ИЗДАЕТСЯ КОМПАНИЕЙ «ИНФОСИСТЕМЫ ДЖЕТ»

Jet Info

WWW.JETINFO.RU

ТЕМА НОМЕРА

НОВАЯ РЕАЛЬНОСТЬ – КАК ИТ УСПЕТЬ ЗА БИЗНЕСОМ

№ 5 (273)/2016

Jet Info

Издается с 1995 года

Редакция:
Акимова Н.А., Дискина А.Л.,
Дмитриев В.Ю., Шедова Е.Л.

Дизайн и верстка:
Саблина М.А.

Корректура:
Макеева Е.И.

Иллюстрации:
www.shutterstock.com

Над номером работали:
С. Ю. Андронов, И.М. Воронин,
Д.В. Глушенок, А.Ю. Лаврентьев
Ю.А. Семенюков, А.Л. Струченко

Издатель:
Компания «Инфосистемы Джет»

Контакты:
тел.: (495) 411-76-01
info@jet.msk.su,

www.jetinfo.ru

ИЛЬЯ ВОРОНИН,
директор Центра проектирования вычислительных комплексов компании «Инфосистемы Джет»

Бизнес сегодня подвержен стремительным изменениям. Это обусловлено как внешними макроэкономическими обстоятельствами, а проще говоря, кризисом, так и естественными факторами – ростом конкуренции, желанием увеличить клиентскую базу, повысить рентабельность. Компании стремятся минимизировать time-to-market для своих новых продуктов и услуг, максимально оперативно менять географию филиальных сетей, осуществлять слияния/поглощения и т.д.

Естественно, информационные технологии должны быть способны поддерживать все происходящие изменения. Причем речь чаще всего идет не о месяцах и даже не о неделях: время, которое дается ИТ-ландшафту на трансформацию в соответствии с поступающими от бизнеса запросами, исчисляется буквально днями. Видя потребности рынка, в сегодняшнем номере мы рассматриваем ряд инновационных технологий, которые позволяют ИТ-инфраструктуре соответствовать этой новой реальности. Мы одними из первых в России протестировали их в собственных лабораториях и делимся с вами полученными результатами.

Содержание

5

8

16

22

3 **От редакции**

5 **Тема номера**
ИТ на полной скорости
ИЛЬЯ ВОРОНИН, СЕРГЕЙ АНДРОНОВ

8 **Тема номера**
Анализ производительности сервера Oracle SPARC T7-2
АЛЕКСЕЙ СТРУЧЕНКО

12 **Тема номера**
Изучение Oracle In-Memory и SQL in Silicon
АЛЕКСЕЙ СТРУЧЕНКО

16 **Тема номера**
Пора ли переходить на All-Flash СХД?
АЛЕКСЕЙ ЛАВРЕНТЬЕВ

22 **Тема номера**
Обзор функционала SDS
ДМИТРИЙ ГЛУШЕНОК

27 **Тема номера**
Серверная платформа Huawei KunLun – выводы тестирования
ДМИТРИЙ ГЛУШЕНОК

30 **Тема номера**
Тестирование функций виртуализации платформы SPARC на Oracle SPARC T7-2
ЮРИЙ СЕМЕНЮКОВ

34 **В тему номера**

ИТ НА ПОЛНОЙ СКОРОСТИ

Сегодня, чтобы идти в ногу со временем, нужно брать ноги в руки

Писатель Леонид Сухоруков

ИЛЬЯ ВОРОНИН,
директор Центра проектирования вычислительных комплексов компании «Инфосистемы Джет»

СЕРГЕЙ АНДРОНОВ,
директор Центра сетевых решений компании «Инфосистемы Джет»

Мы стали жить очень динамично: на работе и на связи практически 24 часа в сутки, совершаем ночные покупки в круглосуточных супермаркетах, проверяем состояние своих финансов через интернет-банк, находясь в другой стране/на другом континенте. Технологии, чье развитие в свое время обусловило высокий темп жизни, теперь вынуждены ему соответствовать, а иногда и догонять. Естественно, бизнес тоже не избежал подобной участи. Новые услуги и сервисы должны выводиться на рынок с минимальными временными задержками, поскольку конкуренты постоянно дышат в затылок и норовят обойти на повороте в гонке за клиента. То же касается технического обеспечения слияний и поглощений компаний: банков, операторов связи, «промышленников» и т.д.

Происходящие экономические кризисы только подливают масла в огонь, особенно характерен в этом смысле последний, в реалиях которого мы живем до сих пор. Резкое изменение цен на нефть и валютные скачки заставляли бизнес оперативно пересматривать линию поведения: банки были вынуждены корректировать процентные ставки по вкладам после каждого повышения/снижения ключевой ставки ЦБ РФ, ритейлеры меняли поставщиков и логистические цепочки в соответствии с политикой импортозамещения. ИТ на своем уровне должны поддерживать эти пертурбации. Другой вопрос – могут ли они это сделать?

С точки зрения крупного корпоративного бизнеса последние 5–7 лет шло постоянное вливание средств в ИТ-инфраструктуру и реализацию ИТ-проектов. Это обеспечивало гарантированный результат – поддержку возникающих

бизнес-задач – с временными задержками от нескольких месяцев до полугода-года. Соответственно, у топ-менеджмента за это время сформировалось ощущение, что ИТ-ландшафт более чем насыщен и способен поддержать возможные, в том числе кризисные, изменения. Но, как говорится, не тут-то было. Временных задержек при классическом подходе к построению и развитию инфраструктуры избежать не удастся. Но лишних нескольких месяцев на ожидание (согласование, поставка оборудования и ПО, инсталляция, миграция и др.) у бизнеса теперь нет – есть максимум несколько дней. Таким образом, первое требование, предъявляемое бизнесом к ИТ, – высокая скорость их изменения.

Теперь вернемся к истории о слиянии/поглощении компаний. Мы в данный момент участвуем в нескольких подобных проектах и первое, что говорят финансисты на общих встречах: «Мы не готовы существенно инвестировать в объединенную инфраструктуру – перестраивать сеть, модернизировать СХД, переходить на новую серверную платформу. Нам нужно максимально быстро навести порядок в объединенном ИТ-хозяйстве – получить единую прозрачную инфраструктуру, растянутую на обе компании». То есть второе требование к ИТ – это сохранность прошлых инвестиций, сохранение КПД – высокой утилизации – используемых технологий.

Отметим, что параллельно с изменением самого восприятия ИТ происходит смена ИТ-управленцев. Старую гвардию, привыкшую мыслить в классической парадигме, постепенно сменяют новые кадры – поколение Y, а в скором времени их сменит поколение Google. Они не вдаются в технические вопросы, в первом приоритете у них – именно бизнес-задачи и их своевременное выполнение. Если ИТ не могут это обеспечить, значит, они не эффективны и нуждаются в изменении. В каком именно?

НОГИ В РУКИ И ВПЕРЕД

Если говорить упрощенно, бизнесу нужен универсальный «интернет-магазин» ИТ-ресурсов под управлением ИТ-подразделения. Бизнес стремится внедрять инновационные изменения для обеспечения своего динамичного развития и конкурентоспособности – минимизации time-2-market

и вывода продуктов/сервисов на рынок раньше «коллег». Инновации, в свою очередь, требуют высокой скорости изменений в ИТ, а это подразумевает использование гибкой и адаптивной ИТ-инфраструктуры. О принципах ее построения и функционирования мы и поговорим ниже.

Это должна быть легко кастомизируемая среда выполнения приложений, позволяющая не думать о «железном» уровне – как устроен вычислительный комплекс, организован backup, осуществляется мониторинг и т.д. Речь идет о программно-аппаратном комплексе, предоставляющем набор базовых инфраструктурных сервисов и создающем среду выполнения, гарантирующую заданный уровень производительности и доступности для размещённых в ней бизнес-приложений. Комплекс должен обеспечивать автоматизацию типовых задач по предоставлению и учёту ИТ-услуг, а также иметь встроенные средства мониторинга и самодиагностики, работающие на базе заложенной в него «модели здоровья».

Для бизнес-пользователей система выглядит как корпоративное частное облако, при размещении в котором приложения автоматически кластеризуются, интегрируются с СРК, «встают» на мониторинг, проходят плановое обслуживание и т.д. При этом пользователи получают доступ ко всем высокоуровневым инструментам взаимодействия с платформой и регулярную отчетность о выполнении SLA, текущей утилизации и трендах роста нагрузок.

Поскольку подобная облачная платформа обеспечивает параллельное выполнение большого количества бизнес-приложений, разделяющих общий набор физических и логических ресурсов, уровень утилизации может составлять 70–75%.

Отметим, что такой комплекс можно построить на основе общедоступного оборудования и ПО ведущих мировых производителей/программных компонент с открытым исходным кодом. Ключевые аппаратные элементы комплекса общеизвестны – это собственно вычислительная платформа, СХД, сетевая инфраструктура и средства обеспечения информационной безопасности.

Набор системного ПО облачной платформы обеспечивает:

1. Базовые инфраструктурные сервисы
2. Сервисы обеспечения высокой доступности и катастрофоустойчивости вычислительной инфраструктуры
3. Сервисы защиты данных от потери, повреждения или недоступности в случаях физического или логического повреждения или ЧС
4. Сервисы автоматизации предоставления ИТ-услуг (к примеру, создания и обновления тестовых сред на основе данных продуктивных окружений с автоматическим обезличиванием)
5. Сервисы мониторинга использования ресурсов и отчётности
6. Сервисы контроля и самодиагностики

Таким образом, инструментарий для оперативного предоставления ИТ-ресурсов бизнес-пользователям представляет собой комплексную систему, включающую десятки различных компонент и подразумевающую наличие сотен связей между ними. Поэтому в данный момент разработка, внедрение и обслуживание такой системы требуют наличия разнообразных компетенций у ИТ-специалистов – «операторов» облачной платформы, осуществляющих поддержку низкоуровневых сервисов.

В части разработки речь в первую очередь идет о создании набора алгоритмов – сценариев, по которым ИТ-услуги из «интернет-магазина» будут предоставляться конечным пользователям, а также web-интерфейса взаимодействия пользователей с системой. На рынке пока нет готовых решений, в которых воплощен подобный функционал. И это поле битвы для системных интеграторов. Так, мы сейчас участвуем в нескольких проектах, в которых требуется реализовать логику работы сценариев предоставления ИТ-ресурсов разным группам пользователей. По нашим ощущениям, в ближайшие пару лет в этой сфере произойдет качественный скачок, который приведет к повсеместному использованию описанной сервисной модели. А пока в статьях нашего номера мы рассмотрим отдельные инфраструктурные решения и компоненты, позволяющие претворить в жизнь описанную нами концепцию. JI

АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ СЕРВЕРА ORACLE SPARC T7-2

АЛЕКСЕЙ СТРУЧЕНКО,
руководитель Отдела
оптимизации СУБД и приложений
компании «Инфосистемы Джет»

Одной из самых важных новостей компании Oracle в 2015 году стал выход нового процессора SPARC M7 (далее – M7) и линейки серверов на его основе: в нее вошли серверы T-серии (T7-1, T7-2, T7-4) и серверы M-серии (M7-8, M7-16).

Помимо уникальных физических характеристик на процессоре M7 заявлена возможность переноса части SQL-логики базы данных Oracle на специальные сопроцессоры DAX (технология SQL in Silicon). Таким образом, процессор M7 позиционируется как первый процессор в истории ИТ, оптимизированный под задачи Oracle Database.

