

IBM POWER

Алексей Бойко
IBM Power Systems Product Manager
Alexey.Boyko@ru.ibm.com



“
**Good Design is
Good Business.**”

Thomas J. Watson, Jr.
CEO, IBM



SYSTEM 360

Немного о процессорах

IBM Power



RICS процессоры

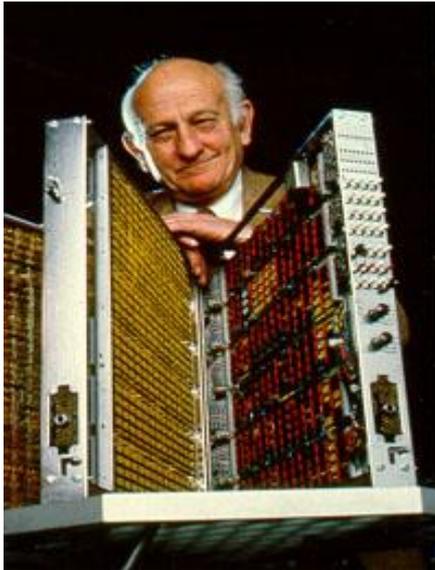
RISC-*Restricted (reduced/regular) Instruction Set Computer*-**сокращенный** набор команд

CISC -*Complex Instruction Set Computer* – **полный** набор команд

Термин «**сокращённый**» означает, что сокращён объём (и время) работы, выполняемый каждой отдельной инструкцией — за один(или менее) цикл доступа к памяти.

Сложные инструкции CISC-процессоров могут требовать сотен циклов доступа к памяти для своего выполнения.

Появление RISC-процессоров



1974-1978 – проект 801.

John Cocke – «отец» IBM RISC архитектуры



<i>CPU</i>	<i>Dhrystones 1.1</i>	<i>Whetstones (millions)</i>	<i>Linpack(dp) Fortran rolled</i>	<i>Livermore loops (geometric mean)</i>	<i>SPEC* Mark</i>
Sun 4/200	19 000	3.9	1.6	0.77	—
DecStation 3100	25 000	8.8	1.6	1.99	10.1
Apollo DN10000	25 461	14.9	5.1	2.50	13.9
MIPS M/2000	43 100	14.1	3.9	3.60	17.6
RISC System/6000 Processor	60 700	25.5	10.9	8.90	28.9

Характерные особенности RISC-процессоров

Фиксированная длина машинных инструкций (например, 32/64 бита) и простой формат команды.

Специализированные команды для операций с памятью — чтения или записи.

Операции вида «прочитать-изменить-записать» отсутствуют.

Любые операции «изменить» выполняются только над содержимым регистров (т. н. архитектура load-and-store).

Большое количество регистров общего назначения (32 и более).

Отсутствие поддержки операций вида «изменить» над укороченными типами данных — байт, 16-битное слово.

Отсутствие микропрограмм внутри самого процессора.

*ARM (Advanced RISC Machines)—

Apple iPods /iPhone/iPad

Palm and PocketPC PDAs

RIM BlackBerry smartphone/email devices

Nintendo

HTC(Qualcom proc)

* PowerPC (Apple–IBM–Motorola alliance)

* Motorola 68000

* MIPS(SGI computers and the PlayStation, PlayStation 2, Nintendo 64, PlayStation Portable game consoles)

* IBM POWER(IBM's supercomputers, midrange servers and workstations, in Nintendo's Gamecube and Wii, Microsoft's Xbox 360 and Sony's PlayStation 3 game consoles, EMC's DMX range of the Symmetrix SAN, and in many embedded applications like printers and cars.

* SPARC, by Oracle/Sun Microsystems and Fujitsu

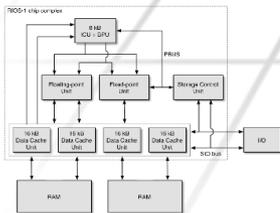
* Hewlett-Packard's PA-RISC, also known as HP-PA

