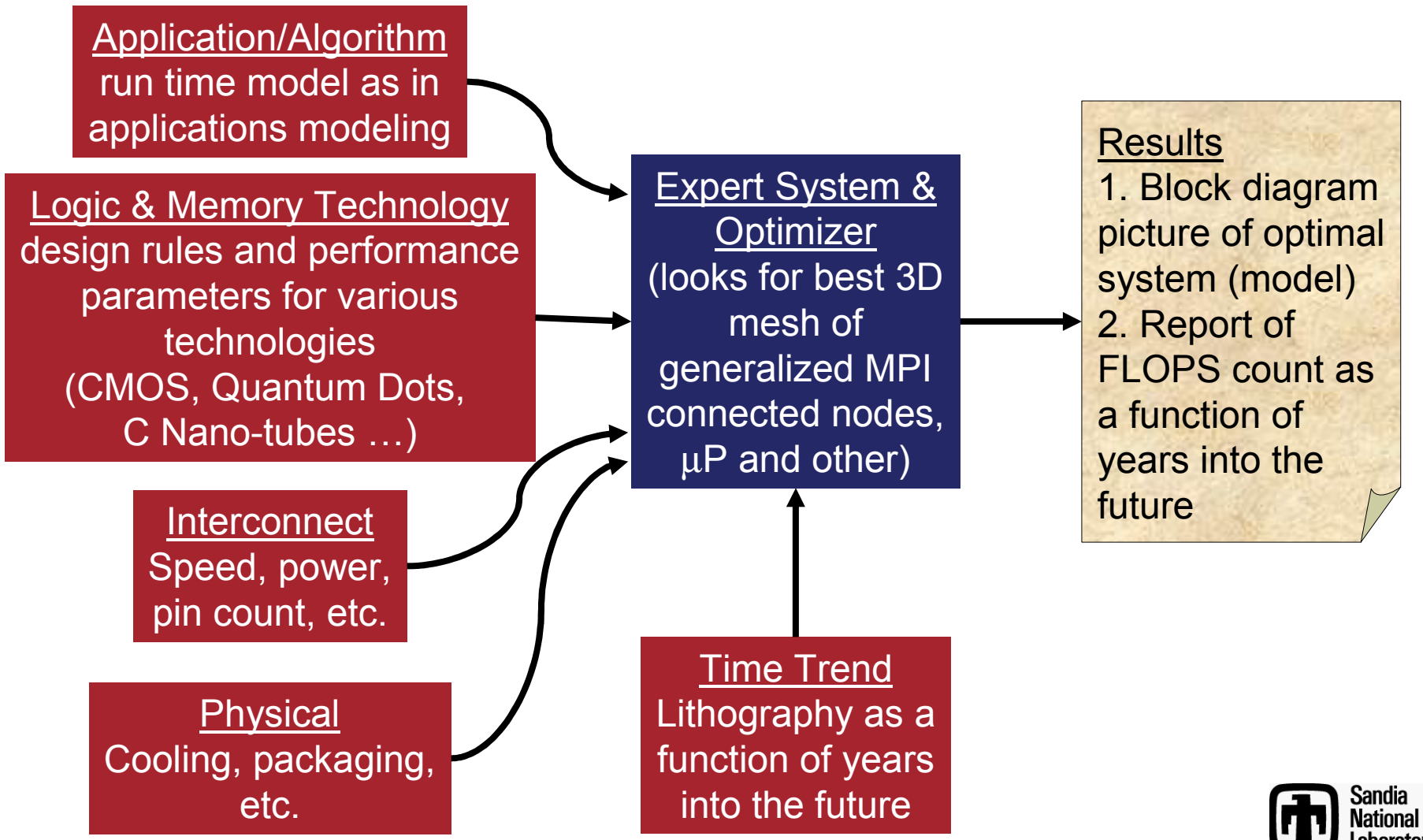
Sandia National Laboratories

Суперкомпьютерная экспертная система



Sandia National Laboratories

Supercomputer Expert System



Sandia National Laboratories

Экспертная система для будущих суперкомпьютеров

- Моделирование программ
  - Проход
    $T_{run} = f_1(n, \text{проект})$
- Технологическая карта
  - Скорость = f_2(год),
  - Плотность чипа = f_3(год),
  - cost = $(n, проект), …
- Масштабирование объективной функции
  - Я имею $$C_1$ и могу ждать $T_{run}=C_2$ секунд. Каково наибольшее n которое я смогу решить в год Y?

- Использование “экспертной системы” для решения:

$$\max_{\text{All designs}} n: \$<C_1,\ T_{run}<C_2$$

- Отчет:

Операции с пл. запятой

T_{run}(n, проект.)

