



ORACLE

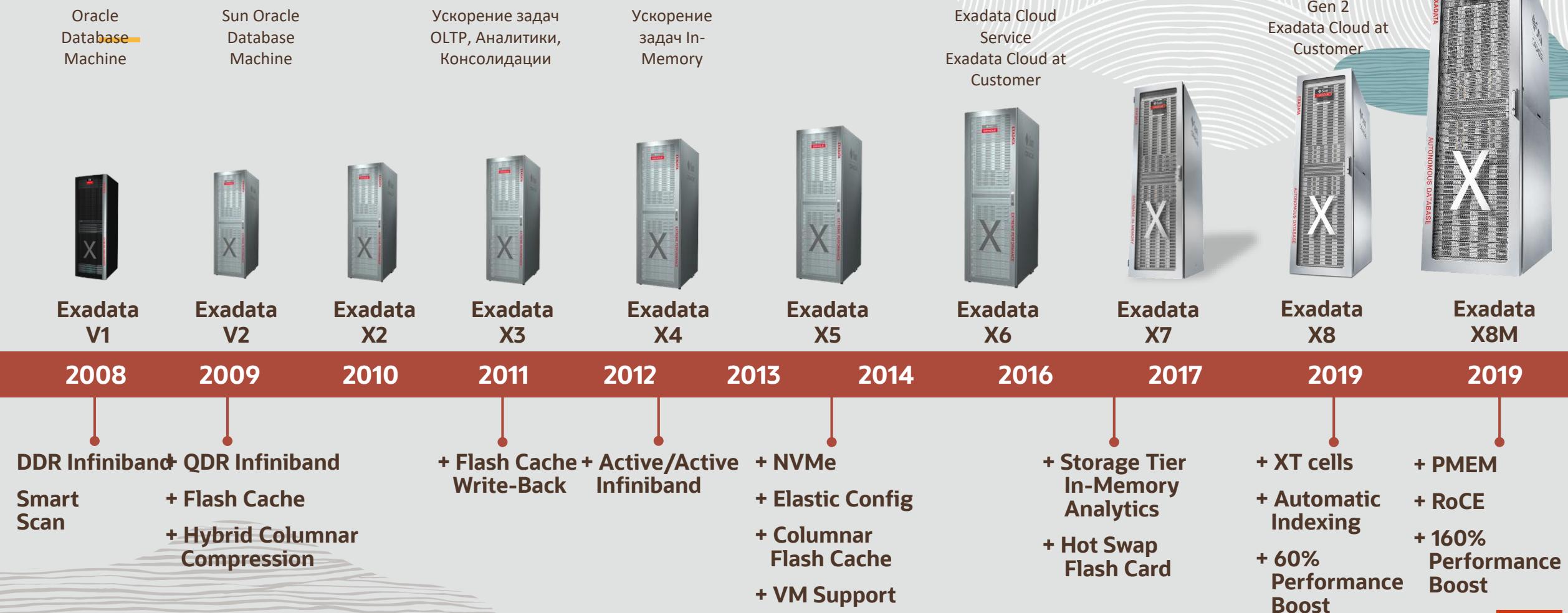
Новости программно-аппаратных комплексов

30 октября 2019

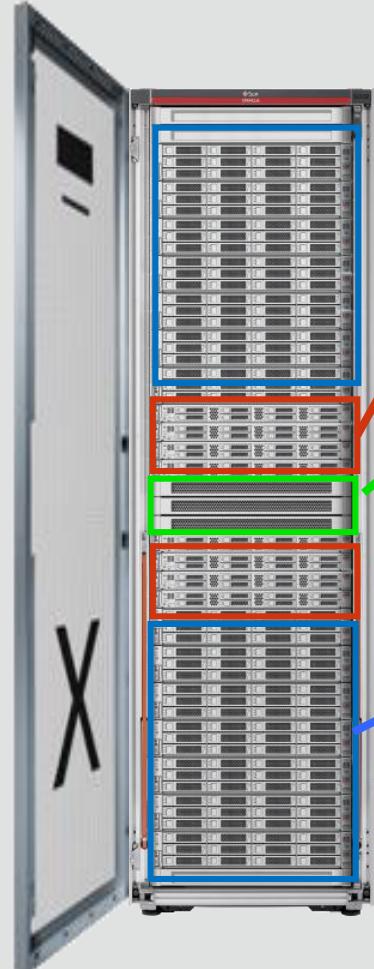


- Exadata X8M*
- ExaCC Gen2*
- ZDLRA*
- PCA*

Oracle Exadata – наращивание технологий



Что нового в Exadata X8M?

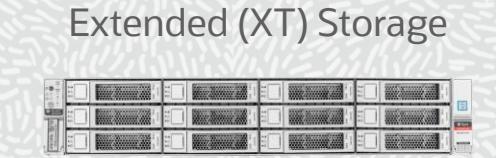
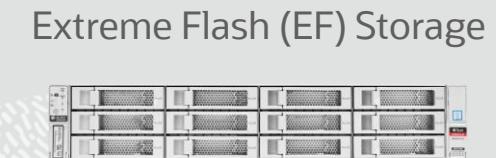
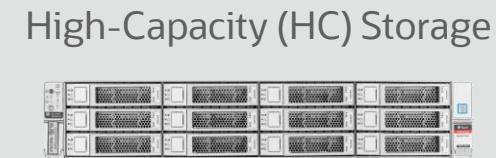


Серверы БД: Гипервизор **KVM**

интерконнект **100 Gb/sec RDMA over
Converged Ethernet RoCE**

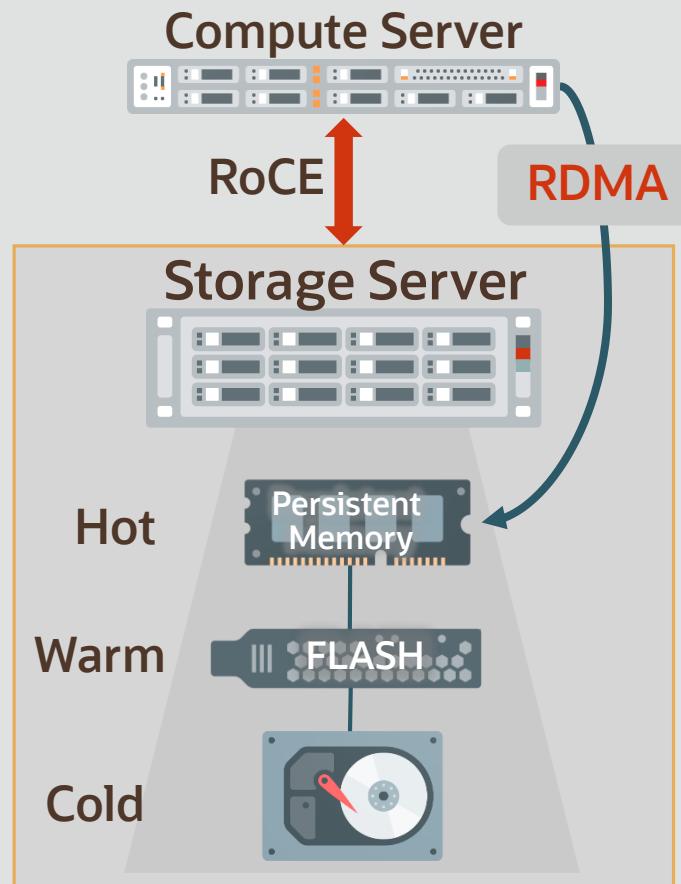
В ячейках появилась **энергонезависимая
память** на базе Intel DC Optane

- **1.5 ТВ РМЕМ в каждой ячейке**



Exadata X8M: RoCE + PMEM

Каков результат применения новых технологий?

