

Школа-семинар

"Поиск эффективных суперкомпьютерных архитектур в пост-Муровскую эру"

Влияние баланса вычислительной производительности и скорости доступа к памяти на эффективность расчетов электронной структуры : сравнение процессоров Intel, AMD и Nvidia

Вечер В.С., Стегайлов В.В.



Популярность задач вычислительного материаловедения

- ARCHER – самый мощный академический суперкомпьютер в Великобритании

Code Usage on ARCHER (2014-15) by CPU Time:

Rank	Code	Node hours	Method	GPU Acceleration
1	VASP	5,443,924	DFT	-
3	CP2K	2,121,237	DFT	<i>Partial</i>
4	CASTEP	1,564,080	DFT	<i>Partial</i>
9	LAMMPS	887,031	Classical	+
10	ONETEP	805,014	DFT	<i>Partial</i>
12	NAMD	516,851	Classical	+
20	DL_POLY	245,322	Classical	+

52% of all CPU time used by Chemistry / Materials Science / Biomolecular Simulation

Мировые тенденции снижения энергопотребления

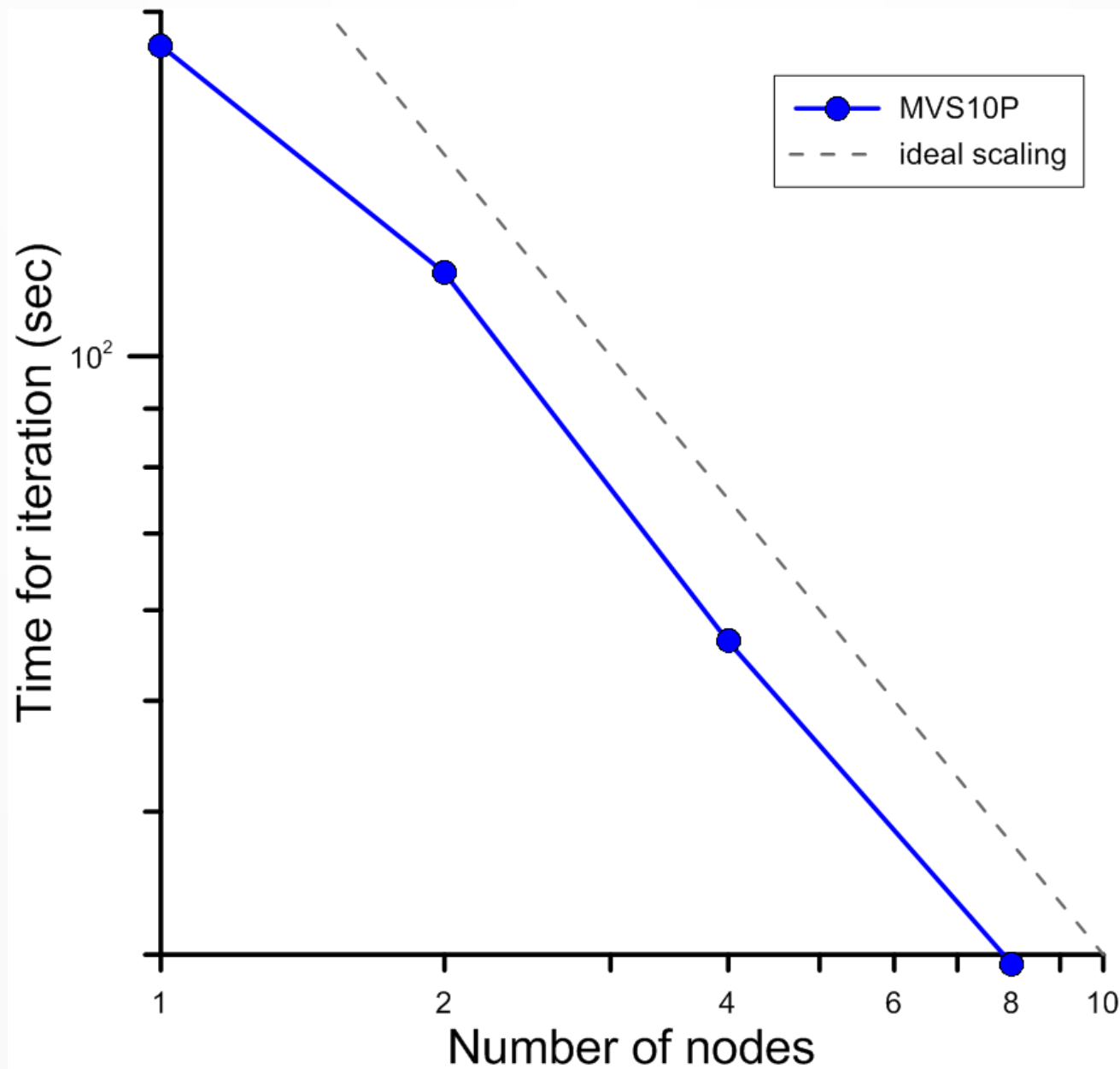
Green500 Rank	MFLOPS/W	Site	System	Total Power(kW)
1	6673.8	Advanced Center for Computing and Communication, RIKEN	ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SCnp	150.0
2	6195.2	Computational Astrophysics Laboratory, RIKEN	ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SCnp	46.9
3	6051.3	National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	15371
4	5272.1	GSI Helmholtz Center	ASUS ESC4000 FDR/G2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR, AMD FirePro S9150	57.2
5	4778.5	Institute of Modern Physics (IMP), Chinese Academy of Sciences	Sugon Cluster W780I, Xeon E5-2640v3 8C 2.6GHz, Infiniband QDR, NVIDIA Tesla K80	65
6	4112.1	Stanford Research Computing Center	Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, Nvidia K80	190
7	3775.5	Internet Service (B)	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110
8	3775.5	Internet Service (B)	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110
9	3775.5	Internet Service (B)	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110
10	3775.5	Internet Service (B)	Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40	110



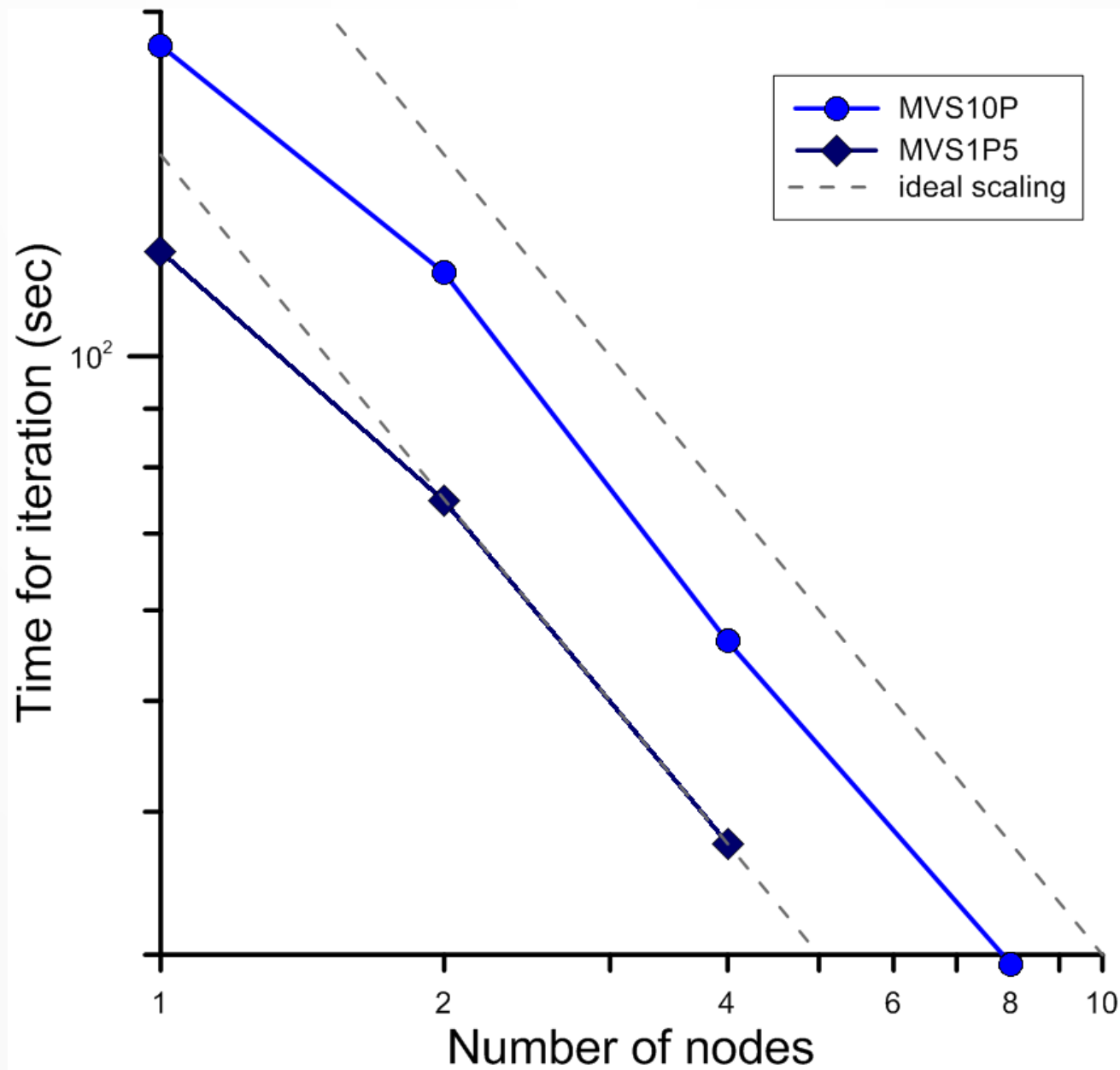
Квантовые ab-initio расчеты (VASP)

- VASP (*The Vienna Vienna Ab initio simulation package*) – самый популярный пакет для квантовых расчетов в мире (15 - 20% всей вычислительной мощности тратятся на расчеты в VASP)
- Является memory bound и compute bound кодом

