

Jet Info

ИНФОРМАЦИОННЫЙ БЮЛЛЕТЕНЬ

№ 11 (198)/2009

Современные подходы к построению вычислительных комплексов

ВЫЧИСЛИТЕЛЬНЫЕ
КОМПЛЕКСЫ

Современные подходы к построению вычислительных комплексов

СОДЕРЖАНИЕ

Новости3

Тема номера

XenoBridge с точки зрения «ораклиста» (А. Струченко)	5
Консолидация – скрытые резервы ЦОД (А. Иванов)	8
Как построить эффективное решение с IBM System Power (Д. Фомичев)	15
Система виртуализации серверов IBM Power (Д. Фомичев, И. Ушкало)	19
Вычислительные комплексы масштаба предприятия на платформе IBM Power (И. Ушкало)	23

Подходит к концу первый кризисный год. Каковы его уроки? Что изменилось на рынке вычислительных систем корпоративного уровня?

Первая тенденция – это экономия. Экономия на всем – и на капитальных затратах, и на операционных расходах. Следовательно – востребованы дешевые решения, младшие модели, консолидация всего и вся.

Вторая тенденция – бизнес все же не стоит на месте, требует развития, реализации новых бизнес-идей, бизнес-проектов. В условиях экономии – непростая задача для ИТ. Хорошо, что на помощь приходит виртуализация – ее технологии (а это, даже в вычислительных комплексах, несколько принципиально различных подходов, независимо развиваемых вендорами) стали достаточно зрелыми для использования в серьезных корпоративных системах.

Третья тенденция – пришло осознание того, что чудес не бывает и за все приходится платить. За экономию операционных расходов – внедрением достаточно сложных и стоящих денег решений, за низкие стартовые затраты – повышением стоимости обслуживания, за переход на новые технологии – затратами на миграцию...

Обо всем об этом наш номер: о консолидации и виртуализации, о технологиях и преимуществах, построении новых систем и миграции со старых.

Открытый Вами журнал – первый из серии, рассказывающих о решениях, основанных на подходах, продуктах и технологиях IBM. И это, кстати, еще одна из тенденций кризисного года: востребованы надежность этой известнейшей ИТ-фирмы, техническое совершенство и широкий спектр предлагаемых продуктов, авторитет консультантов вендора и опыт специалистов-партнеров.

IBM расширяет свои предложения корпоративных систем хранения данных

IBM представляет систему System Storage DS8700, наиболее совершенную модель в линейке DS8000, которая предлагает гораздо более высокие уровни производительности, надежности и устойчивости, а также энергетической эффективности. Система DS8700 оснащается новыми двухпроцессорными контроллерами на базе IBM POWER6, интерфейсом внутренних межсоединений на основе архитектуры Fibre Channel и модернизированными адаптерами устройства, что обеспечивает улучшение производительности вплоть до 150% по сравнению с наиболее мощной моделью предыдущего поколения. Эти нововведения расширяют способность DS8700 поддерживать самые ресурсоемкие требования бизнес-приложений и систем бизнес-анализа.

IBM System Storage DS8700 позволяет клиентам, не прерывая текущую работу, осуществлять модернизацию базовых моделей до наиболее мощных и функциональных конфигураций. Таким образом, организации могут наращивать возможности своих существующих систем по мере роста потребностей бизнеса без необходимости приобретения новых систем. Наряду с этим гибким и масштабируемым конструктивным исполнением, DS8700 дает клиентам возможность продолжать получать отдачу от инвестиций, которые были сделаны в системы предыдущего поколения. Так DS8700 предлагает полную функциональную совместимость с жесткими дисками, корпусами для дисков, программным инструментарием, скриптами и сервисами копирования. Новая система System Storage DS8700 также обладает большей энергетической эффективностью, демонстрируя более чем 50%-ный рост показателя количества операций ввода/вывода в секунду в расчете на потребляемую мощность в ваттах (IOPS/Watt) по сравнению с предыдущими моделями DS8300.