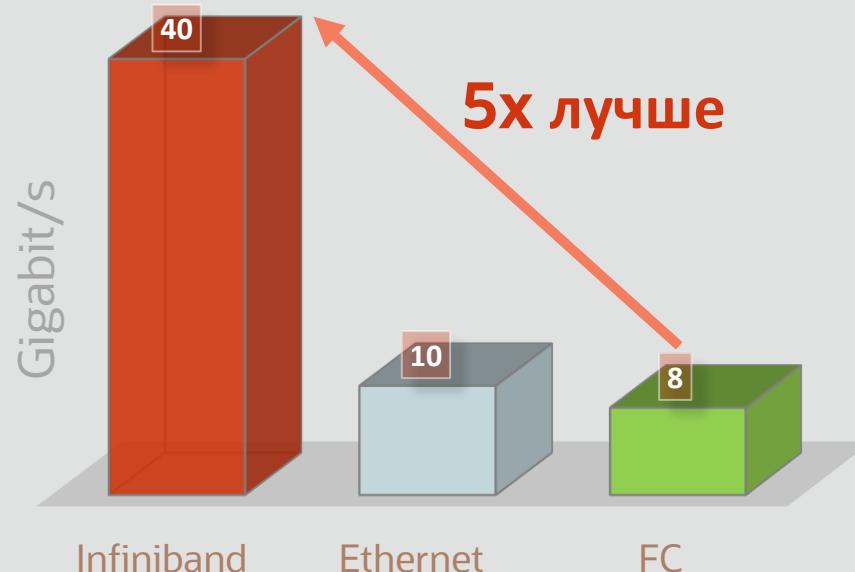


- Новый рекорд по операциям I/O
 - 16 миллионов IOPS – в 2,5 раза лучше показателей предыдущей Exadata X8
- Радикальное улучшение латентности
 - СУБД работает с PMEM по RDMA напрямую минуя системный стек ОС, прерывания
 - <19 μ сес латентность на чтении БД блоками 8К в – 10 раз лучше
- Сверхбыстрая запись журнала транзакций
- Новые технологии Exadata:
 - Exadata PMEM cache
 - Exadata PMEM Log

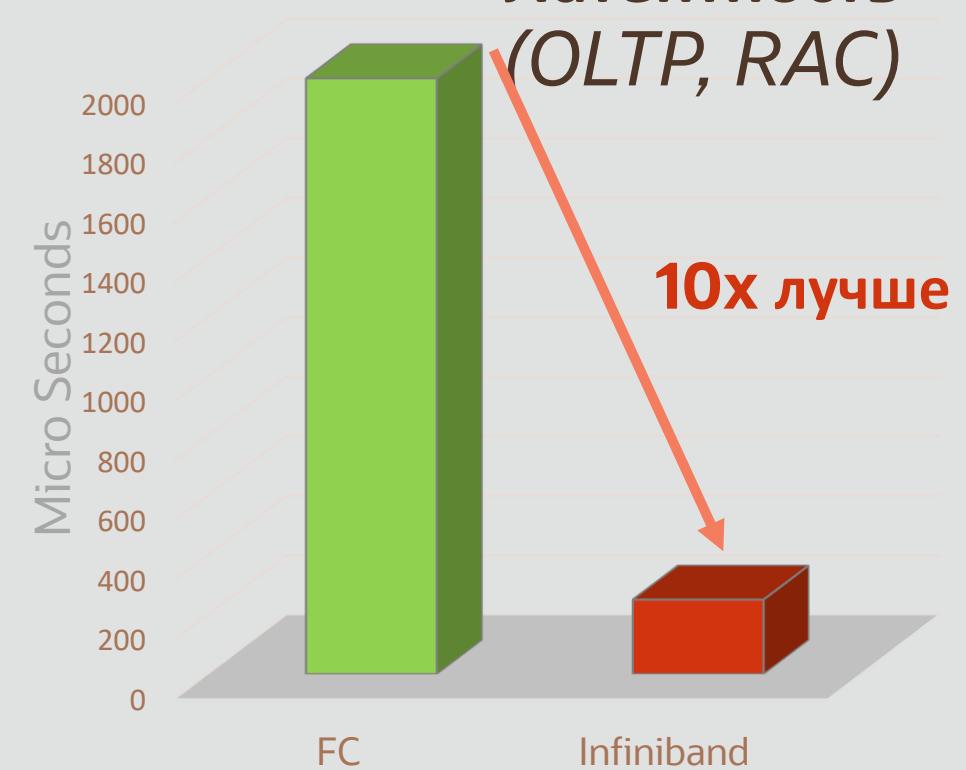
Почему был выбран Infiniband?

вспомним 2008-2009гг

Пропускная способность (Аналитика)