и иллюстрация “проекта”

Sandia National Laboratories

Expert System for Future Supercomputers

- Applications Modeling
  - Runtime
    $T_{run} = f_1(n, design)$
- Technology Roadmap
  - Gate speed = f_2(year),
  - chip density = f_3(year),
  - cost = $(n, design), …
- Scaling Objective Function
  - I have $\$C_1$ & can wait $T_{run}=C_2$ seconds. What is the biggest n I can solve in year Y?

- Use "Expert System" To Calculate:

$$\max_{\text{All designs}} n: \$<C_1,\ T_{run}<C_2$$

- Report:

<u>Floating operations</u>

T_{run}(n, design)

and illustrate "design"

Sandia National Laboratories

Outline - Программа

- Future Applications
- Limits of Moore's Law
- Applications Modeling
- Implications to Design
- Conclusions

- Будущие прикладные программы
- Пределы закона Мора
- Моделирование прикладных программ
- Осложнения с дизайном
- Выводы

Sandia National Laboratories

Example - Пример

- Project to build a 65536 processor supercomputer
- Build first 512 processors
- Project full size with math
- Finally commit to full size
- Проект построения суперкомпьютера с 65536 процессорами
- Построение первых 512 процессоров
- План построения полного суперкомпьютера с математическим обеспечением
- Окончательное решение по построению суперкомпьютера

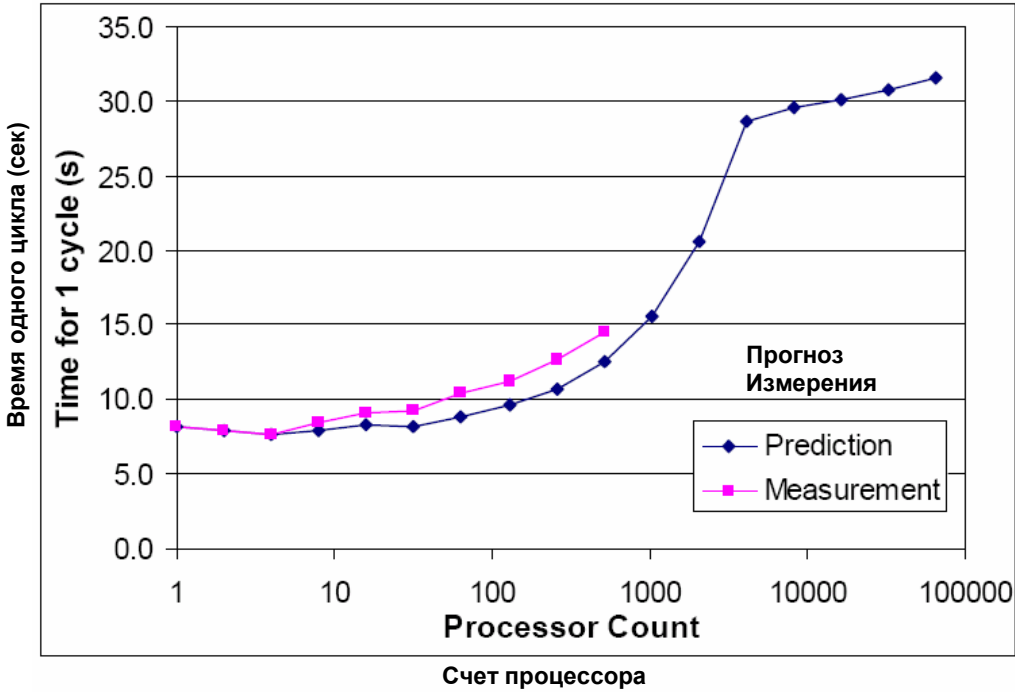
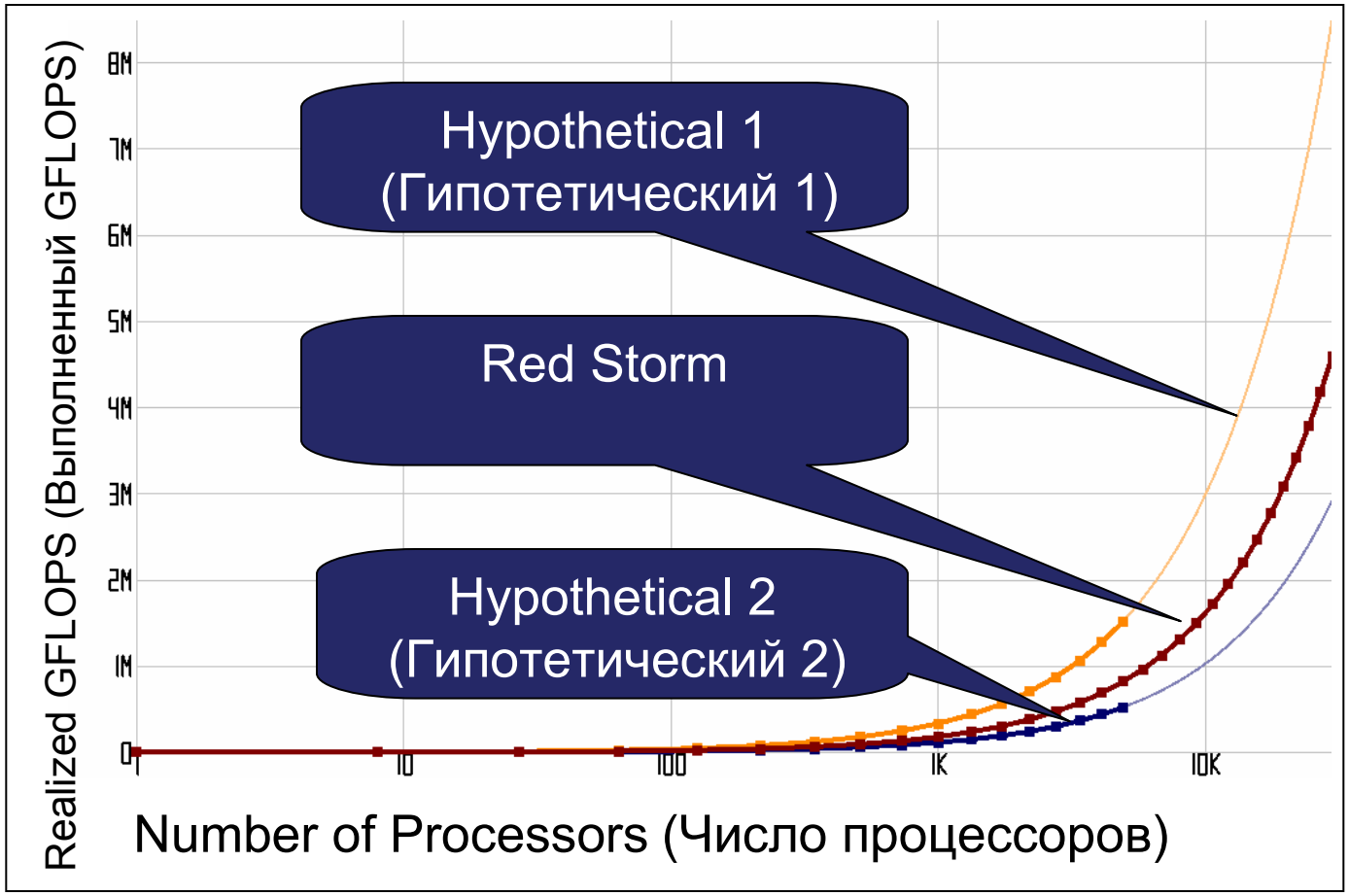