В Табл. 1 приведены основные характеристики процессора M7 в сравнении с процессором T5 (предыдущее поколение SPARC). В новом процессоре при частоте 4,13 гГц 32 ядра и до 256 потоков.

В начале 2016 года стало возможно тестирование серверов T-серии – наша компания одной из первых в России организовала такое тестирование в феврале-марте 2016 года. В тестировании были задействованы 2 сервера T7-2 (один из них – на территории заказчика) со следующими ресурсами:

- два процессора SPARC M7 (всего 64 ядра и до 512 потоков);
- до 1 Тб оперативной памяти;
- Solaris 11.3 в контрольном домене, различные версии Solaris в гостевых доменах, различные версии Oracle.

Тестирование было организовано в несколько этапов:

1. анализ производительности сервера T7-2;
2. исследование новых возможностей виртуализации;
3. использование Oracle In-Memory и изучение технологии SQL in Silicon.

Стоит отметить, что целый класс программного обеспечения разрабатывается с целью проведения синтетических тестирований Hardware, в том числе под характерной нагрузкой со стороны базы данных. Тем не менее выбор средства тестирования является нетривиальной задачей, а зачастую именно этот выбор определяет результаты тестов и дальнейшие выводы.

При тестировании новых процессоров M7 от Oracle мы ставили перед собой задачу исследовать поведение сервера под нагрузками, близкими к максимальным. Для этого нам было необходимо свести к минимуму влияние на результаты тестов как системы ввода-вывода, так и внутренних механизмов базы данных (в данном случае Oracle Database).

Характеристика процессора	T5	M7
Количество ядер	16	32
Максимальное количество потоков	128	256
Частота (гГц)	3,6	4,13

Табл. 1. Основные характеристики процессора M7 в сравнении с процессором T5

Для решения этой задачи нами было значительно доработано ПО с открытым кодом: так, чтобы можно было фиксировать максимальные значения показателя Logical Reads Per Second (логическое чтение или чтение из памяти) экземпляра Oracle. Интенсивность логических чтений из памяти является важной характеристикой нового процессора, а само чтение данных из памяти является неотъемлемой частью функционирования Oracle Database.

На рис. 1 (см. стр. 10) приведены сравнительные результаты, полученные на серверах T5-4 (предыдущее поколение серверов Oracle) и T7-2 (новые процессоры SPARC M7). Результаты зафиксированы в одинаковых по процессорной мощности доменах (32 ядра) при одинаковых настройках экземпляра Oracle и очень близких профилях нагрузки. На этом и последующих графиках по оси *X* отложено количество параллельных сессий, по оси *Y* – максимальное количество логических чтений в секунду согласно статистике AWR.

При одинаковом числе ядер домен сервера T7 оказался в 1,15–1,2 раза производительнее домена сервера T5. Это означает, что новый процессор M7 (в котором ядер вдвое больше, чем в T5) в 2,3–2,4 раза производительнее процессоров предыдущего поколения от Oracle. Заметим, что этот результат зафиксирован на сервере T7 как в контрольном (control), так и в гостевом (guest) домене. При этом на большом количестве сессий (192 и более) в гостевом домене сервера T7 заметно влияние виртуализации: производительность оказывается на 3–5% ниже, чем в контрольном домене при тех же ресурсах.

Максимальное значение показателя Logical Reads Per Second было получено, когда контрольному домену было отдано 64 ядра

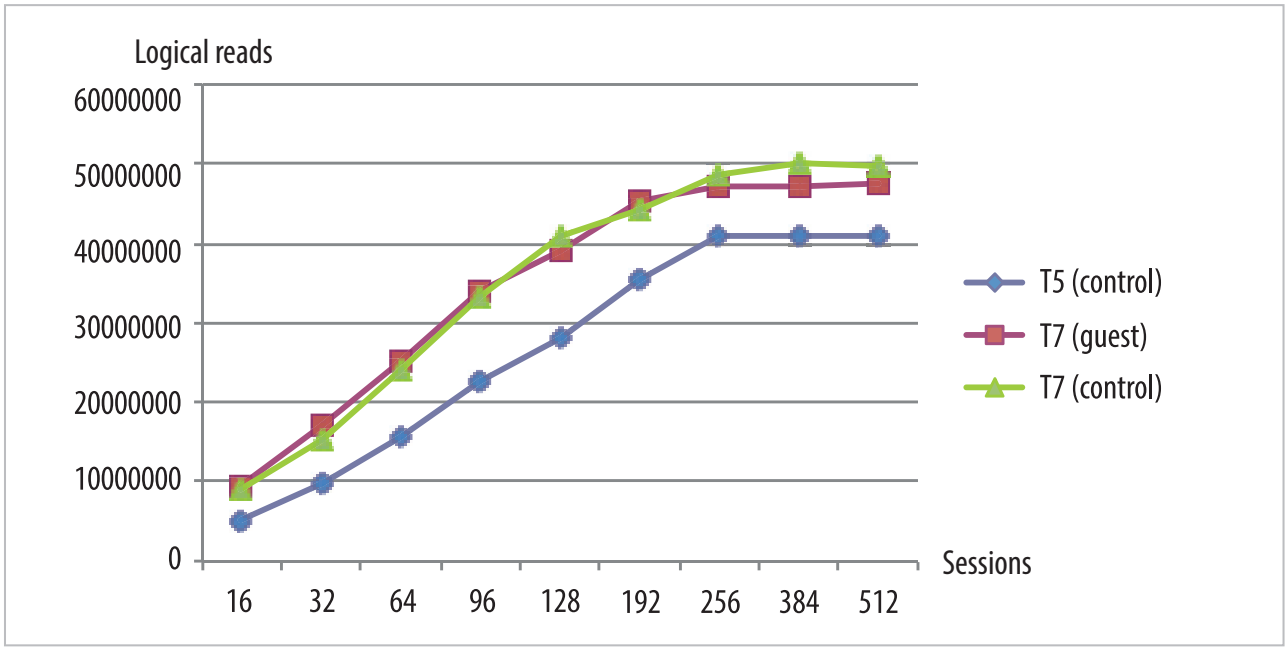


Рис. 1. Результаты в доменах 32 ядра (T5 vs T7)

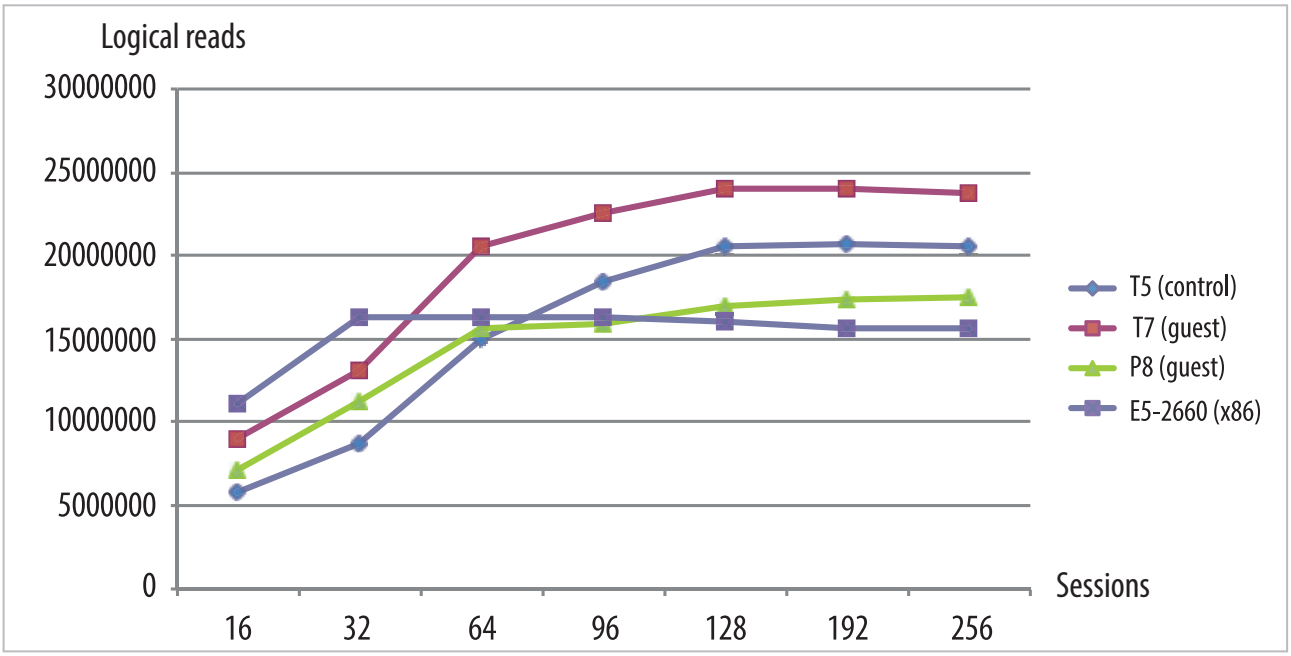


Рис. 2. Сравнительные результаты в доменах 16 ядер (T5/T7/P8/x86)

(всего 512 потоков). В этом случае практически все ресурсы сервера T7-2 (ядра и память) были отданы экземпляру Oracle, и он показал 93 587 823 логических чтений в секунду. За все время тестирования серверов различной архитектуры под нагрузкой Oracle Database мы впервые зафиксировали такие высокие показатели.

Параллельно тестированию сервера T7-2 по той же методике были протестированы актуальные серверы архитектуры IBM Power и x86. На рис. 2 приведены сравнительные результаты, полученные в гостевых доменах серверов T5-4, T7-2, Power8 и E5-2660 (x86). В тестируемые домены было отдано по 16 ядер (следует отметить, что в архитектуре x86 на ядро при-

ходится по 2 потока, притом что у SPARC и Power по 8 потоков). Максимальный результат, полученный в гостевом домене T7-2, – около 24 млн логических чтений в секунду – оказался лучшим. При этом на малых количествах сессий наиболее производительно показала себя архитектура x86.

Также следует отметить, что в ходе тестирования динамического перераспределения ресурсов на сервере T7-2 была обнаружена проблема так называемых «медленных тредов». Она проявляется в том, что некоторые потоки начинают работать в несколько раз медленнее при достаточно интенсивных изменениях ресурсов домена. По этой проблеме в службе поддержки Oracle был заведен Service Request с максимальным приоритетом, мы получили ответ: медленные потоки связаны с работой Power Management, который рекомендуется отключать. В ходе тестирования мы проверили этот Workaround – действительно, при отключении Power Management проблема «медленных тредов» не проявляется.

Мощности сервера T7-2 позволят минимизировать продолжительность технологических окон, во время которых производится расчет данных для аналитических систем. Тем самым обеспечивается доступность BI-системы, близкая к 24x7, что позволяет, например, обеспечить процесс оперативного формирования отчетности в финансовых и страховых организациях.

Можно привести пример высокой востребованности функционала по предоставлению актуальных данных для отчетной системы у одного из наших заказчиков. В «ЛИКАРДе» благодаря ему значительно ускоряется процесс сверки отчетных данных с информацией, получаемой от предприятий нефтепродуктообеспечения (НПО).