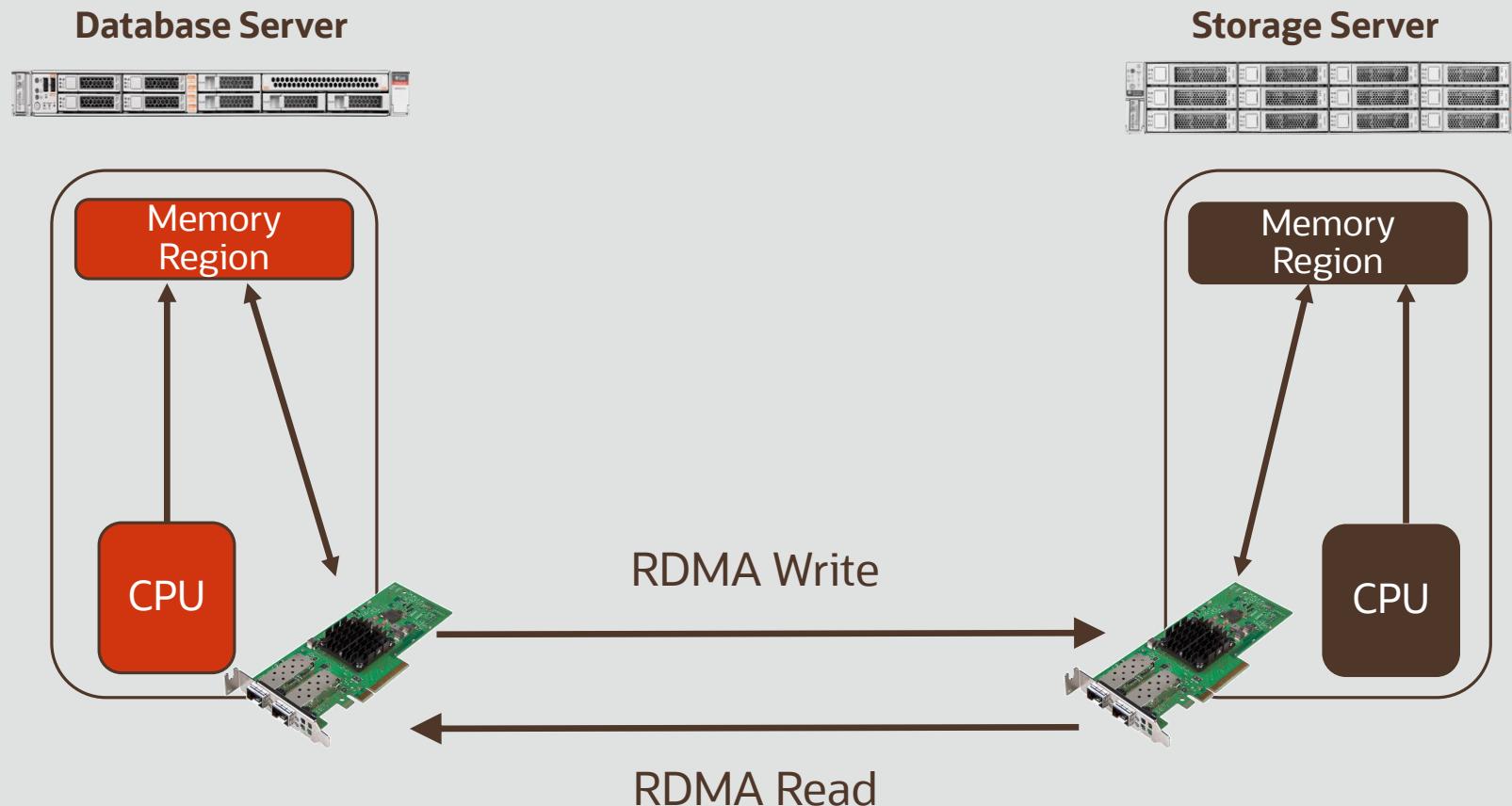


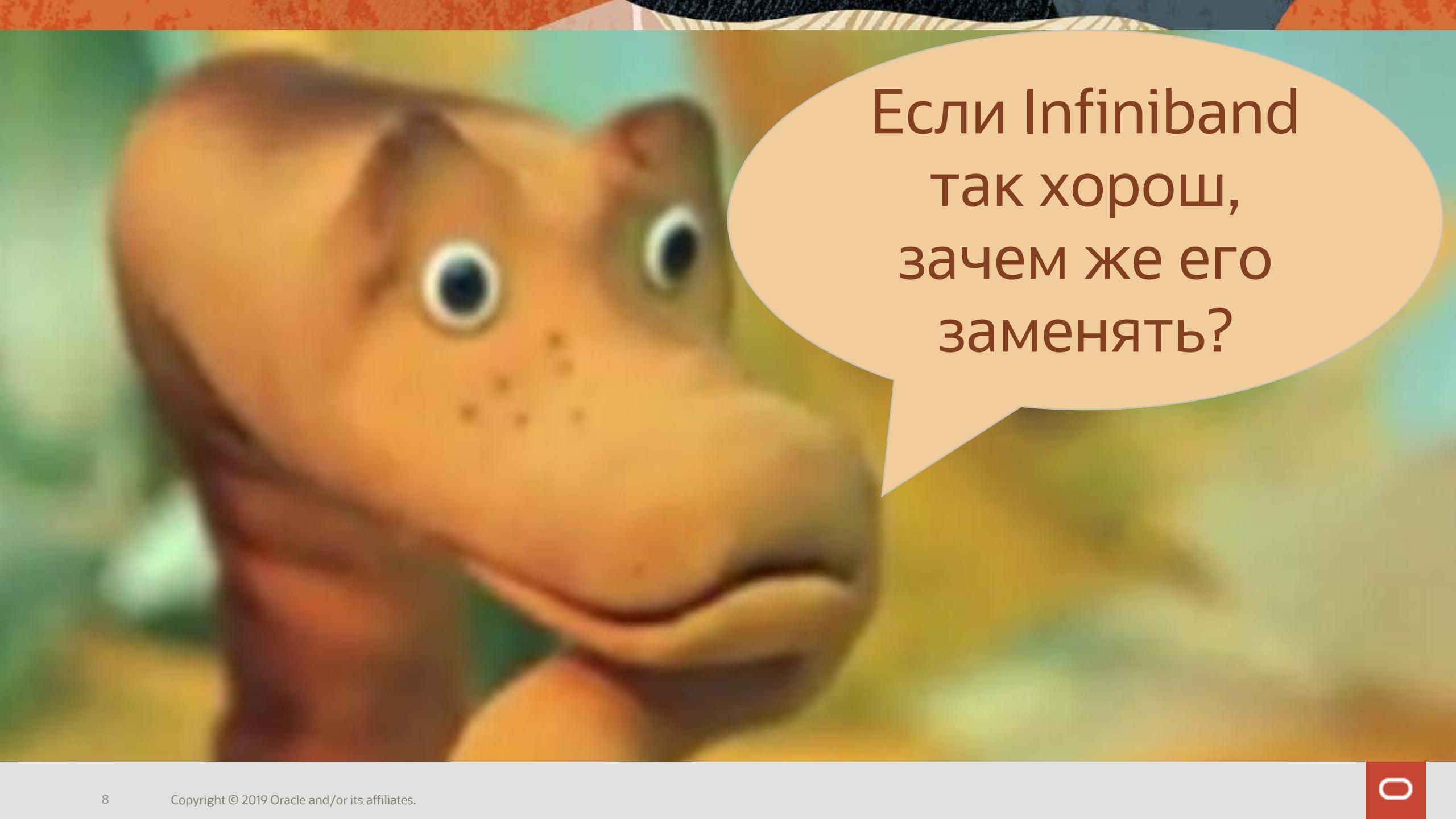
Латентность (OLTP, RAC)



Почему был выбран Infiniband?

сверхнизкая латентность – заслуга RDMA (Remote Direct Memory Access)

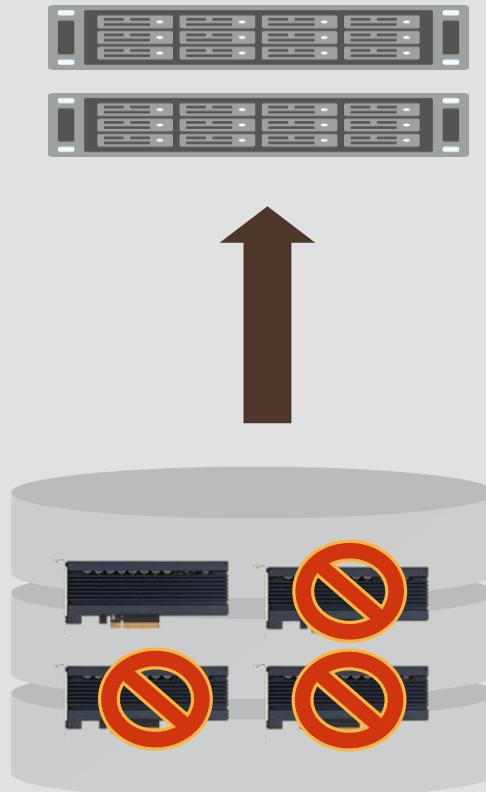




Если Infiniband
так хорош,
зачем же его
заменять?

Сегодня флэш создаёт ГИГАНТСКОЕ узкое место

- Один NVMe-носитель насыщает канал 40 Gbit
- Пропускная способность остальных носителей **теряется**



носителей **теряется**
NVMe Flash



5.8 GB/sec

40Gb/s Link



5 GB/sec

Как удалось добиться прорыва в Exadata X8M?

V1	V2	X2	X3	X4	X5	X6	X7	X8	X8M
									
Sep 2008 Xeon E5430 Harpertown	Sep 2009 Xeon E5540 Nehalem	Sep 2010 Xeon X5670 Westmere	Sep 2012 Xeon E5-2690 Sandy Bridge	Nov 2013 Xeon E5-2697v2 Ivy Bridge	Dec 2014 Xeon E5-2699v3 Haswell	Apr 2016 Xeon E5-2699v4 Broadwell	Oct 2017 Xeon 8160 Skylake	Apr 2019 Xeon 8260 Cascade Lake	Sep 2019 Xeon 8260 Cascade Lake

Flash Cache (TB)	0	5.3	5.3	22.4	44.8	89.6	179.2	358	358	358
Scan Rate (GB/s)	14	50	75	100	100	263	301	350	560	560
Read IOPS (M)	.05	1	1.5	1.5	2.66	4.14	5.6	5.97	6.57	16