Figure 22. SAGE Performance (timing_h) on BG/L

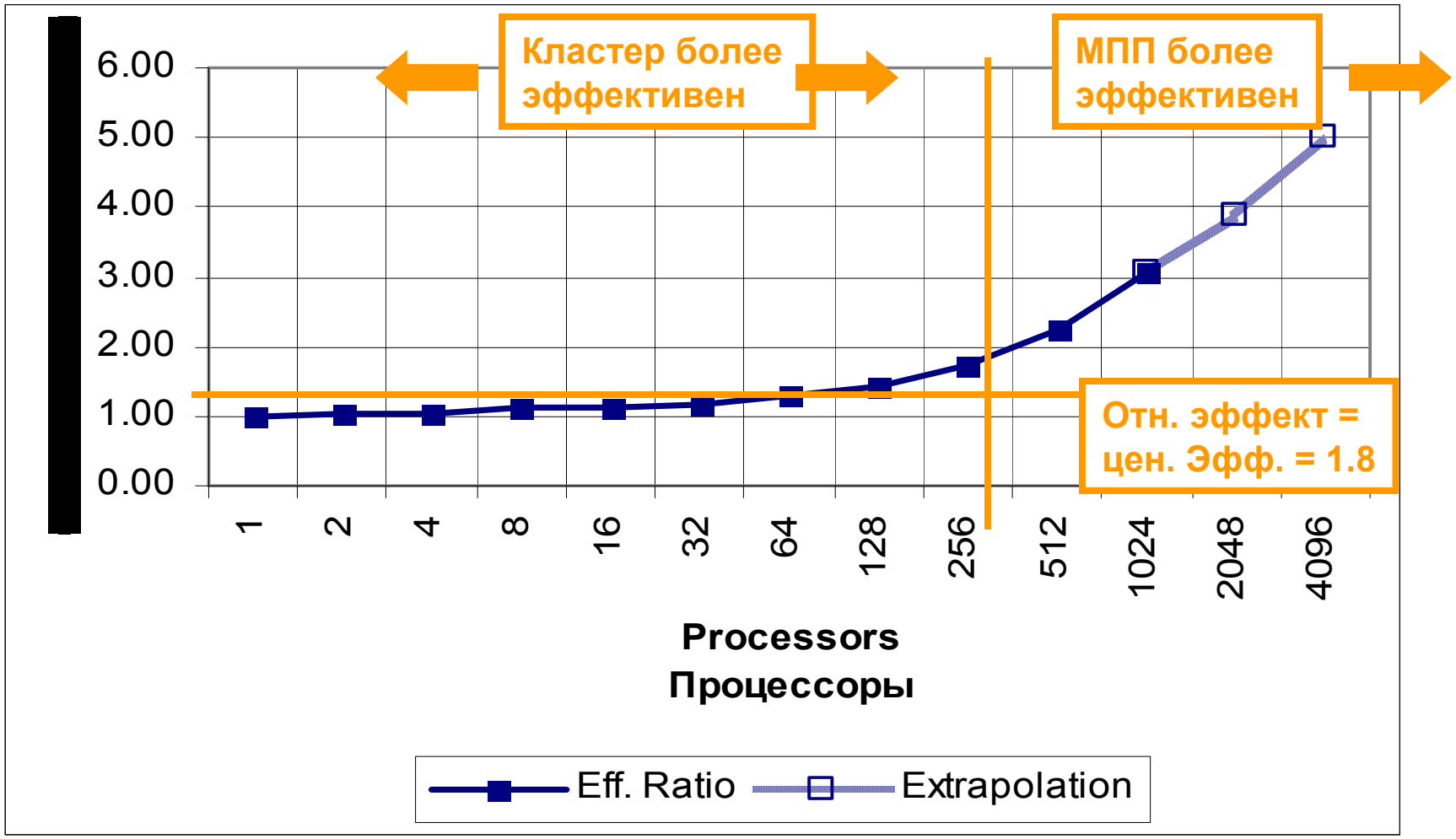
Sandia National Laboratories

New Situations - Новые ситуации