Таким образом, анализ производительности сервера T7-2 под характерной нагрузкой Oracle Database показал:

1. В сравнении с T5 (процессор SPARC предыдущего поколения) производительность ядра M7 оказалась выше в 1,15–1,2 раза, а производительность всего процессора (в котором ядер вдвое больше) – в 2,3–2,4 раза соответственно. Это позволяет говорить о том, что при консолидации баз данных Oracle сервер T7-2 способен заменить более двух серверов T5-2.

2. Абсолютный зафиксированный на сервере T7-2 результат в 93 587 823 логических чтений в секунду более чем в 2 раза превышает аналогичные результаты на других серверах (включая серверы архитектуры IBM Power и x86) – за все время тестирования под нагрузкой Oracle Database такие высокие показатели у нас получены впервые.

3. Динамическое перераспределение процессорных ресурсов (ядра, потоки) и ресурсов памяти в большинстве случаев не требует остановки экземпляра Oracle, однако в ходе изучения управления ресурсами на серверах T7-2 была обнаружена проблема медленных потоков. Служба поддержки производителя предложила для нее эффективное решение.

4. Результаты синтетических тестов позволяют сделать вывод о том, что даже без специальных возможностей SQL in Silicon сервер T7-2 показывает очень высокую производительность на задачах Oracle Database. С учетом достаточной гибкости управления ресурсами LDOM можно смело рекомендовать серверы на основе процессора M7 как платформу консолидации Oracle Database.

***

Итак, характеристики нового процессора M7 могут давать существенные конкурентные преимущества по производительности на задачах с большим количеством активных сессий Oracle Database. Это может быть либо консолидация баз данных Oracle, либо экземпляр базы данных Oracle с крайне высокой степенью параллелизма (например, при выполнении аналитических отчетов DWH).

ИЗУЧЕНИЕ ORACLE IN-MEMORY И SQL IN SILICON

АЛЕКСЕЙ СТРУЧЕНКО,
руководитель Отдела
оптимизации СУБД и приложений
компании «Инфосистемы Джет»

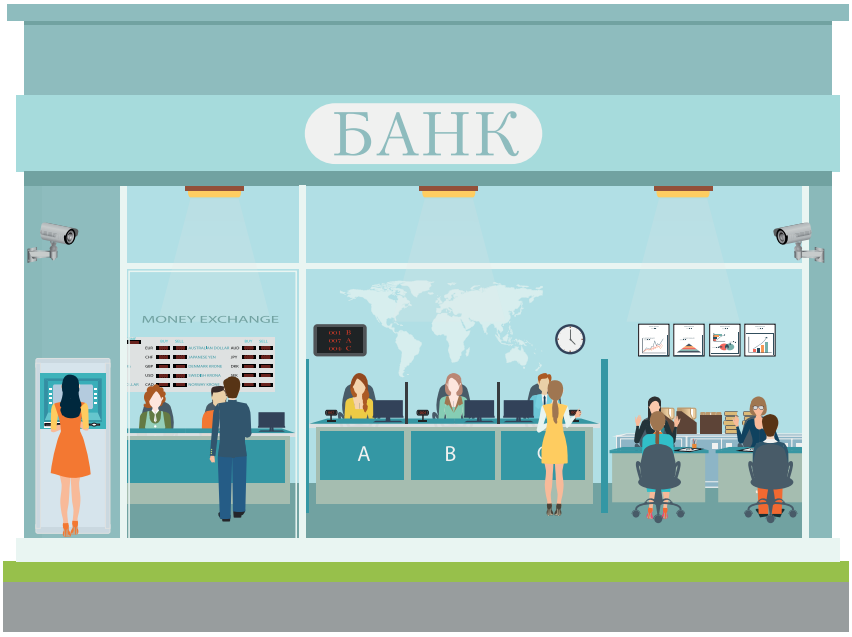
Одной из существенных инноваций процессора M7 является возможность переноса части SQL-логики на специальные сопроцессоры DAX (Data Analytics Accelerator). Эта технология получила название SQL in Silicon. На сегодняшний день доступ к DAX открыт для разработчиков через специальные библиотеки, а база данных Oracle использует эту технологию для ускорения работы Database In-Memory в версии 12c. Отметим, что Database In-Memory – одна из самых ожидаемых и обсуждаемых опций новой версии Oracle Database 12c. Несколько слово о том, что она из себя представляет.

Традиционные базы (в том числе Oracle Database) используют строчное хранение данных (Row Database) как в памяти, так и на дисках. Принято считать, что такое хранение оптимально для транзакционных систем класса OLTP, а вот в хранилищах класса DWH определенные аналитические запросы могут работать гораздо быстрее при колоночном хранении. Базы данных, реализующие такое хранение (Columnar Database), достаточно активно развиваются, в том числе в направлении колоночного хранения в памяти.

Появившаяся в Oracle Database 12c опция Database In-Memory реализует колоночное хранение данных в памяти дополнительно к традиционному строчному хранению. Такое хранение возможно благодаря специальной области памяти, в которой администратор Oracle Database может кэшировать данные как целых таблиц, так и отдельных колонок либо партиций в колоночном формате. Эти данные поднимаются в In-Memory кэш при старте экземпляра Oracle и затем поддерживаются в консистентном состоянии с данными обычного буферного кэша Oracle и с данными на дисках.

Такое дополнительное колоночное хранение данных в памяти прозрачно для приложения, при этом у оптимизатора Oracle появляется возможность выбора из памяти нужных данных как в строчном, так и в колоночном представлении. То есть в Oracle Database реализовано уникальное сочетание строчного и колоночного хранения данных в памяти.

При построении хранилищ и витрин данных мы регулярно применяем «шинный» подход к архитектуре, при котором одни и те же справочники используются для разных витрин и разрезов данных. Использование In-Memory для таких справочников позволит значительно ускорить работу бизнес-приложений, в том числе аналитических систем. Например, в банковском секторе скорость и качество обслуживания клиентов во многом зависят от доступности сервисов – от скорости предоставления информации, будь то выписка по счету, список уже приобретенных клиентом услуг или консультации по новым продуктам. Применение In-Memory гарантирует оперативное предоставление и анализ требуемых данных.



ID	SQL Plan Hash	User	P.	Database Time	IO Requests	Start	Ended	SQL Text
bp463k8ng3kps	986948374	SYS		0.5s		4:26:55 PM	4:26:56 PM	select /*+ monitor */ sum(salary) from dax.persons p where country_id in (select id from dax.countri...
178f19sygj46f	1045588844	SYS		1.9s		4:23:45 PM	4:23:47 PM	select /*+ no_inmemory(p) monitor */ sum(salary) from dax.persons p where country_id in (select id ...

Рис. 1. Enterprise Manager: демонстрация работы In-Memory

```
5  dax0   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax1   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax2   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax3   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax4   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax5   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax6   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
5  dax7   DAX_SCH_query_cmd_sched 0   DAX_QRYO_input_valid 0   DAX_QRYO_output_valid 0
```

Рис. 2. Вывод busstat с отключенным DAX

```
5  dax0   DAX_SCH_query_cmd_sched 80   DAX_QRYO_input_valid 694586   DAX_QRYO_output_valid 1530256
5  dax1   DAX_SCH_query_cmd_sched 80   DAX_QRYO_input_valid 754450   DAX_QRYO_output_valid 2017088
5  dax2   DAX_SCH_query_cmd_sched 79   DAX_QRYO_input_valid 522635   DAX_QRYO_output_valid 758496
5  dax3   DAX_SCH_query_cmd_sched 79   DAX_QRYO_input_valid 672683   DAX_QRYO_output_valid 1529568
5  dax4   DAX_SCH_query_cmd_sched 79   DAX_QRYO_input_valid 589392   DAX_QRYO_output_valid 1073248
5  dax5   DAX_SCH_query_cmd_sched 79   DAX_QRYO_input_valid 635264   DAX_QRYO_output_valid 1502832
5  dax6   DAX_SCH_query_cmd_sched 79   DAX_QRYO_input_valid 615433   DAX_QRYO_output_valid 1080257
5  dax7   DAX_SCH_query_cmd_sched 80   DAX_QRYO_input_valid 810452   DAX_QRYO_output_valid 2295872
```

Рис. 3. Вывод busstat с включенным DAX

При этом опция In-Memory – это не просто работа с данными в памяти в новом для Oracle колоночном формате, это еще и новые математические подходы: доработанные алгоритмы сжатия, механизмы быстрого определения минимальных и максимальных значений, Bloom-фильтры. В результате использование In-Memory в ряде случаев может приводить к серьезному ускорению SQL-запросов (прежде всего аналитических).

Мы неоднократно знакомили орацкловое сообщество с результатами, полученными в наших тестированиях работы In-Memory, в частности мы разработали методику, эмулирующую работу систем класса DWH. В таблицу базы данных Oracle были сложены случайным образом сгенерированные данные о жителях Европы и их зарплатах, а роль аналитика эмулировал запрос, считающий сумму всех зарплат жителей конкретной европейской страны.

На этом запросе нам удалось получить ускорение в сотни и даже тысячи раз при чтении данных из In-Memory кэша по сравнению с чтением данных с дисков. При этом по сравнению с чтением из буферного кэша чтение из кэша In-Memory оказалось быстрее почти в 4 раза (0,51 секунды против 1,9 секунды). Результат значительного ускорения по сравнению с чтением с дисков очевиден, а вот 4-кратное ускорение по сравнению со строчным форматом, на наш взгляд, является заслугой заложенных в In-Memory математических подходов. Этот резуль-

Работа запроса	Время исполнения, сек	OS: busstat -w dax
Без In-Memory (обычный кеш)	1,95	0
С In-Memory и выключенным DAX	0,51	0 (рис. 2)
С In-Memory и включенным DAX	0,39	Восемь строк >0 (рис. 3)

Табл. 1. Ускорение работы In-Memory с помощью сопроцессора DAX

тат демонстрирует Print Screen из Enterprise Manager, приведенный на рис. 1. Для исследования работы In-Memory мы использовали хинт оптимизатора/*+no_inmemory(имя_таблицы)*/, в явном виде запрещающий оптимизатору Oracle использовать In-Memory кэш.

Поскольку заложенная Oracle в новые процессоры SPARC M7 технология SQL in Silicon на сегодняшний день может ускорять работу именно In-Memory, важной частью тестирования серверов T7-2 стала проверка на них работы опции In-Memory по уже отлаженной методике (в ее основе лежит аналитический запрос, считающий суммарную зарплату).

Для таких тестов крайне важно было научиться явно включать возможности сопроцессора DAX на уровне как OS (соответствующие патчи Solaris), так и Oracle Database (специальный параметр экземпляра), а также правильно мониторить работу DAX. Команда busstat вполне адекватно показывает участие сопроцессоров DAX в работе основного процессора с точки зрения OS: на рис. 2 и 3 приведены выводы команды busstat в сравнении с выключенным и включенным параметром экземпляра.

> Опция In-Memory – это не просто работа с данными в памяти в новом для Oracle колоночном формате, это еще и новые математические подходы: доработанные алгоритмы сжатия, Bloom-фильтры и т.д.

Итак, запрос, суммирующий зарплату жителей всех европейских стран на букву R (Россия и Румыния), в наших тестах показал 1,95 секунды при традиционной работе Oracle Database и 0,51 секунды при настройке опции In-Memory и поднятии части данных (колонка с зарплатой) в специальный In-Memory кэш. Включение дополнительных возможностей сопроцессоров DAX (SQL in Silicon) дало результат 0,39 секунды, т.е. дополнительное ускорение на 30% сверх 4-кратного ускорения в In-Memory. Эти результаты сведены в табл. 1.