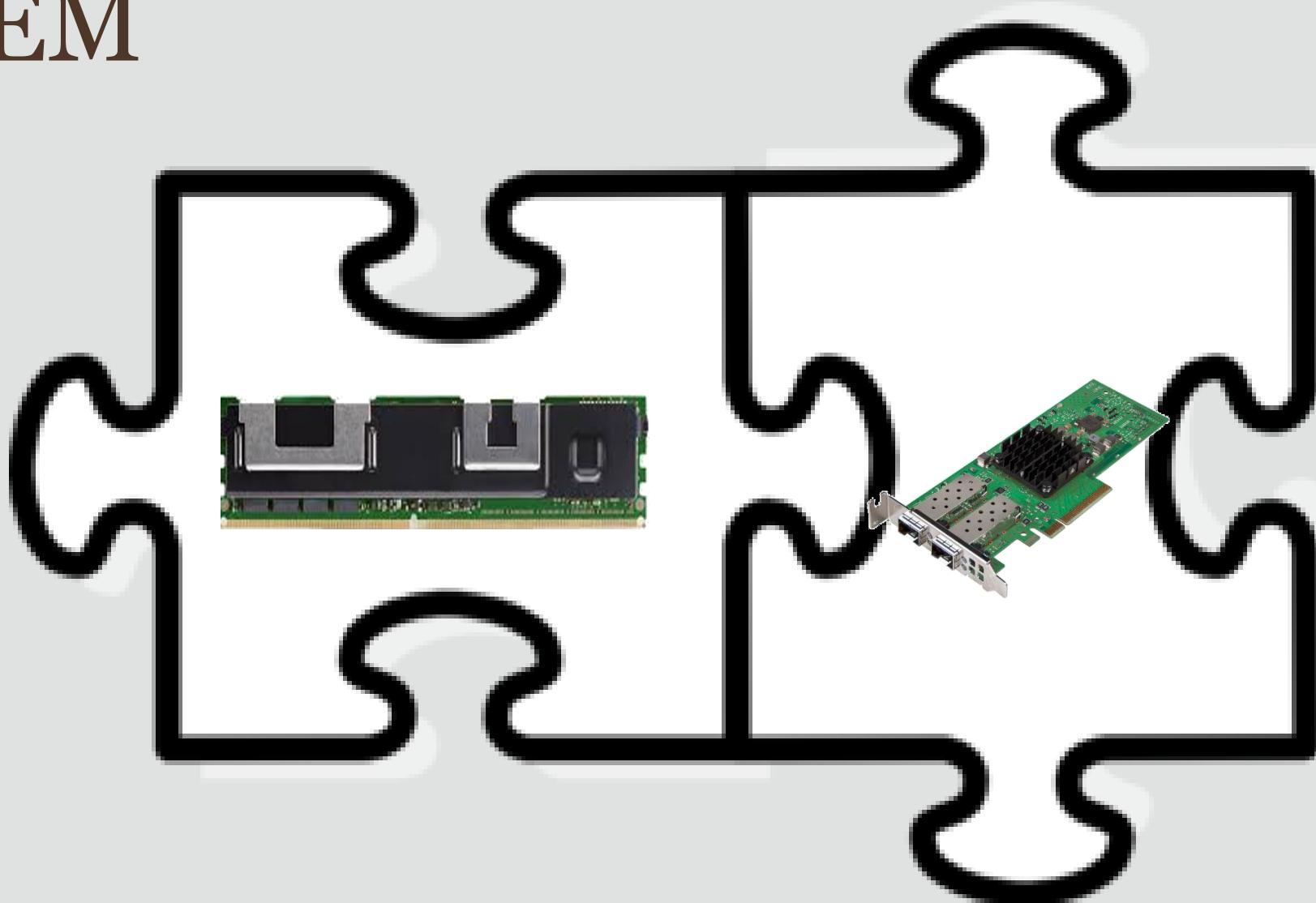


100 GbE Ethernet с RoCE – замена Infiniband

- RoCE = RDMA over Converged Ethernet – протокол с InfiniBand RDMA API на основе Ethernet
 - то же ПО на верхнем уровне сетевого стека
 - сочетает скорость Infiniband и гибкость IP
- Сохраняются все оптимизации Exadata на основе RDMA
- Стандарт разработан InfiniBand Trade Association (IBTA)
 - Разрабатывается как Open Source
 - Широко поддержан производителями сетевого оборудования

Layer	RoCE	InfiniBand
Application	User Application	
	Transport (InfiniBand)	
Network	IP Network	InfiniBand Network
	Ethernet	InfiniBand

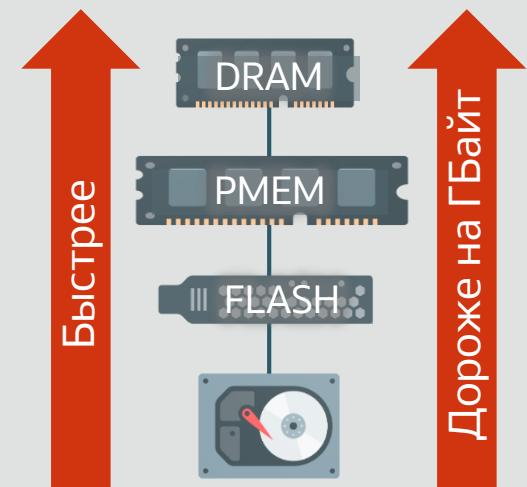
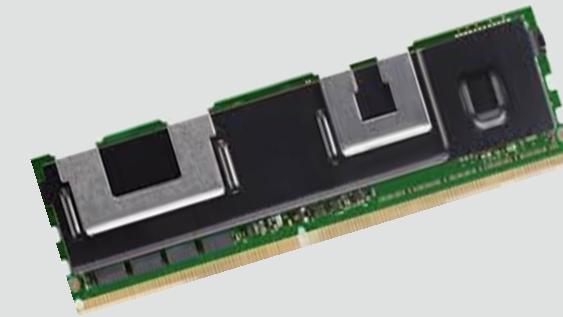
Перейдём к второй части пазла -
РМЕМ



Энергонезависимая память

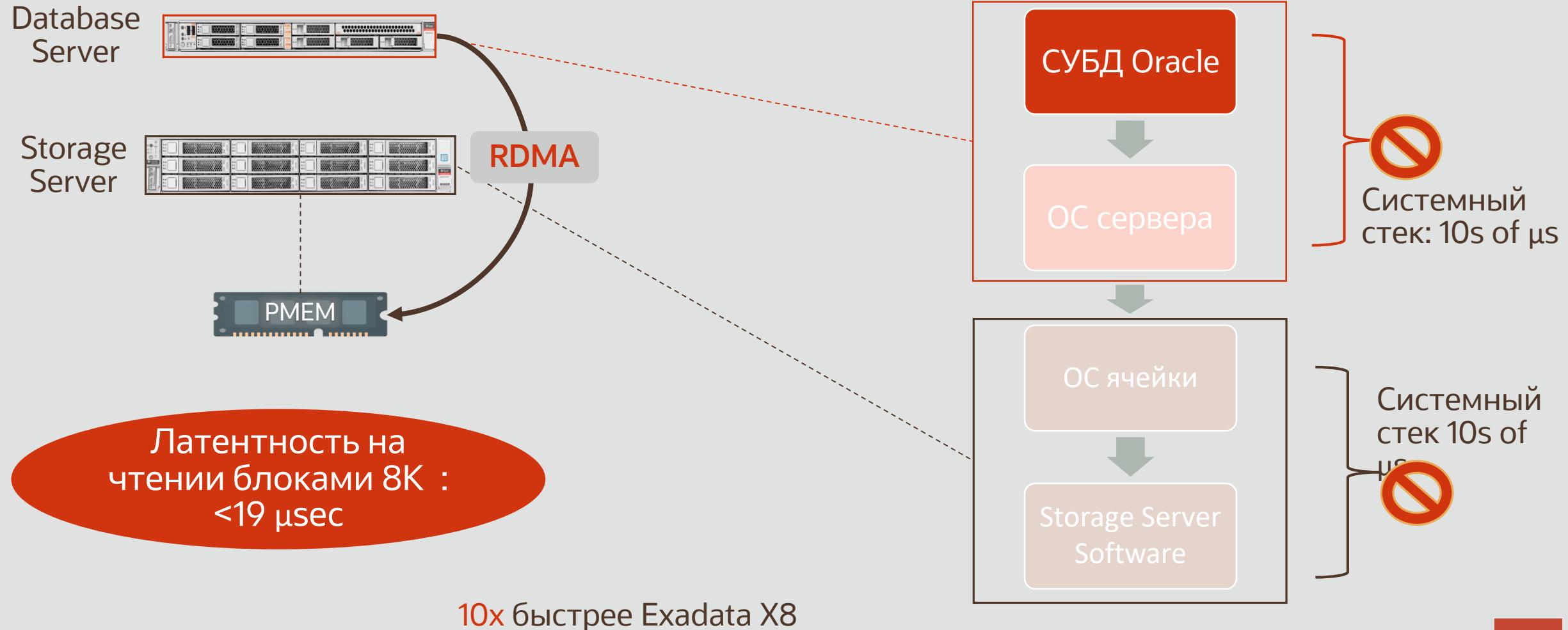
PMEM = Persistent MEMory

- Энергонезависимая память
 - По ёмкости и цене между динамической памятью и флэш
 - По производительности близка к динамической памяти
- Intel® Optane™ DC Persistent Memory:
 - Операции чтения/записи на скорости, почти как у памяти:
~50-100 быстрее флэш
 - Отключение питания переносит как флэш
- Обладает интерфейсом для работы из приложения (СУБД Oracle)