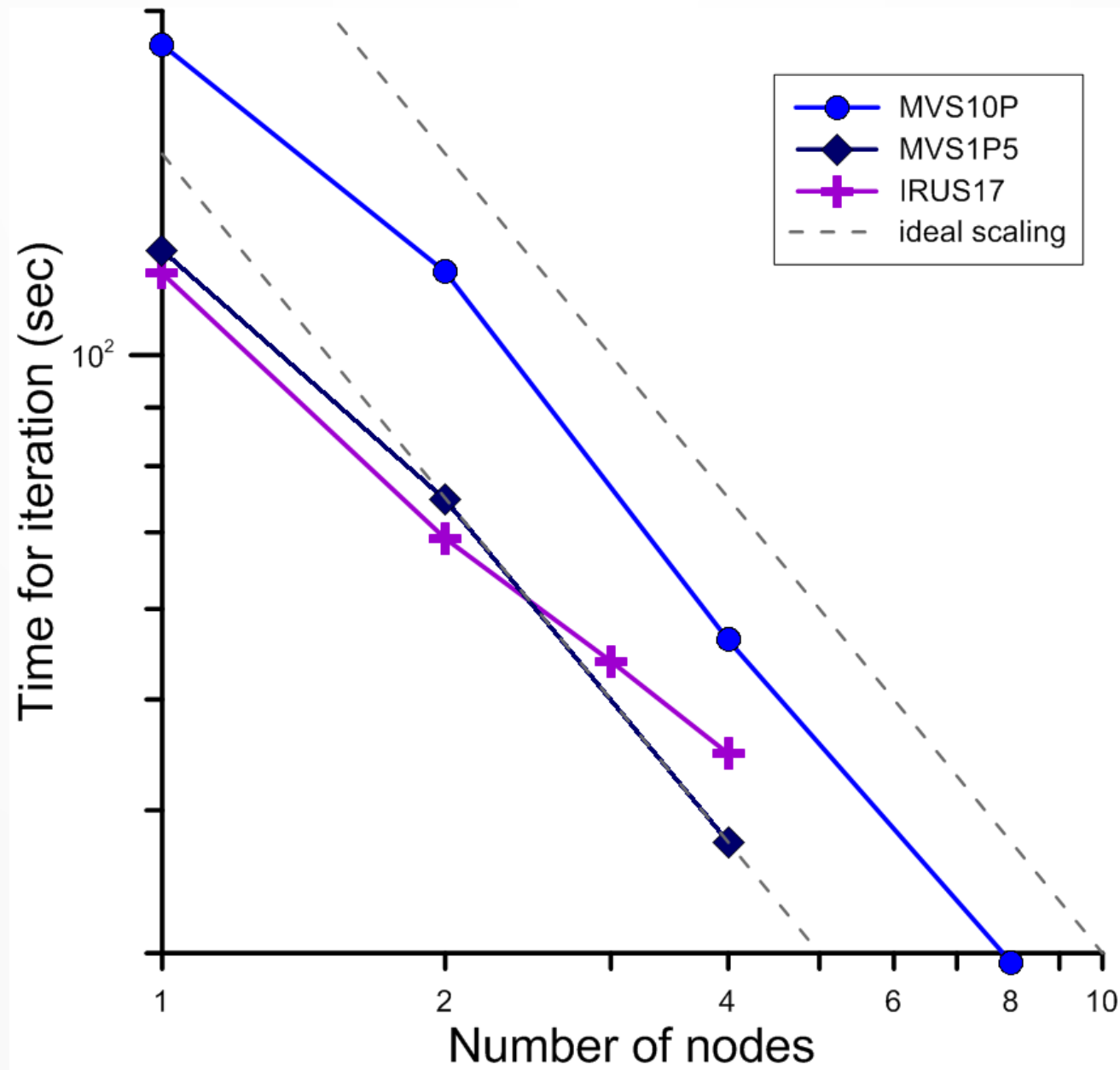
Масштабируемость VASP



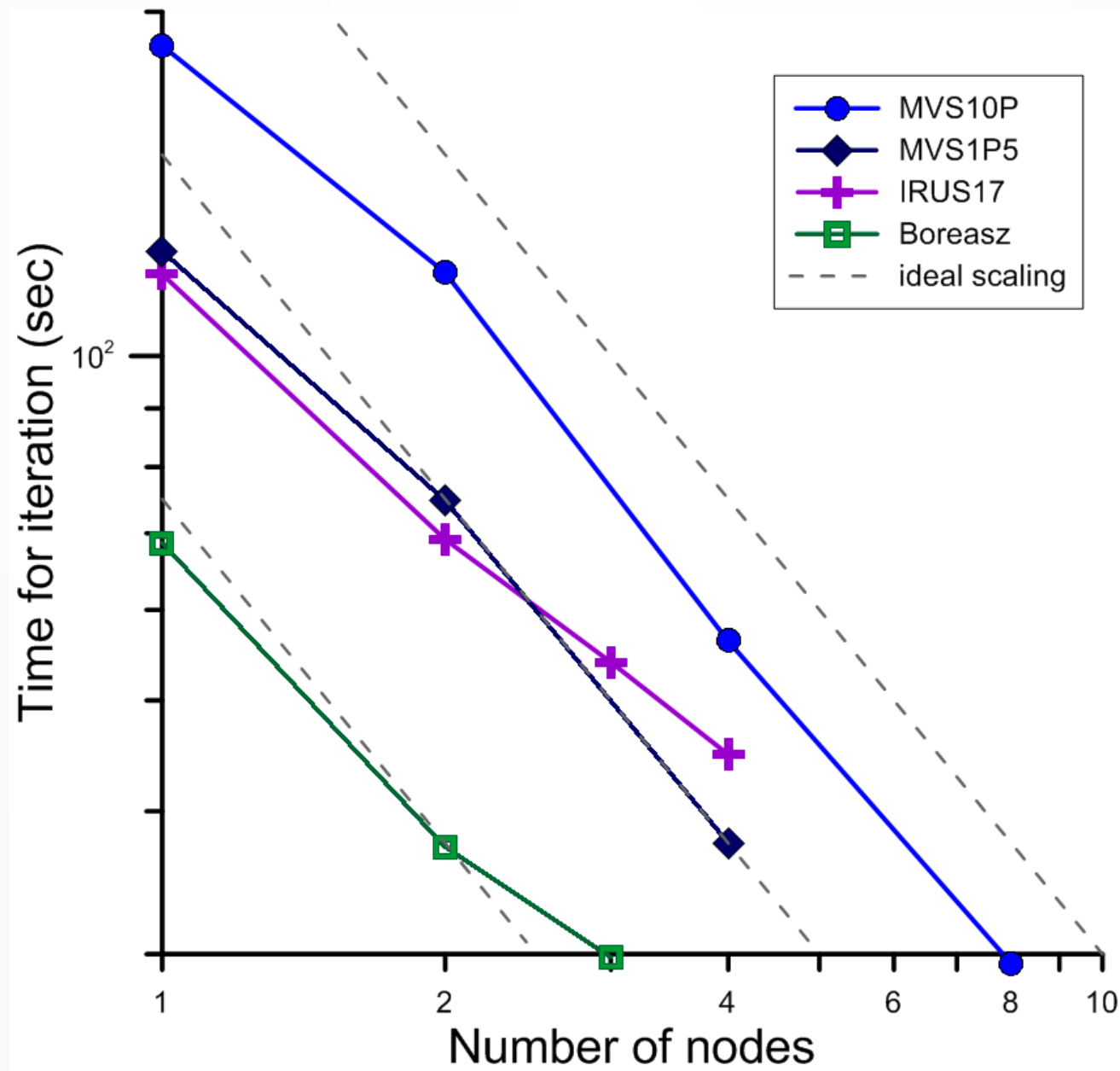
Масштабируемость VASP



Масштабируемость VASP



Масштабируемость VASP

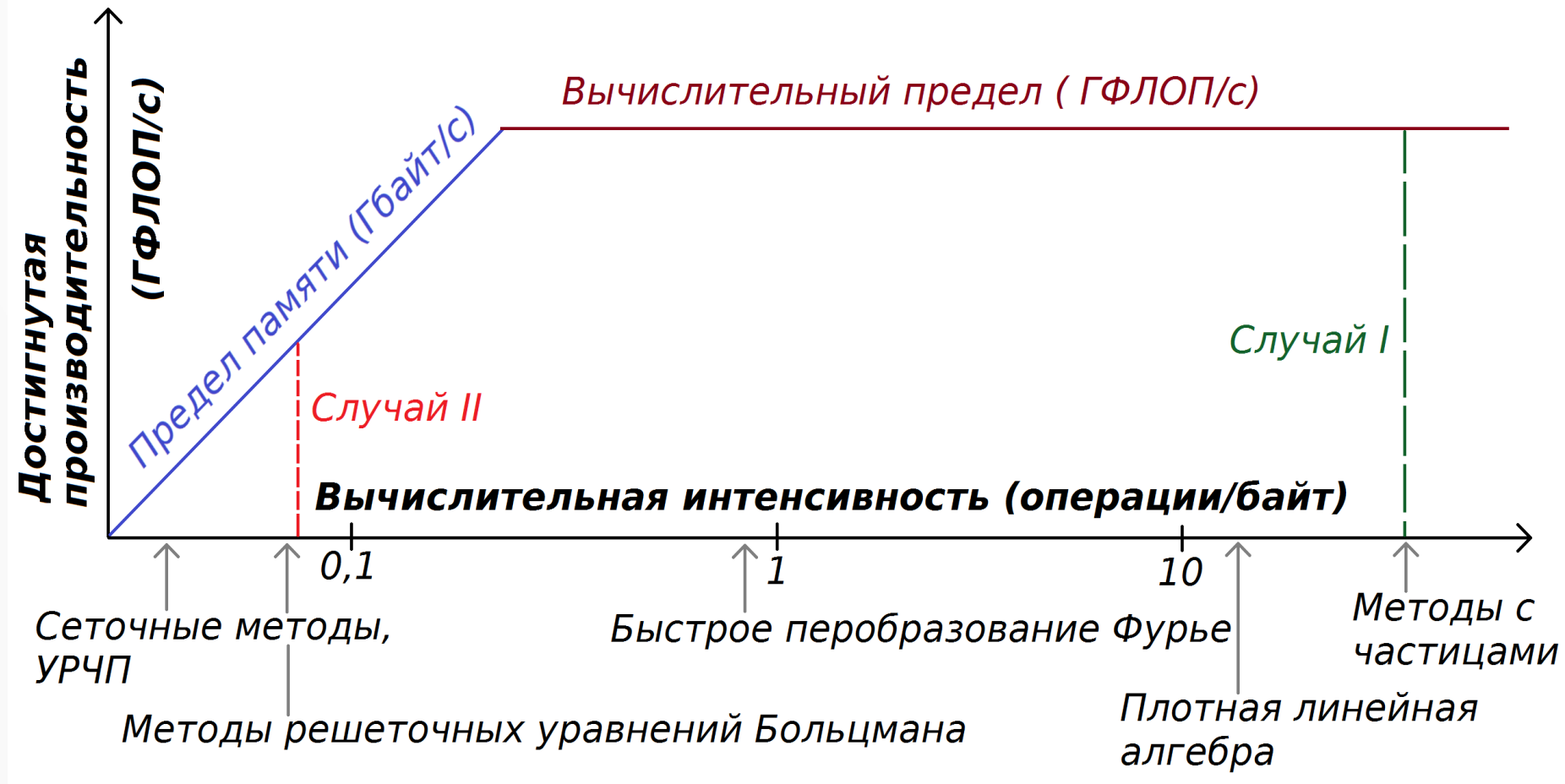


Разнообразие современных архитектур

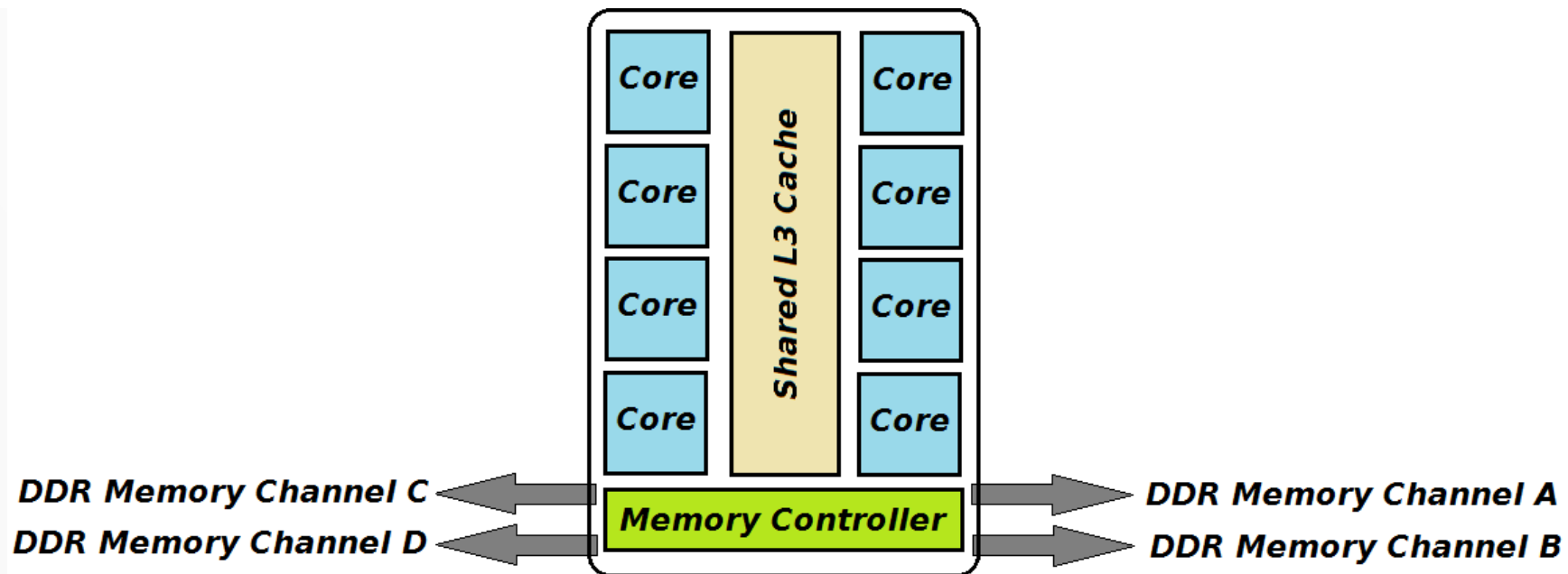
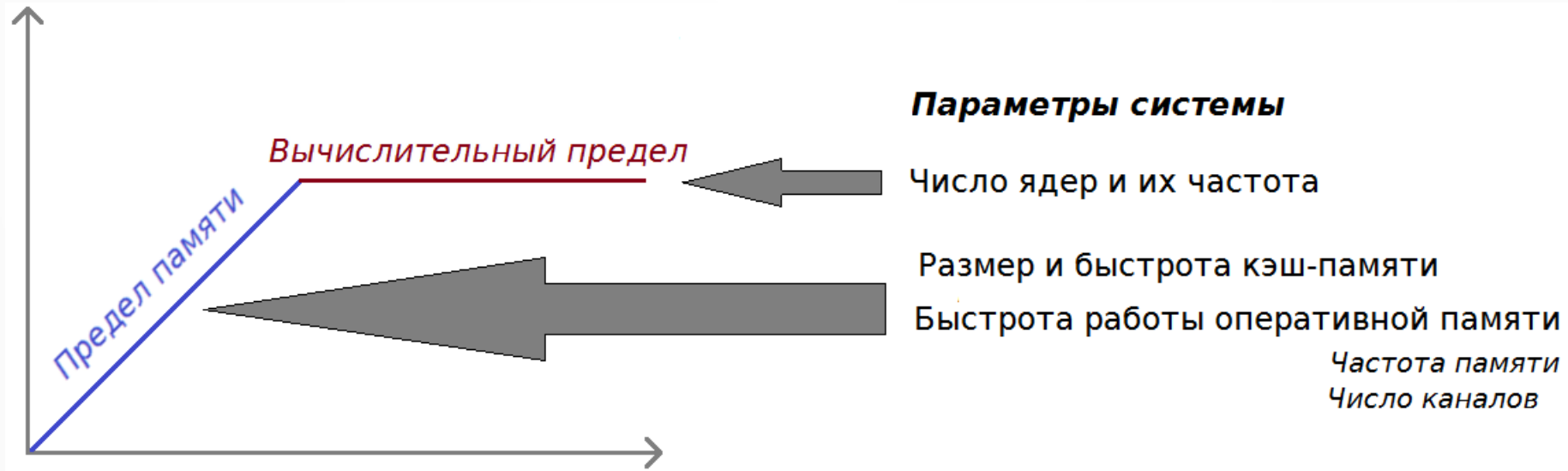
CPU type	N_{cores}	$N_{mem.ch.}$	LLC (Mb)	CPU_{freq} (GHz)	$DRAM_{freq}$ (MHz)
Single socket, Intel X99 chipset					
Xeon E5-2620v4	8	4	20	2.1	2133
Xeon E5-2660v4	14	4	35	2.0	2400
Single socket, AMD B350 chipset					
Ryzen 1800X	8	2	16	3.6	2400
Nvidia Jetson TX2 (2 Denver + 4 Cortex-A57 cores)					
Tegra "Parker"	2+4	4 (32-bit)	2+2	2.0	1866
Dual socket, Intel C602 chipset (the MVS10P cluster)					
Xeon E5-2690	8	4	20	2.9	1600
Dual socket, Intel C612 chipset (the MVS1P5 cluster)					
Xeon E5-2697v3	14	4	35	2.6	2133
Dual socket, Intel C612 chipset (the IRUS17 cluster)					
Xeon E5-2698v4	20	4	50	2.2	2400
Quad socket, IBM Power 775 (the Boreasz cluster [21])					
Power 7	8	4	32	3.83	1600