Sandia National Laboratories

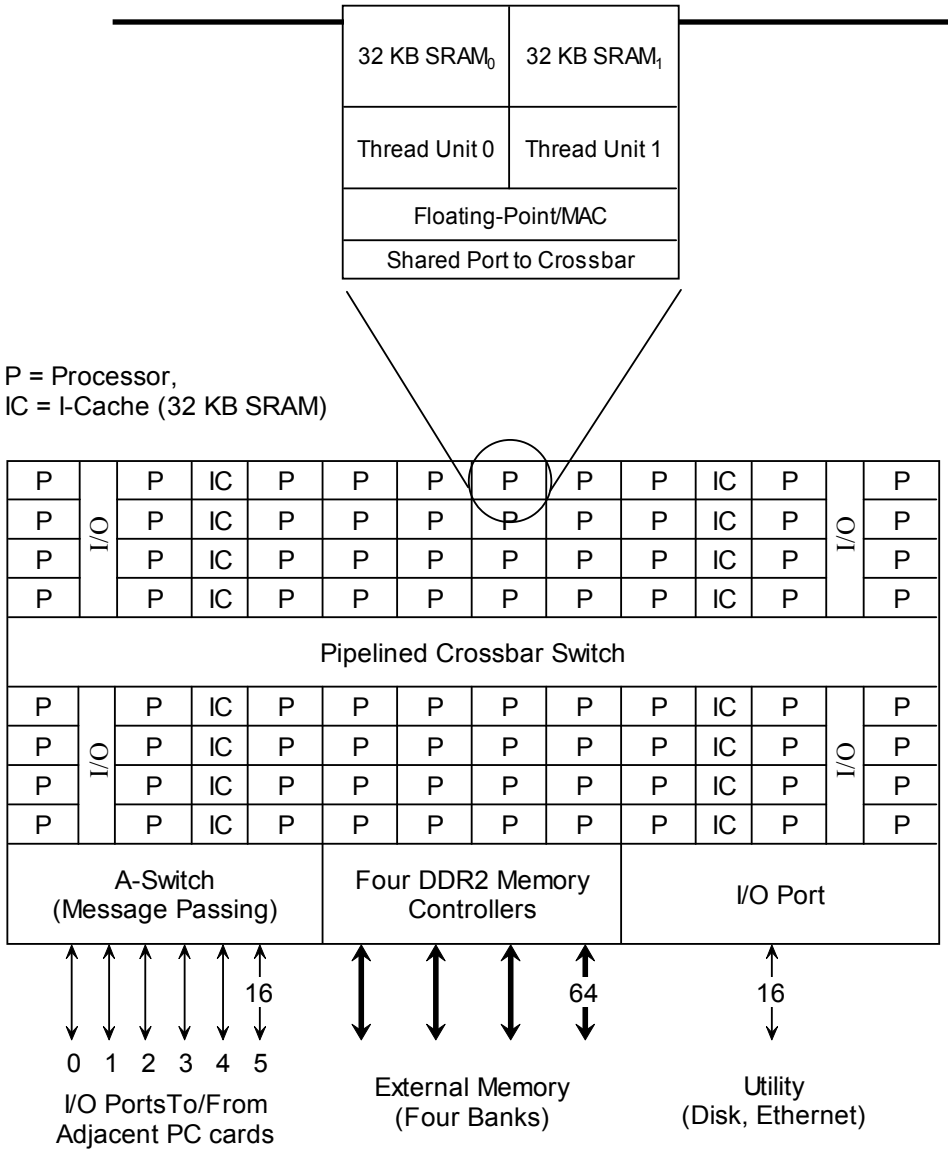
Future portfolio combined efficiency
Совместная эффективность будущего пакета