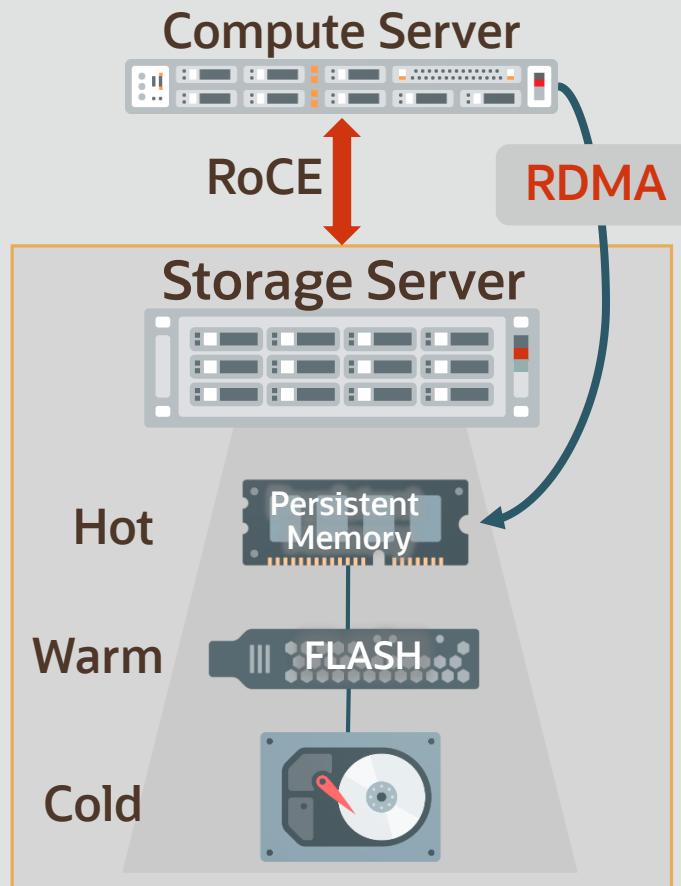


PMEM Cache

Как исключение лишнего системного стека ускоряет I/O



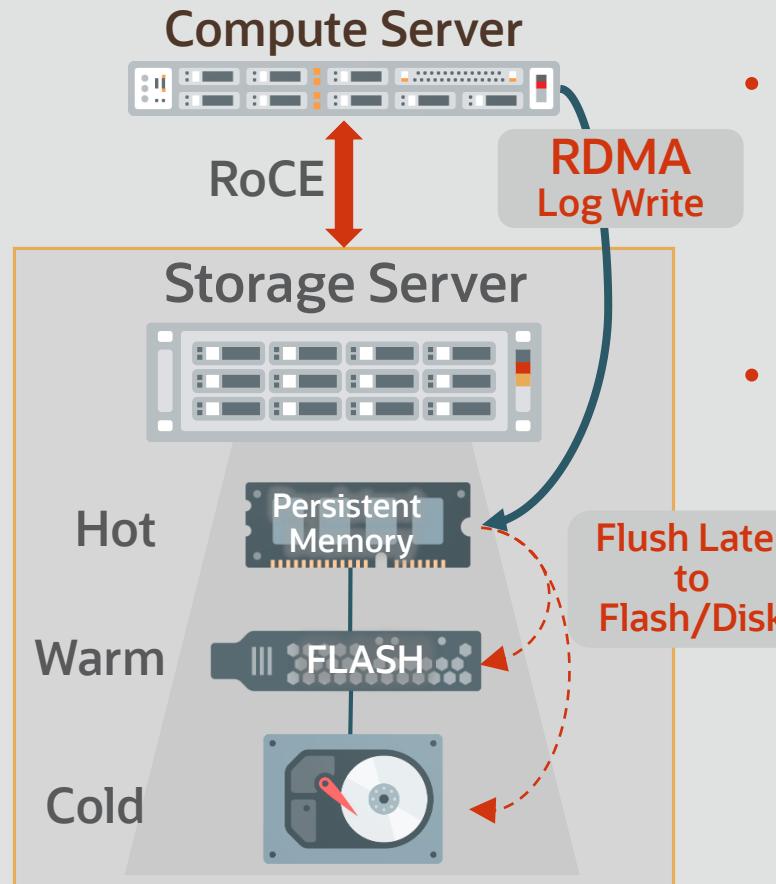
Производительность в зависимости от релиза СУБД



- DB 19 (включая GI 19) + Exadata System SW 19.3 с включенным PMEMCache
 - Латентность ~ 19usec
- DB 18, 12, 11.2 + Exa ПО 19.3 Cell с PMEMCache
 - ~100 usec
- Без использования PMEMCache
 - ~200 usec

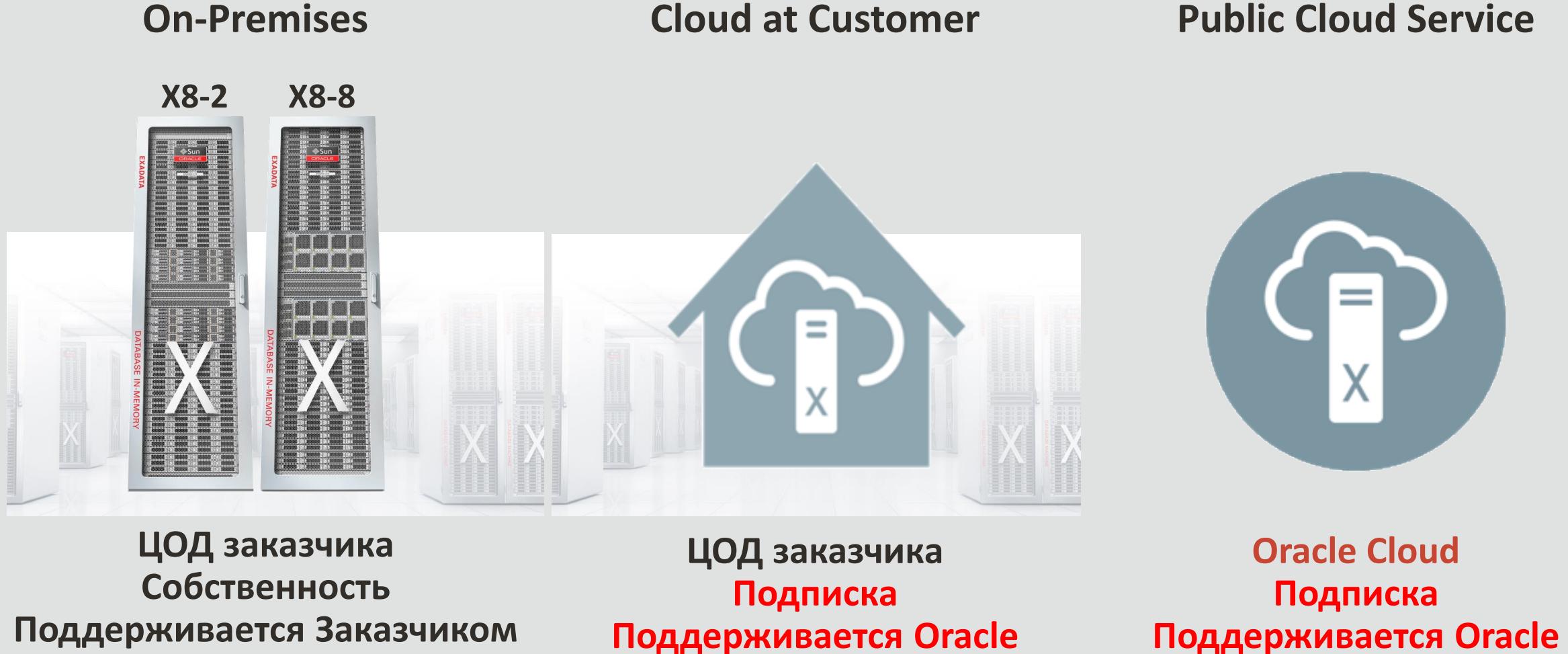
Exadata PMEM Log

ускорение на аналогичных принципах – исключение
системного стека



- Быстрая запись REDO Log критична для производительности OLTP приложений
- PMEM + RDMA ускоряют запись REDO Log
 - СУБД напрямую работает с PMEM, не обращаясь даже к системному ПО ячейки
 - Ускорение ~8 раз

Модели развёртывания Exadata



Представляем: Gen 2 Exadata Cloud at Customer

Простота публичного облака и эластичность в вашем ЦОД

Gen 2 Exadata Cloud at Customer

DB как сервис в
вашем ЦОД

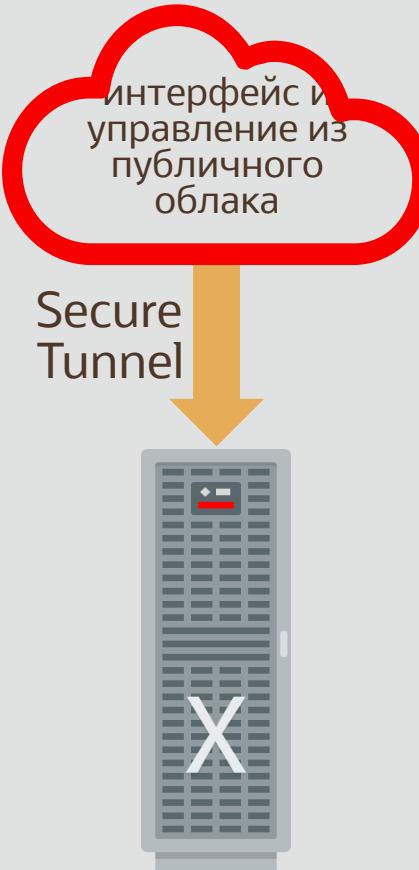
Тоже
оборудование,
ПО, API, что и в
публичном
облаке