IBM укрепляет безопасность виртуальных сред

Корпорация IBM представила IBM Virtual Server Security for VMware® vSphere™, программный продукт, призванный помочь организациям обезопасить свои инфраструктуры виртуальных серверов. Новое решение позволит обеспечить информационную защиту сред виртуальных серверов и предоставит компаниям более безопасный путь для перевода критически важных активов в виртуальные корпоративные центры обработки данных.

Признание и распространение технологий виртуализации растет повсеместно по мере консолидации компаниями своих центров обработки данных. Наряду с тем, что миграция в виртуальные среды обещает множество преимуществ, она также требует от компаний принятия исключительных мер по борьбе с угрозами безопасности следующего поколения и решению проблем соблюдения соответствующих нормативных требований. Данные меры необходимы вследствие ухудшения прозрачности и контроля как результата добавления новых «слоев» информационных технологий. Учитывая эту тенденцию, традиционных средств обеспечения безопасности, созданных для физических вычислительных сред, становится недостаточно в качестве единственного используемого решения.

Новый продукт IBM Virtual Server Security for VMware® vSphere™ помогает справиться с этой проблемой, предоставляя средства защиты для каждого слоя виртуальной инфраструктуры, включая гипервизор, операционную систему, сеть, приложения, виртуальные «рабочие столы» на базе сервера, виртуальную машину, а также трафик между виртуальными машинами. Благодаря интеграции с технологией VMware VMsafe™, новой программный продукт обеспечивает клиентам улучшенный визуальный контроль, большую степень детализации защиты и повышенную масштабируемость в растущих средах их виртуальных центров обработки данных.

КБ «РОСПРОМБАНК» (ООО) подготовлен к введению режима коммерческой тайны

В КБ «РОСПРОМБАНК» (ООО) с помощью компании «Инфосистемы Джет» успешно завершена подготовка к введению режима коммерческой тайны и окончен аудит банка на соответствие требованиям стандарта СТО ИББС Банка России.

В ходе проекта консультанты компании «Инфосистемы Джет» провели оценку системы обеспечения информационной безопасности в соответствии с требованиями стандарта СТО ИББС Банка России, а также выработали дальнейшие шаги по повышению уровня защищенности информационной системы банка. Также был подготовлен комплект документов, регламентирующих режим коммерческой тайны в соответствии с Законом «О коммерческой тайне», и разработаны рекомендации по повышению эффективности защиты этого вида конфиденциальной информации.

Высокий уровень исполнительской дисциплины в банке и хорошо проработанная нормативная база в части обеспечения защиты информации обусловили достаточно высокую оценку соответствия требованиям стандарта СТО БР ИББС, а также готовность банка к введению режима коммерческой тайны.

Проведение аудита и подготовка к введению режима коммерческой тайны в банке изначально были разными проектами, но специалисты компании «Инфосистемы Джет» сумели их объединить и часть работ выполнить совместно. Это позволило не только сократить время проведения проектов, но и существенно снизить финансовые затраты банка.

Расширение портфеля решений для информационной инфраструктуры

IBM расширяет продуктовую линейку систем хранения данных N series, в том числе представляя новые ускорители записи Performance Accelerator Module (PAM II), которые обладают в 16 и 32 раза большей емкостью, чем предыдущая версия. Это позволяет использовать увеличенную кэш-память для улучшения время отклика до 30% для всех систем N-серии, помогая при этом снизить энергопотребление до 50% по сравнению с применением накопителей на традиционных жестких дисках. В IBM также разработаны новые дисковые модули расширения IBM System Storage EXN3000 с интерфейсом SAS (Serial Attached SCSI) для N series, которые отличает высокая емкость и устойчивость работы, а также на 22% лучший показатель эффективности использования пространства в аппаратной стойке, что позволяет разместить больше дисков в секциях стойки. неотъемлемой частью этого выпуска является новая технология виртуализации SnapManager for Microsoft Hyper-V для систем N series, которая предоставляет клиентам инструмент для управления виртуализацией, чтобы обеспечить автоматическую защиту и аварийное восстановление данных для сред Microsoft Hyper-V.