Sandia National Laboratories

Cyclops: Advanced Cellular Architecture
Циклоп: усовершенствованная ячейковая архитектура

Cyclops: Advanced Cellular Architecture
Циклоп: усовершенствованная ячейковая архитектура

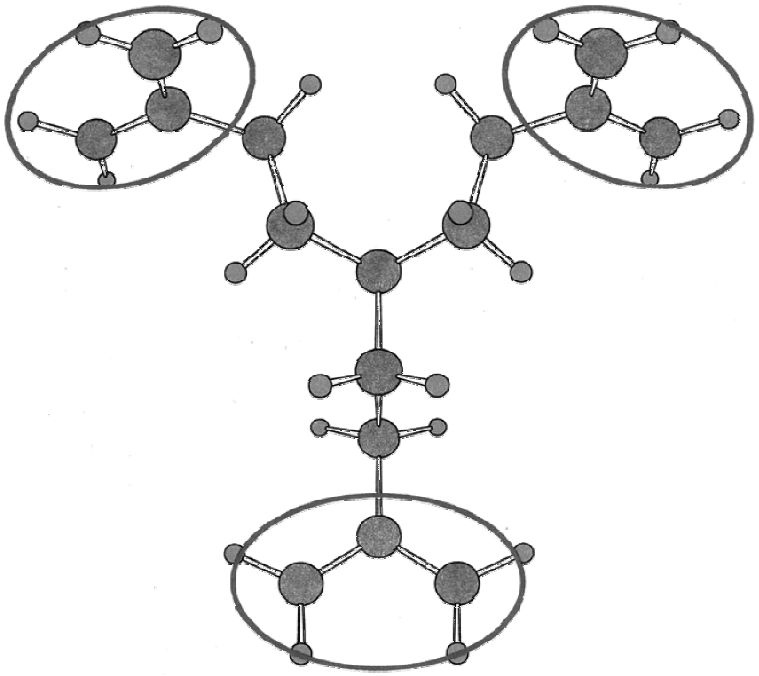


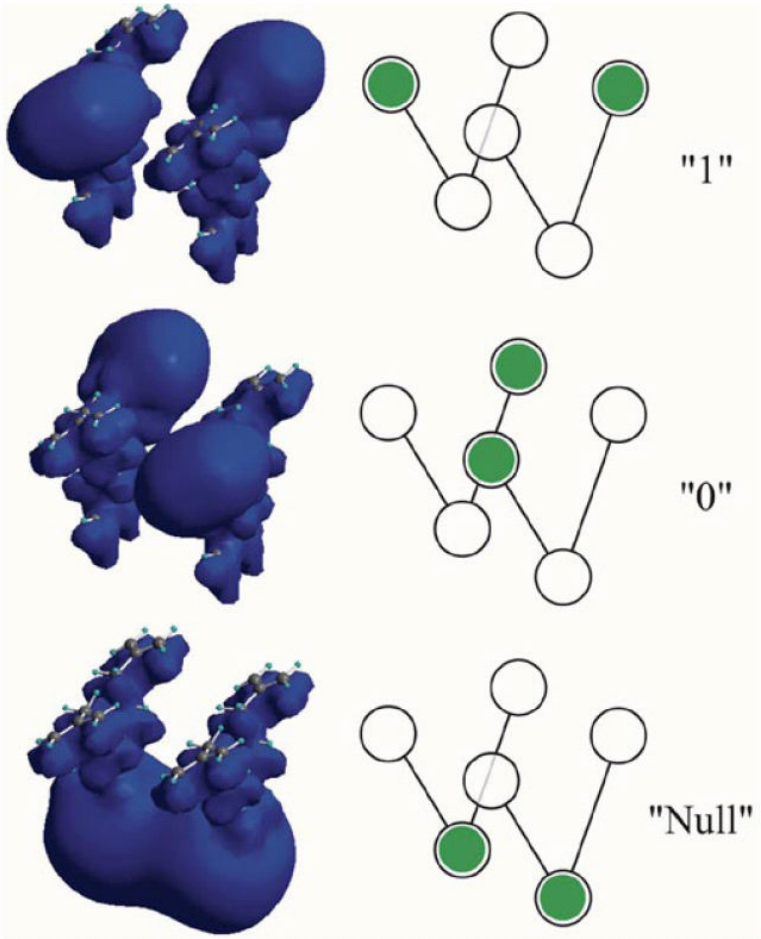
- **Chip Architecture on Left**
  - **80 floating point units per CPU**
- **Full system is 24×24×24 3D mesh**
- **Архитектура микрочипа слева**
  - **80 узлов плавающей запятой на ЦПУ**
- **Полная система: 3х-мерная сетка 24×24×24**