* Alpha

* Hitachi's SuperH

POWER1

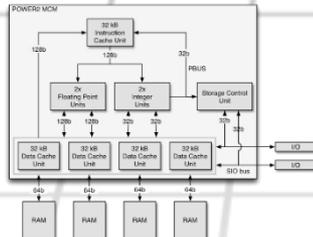
20 - 62 MHz



1990

**POWER2
P2SC**

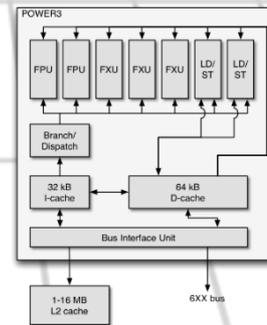
55 - 135 MHz



1995

**POWER3
630**

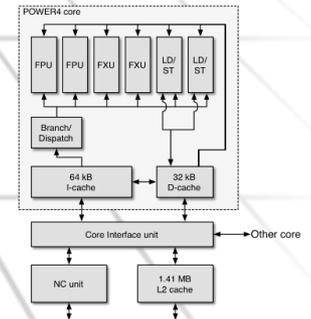
200 - 450 MHz



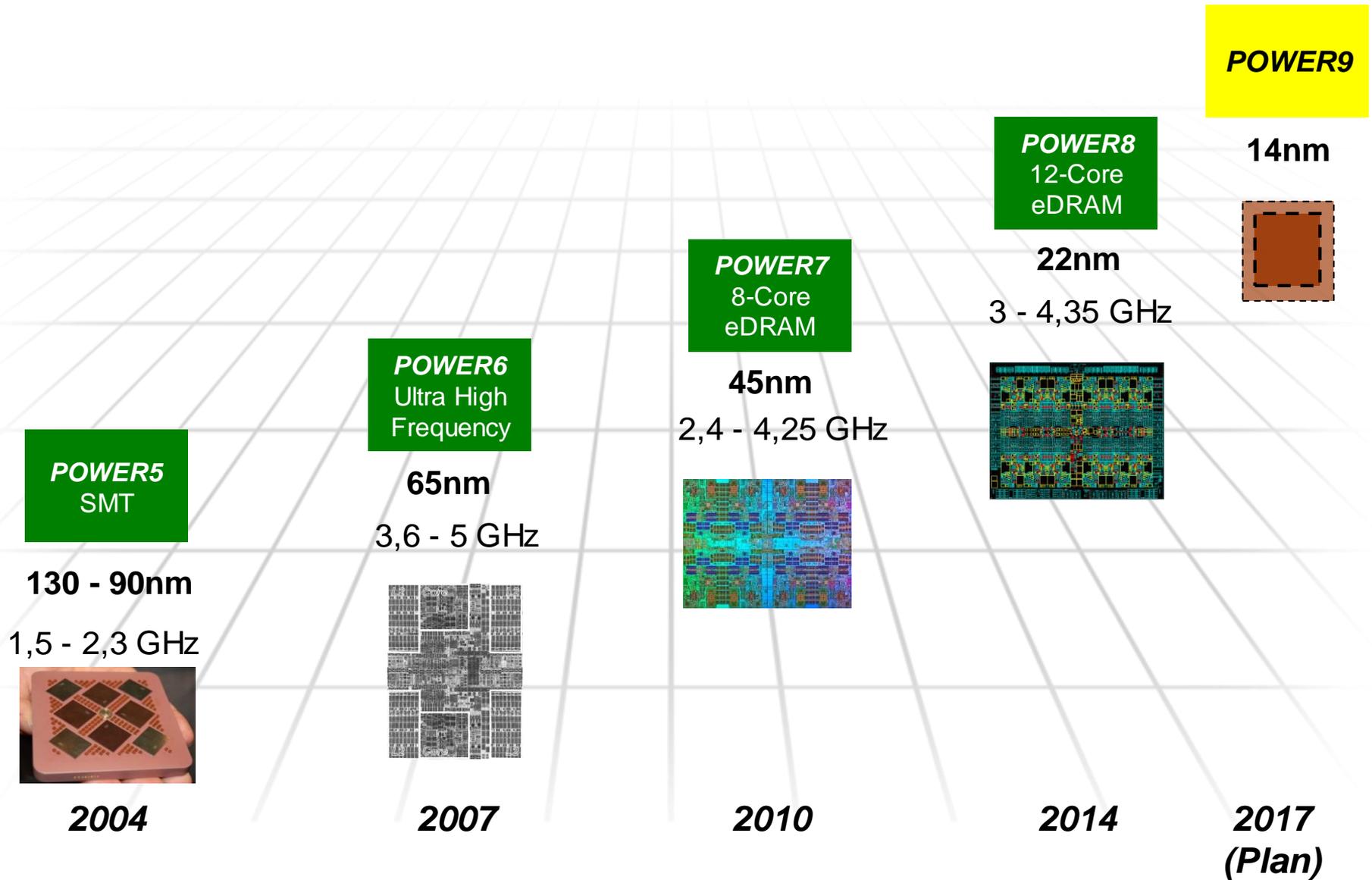
1998

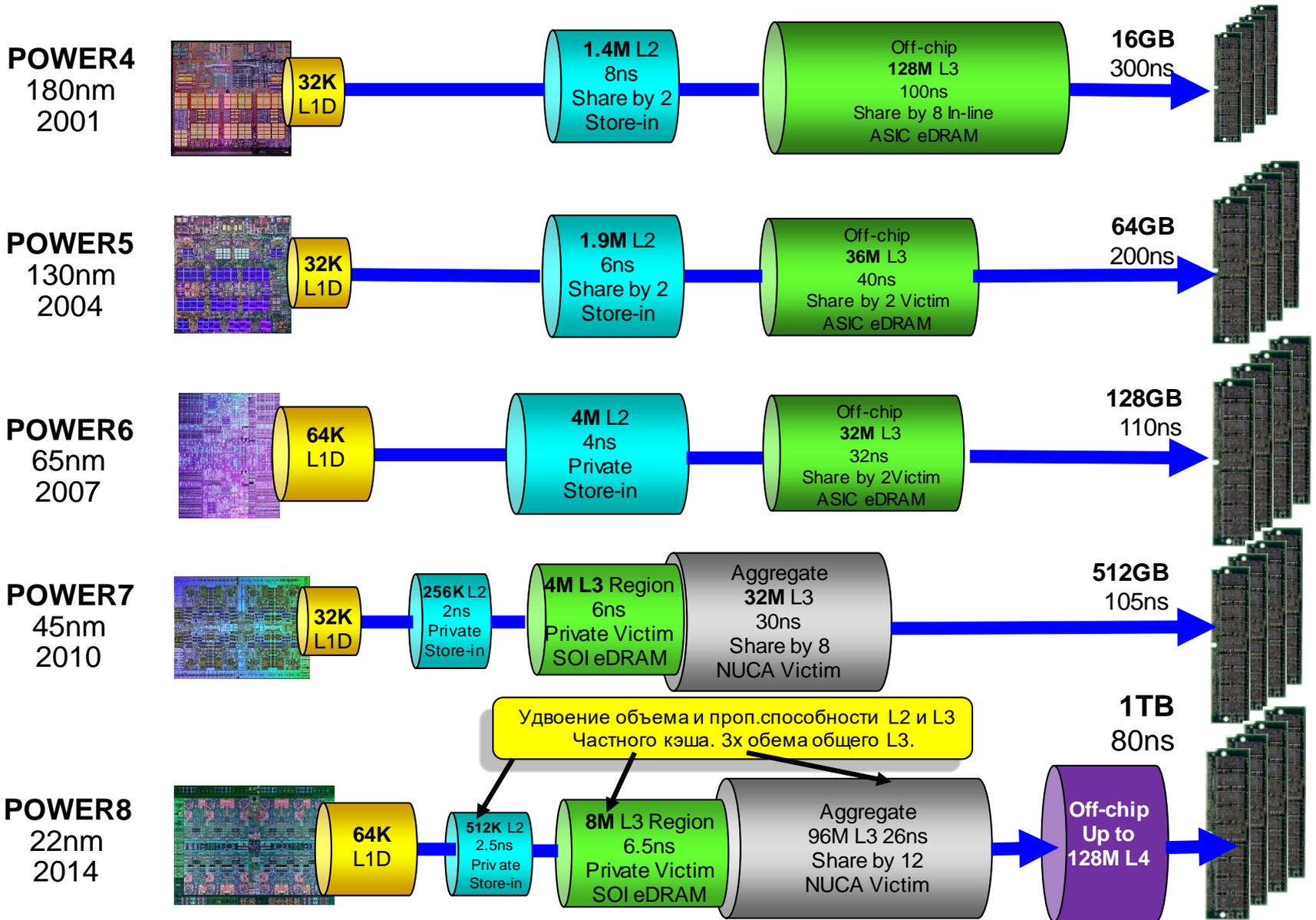
**POWER4
Dual Core**

180nm
1,1 - 1,9 GHz
Первый CPU
> 1 GHz

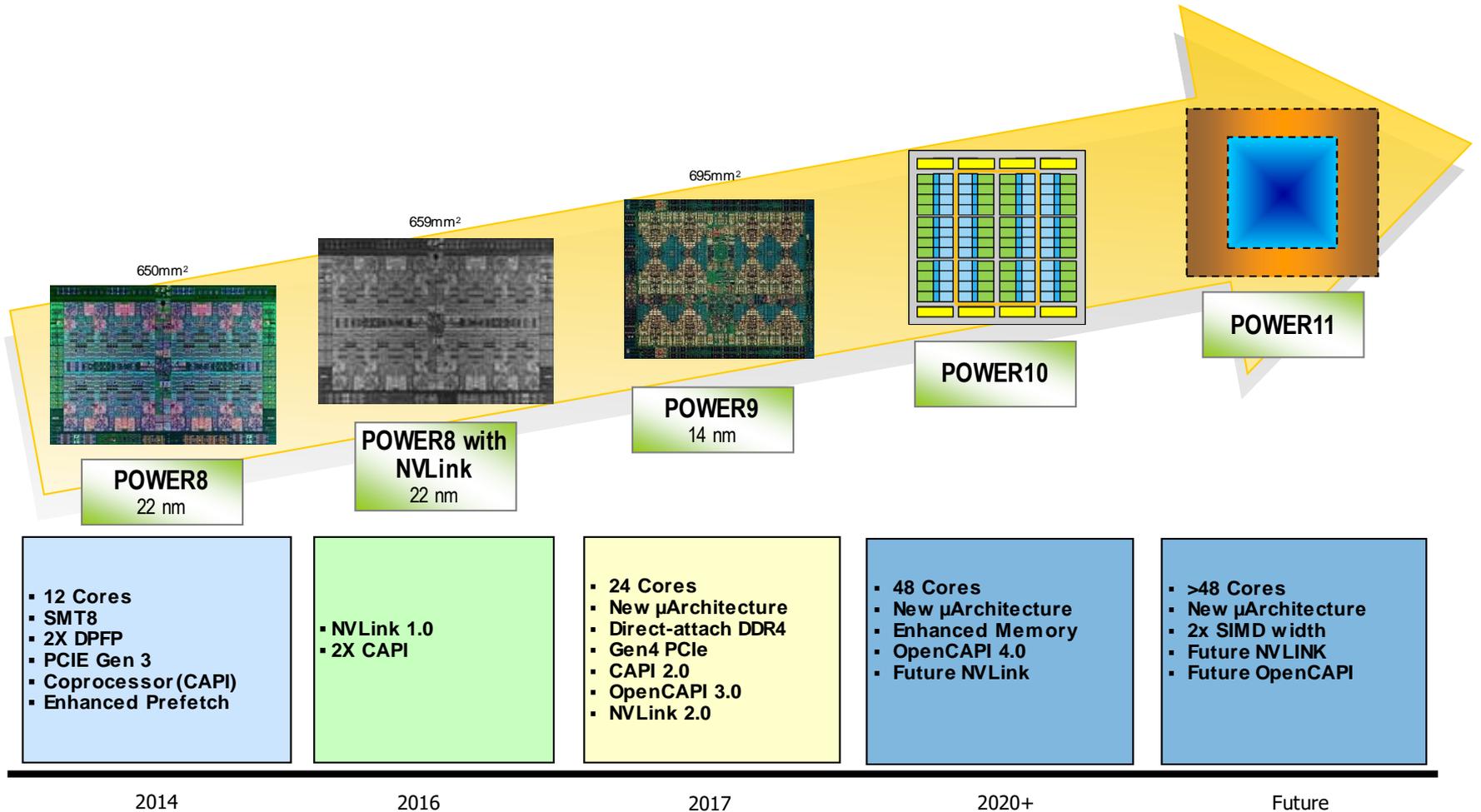


2001





Processor Technology Roadmap



POWER9 Processor – Common Features

New Core Microarchitecture

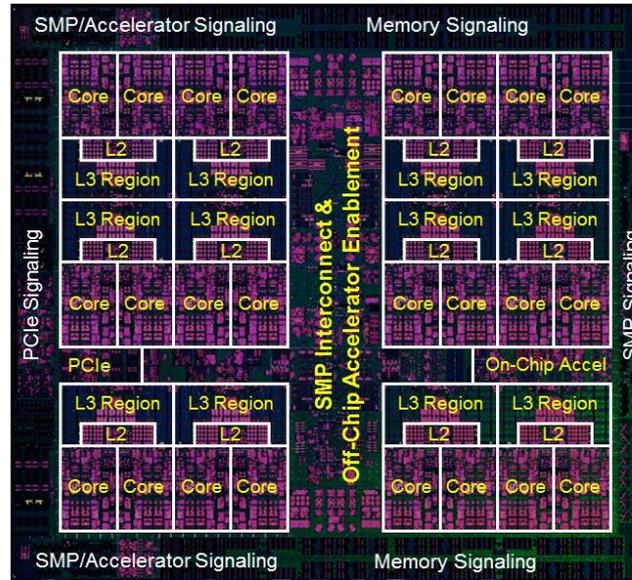
- Stronger thread performance
- Efficient agile pipeline
- POWER ISA v3.0

Enhanced Cache Hierarchy

- 120MB NUCA L3 architecture
- 12 x 20-way associative regions
- Advanced replacement policies
- Fed by 7 TB/s on-chip bandwidth