Операционная
модель
публичного
облака

Финансовая
модель
публичного
облака

Бесшовная
совместимость с
публичном
облаке

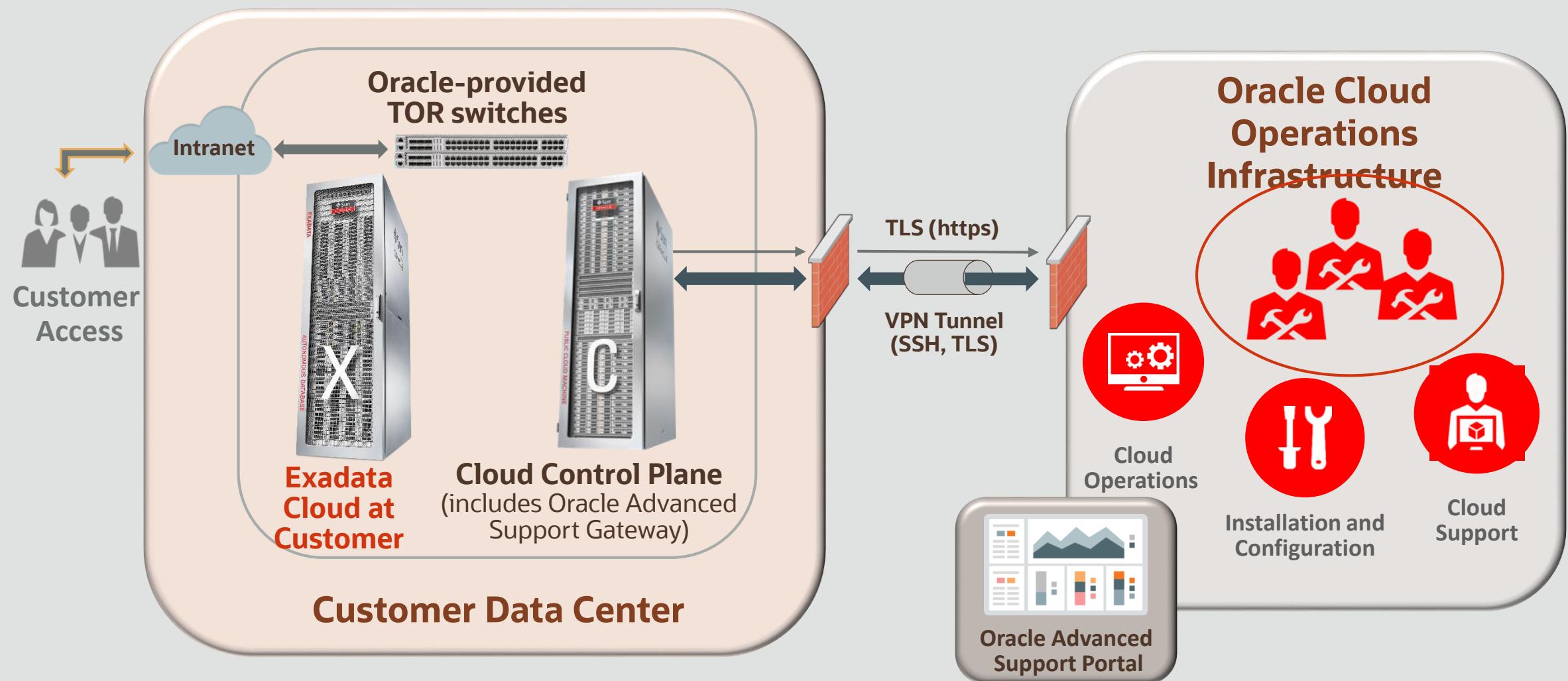
Gen 2 Exadata Cloud at Customer. Что нового



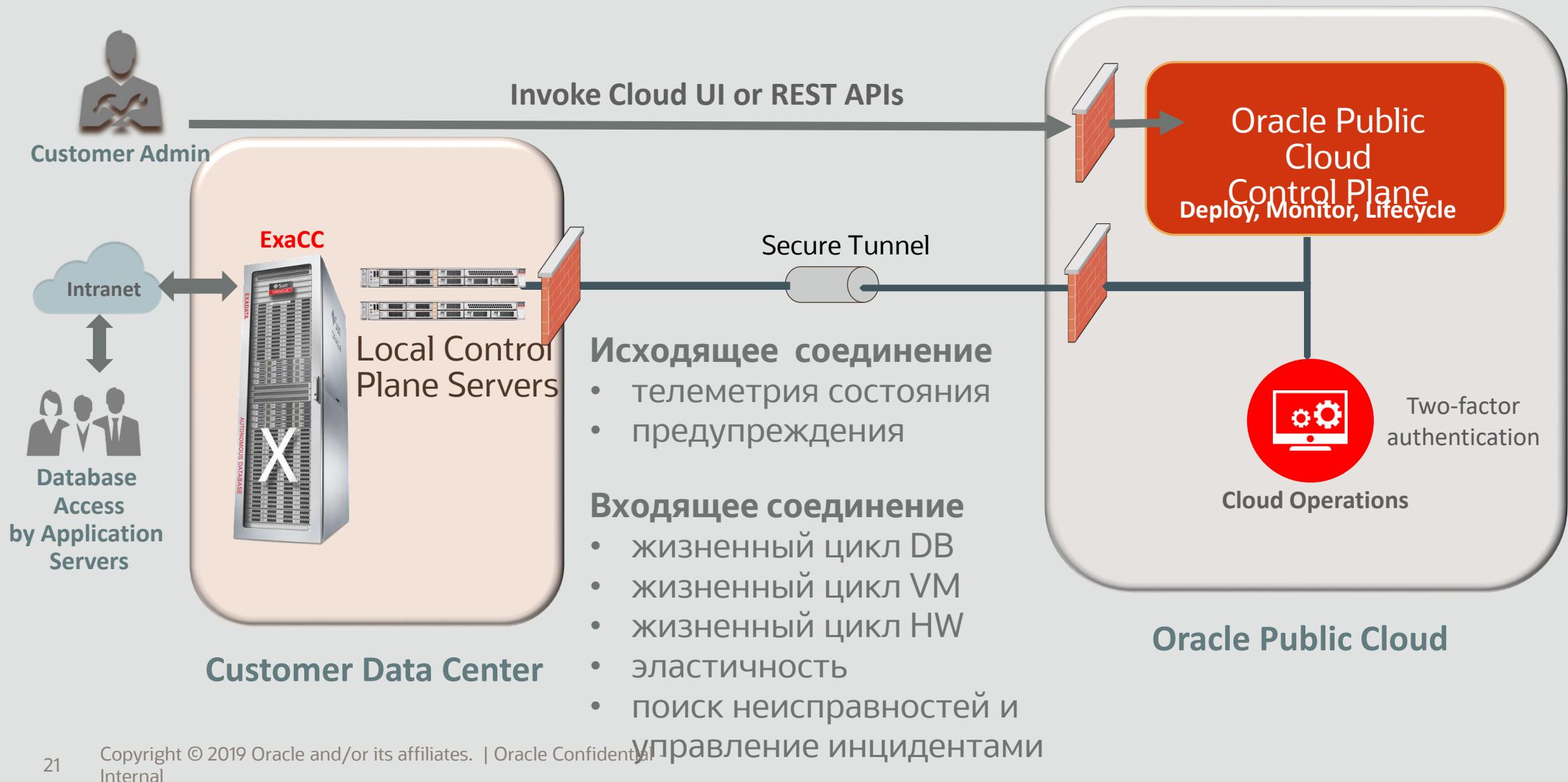
- Управление из public cloud Gen 2
 - Исключается дополнительный шкаф в ЦОД заказчика
 - Проще, ниже стоимость, ниже ТСО
- Новое оборудование
 - Быстрее CPUs, больше ядер, больше емкость, чем в ExaCC X7
- Проще соединение к ЦОД клиента
 - Адаптация к стандартам и требованиям заказчика
- Поддерживается Oracle Database 19c
 - Долгосрочная поддержка 12.2 family
- Готовность к Autonomous Database at Customer