Sandia National Laboratories

Beyond Transistors: Quantum Dot Logic
После транзисторов: логика кванотовых точек

- **Pairs of molecules create a memory cell or a logic gate**
- **Пары молекул создают ячейку памяти или логический переключатель**



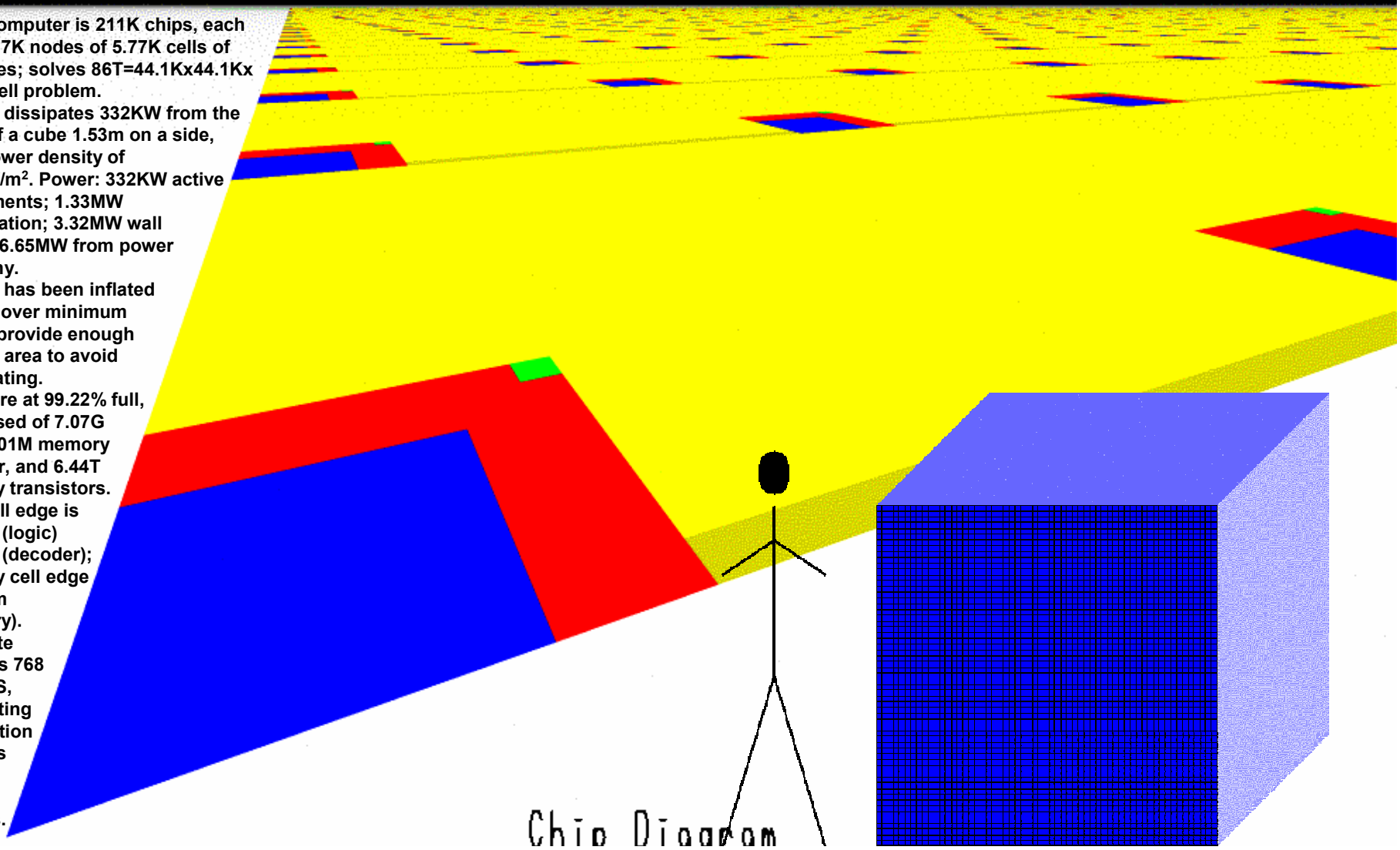


Ref. "Clocked Molecular Quantum-Dot Cellular Automata," Craig S. Lent and Beth Isakse
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 50, NO. 9, SEPTEMBER 2003