Cloud + Virtualization Innovation

- Quality of service assists
- New interrupt architecture
- Workload optimized frequency
- Hardware enforced trusted execution



14nm finFET Semiconductor Process

- Improved device performance and reduced energy
- 17 layer metal stack and eDRAM
- 8.0 billion transistors

Leadership

Hardware Acceleration Platform

- Enhanced on-chip acceleration
- Nvidia NVLink 2.0: High bandwidth, advanced new features
- CAPI 2.0: Coherent accelerator and storage attach (PCIe G4)
- OpenCAPI: Improved latency and bandwidth, open interface

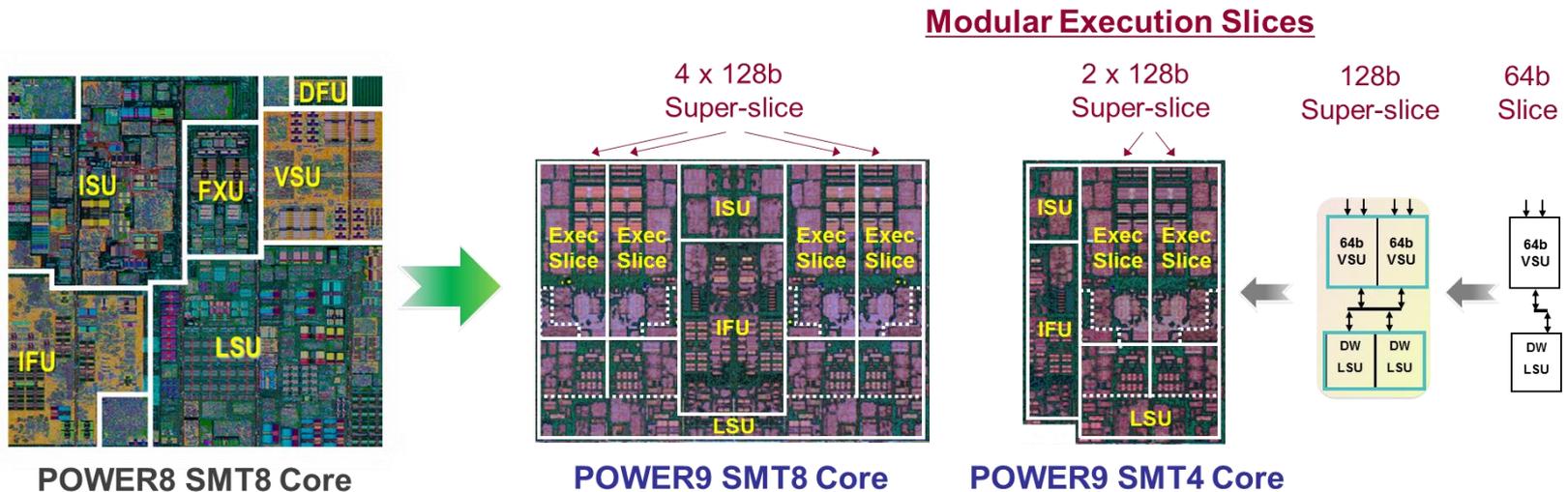
State of the Art I/O Subsystem

- PCIe Gen4 – 48 lanes

High Bandwidth Signaling Technology

- 16 Gb/s interface
 - Local SMP
- 25 Gb/s interface – 25G Link
 - Accelerator, remote SMP

POWER9 Core Execution Slice Microarchitecture



Re-factored Core Provides Improved Efficiency & Workload Alignment

- Enhanced pipeline efficiency with modular execution and intelligent pipeline control
- Increased pipeline utilization with symmetric data-type engines: Fixed, Float, 128b, SIMD
- Shared compute resource optimizes data-type interchange

New POWER9 Cores

Optimized for Stronger Thread Performance and Efficiency

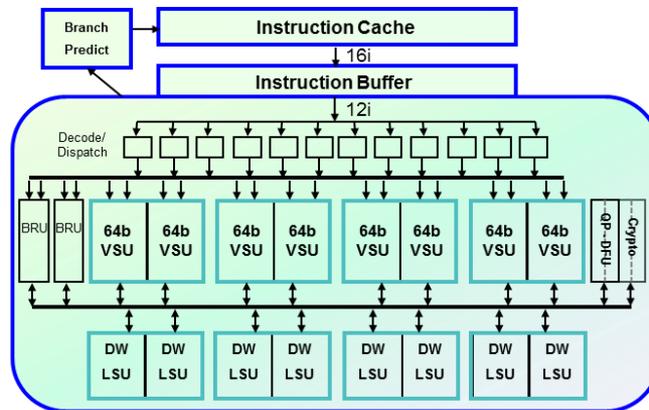
- Increased execution bandwidth efficiency for a range of workloads including commercial, cognitive and analytics
- Sophisticated instruction scheduling and branch prediction for unoptimized applications and interpretive languages
- Adaptive features for improved efficiency and performance especially in lower memory bandwidth systems

Available with SMT8 or SMT4 Cores

8 or 4 threaded core built from modular execution slices

POWER9 SMT8 Core

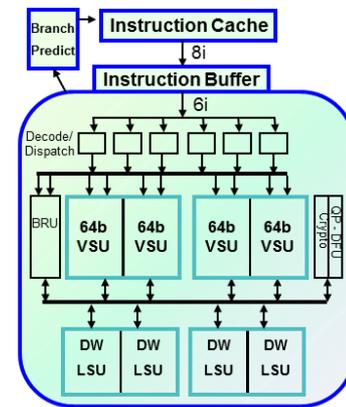
- PowerVM Ecosystem Continuity
- Strongest Thread
- Optimized for Large Partitions



SMT8 Core

POWER9 SMT4 Core

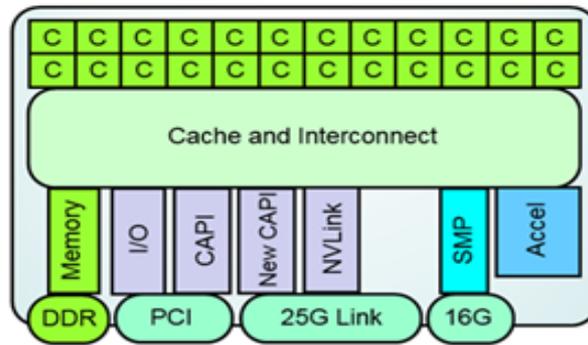
- Linux Ecosystem Focus
- Core Count / Socket
- Virtualization Granularity



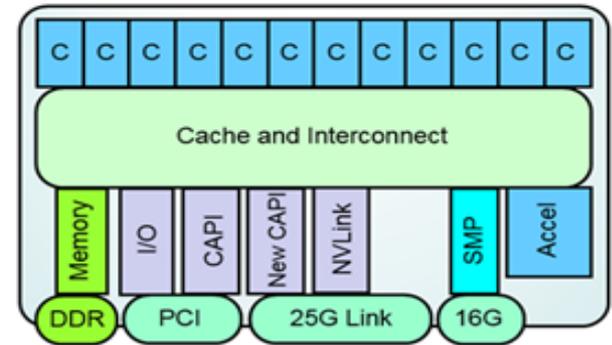
SMT4 Core

Варианты процессоров POWER9

24 ядра SMT4
Linux ecosystem



12 ядер SMT8
PowerVM ecosystem (AIX)

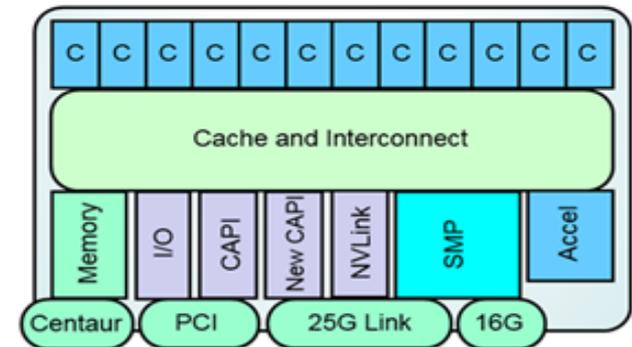


Scale-out

- Оптимизировано под 2S
- Прямое подключение памяти
- Форм-фактор в соответствии со стандартами

Scale-up

- Оптимизировано для больших систем
- Буферизованная память
- Дополнительные возможности ввода/вывода



Эволюция возможностей подключения ускорителей

Ускорители как основа аппаратного обеспечения в области AI



Большой eDRAM кэш для больших датасетов с минимальным latency

Использование **PCIe Gen 4**: 48 линий

Высокая пропускная способность при взаимодействии CPU-GPU, GPU-GPU на базе **NVlink 2.0**