Оценка производительности моделью Roofline

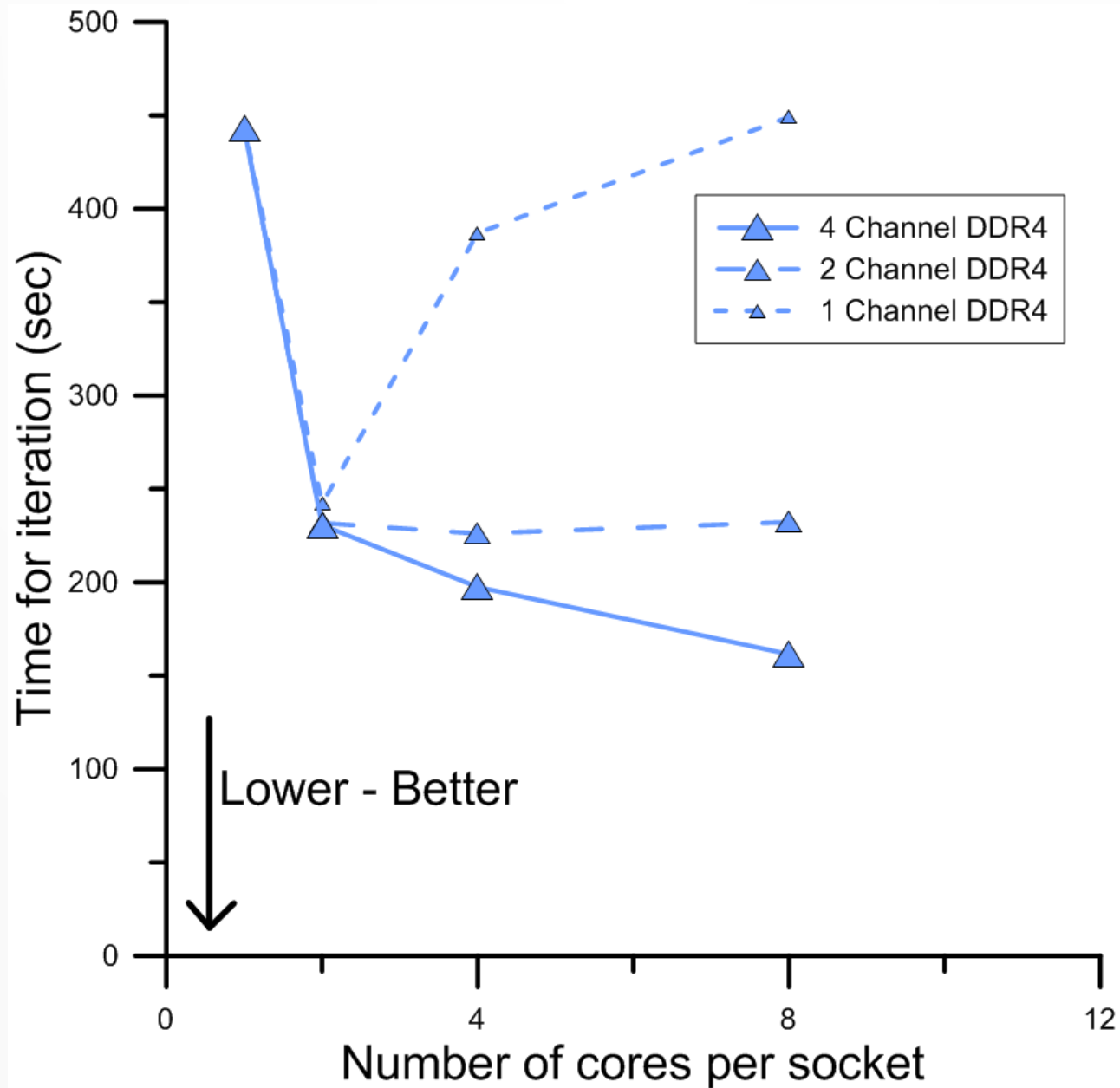
- Разные алгоритмы характеризуются разной **вычислительной интенсивностью** – отношением количества арифметических операций к объему переданных данных.
- Максимальная производительность разных алгоритмов ограничивается пиковой производительностью и пропускной способностью памяти.



Влияние архитектуры узла на производительность в VASP



Пример – влияние числа каналов памяти на скорость расчета (8core Xeon E5-2620v4)



Как учесть все параметры системы и оценить эффективность расчета?

- Разные системы ранжируем по параметру баланса между производительностью процессора и памяти

$$\frac{\text{Теоретическая пиковая производительность } (R_{\text{peak}}) \text{ [FLOP / sec]}}{\text{Пропускная способность памяти [Megabytes / sec]}} = \text{Баланс [FLOP / byte]}$$

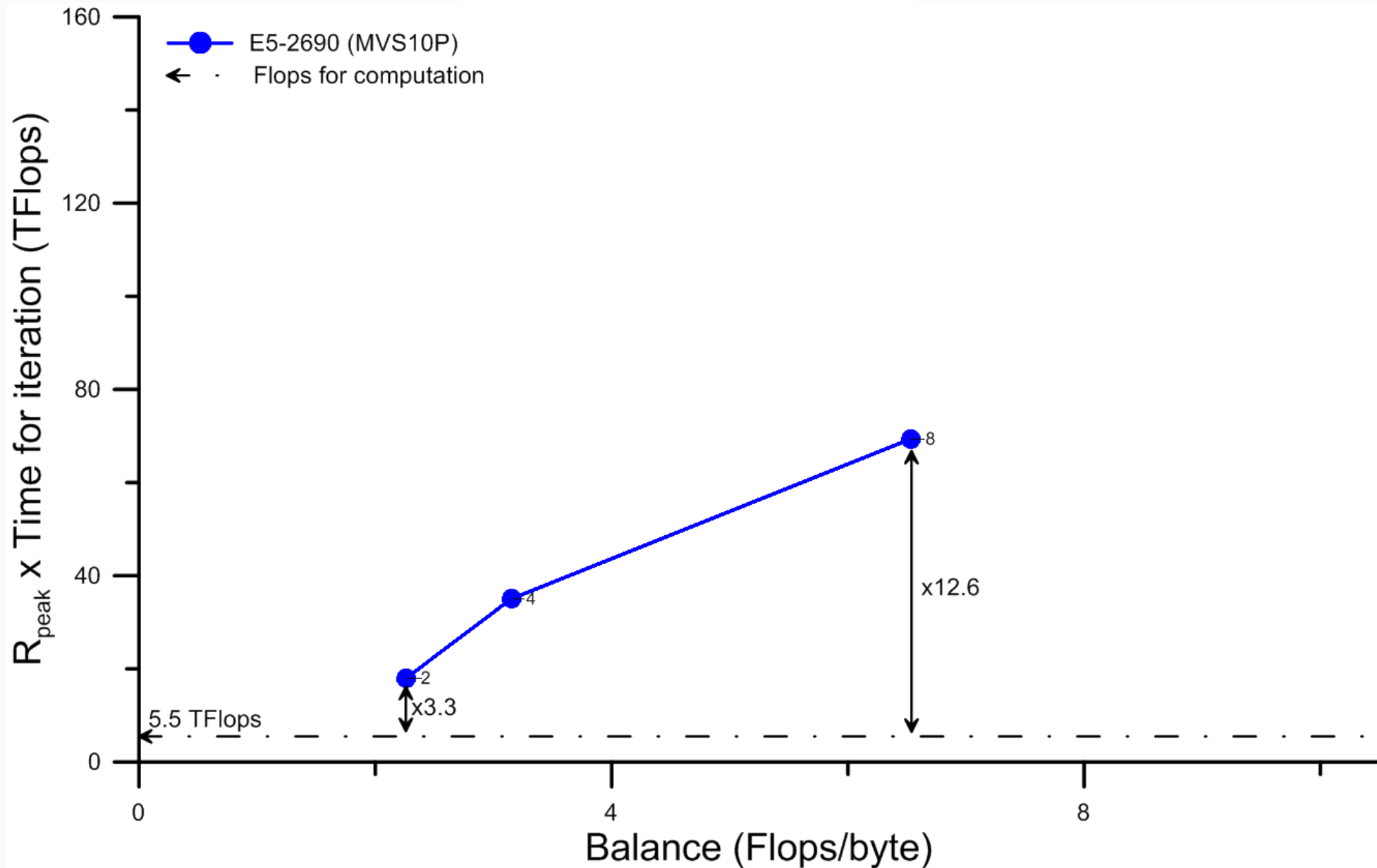
- Не**эффективность расчета на конкретной системе оцениваем как

$$\text{Теоретическая пиковая производительность } (R_{\text{peak}}) \text{ [FLOP / sec]} \times \text{время итерации VASP } (T_{\text{iter}}) \text{ [sec]} = \text{Число операций, которое могло Бы быть выполнено [FLOP]}$$

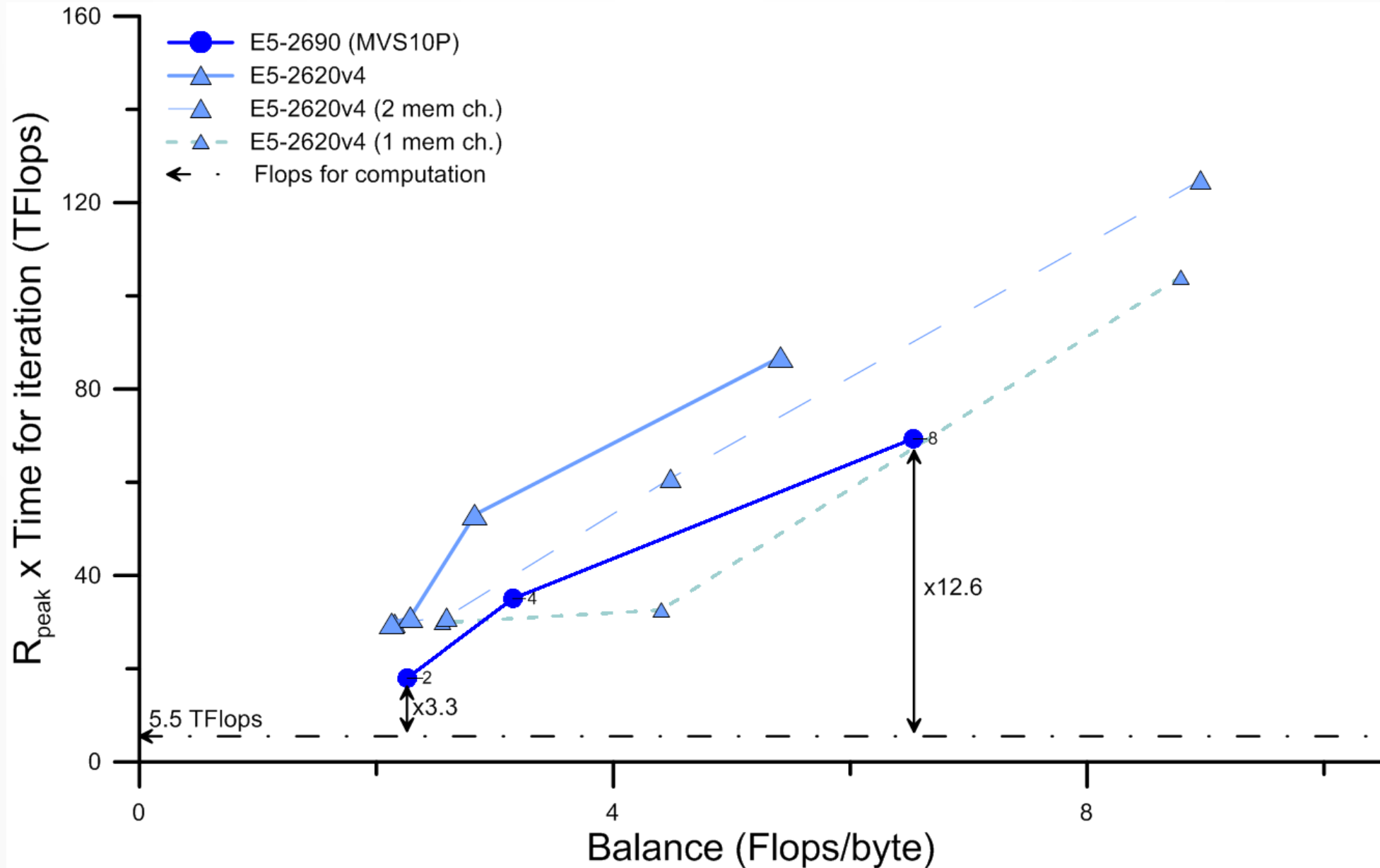
Как учесть все параметры системы и оценить эффективность расчета-2