XenoBridge с точки зрения «ораклиста»

Алексей Струченко,
руководитель группы оптимизации СУБД и приложений,
компания «Инфосистемы Джет»

Большинство проектов в кризисное время так или иначе направлено на сокращение расходов на поддержание ИТ-инфраструктуры — в первую очередь за счет переноса составляющих ИТ-комплекса на более дешевые сервера. В свете этого задача межплатформенной миграции Oracle Database становится все более актуальной.

Данной теме в последнее время уделяют внимание немало компаний, в том числе и IBM. На недавнем семинаре, посвященном XenoBridge (один из продуктов, составляющих IBM Migration Factory), было показано, как он работает с базой данных Oracle. При этом заявляется, что возможна не только работа с другими базами, но и миграция между ними.

Рассмотрим способы межплатформенной миграции, которые предлагает сам Oracle. В первую очередь это проверенные временем способы переноса данных:

- export/import;
- create as select через db link;
- insert select через db link.

Эти способы существовали задолго до выхода Oracle Database 10g и в целом считаются медленными. Скорость миграции тем или иным способом существенно зависит от структуры базы данных, однако в общем можно сказать что скорость переноса 100Гб за 10 часов для export/import считается хорошей. Поэтому в среде администраторов Oracle существует устойчивое мнение, что данные способы не подходят для миграции современных 24x7 баз данных, когда размер базы 1Тб и больше, а окно не превышает 10 часов (одна ночь).

Именно поэтому начиная с Oracle Database 10g появилась возможность конверта-

ции Transportable Tablespace с помощью RMAN. Сами Transportable Tablespace существовали в Oracle и до выхода Oracle Database 10g, они позволяли относительно быстро мигрировать данные в рамках одной платформы. Теперь же появилась возможность конвертации, которая может быть организована как на старой платформе, так и на новой — там, где достаточно дискового пространства. Данным способом можно перенести базу данных размером 1Тб в среднем за 5-10 часов (на это время существенно влияет скорость сети и время отклика дисковой подсистемы).

Однако у межплатформенной миграции через Transportable Tablespace с конвертацией RMAN есть недостаток. База данных переносится физически (на уровне файлов) и нет никакой возможности ее реорганизовать. А в проектах миграции современных систем все чаще возникают дополнительные требования по реорганизации данных — по крайней мере в следующих направлениях:

- изменение фрагментации сегментов (как правило уплотнение таблиц);
- изменение структуры партиций (либо организация, либо реорганизация партицированных таблиц и индексов);
- перестройка индексов и (что реже) ограничений.

В этом случае Transportable Tablespace не помогают, и администраторам Oracle придется возвращаться к старым «медленным» способам. Оказывается, что параллельный поэтапный export/import в сочетании с несколькими db link позволяет перенести 1Тб за 10 часов с попутной реорганизацией данных. Но несмотря на то, что в Oracle Database 10g появился достаточно удобный инструментарий (DataPump), та-