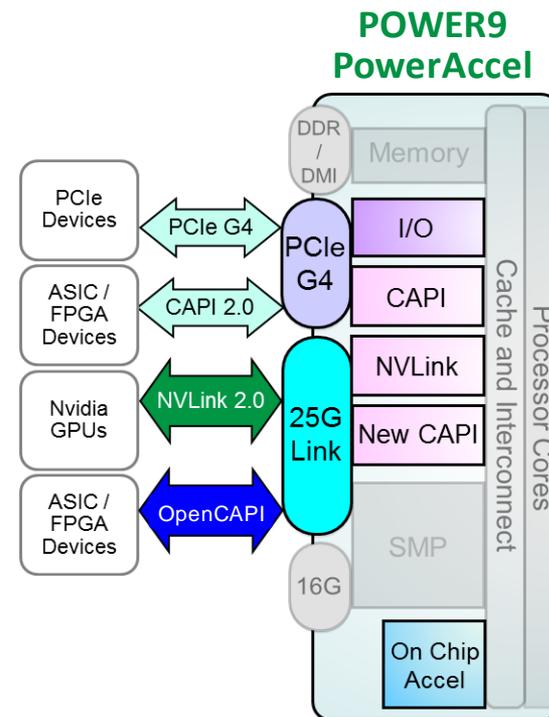
Интерфейс следующего поколения **CAPI 2.0** для когерентного доступа к акселераторам и системам хранения

On-chip компрессия & криптографические ускорители

Новая 25Gb/s шина для присоединения ускорителей

Оптимизация чипа для использования в 2x сокетных scale-out серверах в hyperscale ЦОДах

Прямой доступ к DDR4: 8 каналов памяти, устойчиво >120 GB/s



IBM Power Виртуализация



Лидерство IBM в технологиях виртуализации

Пятидесятилетняя история продолжается в PowerVM™

1967

1973

1999

2004

2008

2010

2013

IBM
разрабатывает
«гипервизор»
в
мейнфреймах

IBM
выпускает
первую
машину с
физическими
партициями

IBM
внедряет
LPAR в
POWER™

IBM внедряет
POWER
Hypervisor™
for System
p™ and
System i™

IBM ships
PowerVM
Editions



PowerVM
delivers
enhanced
storage
virtualization

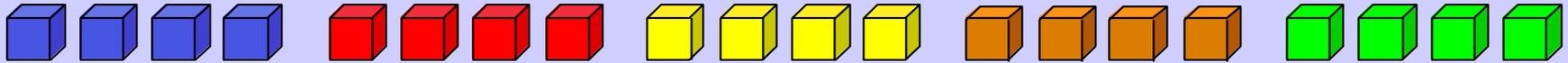
PowerVM
delivers
OpenStack
Enablement



Power is virtualization
without limits

The Power Systems team has integrated PowerVM world-class virtualization into every server – based on best practices gained from IBMers who created the ‘gold standard’ of mainframe virtualization.

Основная концепция виртуализации



Виртуальные ресурсы

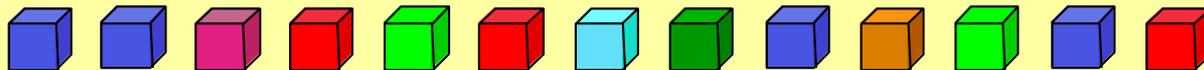
- ☒ Подмена реальных ресурсов: **same interfaces/functions, different attributes.**
- ☒ Часто связано с одним ресурсом, но может быть набором ресурсов.

«Виртуализатор» – процесс подмены

- ☒ Создает виртуальные ресурсы на основе реальных.
- ☒ Осуществляется на основе software и/или firmware.

Ресурсы

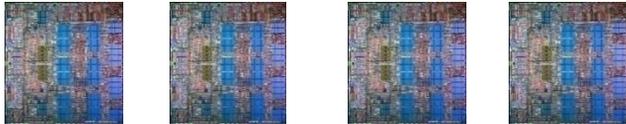
- ☒ Компоненты с **определенными интерфейсами/функциями.**
- ☒ Физические. Централизованные или распределенные
- ☒ Примеры: memory, disk drives, networks, servers.



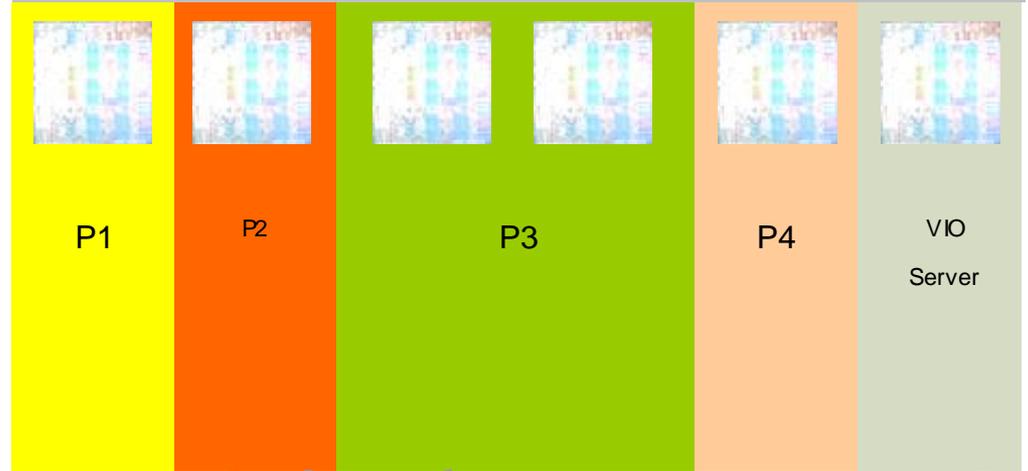
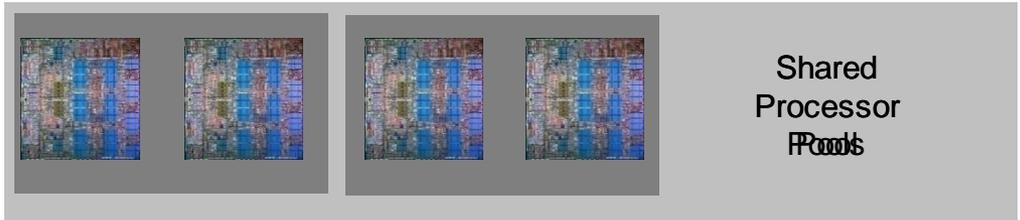
- ☒ Отделение восприятия ресурсов пользователями от реальных ресурсов
- ☒ Консолидация «пулов» ресурсов для выделения их пользователям