- Используя аппаратные счетчики платформ Интел, было измерено **реально выполненное** количество операций с плавающей запятой – 5.5 Терафлоп за итерацию
- Сравним количество выполненных операций, и количество операций, которых можно было бы выполнить.
 - Чем выше это соотношение, тем более эффективно работает подсистема памяти

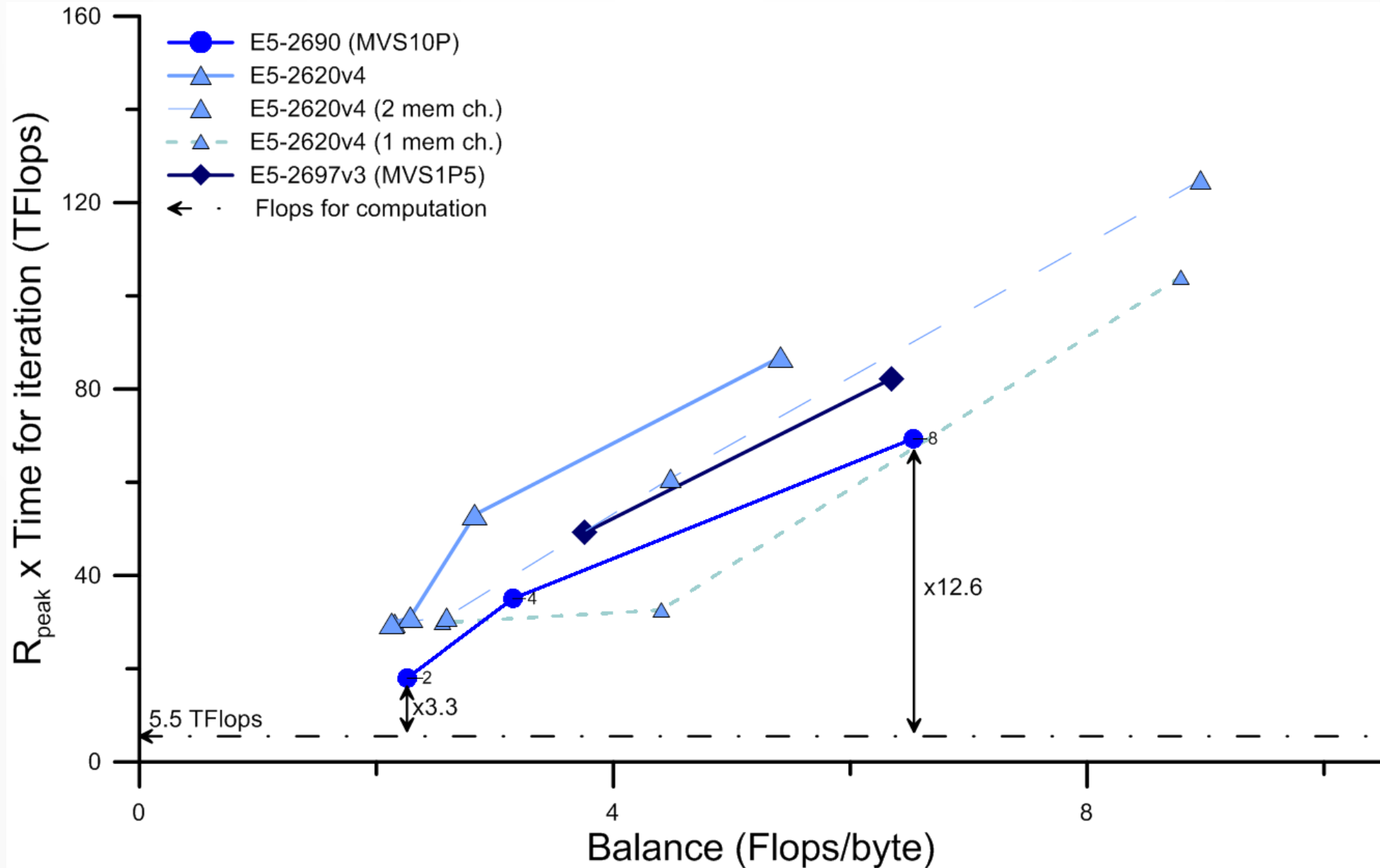
Влияние баланса на скорость расчета



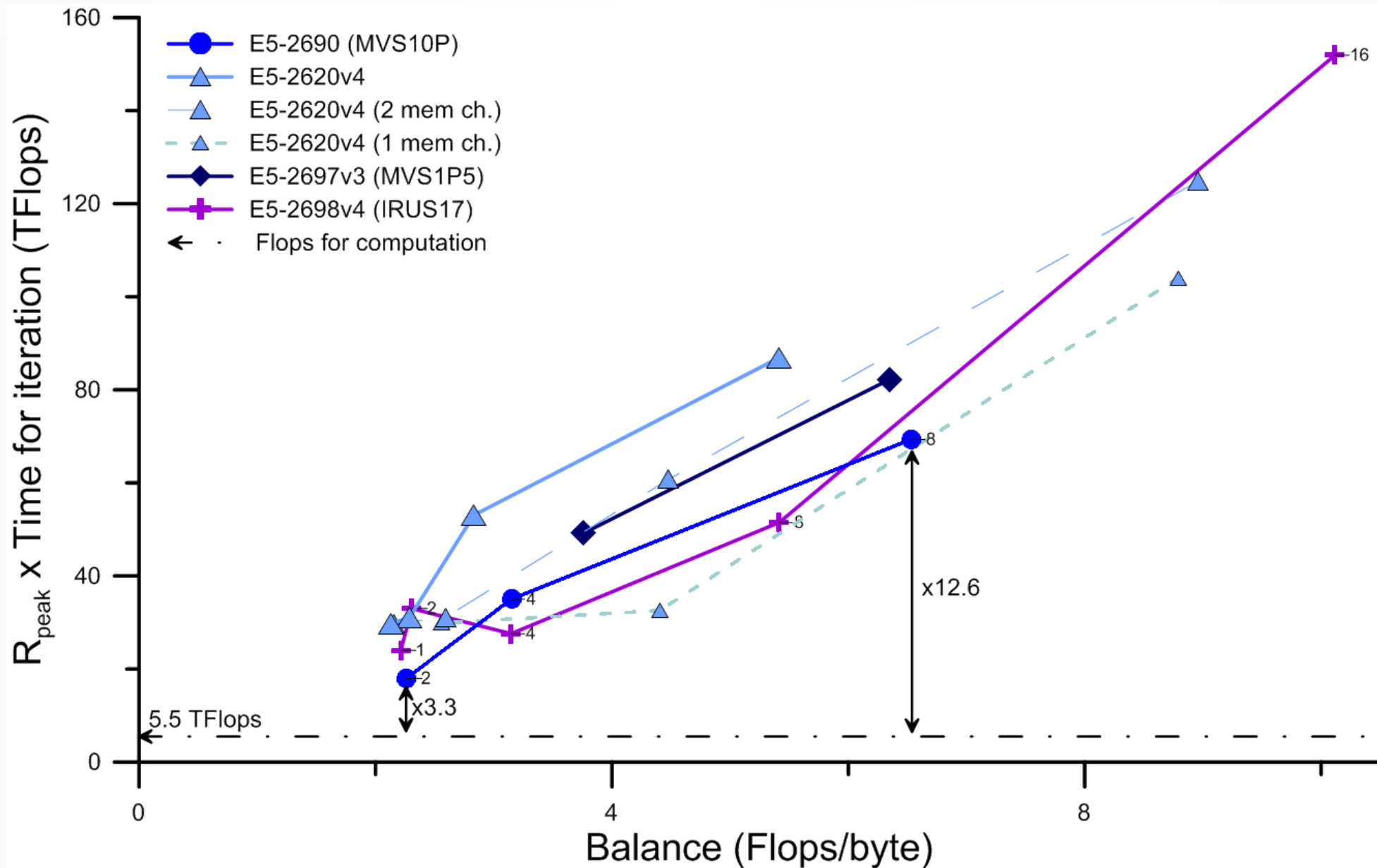