Много это или мало? С одной стороны, параллельно исследованию SQL in Silicon мы провели тестирование работы сопроцессоров DAX с помощью специально созданного приложения. Это приложение использовало DAX напрямую (через открытые Oracle библиотеки) и показало ускорение в 4–6 раз стандартных математических операций по преобразованию множеств. На фоне этого ускорение работы In-Memory выглядит скромнее. С другой стороны, работа опции In-Memory сама по себе достаточно оптимизирована, и ускорение в 4 раза по сравнению с буферным кэшем наглядно это подтверждает.

Тем не менее вполне возможно, что продуктивные аналитические запросы наших заказчиков смогут ускориться за счет DAX более существенно, поэтому технология SQL in Silicon заслуживает тщательного тестирования на реальных данных реальными аналитическими запросами. JI

ПОРА ЛИ ПЕРЕХОДИТЬ НА ALL-FLASH СХД?

АЛЕКСЕЙ ЛАВРЕНТЬЕВ,
инженер-проектировщик
систем хранения данных
компании «Инфосистемы Джет»

В последнее время не проходит и месяца без появления отраслевых новостей, касающихся новых решений для хранения данных. Гонка за бюджеты заказчиков всё чаще разворачивается в плоскости противостояния привычных дисковых массивов и решений, построенных на flash-накопителях. Сегодня большинство производителей выпустили, как минимум, по одной линейке All-Flash массивов (AF). В этой статье мы рассматриваем варианты, когда AF-массивы действительно эффективны, а также случаи, когда им еще нужно время для улучшения своих характеристик. Но сначала скажем несколько слов о стоимости этих решений.

ЭКОНОМИКА ALL-FLASH

Снижение стоимости flash-накопителей в последние несколько лет привело к естественному росту доли SSD в сегменте хранения данных. В большинстве современных систем хранения flash присутствует в том или ином виде:

- для кэширования наиболее часто используемых данных (примеры таких решений – FAST Cache на массивах EMC VNX, FlashPool или FlashCache на массивах NetApp FAS);
- для хранения наиболее часто используемых данных в гибридном пуле, включающем SSD- и HDD-накопители (tiering на массивах Hitachi, HP);
- для наиболее требовательных к производительности задач в виде выделенного пула из SSD-накопителей.

При этом стоимость «сырой емкости» ($/GB) flash-накопителей сегодня в несколько раз превышает стоимость HDD, поэтому повсеместного перехода на All-Flash еще не произошло. Однако производители flash-памяти и СХД усиленно работают над ее снижением. Серьёзным шагом в этом направлении должно стать применение новейших flash-накопителей, произведенных по технологии 3D NAND (множество слоев flash-памяти, расположенных один над другим, как в слоёном пироге). Так, компания Samsung, являющаяся лидером по объему производства flash-накопителей, недавно заявила о выпуске SSD объемом 15 ТБ в форм-факторе 2.5". Такая плотность возможна благодаря применению сразу нескольких технологий – 3D NAND и TLC (3-битные ячейки). Стоит отметить, что по технологии 3D NAND может производиться flash-память как MLC- (2-битные ячейки), так и TLC-типа. Несмотря на относительную новизну накопителей 3D NAND TLC, они уже нашли применение в следующих All-Flash массивах: EMC Unity F, NetApp AFF, NetApp SolidFire, Dell Compellent, HP 3PAR, Pure Storage FlashArray//m.

Intel и Micron также заявили о выпуске flash-накопителей 3D NAND, Toshiba совместно с SanDisk пока еще разрабатывают свой вариант технологии. Мы ожидаем, что повышение конкуренции и насыщение рынка приведет к снижению стоимости flash-накопителей большого объема.

Поскольку при производстве 3D NAND чипов flash-памяти используется более «толстая» технология производства (20–40 нм вместо 16 нм для однослойных чипов), продолжительность жизни ячеек памяти увеличивается. Это практически устраняет проблему износа flash-накопителя даже в самых нагруженных системах.

Что же сегодня предпринимают производители для снижения стоимости своих решений? Многие All-Flash массивы предполагают использование технологий экономии места (дедупликация, «тонкие» клоны, компрессия) благодаря применению современных высокопроизводительных процессоров. И все же периодически приходится идти на компромисс при использовании компрессии и/или дедупли-

О 2D flash-памяти

Если в используемых приложениях преобладают операции записи, правильнее хранить данные на eMLC-накопителях. Когда требуется большая емкость и преобладают операции чтения, можно использовать более экономичные TLC-накопители, т.к. чтение данных не снижает ресурс ячеек flash. В действительности «страшилки» про низкий ресурс flash-накопителей для большинства заказчиков не актуальны. Современные eMLC SSD способны выдерживать до 10 полных циклов перезаписи в день на протяжении всего гарантийного срока. При этом многие производители СХД предлагают бесплатную замену накопителей, находящихся на гарантии или на расширенной поддержке, даже в случае их полного износа и независимо от типа модулей.

кации из-за снижения итоговой производительности массива. Так что при выборе AF-массива необходимо определиться, действительно ли вам нужна скорость 200,000+ IOPS, или будет достаточно 50,000–100,000 IOPS, но с возможностью хранить в 2–4 раза больше данных.

Технологиям дедупликации уже не один год, однако их применение обычно ограничивалось «холодными данными» или резервными копиями. Причиной была низкая производительность механических дисков при случайном доступе к данным. Новое развитие дедупликация получила как раз с появлением flash-накопителей. Однако до сих пор не во всех AF-массивах дедупликация реализована или осуществляется в онлайн-режиме, когда блоки проверяются на уникальность в оперативной памяти еще до записи данных на flash.

КТО ЕСТЬ КТО НА РЫНКЕ

Как уже было сказано, в All-Flash «играют» все основные производители дисковых массивов – EMC, NetApp, IBM, HP, Hitachi, Oracle. Большая часть их решений представляет собой модернизированные дисковые массивы, но есть и несколько очень удачно приобретенных продуктов – All-Flash-стартапов. Также на этом поле присутствует ряд компаний-стартапов, занимающих небольшую долю рынка, например Violin Memory. Один из ведущих мировых игроков на рынке All-Flash, который не представлен в России, – Pure Storage – пока не обозначал планов по выходу на российский рынок. Ряд производителей имеет в своем модельном ряду несколько линеек All-Flash-массивов, отличающихся по функциональности, производительности и решающих разный спектр задач.

Что нас ждет в будущем?

Использование стандарта NVMe вместо SAS для подключения flash-накопителей к контроллерам массива обеспечит снижение задержек и повышение производительности. Это станет возможно благодаря использованию высокоскоростной PCI-Express шины и параллелизации запросов вместо последовательной обработки SCSI-команд SAS-контроллерами. При этом физически NVMe-накопители могут быть выполнены в привычном форм-факторе 2.5", а также в виде PCI-Express- модулей.

Реализация технологии NVMe over Fabric для повышения производительности и снижения задержек при передаче данных между серверами приложений и AF-массивом может стать еще одним драйвером этого рынка.

Стоит сказать несколько слов об All-Flash-продуктах компании Huawei: в данный момент все массивы производителя могут быть сконфигурированы в варианте All-Flash с SSD-накопителями, по своим основным характеристикам они сопоставимы с уровнем развития решений ведущих мировых производителей. В ближайшее время компания планирует выпустить обновление AF-массива OceanStore Dorado, в котором будут применяться многие актуальные технологии, например, NVMe-накопители и сжатие данных в реальном времени.

All-Flash массивы значительно различаются по способу записи и хранения информации, что сказывается на их предрасположенности к использованию для тех или иных задач. Выросшие из стартапов и изначально разрабатываемые именно как flash-массивы EMC XtremIO, Pure Storage FlashArray//m и NetApp SolidFire лучше подходят для большого количества мелких запросов, характерного для сред виртуализации. Они выполняют дедупликацию и компрессию в режиме реального времени. Эти технологии позволяют значительно экономить дисковое пространство и бережнее использовать ресурс ячеек flash-памяти.

Отметим, что дедупликация хорошо работает при использовании виртуализации и VDI, а компрессия может обеспечить уменьшение объема, занимаемого базами данных, до 50% (если она или шифрование не используется в самих БД). Чаще всего компрессия выполняется на процессорах контроллеров flash-массивов, но есть и примеры выделения под эти задачи отдельных ресурсов: в массивах IBM FlashSystem для этого используются специализированные PCI-Express модули, а в Hitachi VSP F – контроллеры, встроенные во flash-модули FMD DC2.

Массивы, создававшиеся для аналитики и высоконагруженных баз данных (Hitachi VSP F, IBM FlashSystem, Violin Memory FSP и Oracle FS1), обеспечивают

высокую производительность в операциях последовательного доступа.

НАПРАВЛЕНИЯ РАЗВИТИЯ ALL-FLASH

Существует 2 направления позиционирования и развития AF-массивов. Первая категория решений рассчитана на замену многофункциональных гибридных массивов, а вторая призвана обеспечить максимальное быстродействие для некоторых задач. К первой группе, реализующей практически весь функционал гибридных собратьев, можно отнести NetApp AFF (блочный/файловый доступ, функционал обычных FAS), Hitachi VSP F (почти весь функционал, кроме Global Active Device), EMC Unity F (функционал гибридных Unity), HP 3PAR All-Flash (функционал гибридных 3PAR), Dell Compellent All-Flash (весь функционал, включая возможность многоуровневого хранения на разных видах SSD-накопителей).

Более специализированные решения для ускорения конкретных задач: EMC XtremIO (виртуализация, VDI, OLTP), IBM FlashSystem и Violin Memory FSP (виртуализация, VDI, базы данных для аналитики и OLTP, антифрод), NetApp EF (базы данных для аналитики и OLTP), Hitachi HFS A (виртуализация, VDI), NetApp SolidFire (Scale-Out для облачных провайдеров).

Подчеркнем, что многофункциональные массивы могут предложить больше сервисов, связанных с хранением и управлением данными, при этом они не всегда уступают по скорости узкоспециализированным решениям.

Отдельно хотелось бы отметить появление новой высокопроизводительной системы

хранения – EMC DSSD. Она является одним из первопроходцев технологии NVMe over Fabrics, используя один из первых и наиболее проработанных вариантов ее реализации – NVMe over PCI-Express. Серверы подключаются к AF-массиву с помощью специализированных адаптеров PCI-Express и кабелей длиной до 4 метров. Использование шины PCI-Express для передачи данных обеспечивает минимально возможные задержки и максимальную на сегодняшний день производительность. Одной из особенностей EMC DSSD является хранение всех данных внутри массива в виде объектов, по этой причине предпочтительно использовать API для связи решения с приложениями. Если отсутствует возможность использовать объектный подход, можно применить специальный драйвер для блочного доступа, однако блоки данных на массиве все равно будут храниться внутри контейнеров, что может несколько снижать итоговую производительность. EMC DSSD целесообразно использовать в случае высоконагруженных баз данных для аналитики, решений задач антифрода, High Performance Computing и Big Data.

ИСПОЛЬЗУЕМЫЕ НАКОПИТЕЛИ

Многие производители используют в качестве flash-накопителей стандартные SSD форм-фактора 2.5" с интерфейсом SAS. Это позволяет применять широко распространенные накопители от разных производителей и легко добавлять новые модели SSD без изменения архитектуры решения.