Виртуализация POWER4/POWER5/POWER6/POWER7/POWER8/POWER9

POWER4

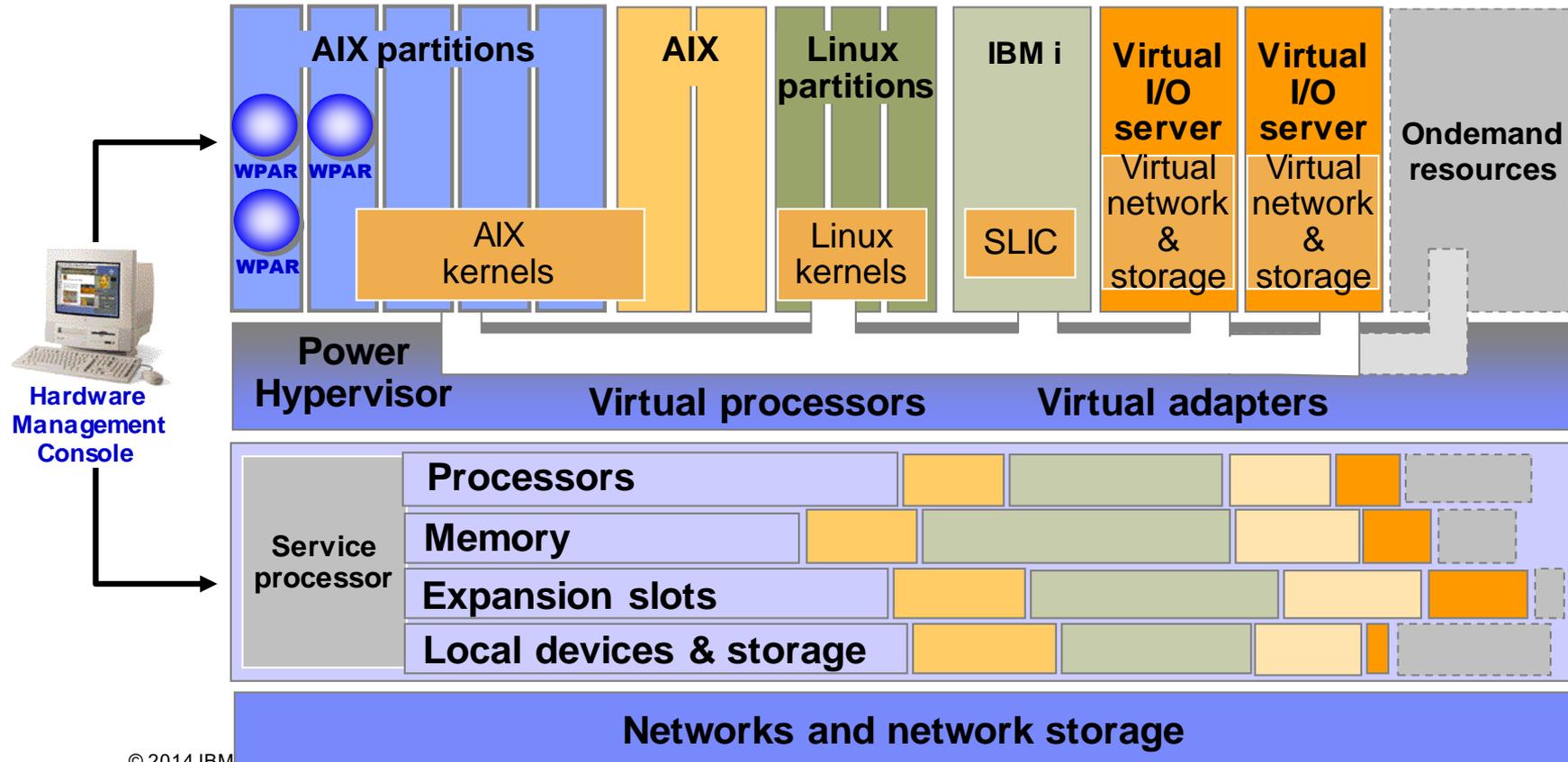


POWER6/7/8/9

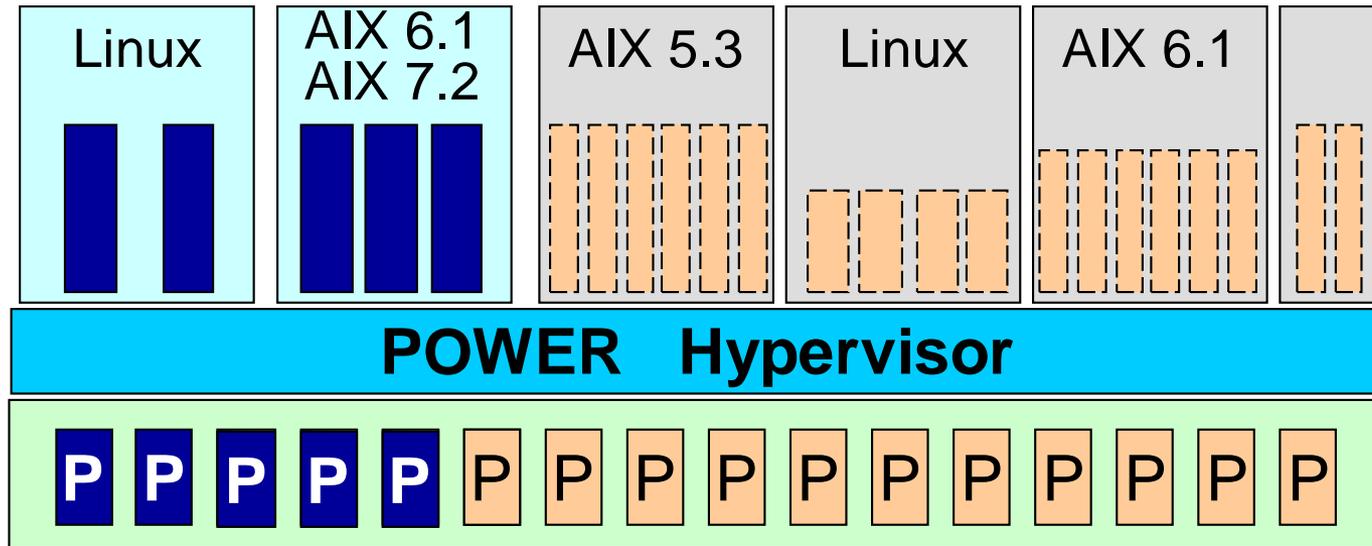


PowerVM key design points

- Designed for high **efficiency** to provide high overall performance
- Designed for high **scalability** – linear from 1/20 to 256 cores
- Designed for **isolation** to provide security and “no compromise” consolidation
- Designed for maximum resource **granularity** to reduce wasting resources



Shared Dedicated Processors

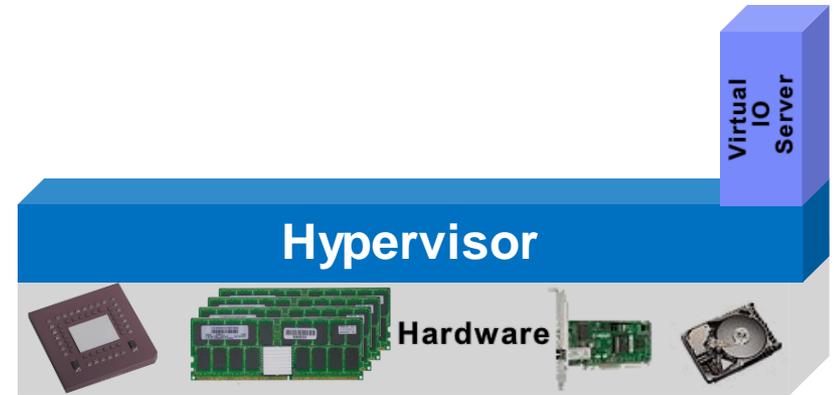
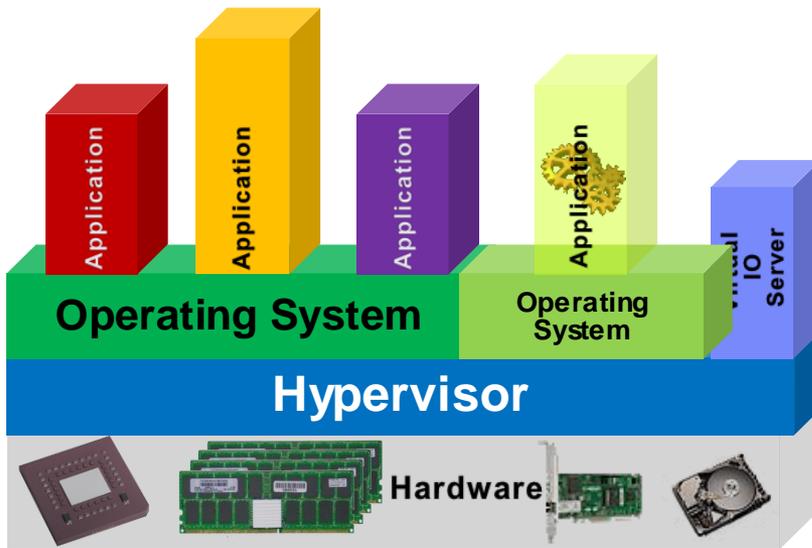


Dedicated Processors

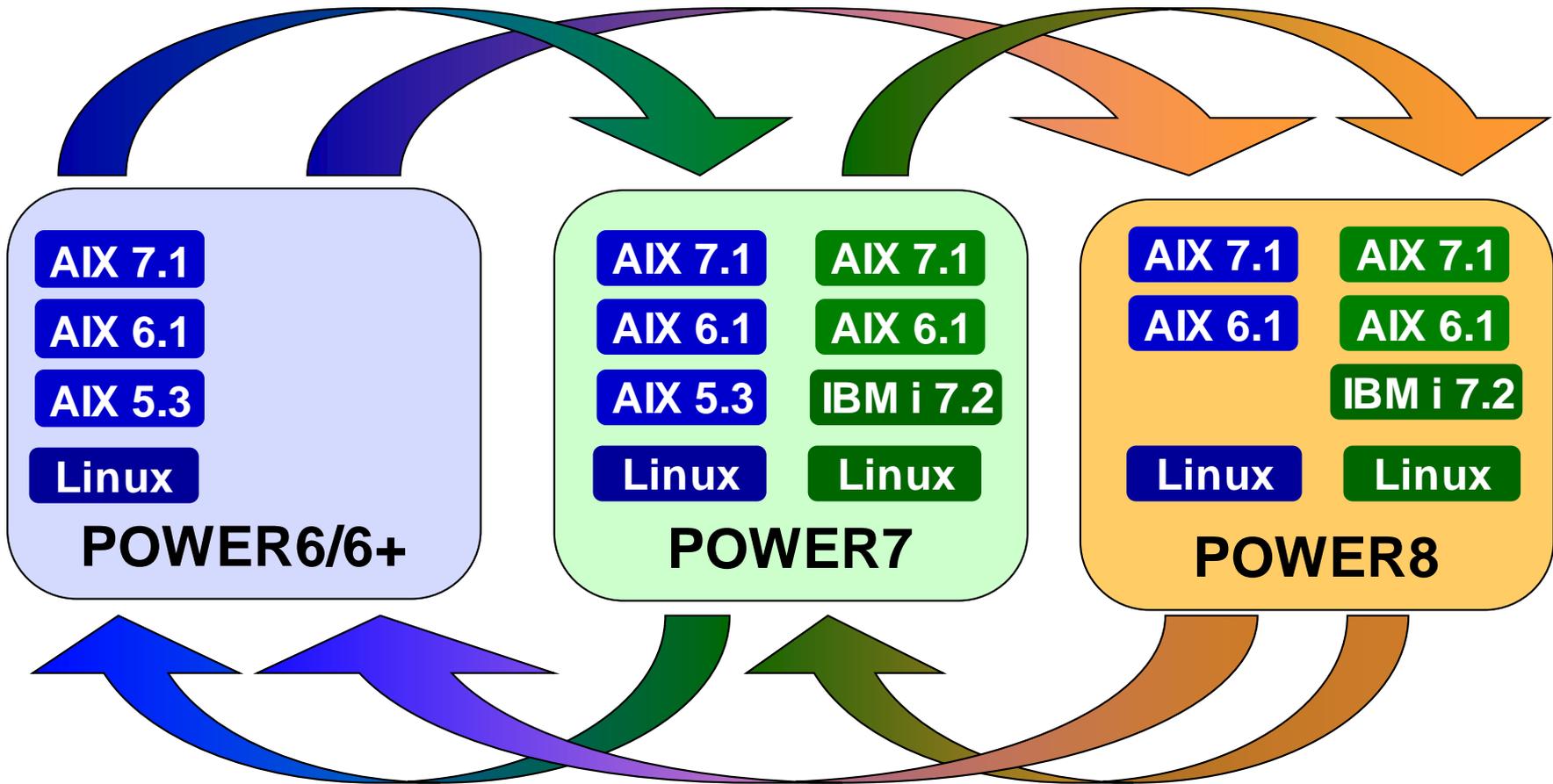
Shared (Non-Dedicated) Processors

Excess Dedicated Capacity Utilization

- Unused capacity in dedicated processor partitions can be “Donated” to shared processor pool
- Excess cycles will only be utilized by uncapped partitions that have consumed all of their entitled capacity.
- POWER6-8 Servers