Лучшая БД на лучшей платформе в лучшем облаке в вашем ЦОД

Gen 1 Exadata Cloud at Customer—управление



Gen 2 Exadata Cloud at Customer— управление

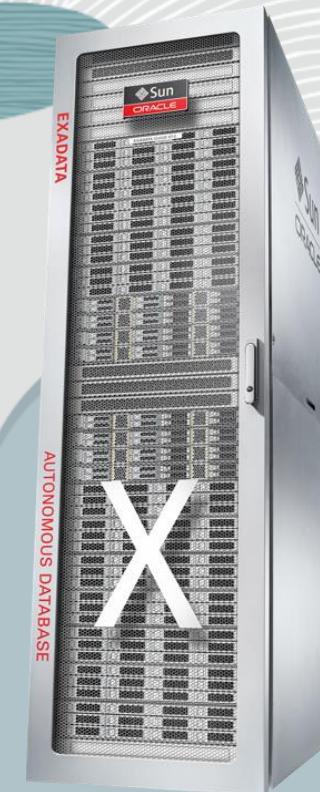


Gen 2 Exadata Cloud at Customer— преимущества перед Gen 1

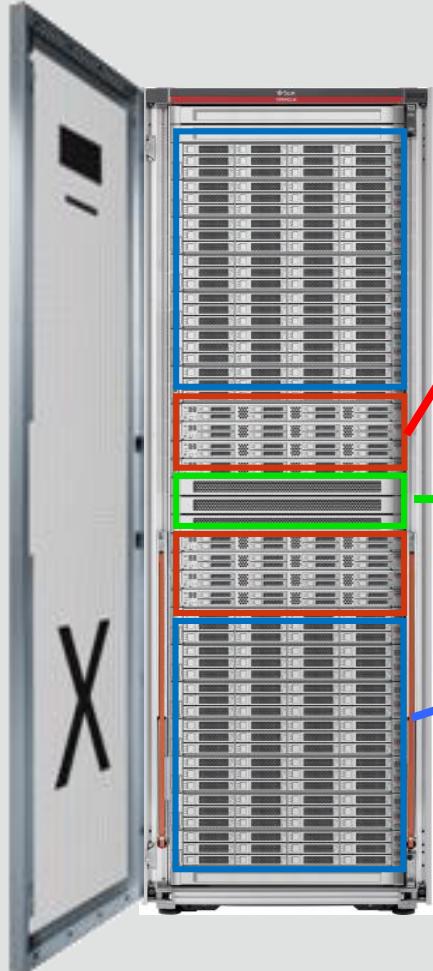
- **Простая инсталляция**
 - нет локальной Control Plane
- **Простое управление**
 - консолидированное управление системами и базами по публичному облаку и C@C
- **Простой контроль безопасности**
 - Консолидированный административный контроль consolidated fine-grained administrative security controls across public cloud and cloud at customer
- **Простой опыт использования**
 - Simpler User Experience - Identical user experience and lifecycle APIs between Gen 2 Public Cloud and Gen 2 Cloud at Customer
- **Exadata X8 оборудование**
 - Больше ядер, больше емкость, больше ядер в сторидж ячейках (даже чем в on-prem)

Gen 2 Exadata Cloud at Customer

аппаратные детали



Gen 2 Exadata Cloud at Customer X8 Hardware*



Масштабируемые 2-х сокетные сервера

Новейшие Intel Cascade Lake **26 ядер CPU**

50 ядер доступно для пользовательской VM

720ГБ памяти доступно на DB сервер для пользовательской VM

По умолчанию в 2 раза больше памяти, чем on-prem Exadata

Сверх быстрая IB фабрика

Масштабируемые умные 2-х сокетные СХД сервера

новейшие **24 ядерные** Intel Cascade Lake CPU

50% больше ядер, чем в Exadata on-prem для вынесения вычислений с DB узлов

25.6TB Flash, 12x 14TB дисков на сторидж сервер

Анонс : Zero Data Loss Recovery Appliance X8M

RPO = 0, RTO -> min

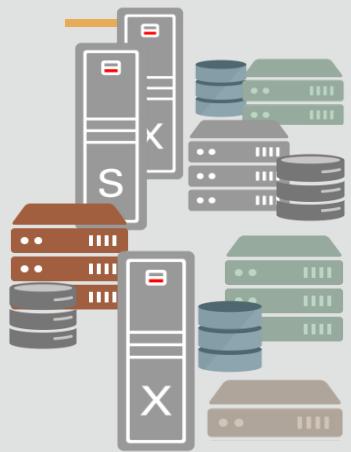
- Единственное СРКиВД созданное, как расширение СУБД Oracle
- ZDLRA X8M построена на базе Exadata X8M и использует 100 Gb/sec RoCE для высокоскоростной передачи данных между вычислительными узлами и ячейками хранения
- Максимально быстрое восстановление в любой точке
- Практически не нагружет продуктивную систему
- Автоматизирует все процессы резервного копирования и восстановления
- Очень выгодная совокупная стоимость владения!



Recovery Appliance X8M

Zero Data Loss Recovery Appliance X8(M)

Защищаемые БД



Delta Push

- БД отправляют только изменения
 - Минимальное влияние на продуктивные БД
- Передача журналов транзакций в режиме

Recovery Appliance



Резервное копирование на ленты



Защита всех БД Oracle в ЦОД

- Петабайты данных
- Версии Oracle 10.2 - 19c, все редакции, любые платформы
- Поддержка архива на Object Storage (S3)**
- Агенты для RMAN

Delta Store

- Хранятся на диске проверенные, сжатые изменения БД
- Быстрое восстановление на любой момент времени используя дельты
- Масштабируемость и эластичность, как в Exadata
- Управление с помощью Enterprise Manager

Репликация на Recovery Appliance в резервном ЦОДе

Oracle Private Cloud Appliance

- ✓ 6 поколение Private Cloud Appliance X8-2
 - на **45%** быстрее чем X7
 - **100 GbE** внутренняя сеть
 - Встроенный ZS7-2 – **17x** больше пространства