Вторым подходом является производство специализи-

рованных flash-накопителей, чаще всего подключаемых к PCI-Express-шине контроллеров. Этот вариант предпочтителен для решений, нацеленных на максимальную производительность, но он снижает гибкость дальнейшего развития продукта, т.к. требует значительных затрат на разработку накопителей и чаще всего завязан на одного поставщика. Подобные накопители применяются в массивах Hitachi VSP, IBM FlashSystem, Violin Memory FSP, EMC DSSD, Pure Storage FlashBlade.

ДЛЯ ЧЕГО СТОИТ ВЫБРАТЬ ALL-FLASH

Наши заказчики всё чаще внедряют AF у себя в дата-центрах. Приведем несколько примеров. Виртуализация – отличный кейс для применения AF-массивов с дедупликацией. Во-первых, в виртуальных машинах обычно содержится много одинаковой информации (файлы операционных систем), а во-вторых, существует так называемый эффект «блендера». Он заключается в том, что множество виртуальных серверов обращаются к одному дисковому ресурсу одновременно, в итоге образуется по большей части случайная нагрузка на диски, что очень хорошо подходит для SSD, но плохо для HDD. Так, один из крупных российских банков, наш заказчик, использует All-Flash в своей виртуальной инфраструктуре для высокопроизводительных виртуальных сред.

В крупных VDI-проектах также целесообразно использовать All-Flash. Эффект «блендера» в случае VDI выражен даже ярче, чем при виртуализации серверов. Утренняя загрузка множества виртуальных рабочих мест характеризуется множеством случайных запросов чтения,

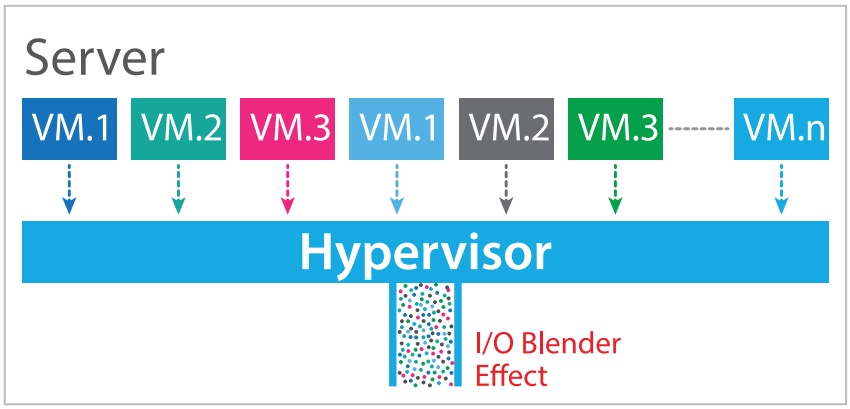


Рис. 1. Эффект «блендера»

а в повседневной работе VDI до 90% нагрузки приходится на мелкие записи. Большое количество российских VDI-проектов первой волны провалилось или потребовало доработки как раз из-за серьезных нагрузок на дисковые массивы, которые были не в состоянии обеспечить требуемую производительность. Это приводило к медленной работе приложений и недовольству конечных пользователей.

All-Flash хорошо подходит для обеспечения работы высоконагруженных транзакционных (OLTP) баз данных, обрабатывающих большое количество небольших запросов и при этом очень требовательных к минимальному времени отклика. Например, до внедрения AF-массива для осуществления процессинга топливных карт в крупной нефтяной компании периодически наблюдались проблемы в работе системы на заправках. Большое количество запросов приводило к значительному увеличению времени отклика со стороны дискового массива и, как следствие, к проблемам в работе всего сервиса. Переход на AF полностью снял этот вопрос с повестки дня, и с тех пор сервис топливных карт не испытывает трудностей, связанных с недостатком производительности СХД.

Кроме того, AF (не все решения) имеет смысл использовать для работы систем бизнес-анализа. Стоит отметить, что при построении аналитических отчетов нагрузка на дисковую подсистему значительно отличается от нагрузки в случае виртуальных сред: на первое место здесь выходит не минимальное время доступа к маленьким блокам информации, а максимальная скорость чтения больших объемов данных. Из-за специфики реализации операций записи некоторые AF-массивы демонстрируют не лучшие результаты для BI. Действительно высокую производительность для бизнес-анализа чаще всего демонстрируют массивы, использующие специализированные flash-модули с PCI-Express интерфейсом. Применение компрессии на All-Flash позволяет уменьшить объем, занимаемый базами данных, до 50%, при этом дедупликация обычно не дает существенной экономии места.

Например, в одной крупной страховой компании после переноса аналитических баз данных на All-Flash значительно сократилось время построения бизнес-отчетов. Это ускорило принятие решений. Также в случае необходимости можно создавать новые отчеты без существенного

снижения скорости выполнения основных задач.

Об областях применения AF-массивов у телеком-операторов стоит сказать отдельно. По сути, их две – это биллинг и антифрод-решения. Биллинг – одна из важнейших и наиболее нагруженных систем. В случае предоплаченных услуг оператор в режиме реального времени ведет учет потребления трафика и оказания различных сервисов для множества абонентов. Для этого системе хранения требуется обеспечить выполнение большого количества запросов с минимальным временем отклика. При постоплате на основе собранных данных с определенной периодичностью (обычно раз в месяц) создается огромное количество счетов на оплату оказанных услуг. Во время их генерации необходима высокая скорость чтения больших объемов информации за относительно короткий промежуток времени.

Защита от мошенничества подразумевает комплекс задач по проверке финансовых транзакций на подозрительные активности. Один из главных критериев работы антифрод-системы – это скорость анализа поведения пользователя на основе различных алгоритмов. Чем быстрее происходит оценка рисков, тем ниже риски финансовых потерь. При большом количестве операций необходимы высокая производительность и низкое время отклика со стороны СХД. И здесь AF-массивы уверенно обходят традиционные решения. Поэтому один из наших заказчиков – телеком-оператор – для систем биллинга и антифрода использует именно All-Flash. Объем хранящихся на этом массиве данных уже приближается к 80 ТБ.

В большинстве приведенных примеров экономическая эффективность использования AF значительно выше, чем у гибридных массивов. За счет этого компенсируется высокая стоимость AF-массива.

> All-Flash подходит для обеспечения работы высоконагруженных транзакционных (OLTP) баз данных, обрабатывающих большое количество небольших запросов и при этом очень требовательных к минимальному времени отклика

ДЛЯ ЧЕГО ALL-FLASH НЕ ПОДОЙДЕТ

По причине высокой стоимости flash-накопителей и систем AF некоторые виды данных выгоднее хранить на медленных, но дешевых HDD или гибридных массивах. Так, хранить резервные копии и архивы на flash-накопителях совершенно неоправданно, т.к. эти данные не требуют высокой скорости доступа, а основным параметром здесь является экономическая эффективность хранения. Возможно, при появлении требований к быстрому восстановлению из бэкапа и снижении стоимости емких 3D NAND TLC дисков эти задачи целесообразно будет перевести в область ответственности All-Flash.

Также не имеет смысла использовать AF-массивы для мультимедийных данных. Они практически не поддаются сжатию и чаще всего представляют собой достаточно большие файлы, которые читаются и пишутся на диски последовательно. С такой нагрузкой хорошо справляются и механические жесткие диски, которые стоят существенно дешевле SSD.

И несколько слов о варианте применения All-Flash для пользовательских файловых ресурсов: часто они не требуют высокой скорости работы, поэтому пока эффективнее использовать для них обычные дисковые или гибридные массивы.

***

All-Flash – это будущее, которое быстро и неотвратимо приближается. Стоимость AF-массивов постоянно снижается, и уже сегодня стоит рассматривать варианты перехода на них для большинства задач, требующих высокой производительности и низкого времени отклика.

При этом важно понимать, что All-Flash не является «серебряной пулей» и не решит всех проблем одним своим присутствием. Очень часто медленная работа приложений может быть обусловлена неэффективными запросами к базе данных, неправильным построением логики приложений и ИТ-систем.

При выборе All-Flash массива нужно обязательно погружаться в задачи и требования заказчика, изучать его инфраструктуру и бизнес-процессы, считать бизнес-кейсы и получаемую выгоду. Наша компания, благодаря огромному опыту проектирования как классических решений, так и новых, основанных на flash-технологиях, готова предложить именно ту архитектуру, из которой компания сможет выжать максимум при сохранении инвестиций. JI

ОБЗОР ФУНКЦИОНАЛА SDS

ДМИТРИЙ ГЛУШЕНОК,
системный архитектор
Центра проектирования вычислительных комплексов
компании «Инфосистемы Джет»

Программно-определяемые СХД в виде абстракции над устройствами хранения появились не вчера. Но относительно недавнее появление высокоскоростных ЛВС и достаточных процессорных мощностей (мы подробнее рассматриваем это ниже) позволило реализовать новые подходы к хранению данных, такие как распределенный между серверами Erasure Coding. А бурное развитие облачных вычислений поставило перед устройствами хранения новые задачи – объектное хранение данных и обеспечение гипермасштабируемости. Ниже мы дадим краткие характеристики используемых сегодня SDS-технологий, разберем составляющие их стоимости и озвучим возможные варианты применения.

ЛОКАЛЬНАЯ СЕТЬ КАК ТРАНСПОРТ

Быстрая скорость передачи данных и низкие задержки всегда были на стороне SAN вообще и Fibre Channel в частности. Но с появлением доступных 10 Гбит/с Ethernet-коммутаторов стало возможным использовать ЛВС для передачи данных между сервером и хранилищем на сравнимых с SAN скоростях. Производители SAN-коммутаторов и HBA (Host Bus Adapter) даже предприняли попытку создания конвергентных решений, использующих Ethernet для транспорта FC.

А с увеличением скоростей Ethernet вплоть до 100 Гбит/с и появлением таких расширений протокола, как iWARP и RoCE, даже для чувствительных к задержкам приложений скорость доступа к SSD-накопителю в соседнем сервере по сети стала сравнимой (а порой даже равной) скорости доступа к локальным накопителям, подключенным к PCIe-шине.

> Программно-определяемые СХД позволяют решать большой пул задач – от архивного хранения до высокоскоростной обработки – за меньшие или сопоставимые с использованием классических дисковых массивов деньги

Программно-определяемые СХД получили возможность стать распределенными по множеству узлов, теперь можно на порядок увеличить объемы хранилища без ущерба для производительности. И если раньше (в пределах сервера) точкой отказа был диск, теперь быстрая сеть позволяет рассматривать сервер целиком (или блейд-шасси, или даже всю стойку) как потенциально сбойный компонент с соответствующей защитой.

СЕРВЕРЫ С ВНУТРЕННИМИ ДИСКАМИ КАК УЗЛЫ РАСПРЕДЕЛЕННОГО ХРАНИЛИЩА

Появление SATA-интерфейсов с возможностью горячей замены и совместимость SATA-дисков с SAS-контроллерами позволили использовать для организации хранилища дешевые SATA-диски в недорогих же серверах. Все это вместе позволяет таким компаниям, как Amazon и Microsoft, предлагать клиентам хранение данных по ценам $10–30 за 1 ТБ/мес.

Появление же таких технологий, как NVMe, NVMe over Fabrics, вместе с ростом объемов SSD-накопителей до HDD позволит и вовсе отказаться от использования дисковых массивов. Таким образом, запросы самых требовательных по скорости и задержкам приложений будут удовлетворены программно-определяемой СХД.