- POWER6 hardware and higher
- Operating System
 - AIX 7, AIX 6.1 or AIX 5.3 TL7
 - Linux RHEL 5-Update 1 or SLES 10-Service Pack 1
- Same network subnet
- Same Hardware Management Console
- Virtualized resources
- SAN storage for boot and data



Leverage POWER6 / POWER7 Compatibility Modes

LPAR Migrate between POWER6 / POWER7 / POWER8 Servers

Can not move POWER8 Mode partitions to POWER6 or POWER7 systems.

Для «больших» серверов – E850, E870, E880:

■ Capacity Upgrade on Demand

- Апгрейд систем активацией неактивных ресурсов: CoD ядра и/или память
- Нет специальных контрактов, не требуется мониторинг – активация **сразу** после закупки

■ Elastic Capacity on Demand

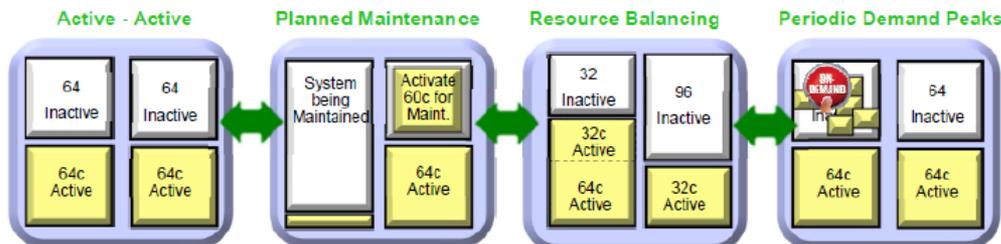
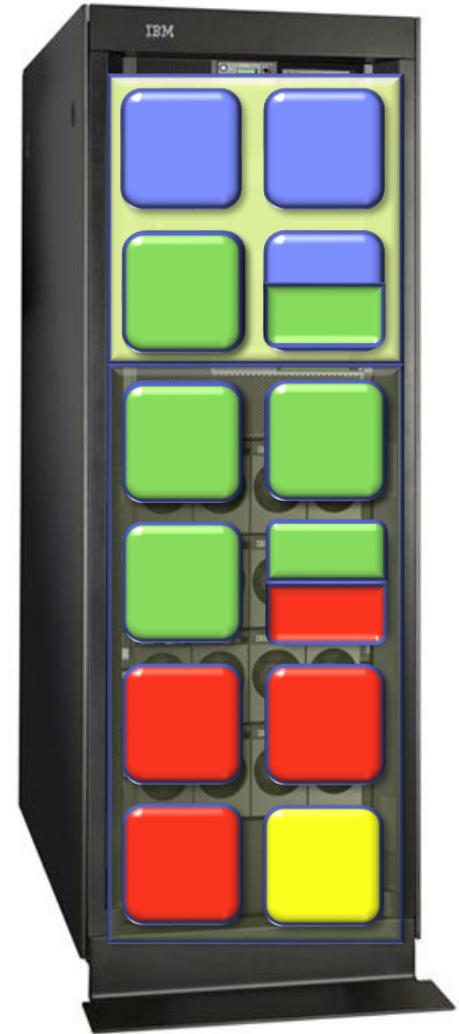
- **Временное** использование CoD ядер и памяти по требованию – ресурсы выделяются **посуточно**
- Ежемесячный мониторинг использования для выставления квартальных счетов

■ Utility Capacity on Demand

- Только CPU – измеряется поминутное использование процессоров. **Временное** автоматическое добавление временно активных ядер
- Предоплатная или постоплатная система счетов на 365 дней

■ Trial Capacity on Demand

- Позволяет клиентам протестировать эффект от добавления ядер и памяти
- **Частичная** или **полная** активации ядер и памяти на **фиксированное** время
- **Бесплатно**



Power Enterprise Pool

Безопасность

никто не хочет быть
целью атаки



TARGET®

Безопасность ограничена «слабым звеном»



b

VMware has 1025 Vulnerabilities
PowerVM has 0 Vulnerabilities
as 02/2019

VULNERABILITIES

SEARCH AND STATISTICS

Q Search Results (Refine Search)

Search Parameters:

- Results Type: Overview
- Keyword (text search): VMware
- Search Type: Search All

1025

There are **1,025** matching records.
Displaying matches **1** through **20**.

VULNERABILITIES

SEARCH AND STATISTICS

Q Search Results (Refine Search)

Search Parameters:

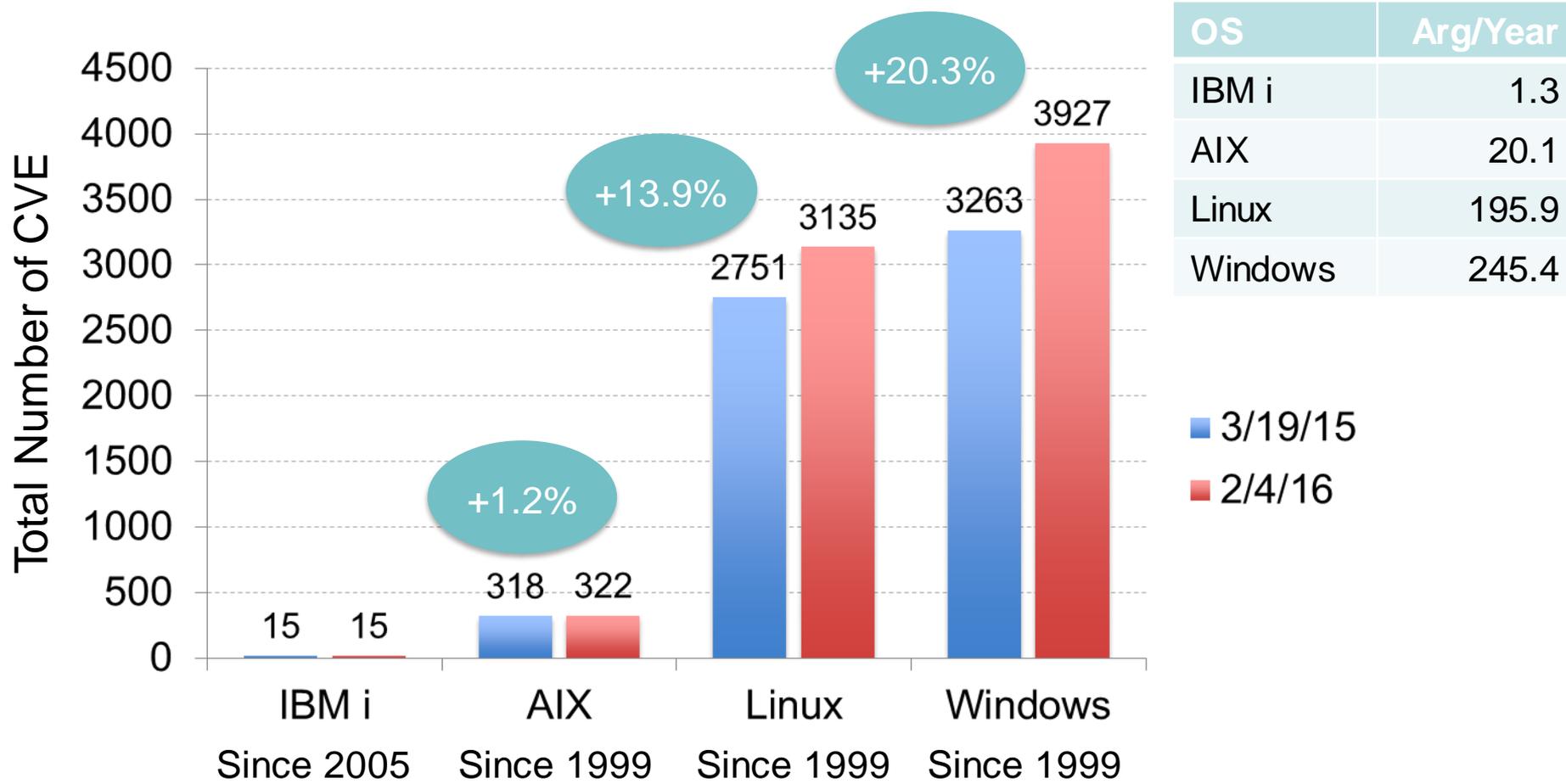
- Results Type: Overview
- Keyword (text search): PowerVM
- Search Type: Search All

There are **0** matching records.