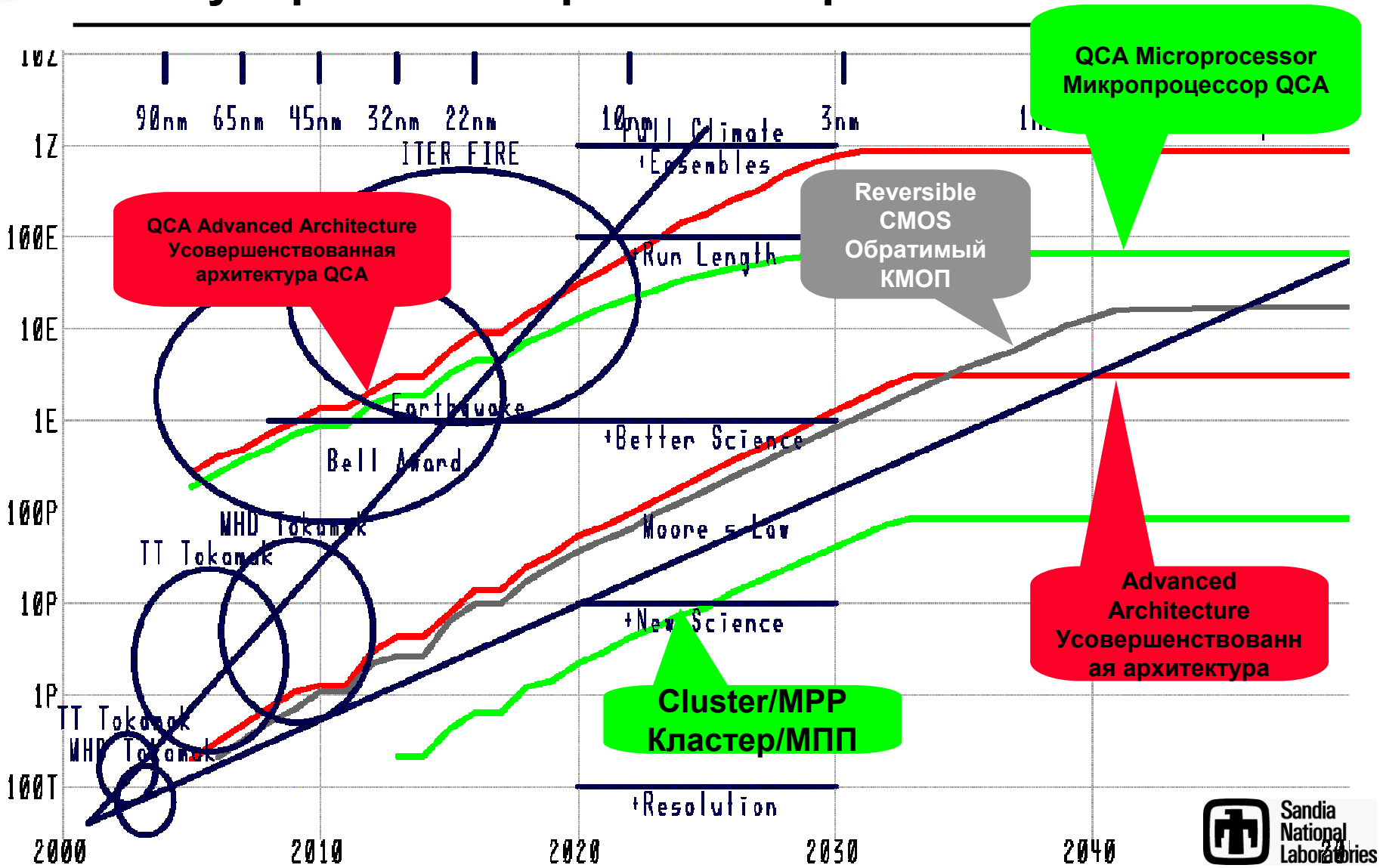
Sandia National Laboratories

CTH at a Zettaflops
CTH на уровне зетафлоп

Supercomputer is 211K chips, each with 70.7K nodes of 5.77K cells of 240 bytes; solves 86T=44.1Kx44.1Kx 44.1K cell problem.
System dissipates 332KW from the faces of a cube 1.53m on a side, for a power density of 47.3KW/m^2. Power: 332KW active components; 1.33MW refrigeration; 3.32MW wall power; 6.65MW from power company.
System has been inflated by 2.57 over minimum size to provide enough surface area to avoid overheating.
Chips are at 99.22% full, comprised of 7.07G logic, 101M memory decoder, and 6.44T memory transistors.
Gate cell edge is 34.4nm (logic) 34.4nm (decoder); memory cell edge is 4.5nm (memory).
Compute power is 768 EFLOPS, completing an iteration in 224μs and a run in 9.88s.