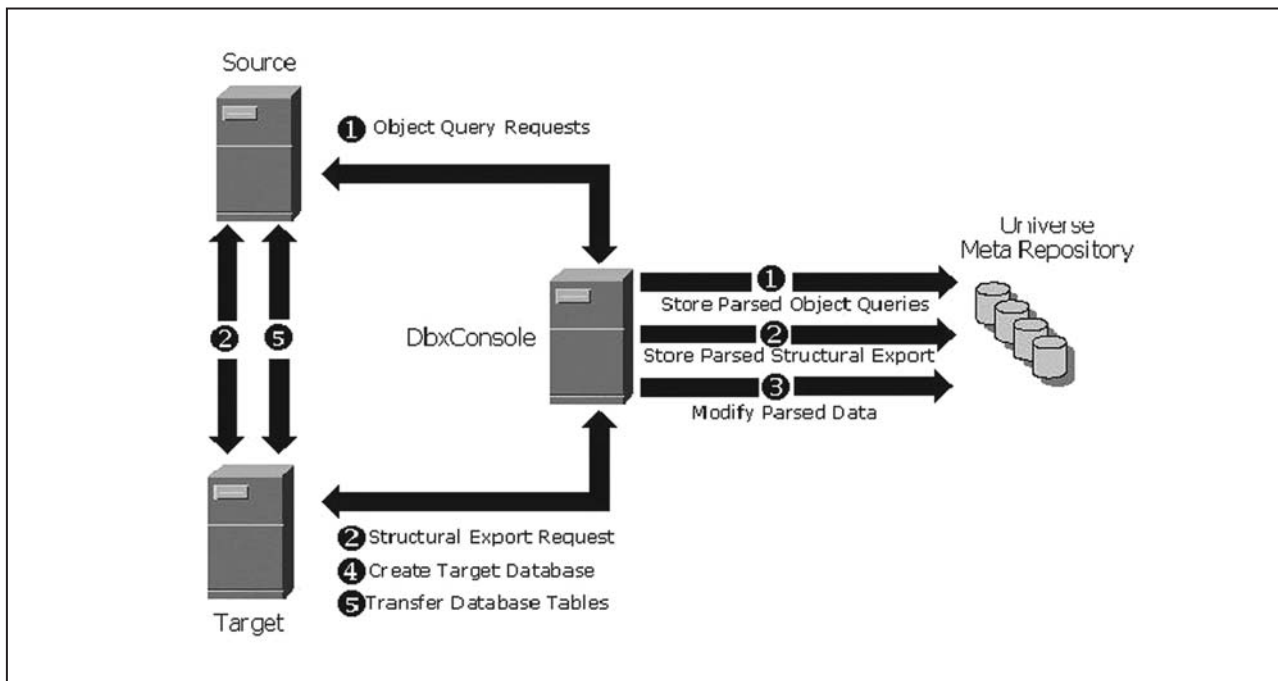


Рис. 1. Архитектура XenoBridge

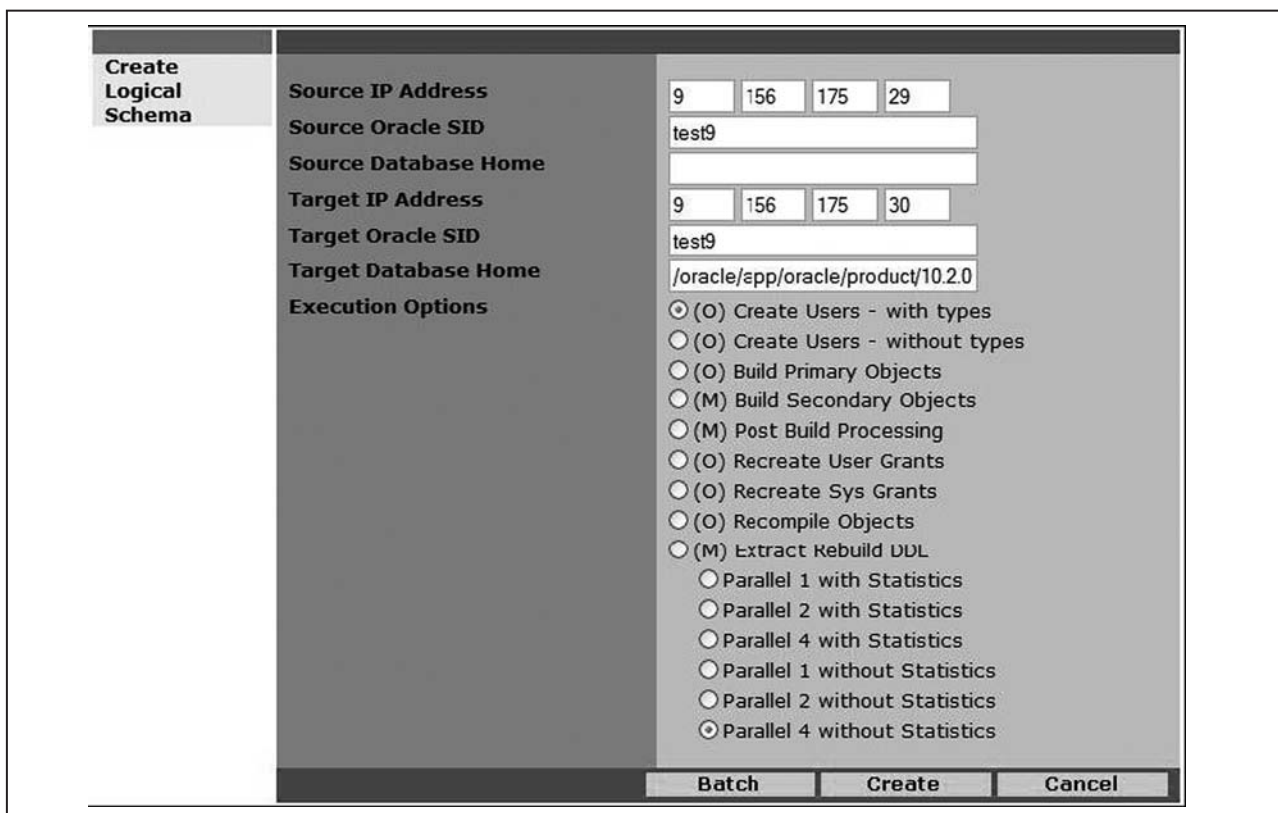


Рис. 2. Создание логической структуры

кой проект получается трудоемким и только тестирование может занять месяцы. Для решения подобных задач как раз и создан XenoBridge, осуществляющий именно такой интеллектуальный алгоритм миграции, который описан выше.

При этом межплатформенная миграция Oracle Database с помощью XenoBridge выглядит следующим образом:

- перенос структуры базы данных с возможностью реорганизации;