IBM PowerVM: 0

https://nvd.nist.gov/vuln/search/results?form_type=Basic&results_type=overview&query=PowerVM&search_type=all

Число обнаруженных уязвимостей операционных систем



Source: cvemitre.org, Note: IBM i numbers include IBM i, OS/400 and i5/OS

R(надежность)
A(доступность)
S(сервисность)



PowerPC Still in the Spirit

**NASA sends IBM chipkill with
RS/6000 on Mars Pathfinder
(1996)**

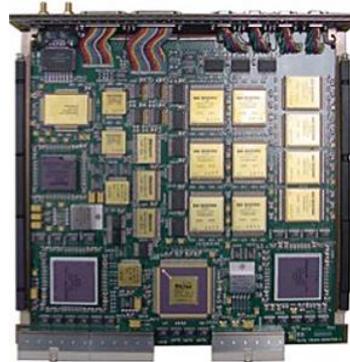


Spirit and Opportunity Rovers Powered by PowerPC

RISC процессоры на Марсе



«Because of this, he added that an Intel Core i7 processor, for example, wouldn't survive a day on Mars because of the extremes in temperature.»



The single-board computer aboard NASA's Curiosity rover is built around a PowerPC RAD750 microprocessor provided by BAE Systems.

ComputerWorld August 8 2012 http://www.computerworld.com/s/article/9230060/NASA_Your_smartphone_is_as_smart_as_the_Curiosity_rover

CBS News http://www.cbsnews.com/network/news/space/home/spacenews/files/081012_msl_update.html

Checkers and Fault Isolation Registers provide the mechanism to ensure **First Failure Data Capture**

FFDC ключевой компонент надежности IBM Power

	Checkers	FIRs
POWER8 S812L	30,068	13,200
POWER8 S824	60,136	26,400
Power 740	37,720	5,800
Power 770 (one drawer)	39,800	6,500
Power 795 (Maximum configuration)	598,000	96,000
Intel Ivy Bridge EX	5,500	none

Power Service 3
Processor Fixes!



FIR: 2
(Fault Isolation Register)
Error data from the checkers are collected and stored here. The service processor monitors the contents



Checkers: 1
Built-in HW error diagnostic probes to identify error conditions



Power HW / FW



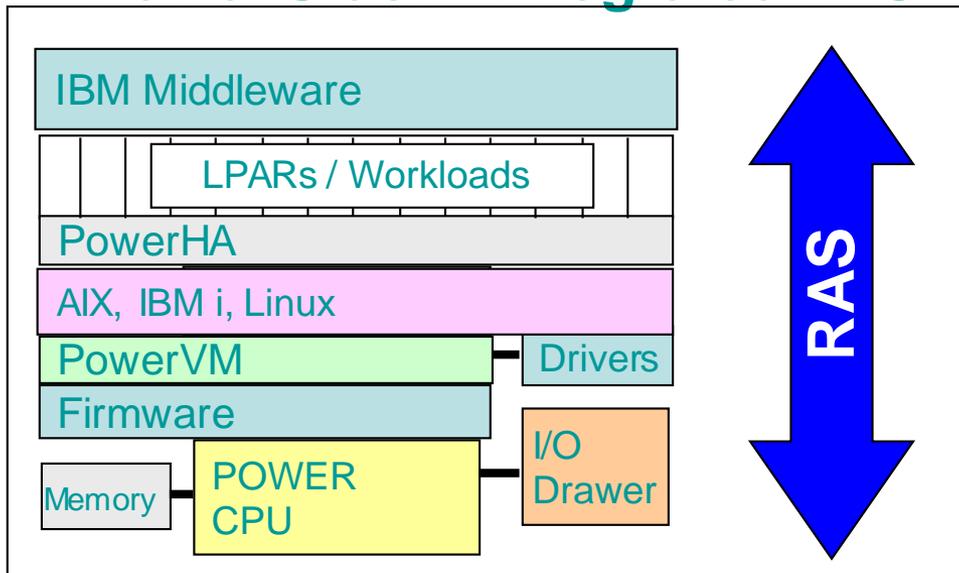
Средства отказоустойчивости заложены в архитектуру серверов

RAS Feature	POWER7/7+	POWER8 (PowerVM)	Intel Ivy Bridge EX
Application/Partition			
Move running workload partitions to another server	Yes	Yes	Yes
Partition Availability priority	Yes	Yes	No
System			
First Failure Data Capture	Yes	Yes	No
Memory Keys ensure secure hardware separation of memory	Yes	Yes	No
Processor			
Processor Instruction Retry	Yes	Yes	No
Alternate Processor Recovery	Yes	Yes	No
Dynamic Processor Deallocation	Yes	Yes	No (1)
Dynamic Processor Sparing	Yes	Yes	No (1)
Memory			
Chipkill - Redundant Memory Capabilities	Yes	Yes	Yes
Chipkill - Survives Double Memory Failures with x8 DRAMS	Yes(2)	Yes(2)	Yes (4)
I/O			
VIOS - Dynamic adapter failover	Yes	Yes	No

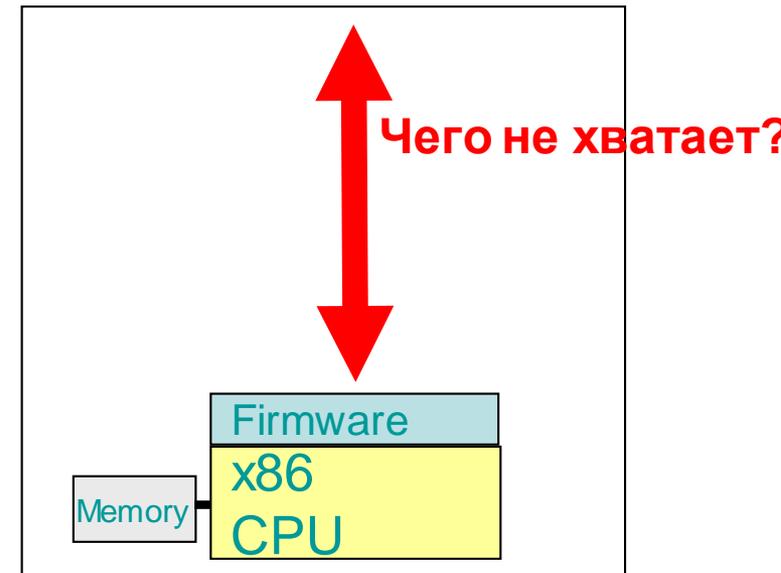
RAS не только на уровне процессора

- Разработка, тестирование и интеграция всего комплекса для достижения наивысших характеристик RAS
- Для того, чтобы RAS x86 приблизить к RAS Power нужно, как минимум использовать дорогостоящее кластерное ПО (например, Oracle RAC)

Power Stack - Integrated RAS



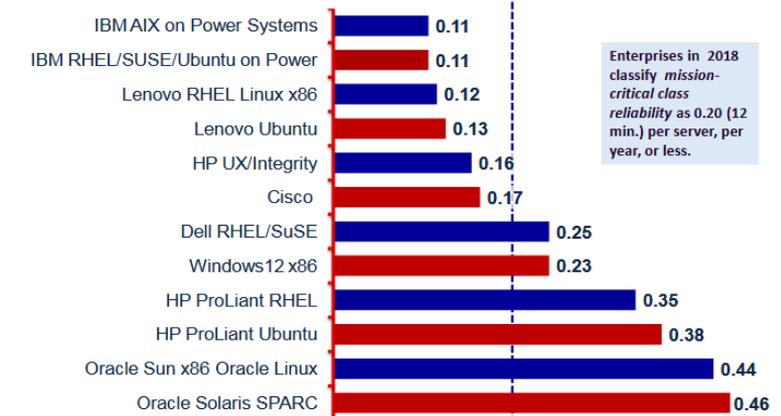
Intel Processor RAS





#1 в каждой категории доступности согласно рейтингу ITIC

Unplanned Downtime in 2017 - 2018 (Hours per Year)