При этом смешивание в SDS устройств хранения с различными характеристиками (например, NVDIMM, NVMe и HDD) с помощью многоуровневого хранения данных решает проблему избытка емкости при нехватке скорости, и наоборот.

ХАРАКТЕРИСТИКИ ТИПОВЫХ SDS

Объектное хранение данных и масштабируемость

Большинство программно-определяемых СХД в первую очередь представляют собой объектное хранилище, позволяющее по названию объекта получить его содержимое. Преимуществом такого подхода является отсутствие необходимости (либо значительное упрощение) индексирования содержимого. При заранее известном количестве узлов СХД нахождение объекта устанавливается путем хэширования его названия и, грубо говоря, деления результата на количество узлов. Полученный остаток от деления и будет указывать на ноду, содержащую искомый объект. Подобный подход позволяет равномерно распределять данные между множеством узлов хранения без создания узких мест вроде выделенных серверов метаданных. При необходимости хранения объектов большого размера их можно делить на блоки фиксированной длины, распределяя эти блоки между всеми узлами СХД.

Иногда объектное хранилище используется для предоставления блочных устройств или

✓ Хранилище для облачного века

Драйвером развития программно-определяемых СХД можно считать Amazon Web Services с его хранилищем S3 (Simple Storage Services), появившимся в 2006–2007 годах. S3 – это распределенное объектное хранилище, доступное поверх протокола HTTP.
В настоящий момент S3-интерфейс де-факто стал стандартом для работы с объектами и поддерживается практически всеми объектными хранилищами других вендоров.

Популярность S3 обусловлена его гипермасштабируемостью – буквально в масштабах интернета. Это то, что требуется поколению приложений, запускаемых в облаке, – отделенное от самого приложения, доступное из любого уголка интернета, надежное, быстрое, безопасное и недорогое хранилище. Масштабирование облачного приложения заключается в увеличении или уменьшении количества его запущенных экземпляров, при этом хранилище его данных масштабируется автоматически при возросшем или уменьшившемся количестве запросов.

файловых систем. И то, и другое будет преобразовано в объекты с распределением между всеми узлами СХД. Здесь нужно отметить, что блочный/файловый доступ для объектного хранилища – это все-таки дополнительный уровень абстракции и не основной функционал. В некоторых объектных хранилищах их нет вовсе, где-то этот функционал только развивается.

Erasure Coding вместо RAID

Для защиты от выхода из строя устройства хранения или целиком узла в программно-определяемых СХД обычно используются 2 механизма – репликация данных между несколькими узлами (обычно хранятся 2 или 3 копии данных) и Erasure Coding. Репликация обеспечивает лучшую производительность, но ценой кратного сокращения полезного дискового пространства. Например, при защите от одновременного выхода из строя двух узлов полезная дисковая емкость составит всего 33% от «сырой».

Erasure Coding представляет собой математический алгоритм генерации из *K* исходных блоков данных *M* блоков избыточной информации таким образом, что исходные данные могут быть восстановлены из *K* любых получившихся блоков. При *M*=1 и *M*=2 получаются конструкции, похожие на RAID5 и RAID6 соответственно. При *M*=3 (и *K*=7, например) система защищается от одновременного выхода из строя трех узлов, при этом полезная дисковая емкость составит уже 70% от «сырой» в отличие от 33% при 3-кратной репликации.

Как репликация, так и кодирование в программно-определяемых СХД обычно реализованы таким образом, что данные распределяются по всем узлам СХД независимо от кратности репликации или соотношения *K*/*M*. Из-за этого восстановление избыточности при выходе узла из строя происходит не на выделенный spare-диск (это долго и с дисками больше 4 ТБ может быть опасно), а на свободное пространство всех оставшихся узлов. В итоге время восстановления сокращается во много раз.

Адаптивная производительность

Ранее приходилось вручную рассчитывать размер каждого уровня многоуровневого хранилища, но сейчас появляются технологии, позволяющие на основе декларативных требований к скорости и надежности хранения автоматически размещать данные на нужном количестве узлов и носителей. В будущем приложения будут запрашивать у СХД необходимые ресурсы без вовлечения в процесс человека.

Также некоторые программно-определяемые СХД могут запоминать последовательности обращения к различным данным в разное время и переносить на быстрый уровень данные, которые скоро понадобятся, в том объеме, в котором они понадобятся. Это позволяет удешевить СХД, храня на быстром уровне не весь объем «горячих» данных, а лишь его часть (необходимую сейчас и в течение ближайшего времени) без существенного влияния на производительность.

СОСТАВЛЯЮЩИЕ СТОИМОСТИ

Идея создания SDS обычно приходит в ответ на требование о снижении стоимости хранилища. Но здесь есть подводные камни.

Сеть

Для возможности быстрого восстановления избыточности при выходе из строя каких-либо компонент требуются, как мини-

мум, 10 Гбит/с сетевые соединения. Поэтому в определенных конфигурациях предпочтительнее создать даже выделенную межузловую сеть, чтобы снизить влияние ребалансировки на клиентские запросы. Но зачастую в имеющейся инфраструктуре порты доступа только гигабитные, и стоимость модернизации участка сети с добавлением 10 Гбит/с коммутаторов может ощутимо увеличить общую стоимость программно-определяемой СХД. Особенно в случае приверженности к дорогим сетевым брендам.

Серверы и диски

С выбором серверов все довольно просто: они в значительной мере уже стали потребительским товаром, их стоимость у разных производителей примерно одинаковая. Но есть 2 нюанса.

Первый – серверов нужно много (десятки, если не сотни), и даже в минимальной конфигурации их стоимость может оказаться больше стоимости дисков. Одним из решений проблемы является размещение приложений на тех же серверах, что «участвуют» в СХД. В некоторых конфигурациях приложение от этого только выиграет, получив возможность читать данные с локального хранилища без использования сетевой передачи.

Второй нюанс – многие производители серверов ревниво относятся к установке в них «чужеродных» дисков. Некоторые прямо это запрещают, другие просто не продают салазки для установки дисков в слоты отдельно от самих дисков. Все бы ничего, но подобная наценка на диски в 100–200% и более удорожает хранилище в разы. Другими словами, если ваша цель – сделать SDS максимально дешевой, то диски (и, возможно, серверы) нужно приобретать у лояльных к этому компаний.

> Большинство программно-определяемых СХД, представляют собой объектное хранилище, позволяющее по названию объекта получить его содержимое

Программное обеспечение

Большинство ПО, позволяющего создавать программно-определяемые СХД, является коммерческим. Многие программы по факту являются «швейцарскими ножиками», способными организовать и объектное, и блочное, и файловое хранилища, причем сразу по нескольким протоколам. Здесь стоит сказать несколько слов о рынке производителей этого ПО. Часть компаний-стартаперов в сфере SDS были приобретены ИТ-гигантами, и теперь их продукты поставляются только в виде комплексных решений вместе с оборудованием. Другие договорились о партнерстве с производителями серверов, которые тоже стали предлагать комплексные решения на своем оборудовании. Часть коммерческих решений пока еще относительно свободна, но сложность используемых технологий и статус амбициозного стартапа не добавляют нам уверенности в их надежности. По стоимости же все коммерческое ПО ориентировано на решение широкого спектра задач, и дешевое хранилище в него не входит.

Некоммерческого ПО не так много, и оно обычно менее функционально по сравнению с коммерческим. Но здесь вполне можно найти хорошее решение под конкретную задачу, даже с возможностью поддержки от вендора. Таким образом, вы гарантируете себе низкую стоимость хранилища вместе с отсутствием жесткой привязки к конкретному производителю (как ПО, так и «железа»).

РЕШАЕМЫЕ ЗАДАЧИ

Хранилище как услуга

Для организации хранилища в виде услуги СХД должна обеспечивать такие базовые функции, как:

- изоляцию пользователей услуги (компаний/подразделений), их внутренних пользователей (отдельных бизнес-пользователей) и приложений друг от друга;
- изоляцию нагрузок между пользователями услуги и «внутри» них;
- разделение сетевых подключений из пересекающихся адресных пространств;

- масштабирование хранилища без изменения конфигурации приложений.

Среди классических СХД подобные решения придется поискать. В SDS же эти функции обычно заложены изначально, либо они несложно реализуются.

Гиперконвергентные системы

В случае если программное решение занимает не один сервер (с виртуализацией или без) и ему требуются внушительные дисковые емкости, программно-определяемая СХД позволяет сократить общую стоимость системы. К тому же использование SDS позволит масштабировать приложение и хранилище, добавляя (удаляя) узлы с необходимым количеством вычислительных или дисковых ресурсов.

Облачное хранилище для облачных приложений

Облачное приложение – в первую очередь мобильное и масштабируемое. Хранилище для него должно быть соответствующим – доступным для приложения из любой точки его запуска, масштабируемым, без явных узких мест. Яркий пример – контейнеры. Все данные контейнерного приложения располагаются во внешнем хранилище, которое либо доступно приложению по сети, либо динамически подключается к контейнеру при его запуске. Современные программно-определяемые СХД успешно справляются с такими задачами, так как, по сути, для них и создавались.

Архивное/холодное хранилище петабайтного масштаба

Дисковый массив на пару петабайт для холодного хранения данных? Он получится «золотым» и будет лишь греть воздух. А через несколько лет стоимость его поддержки станет «платиновой», и будет открыт проект по миграции данных на новый «золотой» массив. Файловая система подобного объема после сбоя сервера будет проверяться не один день, при этом хранилище будет недоступно.

Объектное программно-определяемое хранилище на таких объемах выходит значительно дешевле классических СХД. С помощью Erasure Coding (12+6, например) достигается необходимая отказоустойчивость (показатель сохранности данных – 99,9999999999999%). А при распределении этих данных между тремя и более ЦОД отключение одного из них не вызовет недоступности СХД. При этом отсутствуют недостатки классической репликации в виде кратного дублирования дискового пространства.

> Технологии SDS как нельзя лучше подходят для создания систем холодного/архивного хранения данных, например архивов медиа-контента, записей переговоров в call-центрах, записей систем видеонаблюдения, почтовых архивов и т.д. Одним словом, использование программно-определяемых СХД оптимально, когда речь идет о длительном хранении больших объёмов информации, которые не подвергаются постоянной оперативной обработке, но должны быть легко доступны при необходимости.

Хранилище для СРК

Производители СРК уже включили поддержку объектных хранилищ в свои продукты. И Veritas NetBackup, и CommVault Simpana научились работать с хранилищами как по протоколу S3, так и по специальным протоколам отдельных объектных хранилищ.

Отдельно стоит отметить, что использование программно-определяемого хранилища возможно даже для Oracle ASM с субмиллисекундным доступом. На стенде в нашей компании мы протестировали конфигурацию программно-определяемой СХД, в которой SSD-диски каждого сервера по выделенной сети были предоставлены остальным серверам в виде блочных устройств. На протестированных дисках разницы между сетевым и локальным обращением к SSD для приложения не было. То есть для Oracle ASM (и Oracle RAC в частности) с современными объемами SSD-накопителей необязательно покупать разделяемый дисковый массив. Достаточно сделать внутренние диски серверов доступными другим серверам кластера по сети.