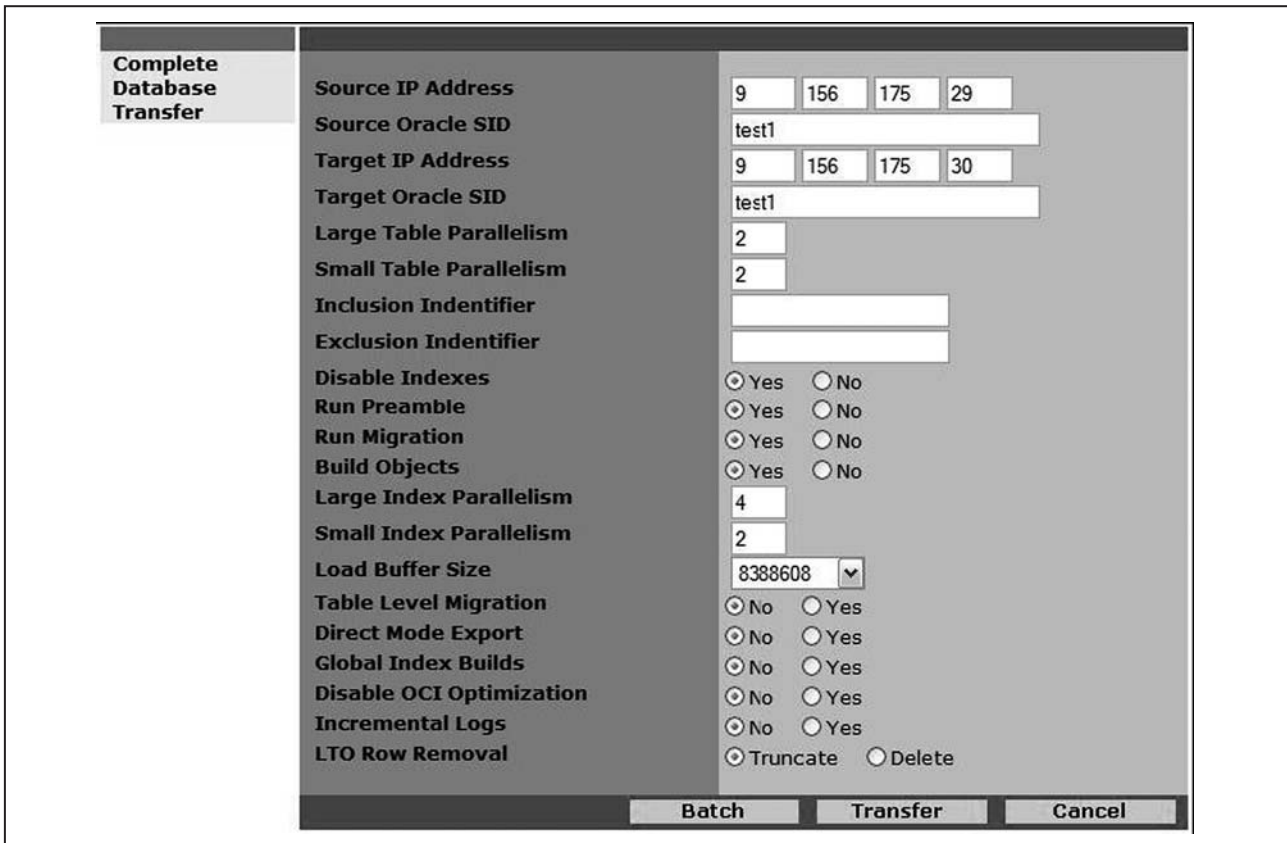


Рис. 3. Основной этап миграции

- проверка перенесенной структуры в сравнении с изначальной;
- отключение индексов/триггеров/ограничений;
- интеллектуальный перенос самих данных — в зависимости от размера таблицы либо через db link, либо export/import параллельно для разных групп таблиц/партиций;
- включение триггеров/ограничений и пересоздание индексов с заданной степенью параллелизма.

На рис. 1 схематично представлена архитектура продукта XenoBridge — консоль имеет свой собственный репозиторий, в частности необходимый для реорганизации данных в процессе миграции.

Заявленная скорость миграции Oracle Database с помощью XenoBridge при средней ско-

рости сети и среднем времени отклика дисковой подсистемы — 100Gb/час, что дает для 1Тб приемлемое время 10 часов. Наглядным примером работы XenoBridge является опыт специалистов IBM по миграции базы данных одного из крупных казахским банков (размер около 1Тб). Данные были успешно перенесены в выделенное окно (порядка 12 часов), при этом все процессоры в ходе миграции были постоянно загружены на 90-95%, что является хорошим показателем утилизации и едва ли достижимо при «ручной» миграции «медленными» способами.

В завершение следует сказать, что продукт XenoBridge имеет дружелюбный интерфейс, существенно облегчающий процесс межплатформенной миграции Oracle Database. На рис. 2 и рис. 3 приведены примеры этого интерфейса для двух важнейших этапов миграции — создания логической структуры и собственно переноса данных.