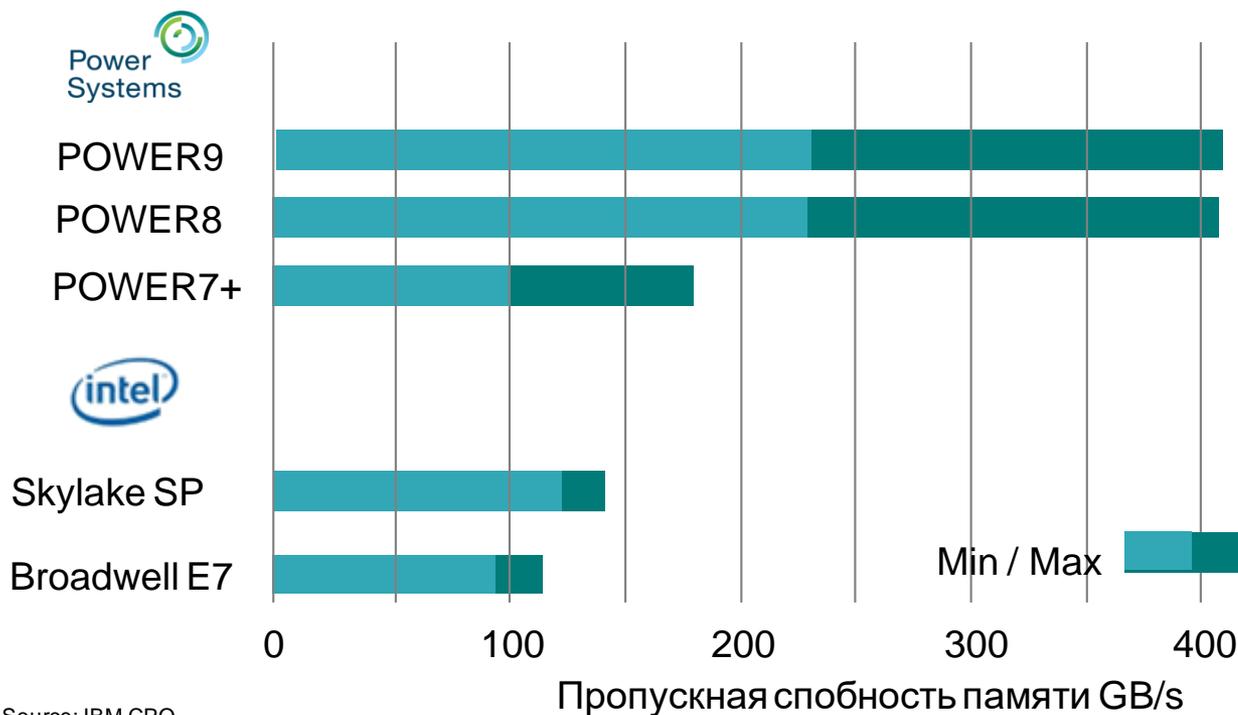


“IBM POWER8-based processor systems and the latest POWER9 servers provide several key feature/function advantages that advance reliability and enable customers to lower Total Cost of Ownership (TCO) and achieve near-immediate ROI.”

Производительность



Быстрая память серверов POWER идеальна для «in-memory» приложений типа SAP HANA



1.8x

POWER против Intel
(до 32TB в E880C и 64TB в E980)

Память
быстрая, большие
объемы для обработки
данных

Source: IBM CPO

SAP TPC-C для 8ми сокетных серверов - апрель 2019
Intel Skylake и Cascade Lake – (28 ядер на кристалл)

SAP TPC-C

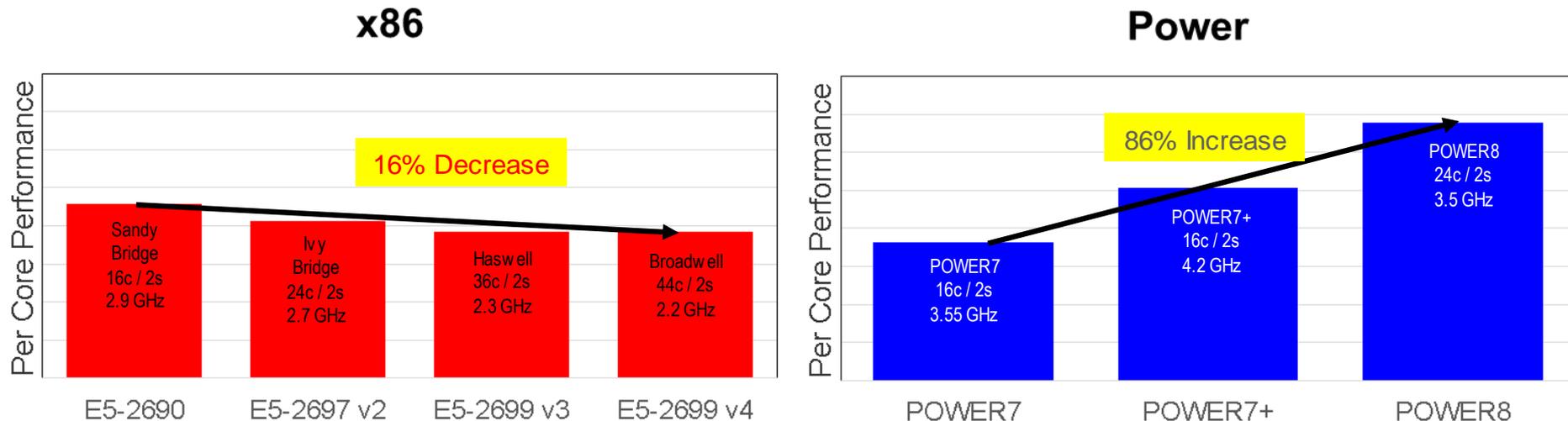
Cascade Lake Server (Fujitsu)	2712 SAPS на ядро
Skylake (Lenovo)	2658 SAPS на ядро

IBM POWER9 более **6800 SAPS** на ядро

Новое поколение процессоров Intel показывает 2% прироста
производительности на ядро
Это простое увеличение ядер в соquete

Производительность ядер процессоров Intel не увеличивается, а уменьшается

Производительность ядер IBM Power увеличиваются при этом требуется меньше лицензий ПО в случае поядерного лицензирования



RPE2** numbers are derived from the following six benchmark inputs:
 SAP SD Two-Tier, TPC-C, TPC-H, SPECjbb2006 and two SPEC CPU2006 components

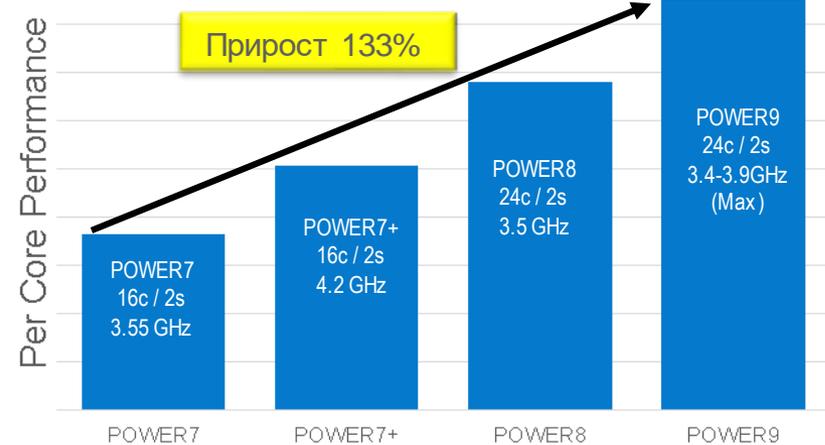
Сравнение производительности 2-хsocketных систем POWER против x86

Процессоры IBM Power показывают прирост производительности на ядро при переходе от поколения к поколению, что снижает количество лицензий ПО при поядерном лицензировании

Intel x86



IBM POWER





IBM POWER
МОДЕЛЬНЫЙ ряд

Power
Systems

Семейство систем IBM POWER9

**BigData
& Analytics**

LC922



9006-22P



**High Performance
Computing/
Machine Learning**

AC922



8335-GTH/GTX
„Newel“

**Commercial
Computing & Cloud**

LC921



9006-12P



Новый спектр задач

**Традиционные
задачи**

9009-22A



S922

1 or 2 socket, 2U
До 20 ядер

9009-41A



S914

1 socket, 4U
До 8 ядер



9009-42A



S924

2 socket, 4U
До 24 cores

9008-22L



L922

2 socket, 2U, Linux
До 24 ядер

9223-22H



H922

2 socket, 2U, Linux
До 20 ядер

8247-42L



H924

2 socket, 4U, Linux
До 24 ядер



Традиционные задачи на Linux

„High-End“



9080-M9S

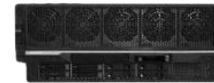


E980C

8 - 192 Cores
64 TB Memory
32 PCIe Adapters



9040-MR9



E950C

32 - 48 Cores
16 TB Memory
11 PCIe Adapters

Z-series



Наилучшая платформа для потребностей большого бизнеса.
Наивысшая отказоустойчивость
Наивысшая степень консолидации нагрузок

Power Systems



Универсальная платформа для бизнеса любого уровня.
Наивысшие показатели производительности
Высокая степень отказоустойчивости и консолидации нагрузок

PowerLinux



Наиболее подходящая платформа для CAMS-приложений и потребностей малого бизнеса



OEM-системы POWER8.
Для нового класса задач
Самые привлекательные по цене

Power Systems



- ✓ Разработаны для Big Data
- ✓ Экономически выгодны для облаков
- ✓ Открытая инновационная платформа



Виртуализация без ограничений

PowerVP



PowerVC



PowerSC



PowerHA



PowerVM

Thanks! Спасибо!



Вопросы?

Alexey.Boyko@ru.ibm.com

+7(495)-775-8800 ext 2904