ORACLE®

Linux

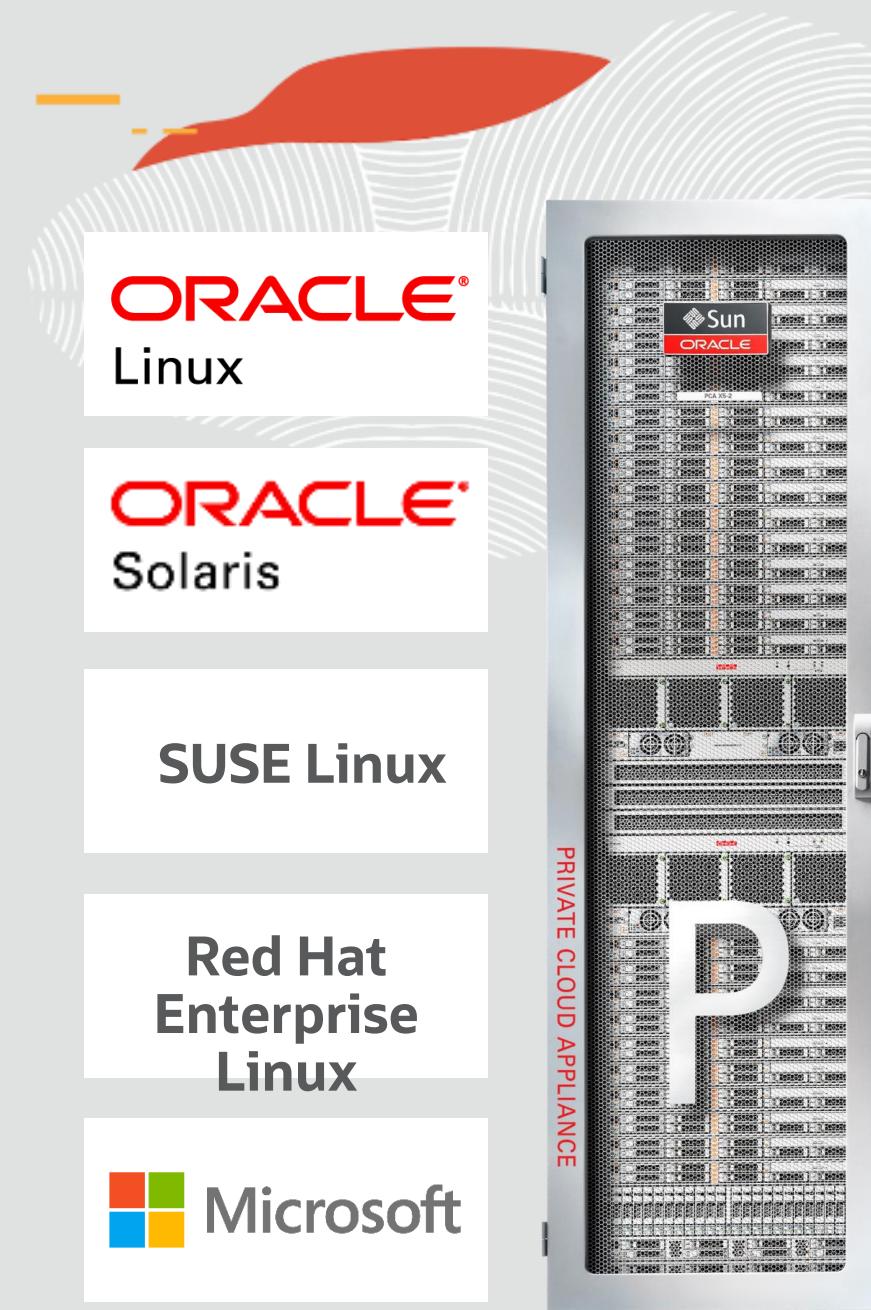
ORACLE®

Solaris

SUSE Linux

Red Hat
Enterprise
Linux

Microsoft



Oracle Private Cloud Appliance X8-2

Вычислительные узлы

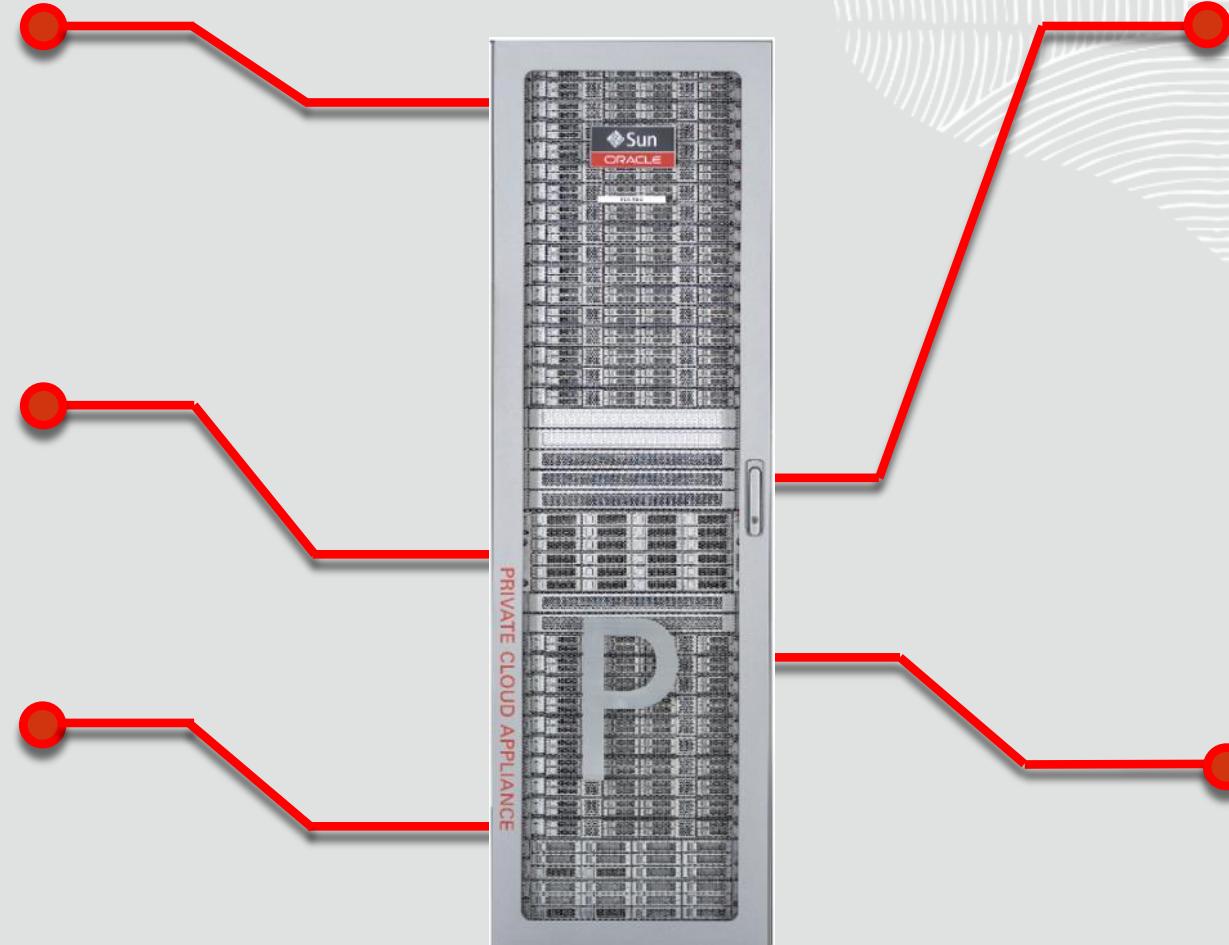
До 25 x Oracle X8-2 Server
2 x Intel Xeon 8260 24 Core 24 GHz
384 – 768 - 1536 GB Memory
1 x Dual Port 100 Gb Ethernet Port

Интегрированная СХД

2 X ZS7-2 Clustered Controllers
1 x DE3-24C полка с 100 TB usable
Max 2PB

Серверы управления

2 x Oracle X8-2 Server
2 x Intel Xeon Gold 5218 16 Core 2.3 GHz
384 – 768 - 1536 GB Memory
2 x 100 Gb Ethernet



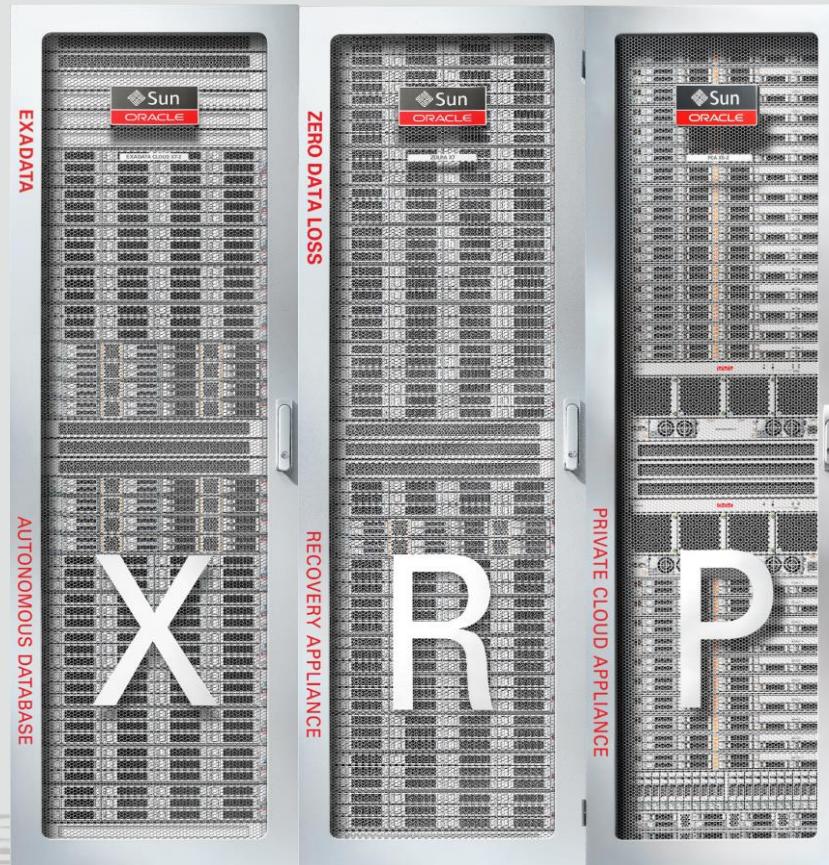
Сеть SDN

100/40/25/10 Gb Ethernet data center connectivity
2 x Leaf Switches
2 x Spine Switches
Fully orchestrated Software defined Networking

Гипервизор & управление

Oracle VM 3.4.6
Enterprise Manager 13c

Oracle Systems: как в облаке, так и в ЦОД



ates
Con
fide



Стройте своё
облако

Задавайте вопросы, пожалуйста!

Больше информации:

- oracle.com/exadata
- blogs.oracle.com/exadata



ORACLE