Консолидация – скрытые резервы ЦОД

Александр Иванов,
инженер-проектировщик в группе вычислительных систем,
компания «Инфосистемы Джет»

Сегодня многие центры обработки данных (ЦОД) содержат большое число разнотипного оборудования различных вендоров, что с одной стороны, избавляет бизнес от рисков, связанных с привязкой к одному производителю, а с другой – удорожает эксплуатацию и ухудшает управляемость всего комплекса разнообразных (подчас слишком) информационно-технических средств (ИТС). Тенденция по формированию парка оборудования из соображения «одна задача – один сервер» сложилась и укоренилась в умах ИТ-сообщества (и использующего его «бизнеса») около 7-9 лет тому назад, что в текущий момент привело к насыщению ЦОД-ов, даже в среде SMB-компаний, менеджмент которых никогда не сталкивался с проблемами подобного рода. Часто проблема стоит даже не в физических ограничениях по размещению оборудования, а в отсутствии дополнительных мощностей по питанию и отводу тепла от оборудования в ЦОД. При этом потребляющее предельное количество энергии оборудование часто не имеет адекватной рабочей нагрузки, что приводит к неэкономичным показателям производительности на ватт потребляемой мощности. А такая ситуация крайне болезненно отражается на финансовых показателях бизнеса, особенно в текущей экономической обстановке.

Однако бизнес, пытаясь преодолеть финансовые трудности, ищет все новые возможности для развития, новые проекты, способы заработать. Это часто требует увеличения нагрузки на существующую ИТ-инфраструктуру, и в некоторых случаях, особенно когда начинают экономить на ИТ, происходит насыщение (перенасыщение) бизнес-задачами текущей ИТ-инфраструктуры.

Основная цель, стоящая перед ИТ-подразделениями в описанных ситуациях в частности и в текущий период в целом, – обеспечить бизнес-приложениям возможности роста при использо-

вании всех инструментов для достижения экономической выгоды как от прямой отдачи бизнес-процессов, так и от экономии на накладных эксплуатационных расходах (ТСО). Попутно необходимо всеми способами доносить до людей, принимающих решения, что сиюминутная экономия на закупке новой ИТ-инфраструктуры или внедрении новых ИТ-проектов приведет к необоснованному росту эксплуатационных расходов на устаревающее оборудование. Весь смысл этой необходимости сформулирован в фразе: «Не экономьте на ИТ, экономьте вместе с ИТ».

Консолидация и виртуализация

Для разрешения вышеописанных ситуаций наилучшим направлением и способом развития ИТ-инфраструктуры является консолидация вычислительных ресурсов, для уменьшения занимаемых площадей и потребляемой мощности на единицу производительности, а также консолидация программных средств реализации бизнес-процессов на единую аппаратную платформу с выделением каждой задаче своего окружения.

Так, например, компания IBM предлагает одни из лучших передовых решений в области консолидации разнородных приложений на базе серверов x86-ой архитектуры, в частности – на базе блэйд-серверов высокой плотности. Для консолидации и виртуализации программной среды IBM успешно сотрудничает с несколькими производителями профильного ПО, в частности с компанией VMware, также выпускающей передовые решения по виртуализации на платформе x86.

Интеграцией всех преимуществ такого аппаратно-программного комплекса с успехом за-

нимаются крупные системные интеграторы, в том числе и компания «Инфосистемы Джет». Опыт реализации подобных проектов позволяет нашим специалистам разрабатывать типовые решения на базе серверов IBM BladeCenter System X и BladeCenter Power System для применения в раз-

личных отраслях: банковские системы, карточный процессинг, телекоммуникационная отрасль, биллинговые системы и т.д.

Основой такого решения, ядром платформы консолидации может являться, например, IBM BladeCenter H, обладающий выдающимися воз-

Переход на платформу IBM BladeCenter






Упрощение. Сокращение расходов. Повышение продуктивности. Экологичность. Эти первоочередные ИТ-задачи вынуждают организации пересмотреть свои стратегии построения серверной инфраструктуры. Blade-системы – это решение следующего поколения, которое обещает существенные улучшения по нескольким направлениям.

- Платформа BladeCenter основана на IBM X-Architecture® – базовой стратегии IBM для

реализации инноваций в секторе x86-систем.

- Поддержка широкого спектра интегрированных коммутационных устройств ведущих поставщиков: Cisco, Nortel, Brocade, QLogic и McDATA.
- Системы BladeCenter поддерживает широкий спектр возможностей подключения, включая интерфейсы: InfiniBand (QDR), SAS, Ethernet, Fibre Channel (FC), Fibre Channel over Ethernet (FCoE) и SCSI over IP (iSCSI).