***

Программно-определяемые СХД пока еще кажутся экзотикой, но уже сейчас они позволяют решать большой пул задач – от архивного хранения до высокоскоростной обработки – за меньшие или сопоставимые с использованием классических дисковых массивов деньги. При этом стандартные СХД все еще с трудом выходят за пределы ЦОД, в то время как SDS уже масштабируются до размеров интернета. JI

СЕРВЕРНАЯ ПЛАТФОРМА HUAWEI KUNLUN – ВЫВОДЫ ТЕСТИРОВАНИЯ

ДМИТРИЙ ГЛУШЕНОК,
системный архитектор
Центра проектирования вычислительных комплексов
компании «Инфосистемы Джет»

Для начала кратко скажем об архитектуре KunLun, поскольку эта информация практически не представлена в русскоязычном сегменте. KunLun создавался как High-End-сервер, отсюда следует дублирование всех его компонентов (в том числе модулей управления и контроллеров NUMA-узлов). Впрочем, одним дублированием компонентов High-End'ность сервера не ограничивается – в нем можно без остановки ОС заменять не только PCIe-платы (что в принципе не ново), но и процессоры с памятью. Система проактивно отмечает компоненты, которые скоро могут выйти из строя, не дожидаясь самого сбоя. Инженеры могут заменить их без остановки ОС. Пока горячая замена процессоров и модулей памяти работает только в ОС EulerOS (CentOS от Huawei), скоро обещают поддержку «из коробки» для RHEL и SLES.

Системные платы сервера, состоящие из 1 процессора и 24 модулей памяти, через систему коммутации объединяются в физические партиции с 4, 8, 16 или 32 процессорами. Меньшую гранулярность можно получить лишь при использовании логического партиционирования (гипервизора).

В сервере также имеются встроенные диски – до 4 корзин по 12 дисков в каждой с возможностью создания аппаратного RAID внутри каждой корзины. В некоторых случаях это поможет отказаться от использования внешнего дискового массива.

Основная особенность KunLun заключается в возможности объединения до 32 процессоров Intel Xeon и до 24 ТБ памяти в одной партиции. Так, в HP Superdome X помещается только 16 процессоров, в SGI UV 300 – тоже 32, но сама SGI ориентирована на HPC. Сопоставимое количество процессоров Intel Xeon не объединено больше ни в одном промышленном решении.

И еще одно нововведение: в системе используется BIOS производства Huawei, и производитель готов предоставить исходные коды для сертификации ПО.

EulerOS

Об этой ОС мы узнали при демонстрации замены процессора «на лету». Информации о ней в интернете очень мало, разве что есть данные о ее сертификации по последней версии Linux Standard Base. На поверку оказалось, что EulerOS – это собранная из исходников RHEL – Red Hat Enterprise Linux (по аналогии с CentOS). Huawei кастомизирует ее для своего оборудования, добавляя драйверы горячей замены ЦПУ/ОЗУ.

Помимо EulerOS, в KunLun заявлена поддержка RHEL, SLES 11&12, Windows Server 2012.

Тест производительности SPECint/SPECfp

Образно говоря, с арифметикой в KunLun все хорошо. При проведении SPECint процессы привязываются к конкретным ядрам и работают только с локальной памятью.

Интересным получилось сравнение KunLun с топовым IBM Power E880 (также 16-процессорным) – он не так уж сильно

✓ Почему не каждый производитель может предложить 32-процессорную систему

Реальность такова, что штатные средства процессоров Intel позволяют объединять не более 8 процессоров в одном сервере. Объединение большего количества требует создания специальных устройств – контроллеров NUMA-узлов (node controller). Intel их не производит, но сама возможность использования такого функционала заложена в шину QPI. Этим воспользовались HP, SGI и Huawei – каждый вендор сделал свой контроллер. Здесь нужно понимать, что создание подобного контроллера подразумевает масштабную научную деятельность с вытекающими отсюда затратами. У Huawei, например, разработка заняла 8 лет.

В то же время остальные вендоры (в том числе сам Intel) не стали разрабатывать свои контроллеры. И на это есть веские причины. Первая – это то, что увеличение количества процессоров ведет к замедлению скорости работы с памятью. В большой степени это связано с необходимостью синхронизации состояния кэшей процессоров: чем больше процессоров закэшировали участок памяти, тем больше потребуется оповещений при его изменении одним из процессоров. Вторая причина – для подавляющего большинства вычислительных задач хватает серверов с количеством процессоров от одного до четырех.

Сервер	SPECint	SPECfp
SGI UV 300 (32x, Intel Xeon E7-8890 v3)	22600	15700
KunLun 9032 (32x – Intel Xeon E7-8890 v3)	**22900**	**16300**
IBM Power E880 (16 x Power8 4.0 GHz, 192 core)	14400	11400
KunLun 9016 (16x – Intel Xeon E7-8890 v3)	**11700**	**8050**
SGI UV 300 (16x, Intel Xeon E7-8890 v3)	11400	7880
Integrity Superdome X (16x, Intel Xeon E7-8890 v3)	11100	7670

оторвался от KunLun. То есть для вычислений Intel Xeon сервер Huawei вполне может конкурировать с Power8.

Тест производительности SLOB (Oracle)

Этот тест в большей степени измеряет скорость не самих вычислений, а доступа к памяти. Привязки процессов СУБД к NUMA-узлам нет, для теста вся память считается равноудаленной от процессоров. Данный тест подкрепил цифрами теорию о нелинейном росте производительности сервера при добавлении ресурсов.

Семикратное увеличение процессорных мощностей (с 16 до 144 ядер, с учетом понижения частоты) приводило к пятикратному росту производительности сервера (КПД 71%). Четырехкратное увеличение количества ядер – с 16 (4 ЦПУ) до 64 (16 ЦПУ) – обусловило рост производительности в 2,7 раза (КПД 68%).

Области применения KunLun

По нашему мнению, основным достоинством KunLun является огромный объем памяти на его борту (24 ТБ сейчас, 32 ТБ в будущем). Поэтому основной областью его применения будет In-Memory аналитика, когда вся БД помещается в оперативную память. Использование KunLun позволяет на 3 порядка сократить время доступа к данным по сравнению с жесткими дисками и, соответственно, ускорить время выполнения запросов к БД. KunLun хорошо подойдет для задач SAP HANA и SAP S/4HANA. Объем памяти позволяет использовать HANA даже в одноузловой конфигурации KunLun. Oracle Database (особенно с In-Memory option) и QlikView также хорошо смотрятся в китайском суперсервере.

Соответственно, подобный комплекс найдет свое применение в качестве платформы для SAP HANA у ритейлеров для анализа больших объемов данных и формирования ценной для бизнеса «выжимки» из них: какие товары пользуются наибольшим/наименьшим спросом, каково их наличие на складах и т.д. Банкам связка Oracle In-Memory Option и KunLun даст возможность «на лету» осуществлять анализ кредитоспособности клиентов, рассчитывать нормативы по достаточности капитала банка и т.д. Телеком-операторы могут использовать это решение для задач по управлению лояльностью абонентов – формирования их профилей, таргетирования.

Другим классом задач для KunLun может стать замещение RISC-систем на x86. Для ряда заказчиков актуальны вертикально масштабируемые задачи, переросшие x86-серверы прошлого и выполняемые на RISC. В то же время стоимость обслуживания и модернизации этих платформ такова, что между стоимостью KunLun и годового обслуживания RISC-системы можно поставить знак равенства. По заявленному уровню надежности KunLun им не уступает, но при этом выигрывает по разнообразию прикладного ПО. Отметим, что у себя на родине KunLun активно используется для импортозамещения, в основном как платформа для миграции с RISC-систем. JI

ТЕСТИРОВАНИЕ ФУНКЦИЙ ВИРТУАЛИЗАЦИИ ПЛАТФОРМЫ SPARC НА ORACLE SPARC T7-2

ЮРИЙ СЕМЕНЮКОВ,
руководитель отдела корпоративных решений
Центра проектирования вычислительных комплексов
компании «Инфосистемы Джет»

В начале этого года у нас на тестировании побывал сервер Oracle SPARC T7-2. Основной целью тестирования было знакомство с новыми технологиями аппаратного ускорения работы СУБД Oracle средствами нового процессора Oracle M7, на базе которого построен сервер. Об этих технологиях и результатах их тестирования уже подробно рассказали мои коллеги в других статьях номера. Параллельно с этим мы провели тестирование функций виртуализации платформы SPARC на сервере, о чем и пойдет речь ниже.

Серверы линейки Oracle SPARC T-series позиционируются как машины Enterprise-уровня для консолидации нескольких систем на одном физическом сервере. Для этого они имеют встроенную платформу виртуализации на базе Oracle VM for SPARC, развитые возможности ОС Solaris 11 по поддержке виртуальных сред, а также многоядерную многопоточную архитектуру. Предыдущие модели линейки – серверы Oracle T4/T5 – применяются для похожих задач. T-series довольно часто используют в качестве замены нескольких устаревших серверов платформы SPARC. Именно поэтому Oracle SPARC T7-2 в первую очередь был интересен с точки зрения возможностей виртуализации и консолидации приложений.

Принципиальных отличий в части виртуализации серверы линейки T7 по сравнению с T5 не имеют. Используется тот же гипервизор Oracle VM for SPARC, но более свежей версии. Архитектурно сервер T7-2 позволяет несколько более гибко распределять свои ресурсы ввода-вывода между отдельными виртуальными машинами в связи с появлением выделенных I/O-контроллеров (которые ранее были интегрированы непосредственно в CPU). Однако по сути общая конфигурация виртуальной среды остается без серьезных изменений, при этом имеющийся функционал актуален и полностью работоспособен, что наше тестирование полностью подтвердило. А вот от устаревших версий ОС придется постепенно отказываться: T7-2 накладывает определенные ограничения на версии используемых ОС Solaris 10 и 11.