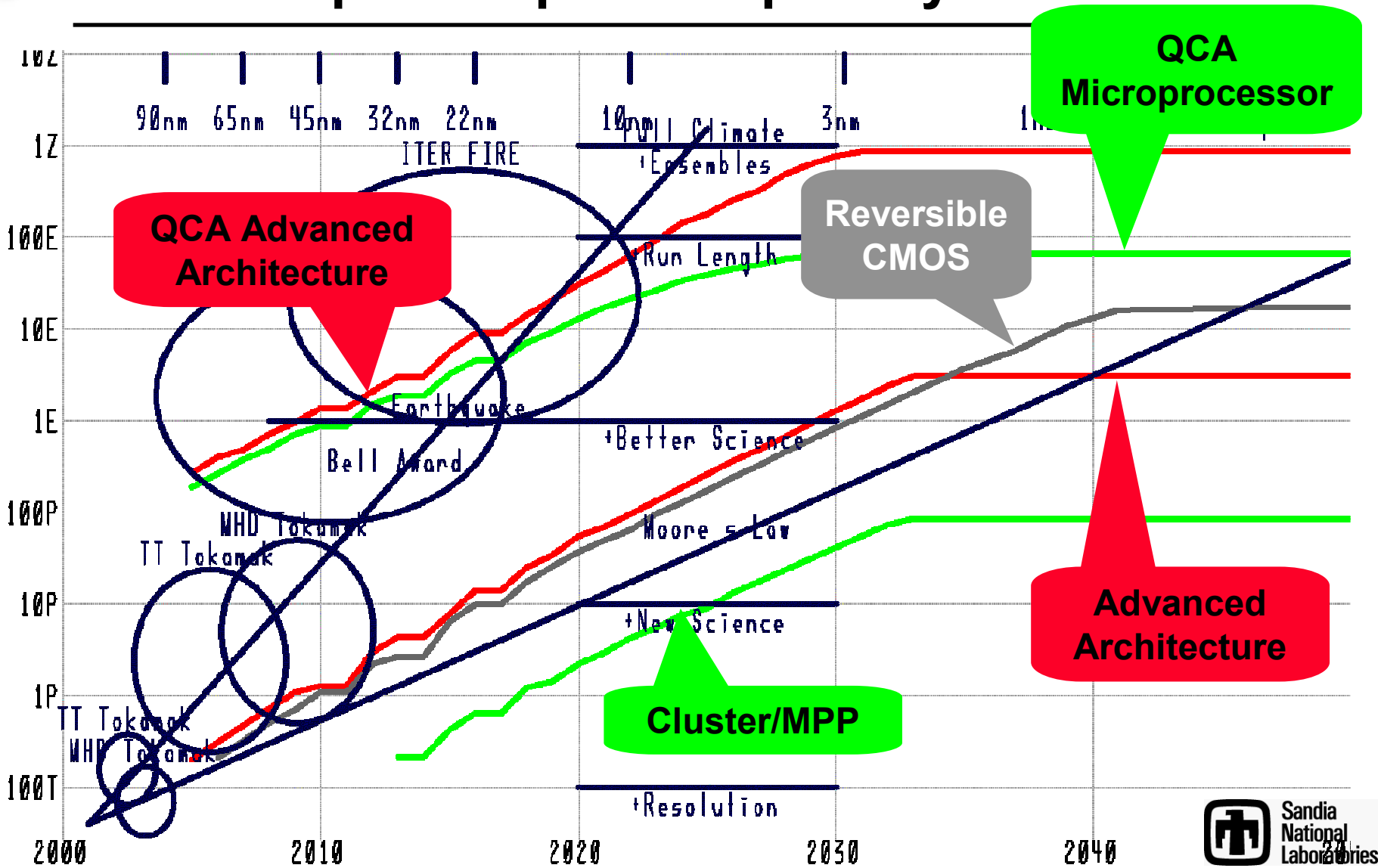


Supercomputer Expert System
Суперкомпьютерная экспертная система



TRANSLATE TO RUSSIAN+ENGLISH

Supercomputer Expert System



Outline - Программа

- Future Applications
- Limits of Moore's Law
- Applications Modeling
- Implications to Design
- Conclusions

- Будущие прикладные программы
- Пределы закона Мора
- Моделирование прикладных программ
- Осложнения с дизайном
- Выводы

Sandia National Laboratories

Conclusions
Выводы

- Faster computers enable more science
  - Applications appear to need up to 10^{21} FLOPS
  - Supercomputer plans can supply up to 10^{15} FLOPS
  - Conclusion: Need more
- Applications Modeling can predict performance
- Generations
  - Cluster/MPP
  - Advanced architecture
  - New devices

- Повышение скорости компьютеров позволяет решать более сложные научные задачи
  - Программам нужна скорость до 10^{21} флоп
  - Суперкомпьютеры планируют дать до 10^{15} флоп
  - Вывод: нужно больше
- Моделирование программ может предсказать производительность
- Поколения
  - Кластер/МПП
  - Усовершенствованная архитектура
  - Новые устройства

Sandia National Laboratories