Влияние баланса на скорость расчета



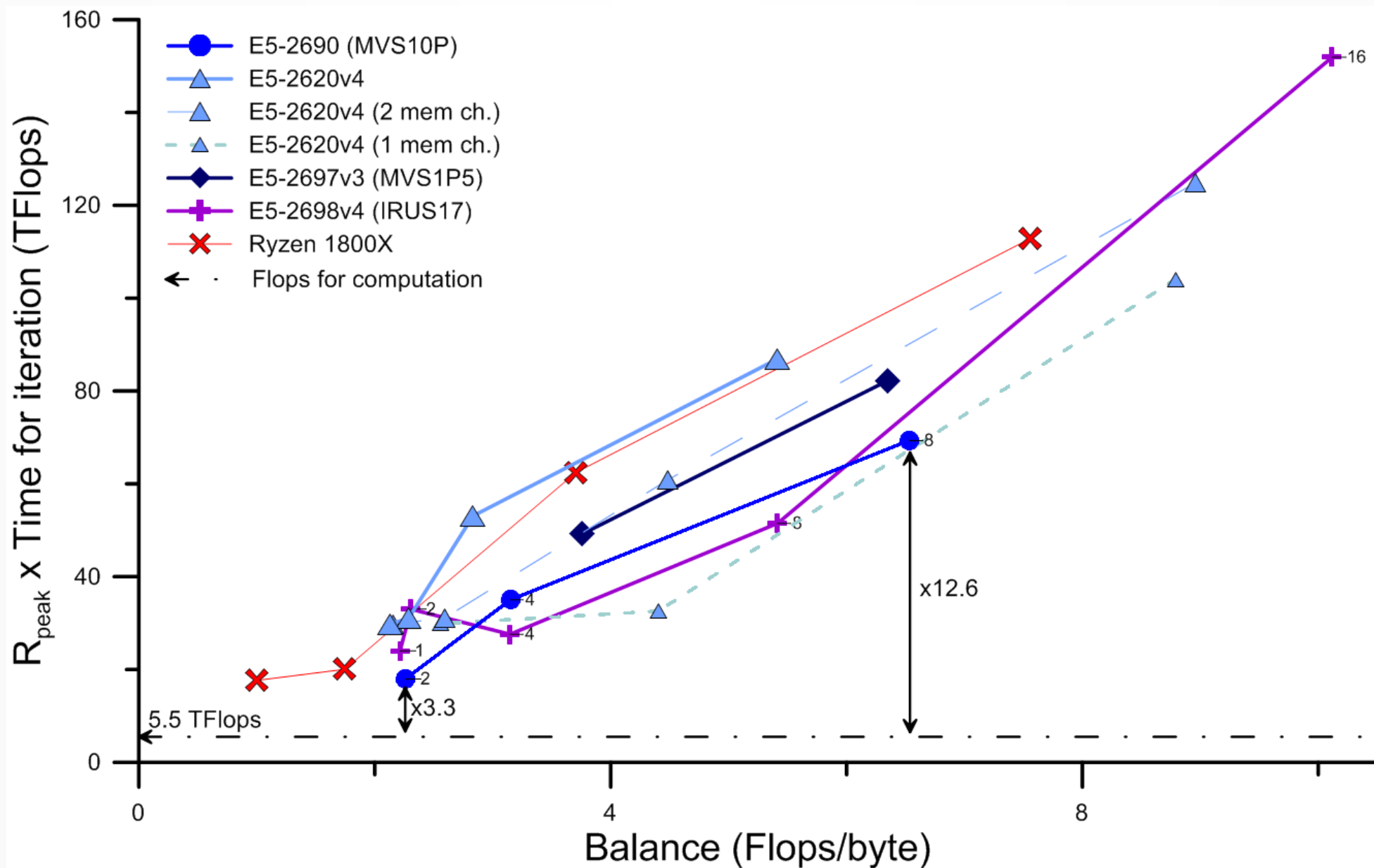
Влияние баланса на скорость расчета



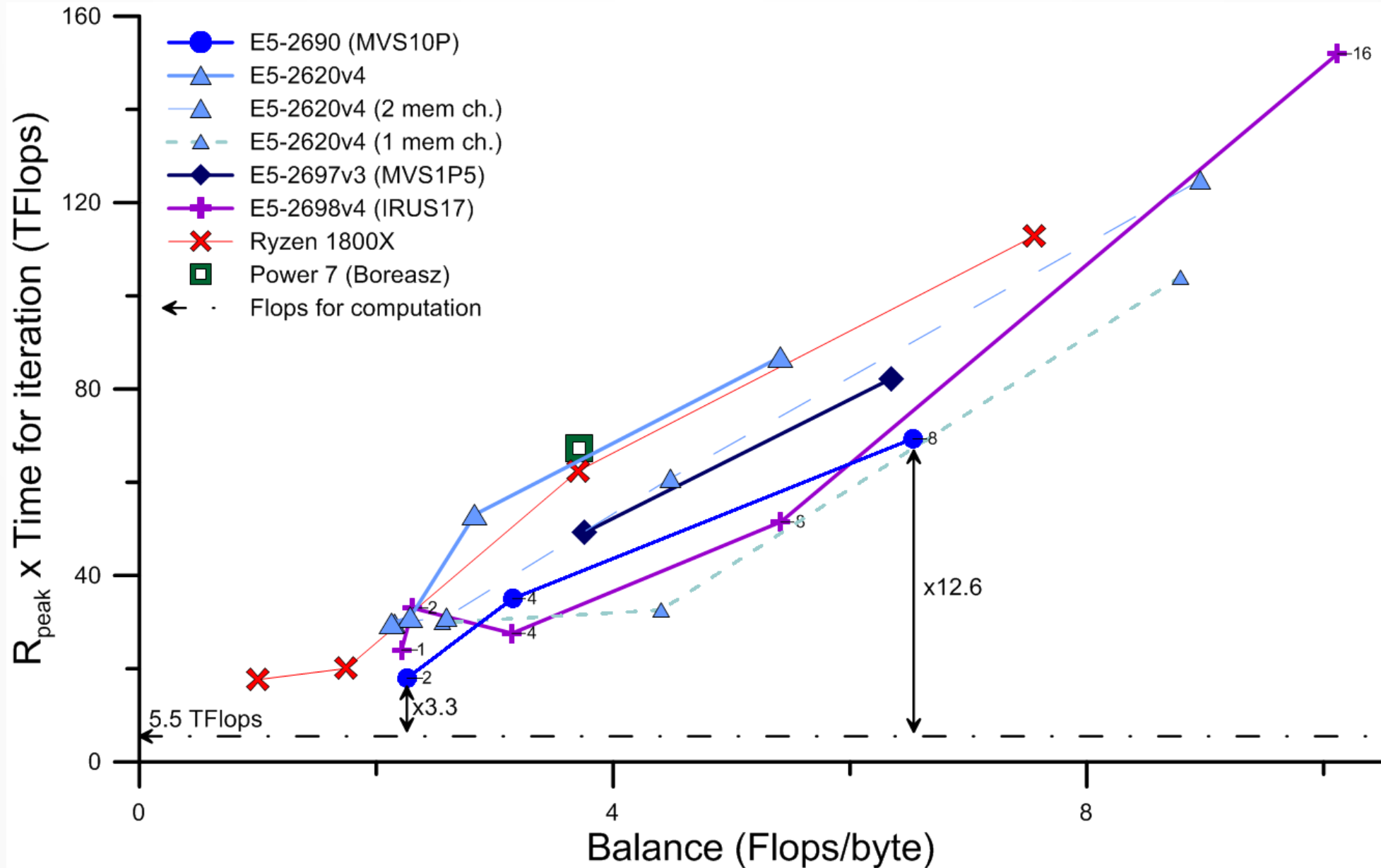
Влияние баланса на скорость расчета



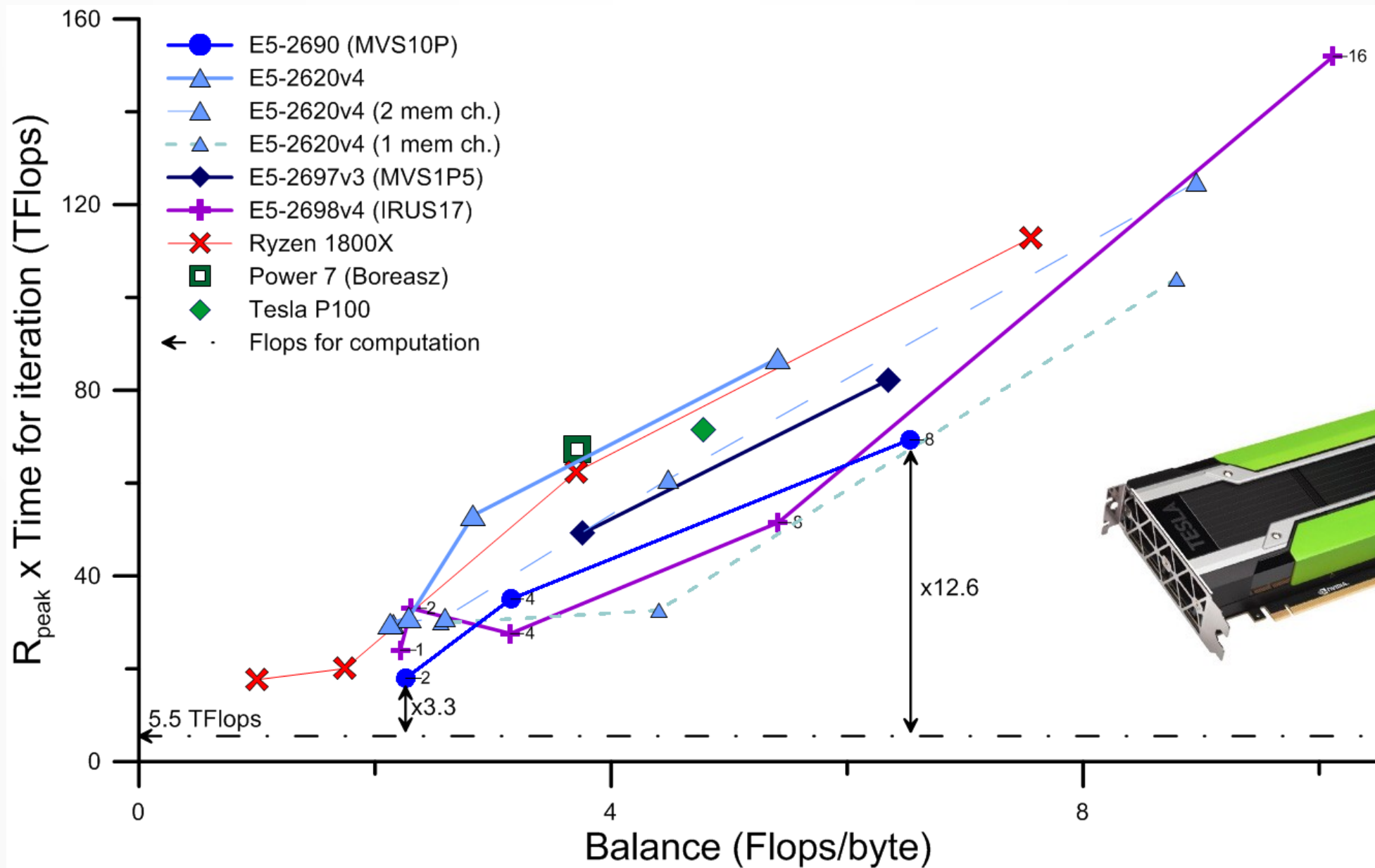
Влияние баланса на скорость расчета



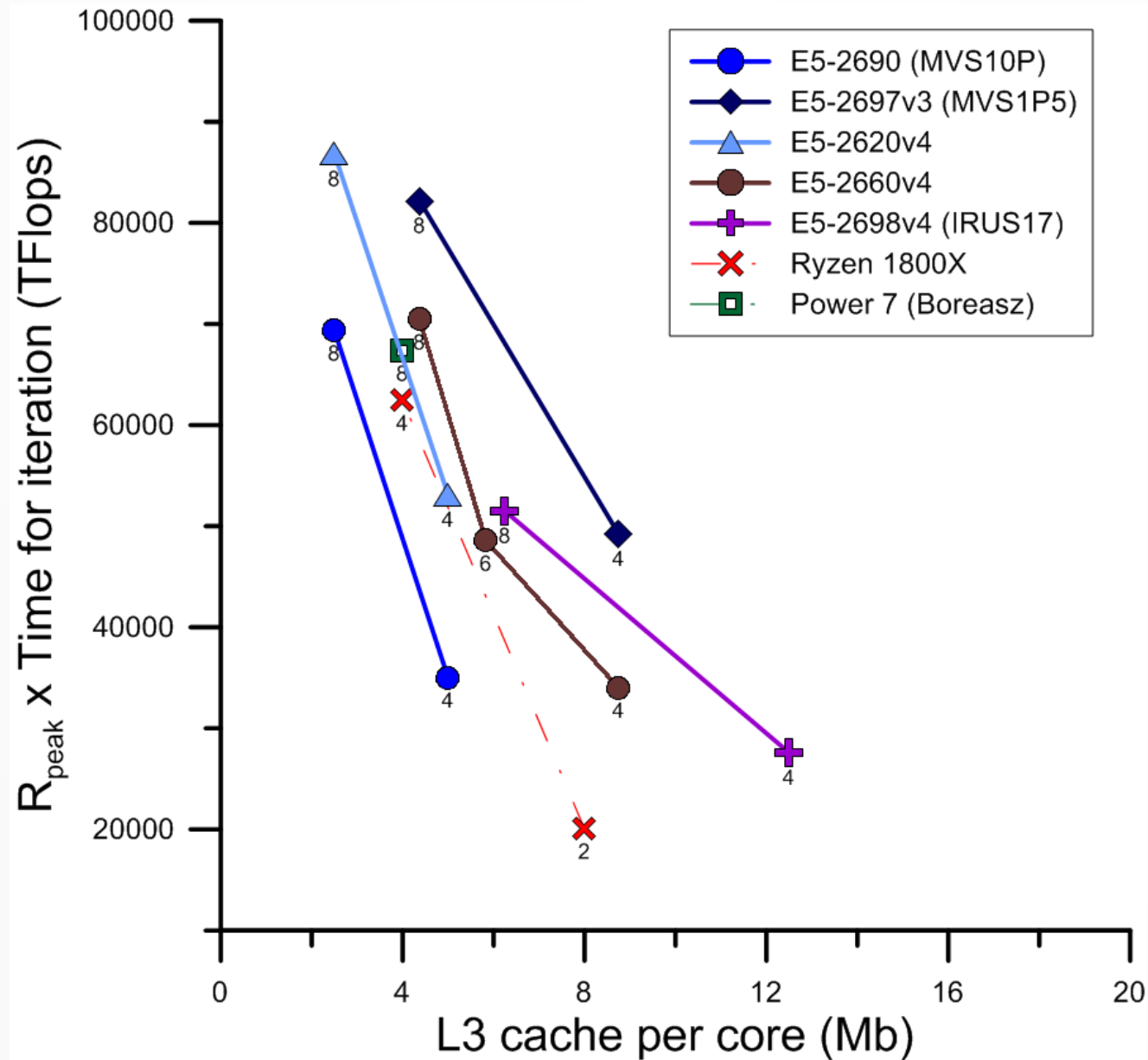
Влияние баланса на скорость расчета



Влияние баланса на скорость расчета



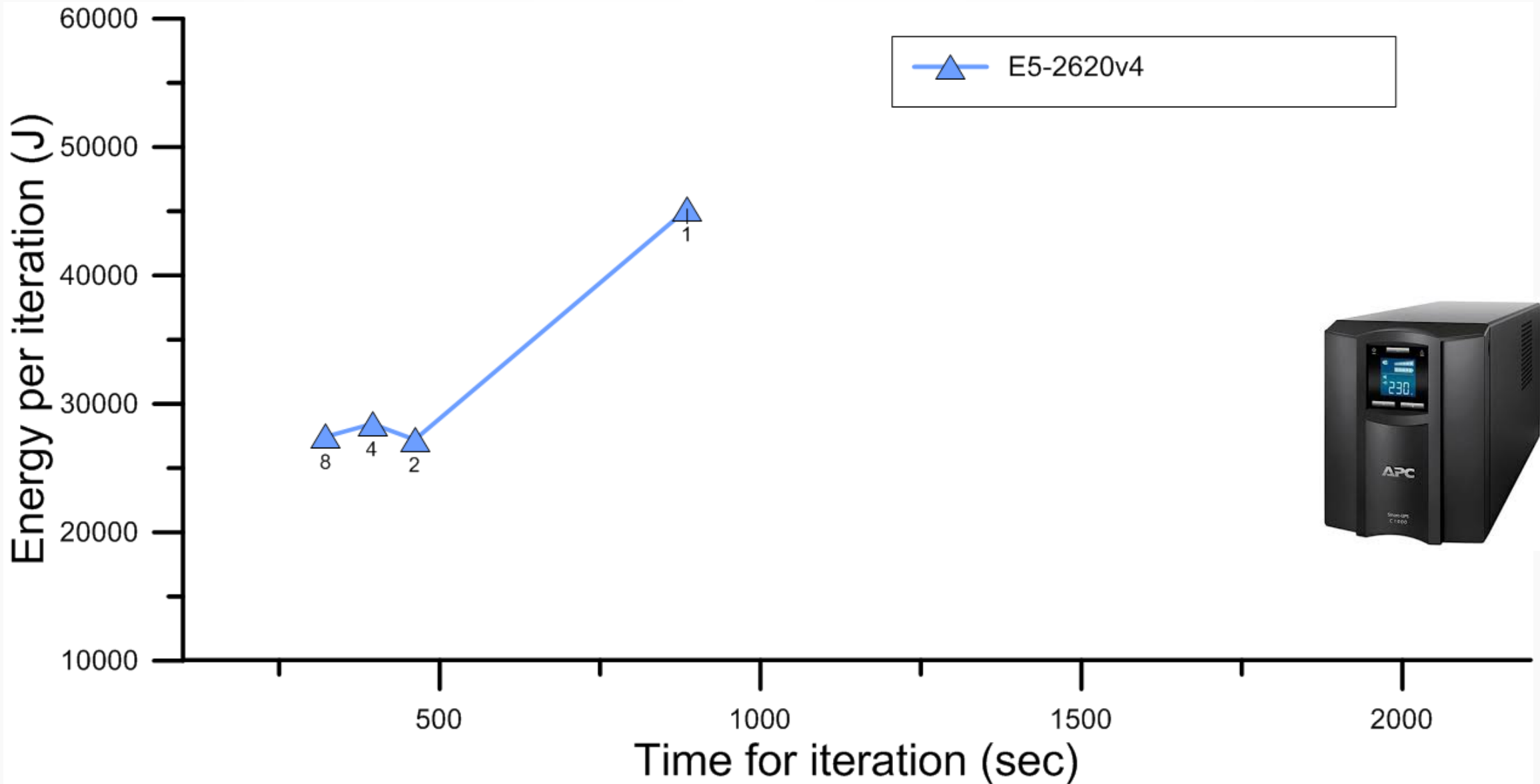
Пример 2 – влияние кэша на скорость расчета