В ходе тестирования мы также проверили ряд новых функций, которые появились с выходом этих серверов и соответствующих версий гипервизора. Одной из них является OVM templates. Это так называемые шаблоны для быстрого создания виртуальных машин. Пользователь может создавать

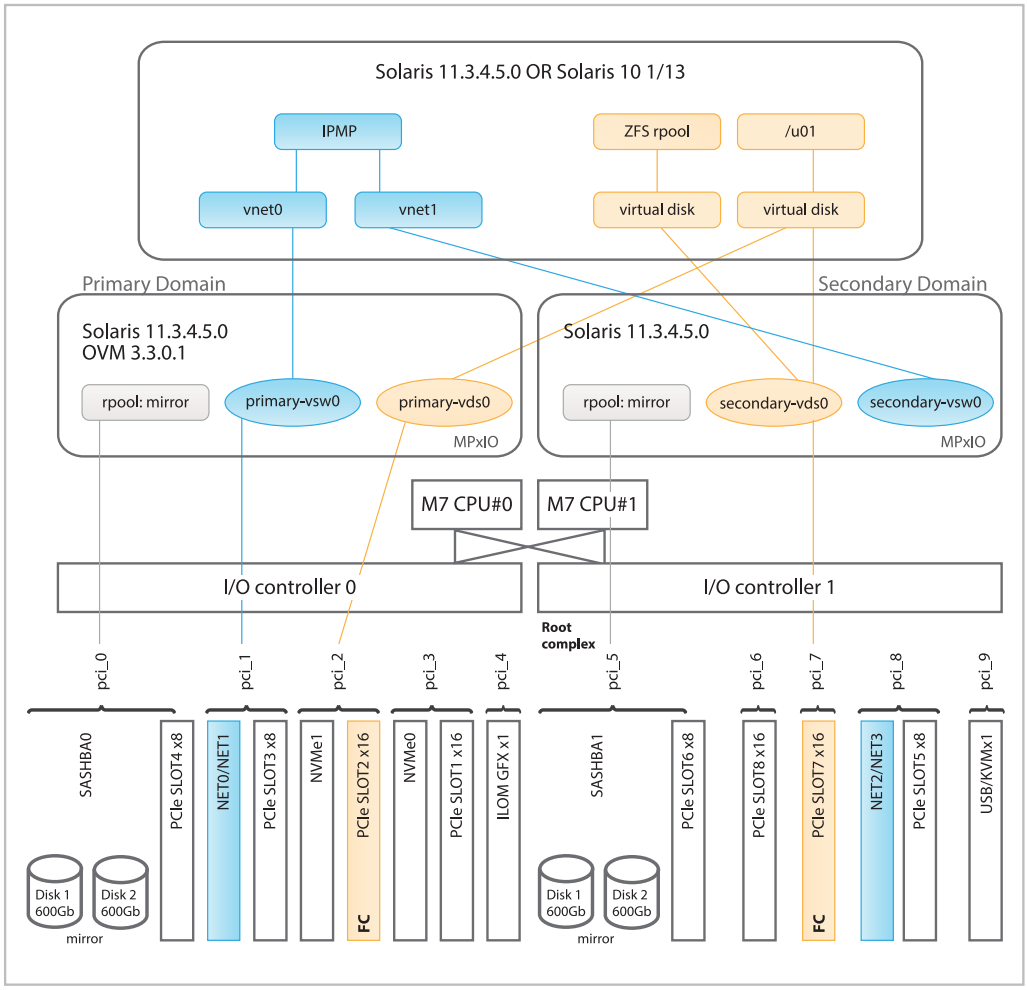


Рис. 1. Общая конфигурация виртуальной среды сервера Oracle SPARC T7-2

Рис. 2. Архитектура Oracle VM Manager

свои шаблоны либо использовать уже имеющиеся. Шаблон включает настройки виртуальной машины, необходимую версию ОС, пакеты предустановленного ПО, пользовательские окружения и т.д. Единожды создав такой шаблон, можно разворачивать аналогичные виртуальные машины гораздо быстрее и эффективнее. Таким образом, эта функция может значительно облегчить задачу развертывания большого количества похожих виртуальных машин. Это актуально для тех компаний, которым необходимо перенести в виртуальную среду большое количество небольших сред и систем в сжатые сроки.

Другое нововведение – технология virtual HBA (vHBA). Этот механизм позволяет полностью эмулировать на виртуальной машине (гостевом домене) SCSI-адаптер (FC HBA), который физически принадлежит одному из управляющих доменов (например, primary-домену). Таким образом, внутри виртуальной машины напрямую доступны все те устройства, которые подключены к этому адаптеру. Т.е. дисковые устройства, ленточные накопители и т.д. могут быть напрямую назначены конкретному гостевому домену. Необходимо отметить, что механизм на текущий момент имеет ряд ограничений. Тем не менее наличие подобной функции является еще одним шагом к упрощению настройки общей инфраструктуры виртуализации на серверах T-series. Этот механизм имеет смысл использовать в тех случаях, когда нужна простота в настройке дисковых устройств либо необходим доступ к специфичным SCSI-устройствам напрямую из гостевых доменов (как пример – резервное копирование напрямую на ленточные приводы).

> Архитектурно сервер Oracle SPARC T7-2 позволяет более гибко распределять свои ресурсы ввода-вывода между отдельными виртуальными машинами в связи с появлением выделенных I/O-контроллеров

Отдельное внимание при тестировании мы уделили функционалу автоматизированного управления сервером и его средой виртуализации. В первую очередь нас интересовало наличие графических средств управления настройками гипервизора OVM for SPARC. Эта задача может быть интересна заказчикам, широко использующим виртуализацию Oracle SPARC и желающим уменьшить объем процедур, исполняемых только средствами командной строки. В настоящее время для графического управления сервером Oracle SPARC T7-2 и средствами виртуализации OVM for SPARC доступны 2 варианта. Мы обозначали их для себя как «легкий» и «тяжелый». «Легкий» вариант – это ПО Oracle VM Manager. «Тяжелый» – Oracle Enterprise Manager Ops Center. Оба инструмента позволяют из графической консоли выполнять операции по графической настройке виртуальных машин на сервере. При этом указанные продукты имеют принципиально разную парадигму управления.

ПО Oracle VM Manager изначально развивалось как консоль управления виртуализацией x86 – платформами от Oracle, функциональность по части SPARC была добавлена позднее. OVM Manager хранит всю конфигурацию в своей встроенной базе данных, при этом сам гипервизор сервера фактически ничего не знает о виртуальных машинах. В случае выхода из строя или утери этой БД конфигурация будет утеряна. Кроме того, ряд операций по изначальной настройке сервера (создание управляющих доменов, разделение физических ресурсов) средствами OVM Manager невозможен в принципе, потребуется выполнить их руками. В целом применение OVM Manager вполне оправданно для тех компаний, которым необходимо построение решений вида «частное облако на SPARC», т.е. там, где требуется большое количество легких сред и приложений, при этом отказоустойчивость обеспечивается на уровне сервиса, а не гипервизора.

ПО Oracle Enterprise Manager Ops Center предлагает другой подход. По своей сути Ops Center – это комплексное конвергентное решение для управления всей ИТ-инфраструктурой, включая серверы, ОС, сети, системы хранения данных, кластерные конфигурации, СУБД и т.д. Оно имеет гораздо более широкий перечень поддерживаемого оборудования и ПО, чем OVM Manager. Управление виртуализацией является лишь небольшой его частью.

Ops Center целиком использует конфигурацию из гипервизора при создании и управлении виртуальными машинами. Количество ручных операций фактически сводится к нулю. Таким образом, это решение может использоваться теми компаниями, которым нужные сложные, нестандартные конфигурации виртуальных машин. Отметим, что обширность функциональных возможностей Ops Center является отчасти и его недостатком: продукт достаточно «развесистый», его изучение требует определенного времени и настойчивости. Тем не менее решение позволяет свести представление всей ИТ-инфраструктуры компании к одной консоли.

По результатам нашего тестирования мы можем заключить, что сервер Oracle SPARC T7-2 хорошо подходит для задач консолидации и виртуализации сред, функционирующих на платформе SPARC. Задачи построения надежных распределенных ферм виртуализации с динамическим перераспределением ресурсов и обеспечением высокой надежности могут решаться с применением этого оборудования весьма эффективно. Последовательное развитие средств виртуализации повышает его стабильность, функциональность и удобство применения для конечных пользователей. Наличие автоматизированных средств управления также позволяет упрощать повседневные задачи по эксплуатации и повышать отдачу от использования. С учетом новых возможностей в части производительности, встроенных в CPU (Oracle Database In-Memory), сервер представляется весьма выгодным вложением в ИТ-инфраструктуру. JI

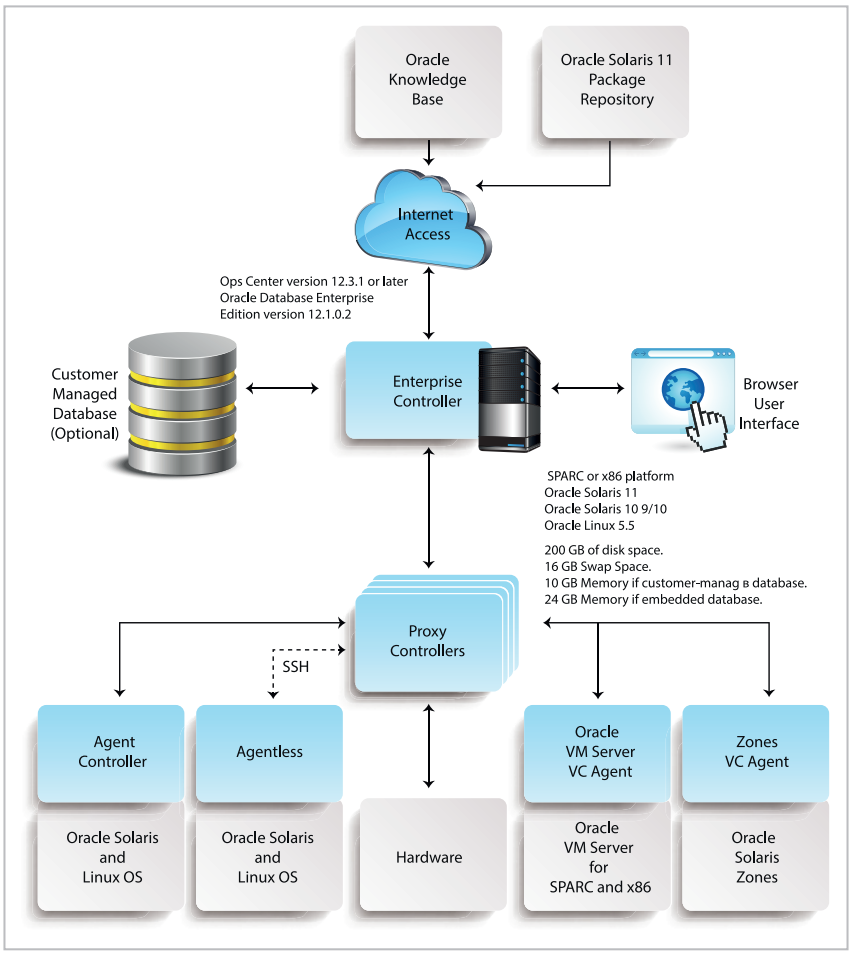


Рис. 3. Архитектура Oracle Enterprise Manager Ops Center

ПОЛНОСТЬЮ ТВЕРДОТЕЛЬНЫЕ СХД – ТЕХНОЛОГИЯ ЗАВТРАШНЕГО ДНЯ

АВТОР: ИЛЬЯ ВОРОНИН

Технология All-Flash – горячая тема на рынке. В чем основные преимущества и недостатки полностью твердотельных СХД? Какие задачи они позволяют решать и насколько эти решения экономически оправданны? На эти и другие вопросы отвечает **руководитель Центра проектирования вычислительных комплексов компании «Инфосистемы Джет» Илья Воронин.**

Источник: TAdviser, апрель 2016 г.

ТЕНДЕНЦИИ РЫНКА СХД И BIG DATA

АВТОР: АРТЕМ ДЕМАРИН

Какие задачи определяют развитие рынка СХД? Как развиваются технологии обработки Больших Данных и в чем подводные камни при реализации проектов в данном направлении? На эти и другие вопросы отвечает **Артем Демарин, руководитель группы развития СХД компании «Инфосистемы Джет».**

Источник: IT News, октябрь 2015 г.

Jet Info infosystems

Jet Info
ИЗДАЕТСЯ КОМПАНИЕЙ «ИНФОСИСТЕМЫ ДЖЕТ»

Главный редактор Дмитриев В. Ю.

Россия, 127015, Москва, Б. Новодмитровская, 14/1,
тел. (495) 411 76 01, факс (495) 411 76 02, e-mail: Jetinfo@jet.msk.su, www.jetinfo.ru

Полное или частичное воспроизведение материалов, содержащихся в настоящем издании, допускается только по согласованию с издателем