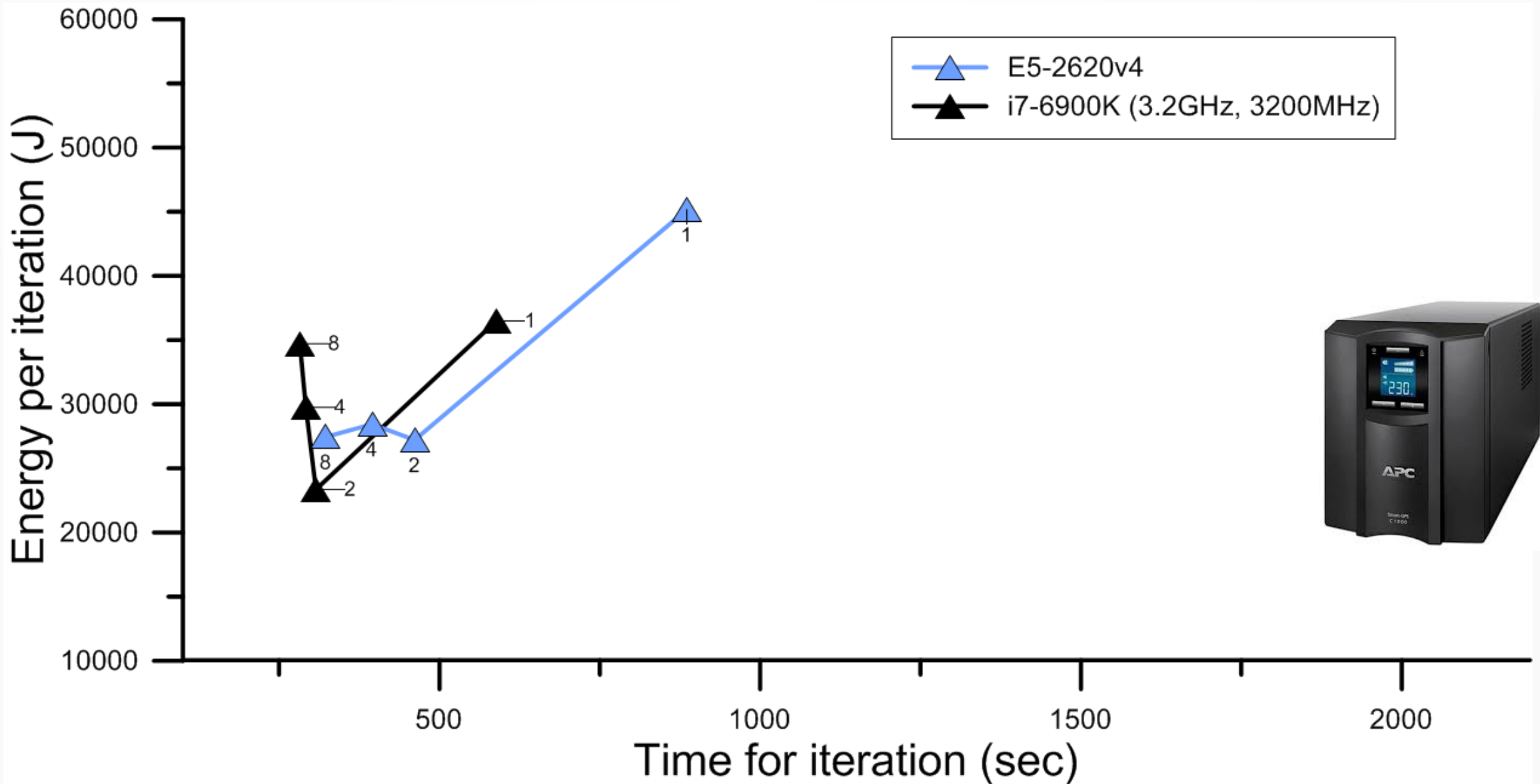
Энергопотребление

- Большинство работ посвящено потреблению одного компонента системы – центрального процессора или памяти.
- В этой работе мы аппаратно измеряли потребление всего вычислительного устройства.

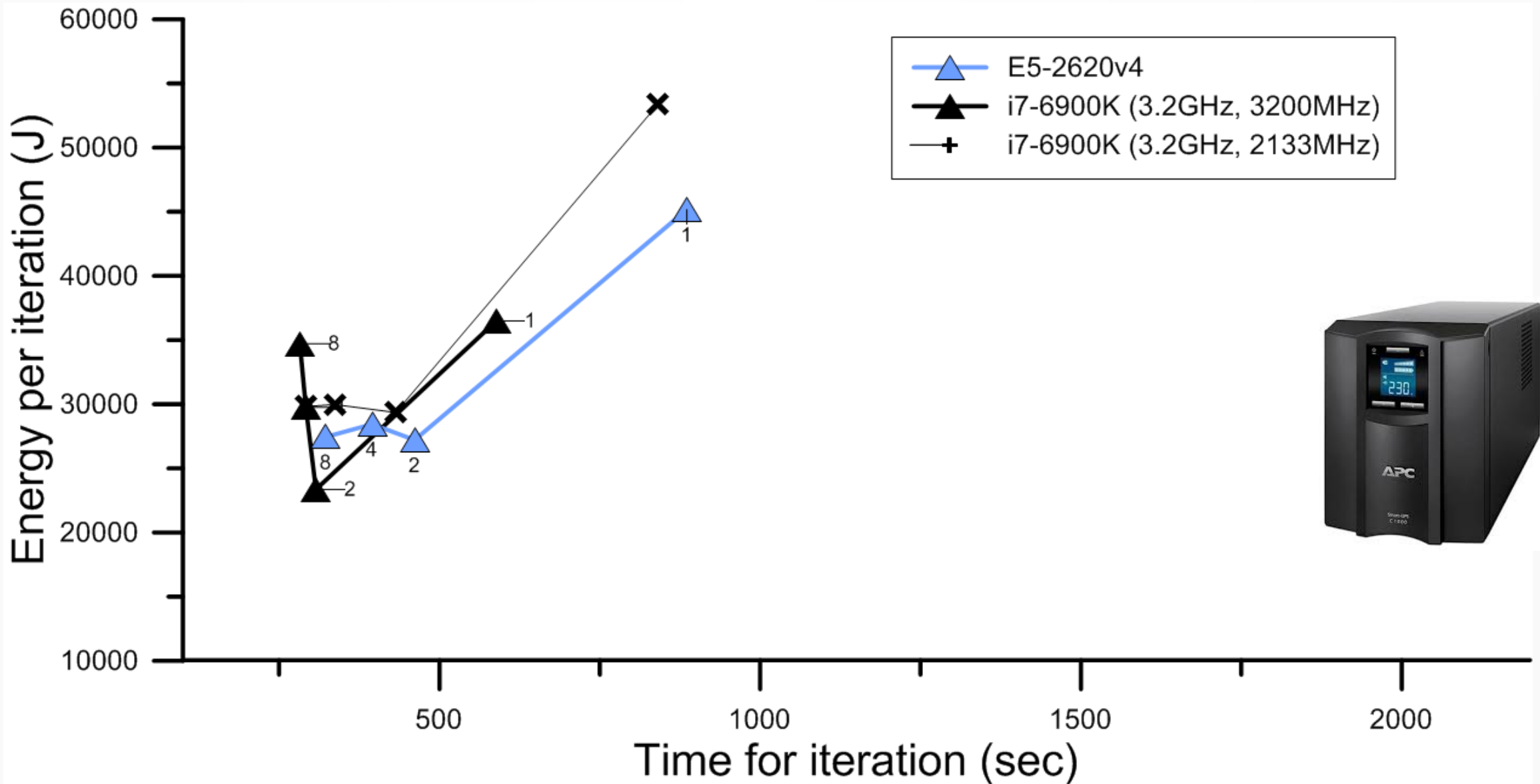
Энергопотребление разных платформ



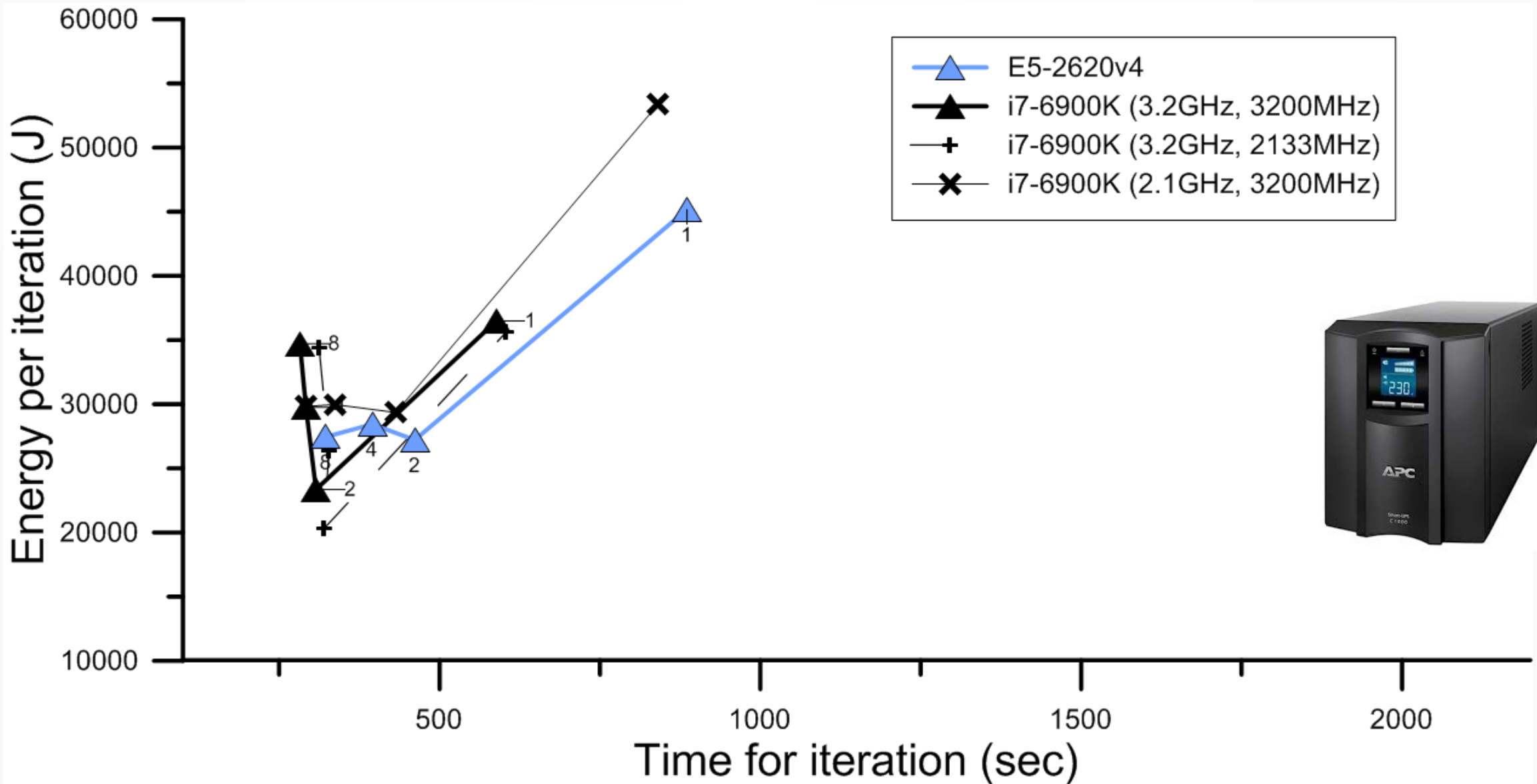
Энергопотребление разных платформ



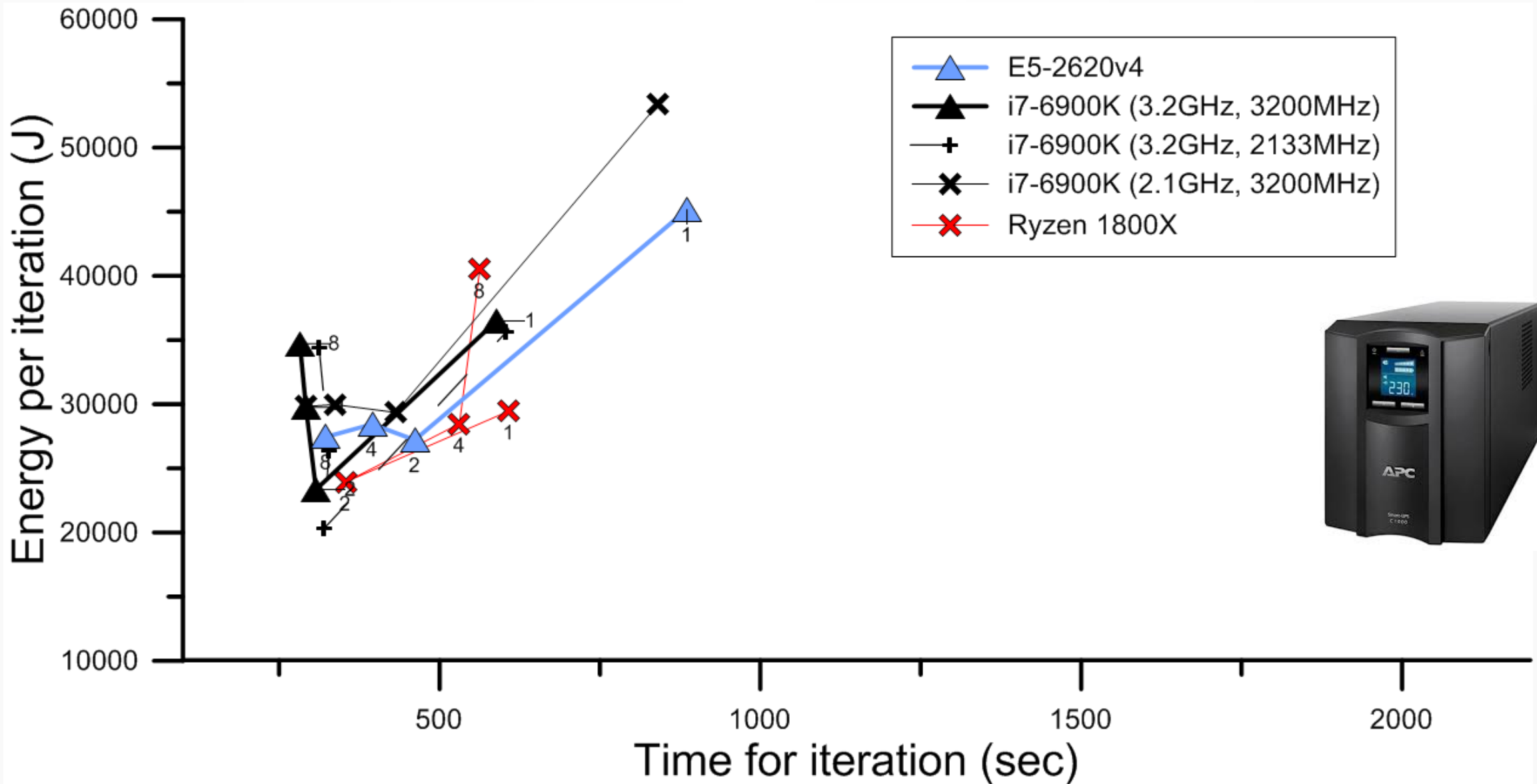
Энергопотребление разных платформ



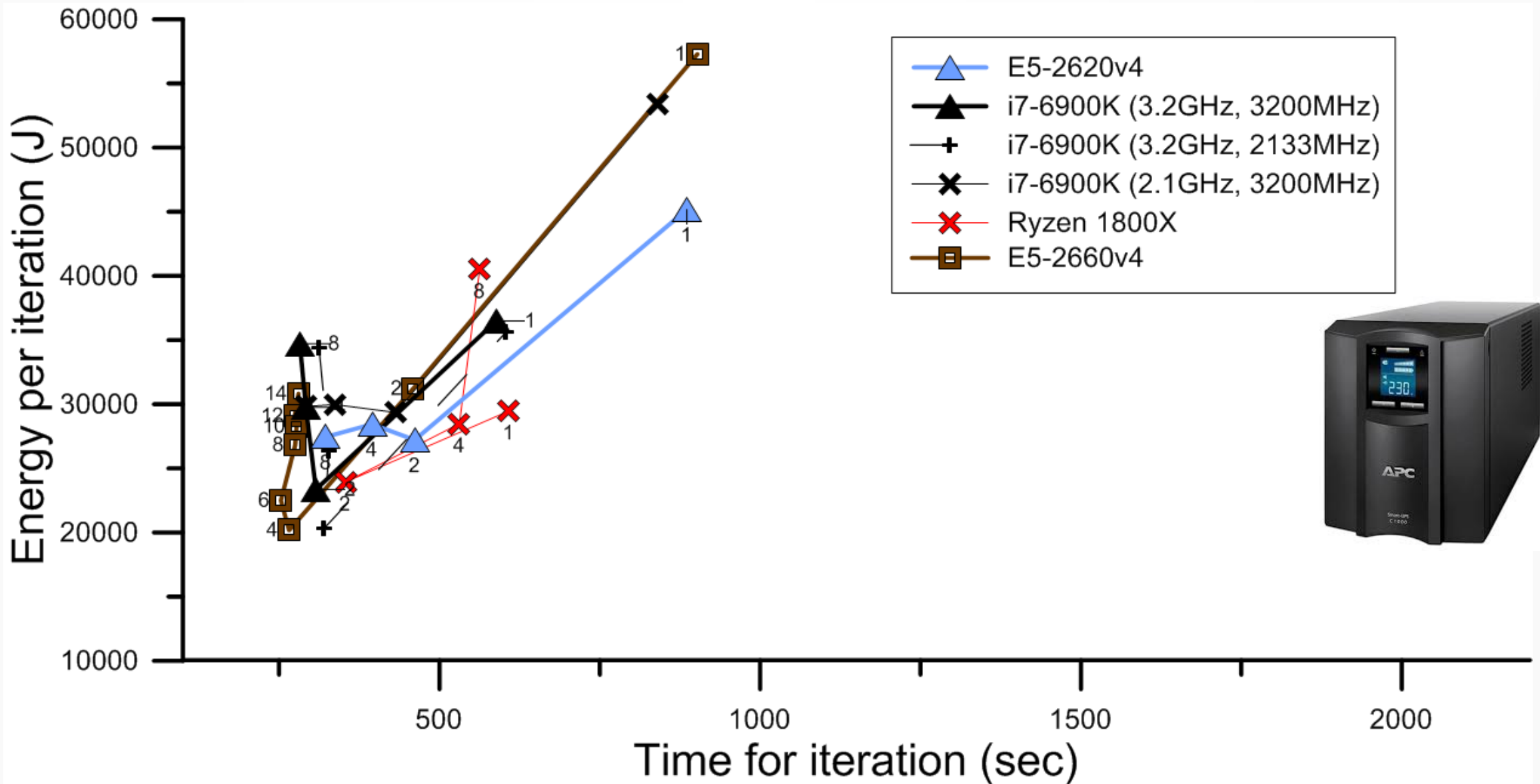
Энергопотребление разных платформ



Энергопотребление разных платформ

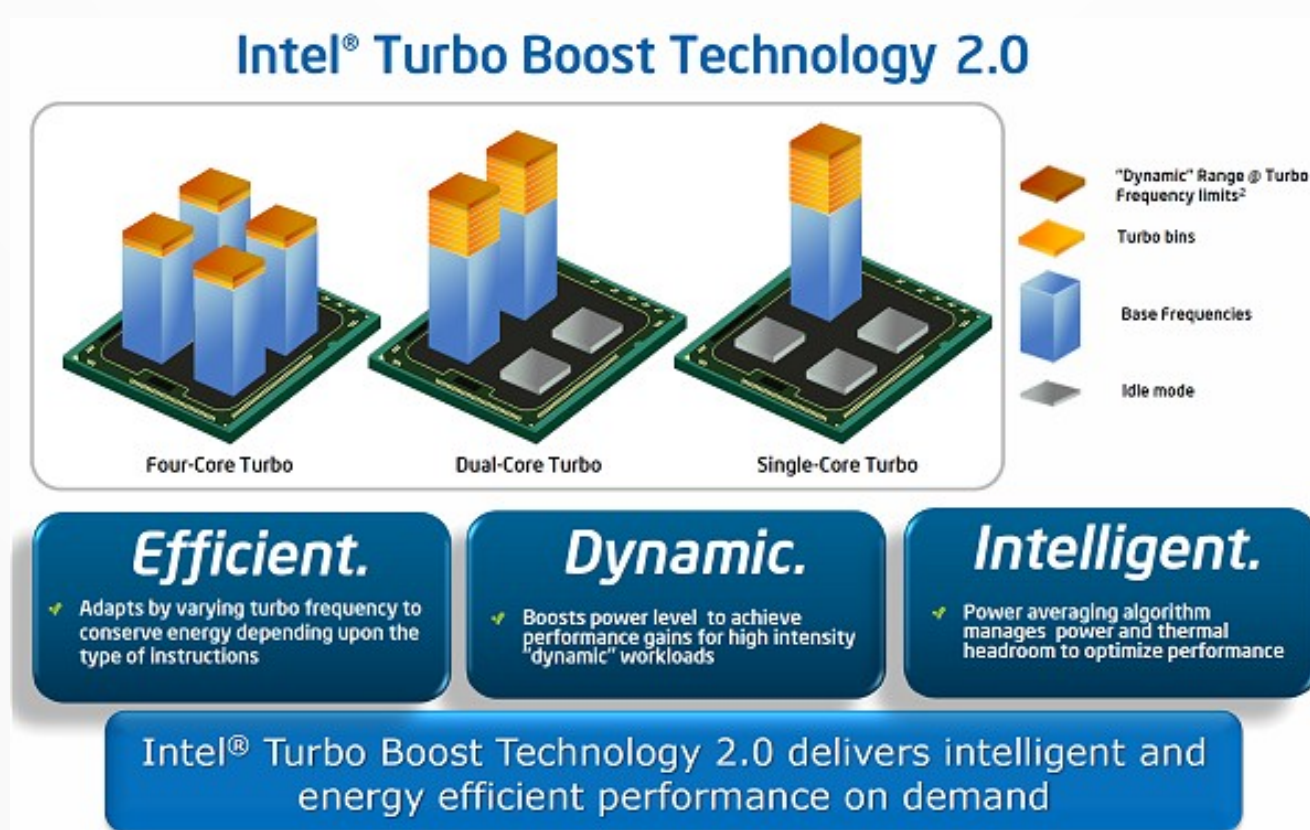


Энергопотребление разных платформ

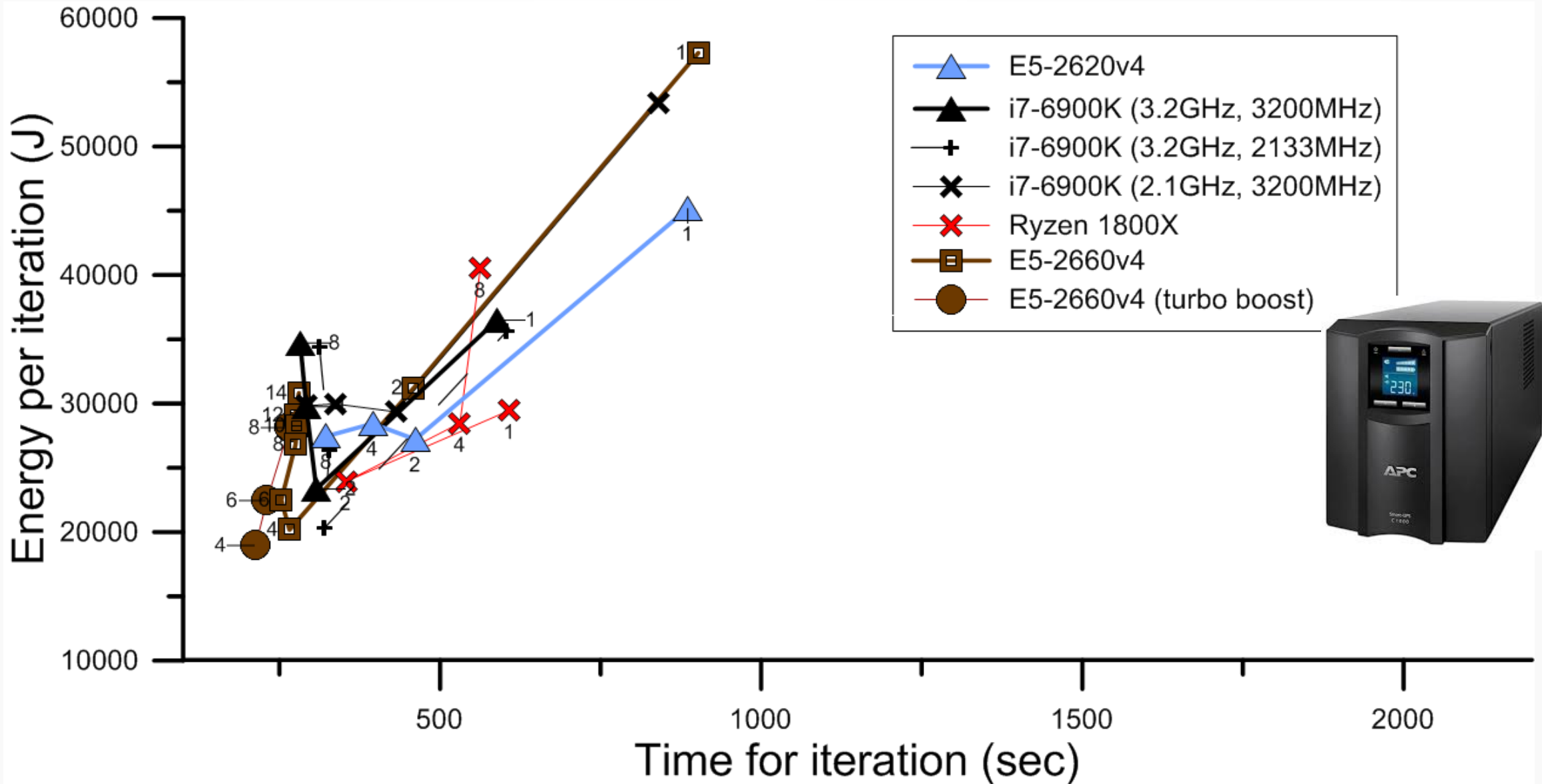


Технология Turbo Boost

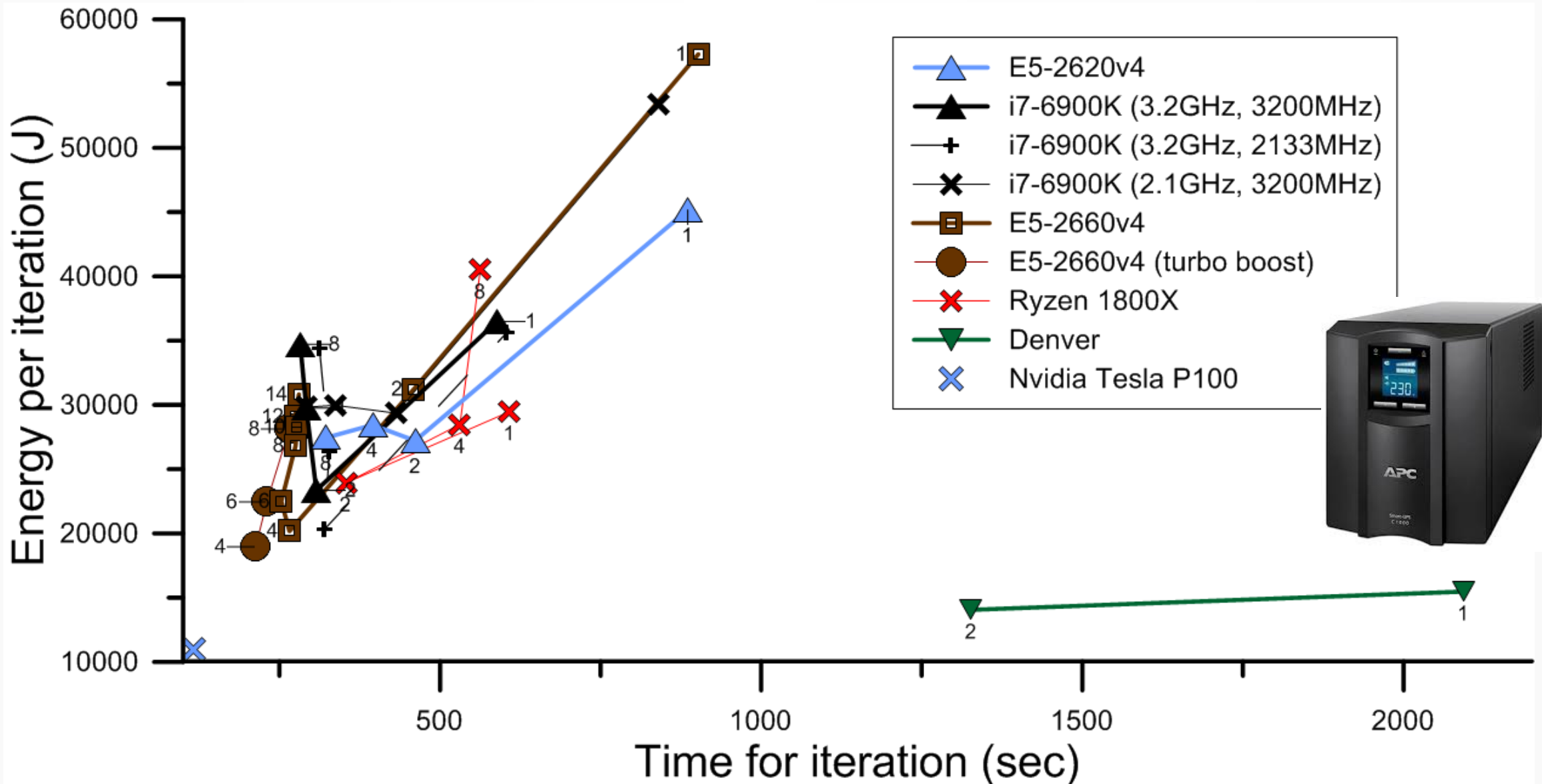
- Технология Turbo Boost позволяет автоматически увеличивать тактовую частоту процессора выше номинальной, если позволяет система охлаждения
- Чем меньше задействовано ядер – тем больше авторазгон



Энергопотребление разных платформ



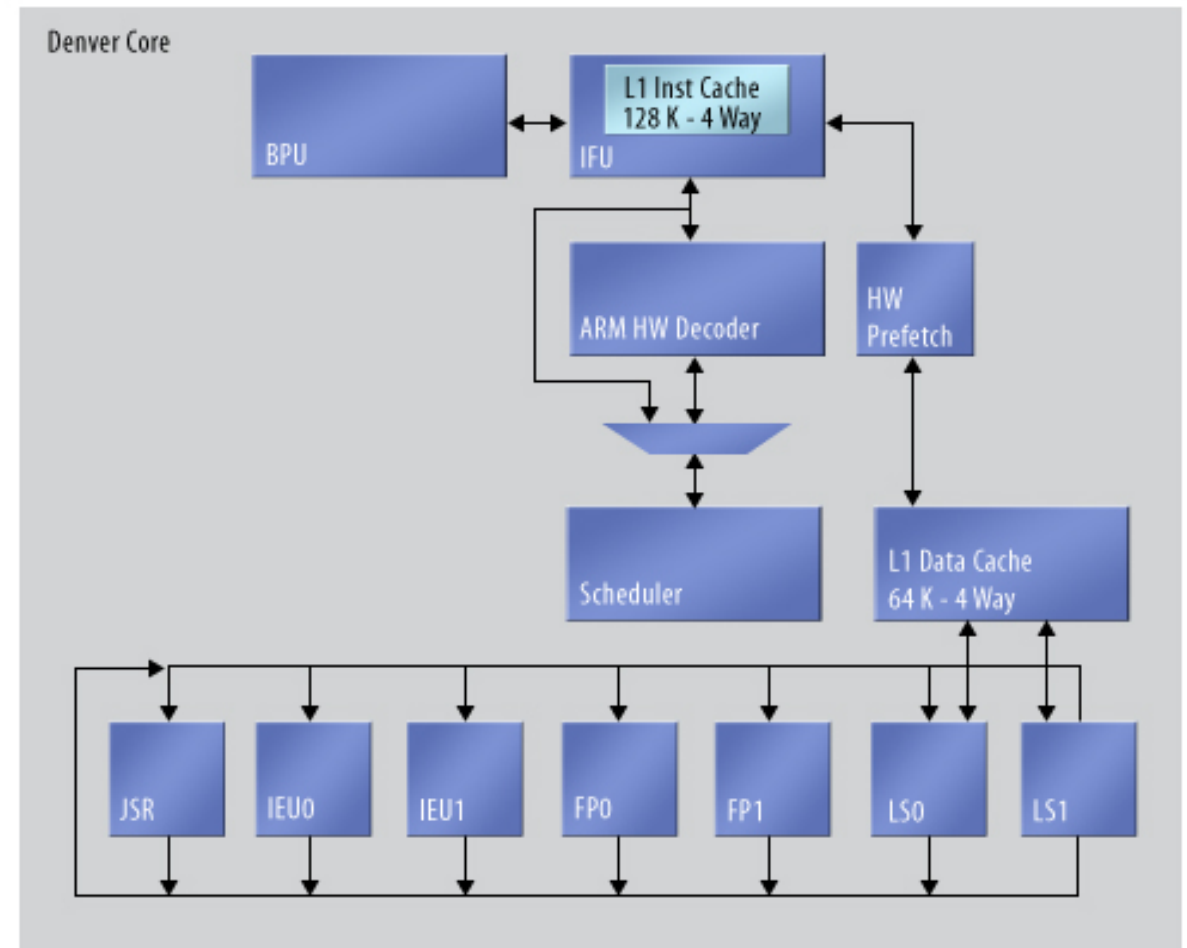
Энергопотребление разных платформ



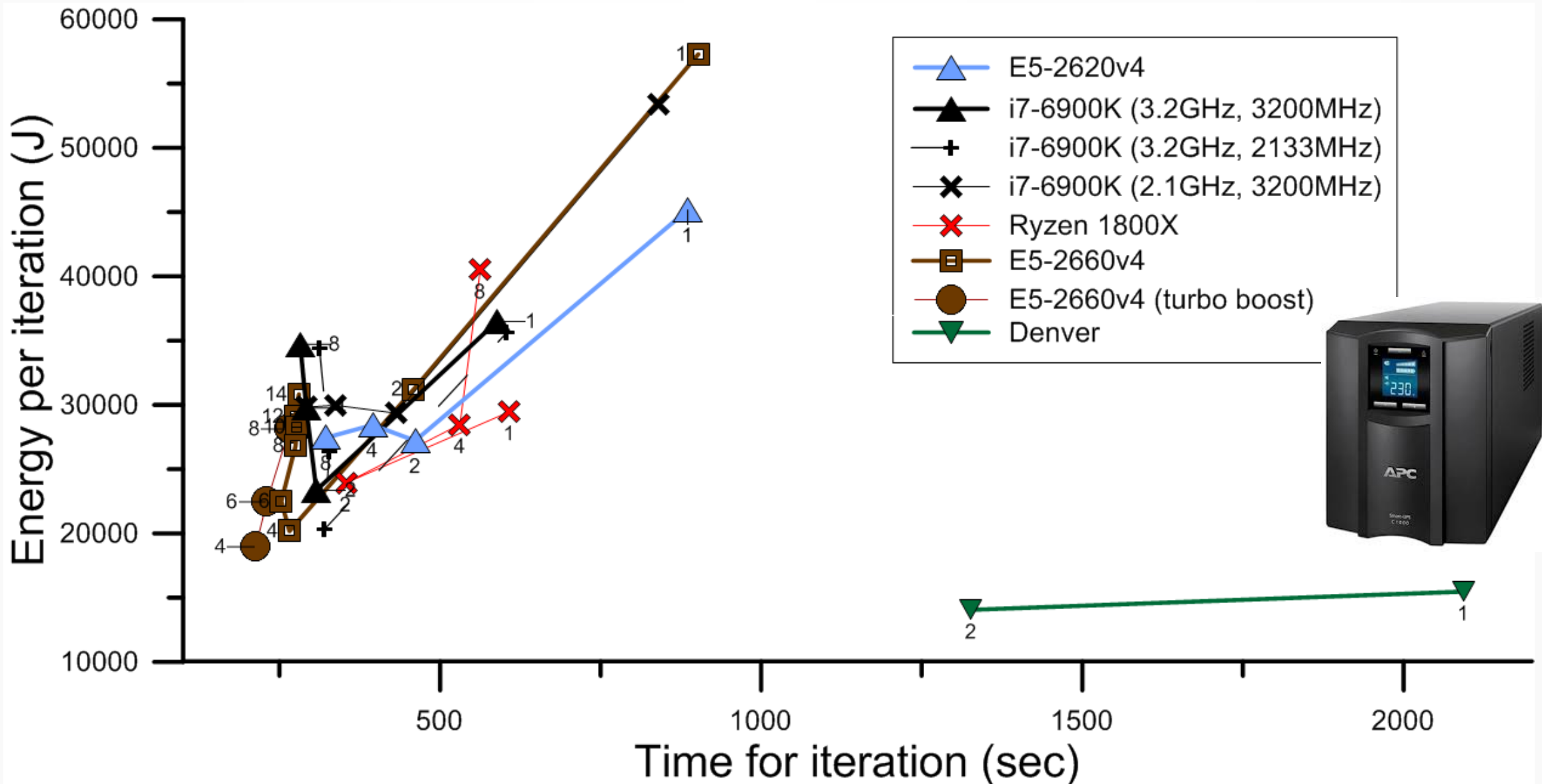


Nvidia Denver

- Nvidia представила свою архитектуру CPU под наименованием Denver
 - Совместима с набором инструкций ARM-32/64bit
 - Отсутствует out-of-order execution, присутствует бинарный рекомпилятор для оптимизации выполнения кода



Энергопотребление разных платформ



Выводы для VASP

- 1. На одной тестовой задаче проведено сравнение 8 типов процессоров различных производителей (Intel, AMD, IBM и Nvidia).
- 2. Для Intel Xeon Broadwell проведено варьирование частот процессора и памяти, а также числа каналов DRAM.
- 3. Установлено наличие универсального тренда $R_{\text{peak}} * \tau_{\text{iter}}$ от значения баланса Флоп/Байт.
- 4. Для расчетов VASP оптимальным является использование 1-2 ядер на один канал памяти. Больше число ядер не приводит к ускорению, но повышает потребление энергии.
- 5. Существенный выигрыш в эффективности дает большее значение L3 кэша на ядро и возможность разгона ядер за счет технологии TurboBoost.

1. Stegailov V., Vechev V. Efficiency analysis of Intel and AMD x86_64 architectures for ab initio calculations: a case study of VASP // Communications in Computer and Information Science, Springer. **Принято в труды конференции RuSCD 2017.**

2. Stegailov V., Vechev V. Efficiency analysis of Intel, AMD and Nvidia 64-bit hardware for memory-bound problems: a case study of ab initio calculations with VASP // Lecture Notes in Computer Science, Springer. **Принято в труды конференции РРАМ 2017.**