Таб. 1. Основные характеристики линейки IBM BladeCenter

Область применения	Малый и средний бизнес, филиалы		Корпоративные решения	Телеком и специальное применение	
	BladeCenter S	BladeCenter E	BladeCenter H	BladeCenter T	BladeCenter HT
    					
Высота при установке в шкаф, U	7	7	9	8	12
Количество отсеков для серверов	6	14	14	8	12
Количество портов ввода/вывода на отсек лезвия	4	4	8	4	8
Количество отсеков для коммутаторов	4	4	8	4	8
Количество отсеков для дисков	12, 3.5" S/SATA	-	-	-	-
Интеграция коммутаторов и кабелирования	Сеть, хранилище, консоль (KVM)	Сеть, хранилище, консоль (KVM)	Сеть, хранилище, консоль (KVM)	Сеть, хранилище, консоль (KVM)	Сеть, хранилище, консоль (KVM)
Блоки питания	4 (N+N) по 950 или 1450 Вт	4 (N+N) по 2000 или 2320 Вт	4 (N+N) по 2900 Вт	4 (N+N) по 1300 Вт (AC или DC)	4 (N+N) по 3160 Вт AC или по 2535 Вт DC
Вентиляторы	4	2	2	4	4
Модули управления	1	2	2	2	2
Серверы-лезвия	Intel Xeon, AMD Opteron, IBM Power6, IBM PowerXCell 8i	Intel Xeon, AMD Opteron, IBM Power6, IBM PowerXCell 8i	Intel Xeon, AMD Opteron, IBM Power6, IBM PowerXCell 8i, IBM BladeCenter PN41 (Intel IXP2805 network processor)	Intel Xeon, AMD Opteron, IBM Power6, IBM BladeCenter PN41 (Intel IXP2805 network processor)	Intel Xeon, AMD Opteron, IBM Power6, IBM PowerXCell 8i, IBM BladeCenter PN41 (Intel IXP2805 network processor)
Технологии ввода/вывода	Ethernet, SAS, Fibre Channel, iSCSI	Ethernet, SAS, Fibre Channel, iSCSI	Ethernet, SAS, Fibre Channel, iSCSI, InfiniBand, FCoE	Ethernet, SAS, Fibre Channel, iSCSI	Ethernet, SAS, Fibre Channel, iSCSI, InfiniBand, FCoE

возможностями по расширению подсистемы «ввода\вывода» модулями с использованием различных технологий передачи данных: Gigabit Ethernet, 10 Gigabit Ethernet, FibreChannel 4/8 Gb (FC), FibreChannel over Ethernet (FCoE), Serial Attached SCSI (SAS), iSCSI, InfiniBand QDR (40 Gb).

Также, как типовое решение для филиалов и удаленных, возможно, автономных офисов, может быть использован интегрированный комплекс IBM BladeCenter S, включающий в себя помимо 6-ти блэйд-серверов любого типа, стандартных модулей коммутации FC и Ethernet, еще и разделяемую между серверами дисковую систему хранения (до 12 дисков SAS/SATA). Возможно применение блэйд-серверов со специализированными процессорами IBM PowerXCell 8i, Cell Broadband Engine для ускорения обработки интенсивных задач высокопроизводительных вычислений, таких как финансовый анализ, обработка сейсмических данных, обработка изображений, обработка сигналов и визуализация изображений, и специализированного сетевого процессора Intel® IXP2805 для обработки пакетов и анализа сетевого трафика.

Для повышения утилизации аппаратных ресурсов IBM BladeCenter специалисты компании «Инфосистемы Джет» считают целесообразным использовать в типовом решении программное обеспечение VMware vSphere 4, позволяющее разместить несколько различных операционных систем на одном аппаратном сервере. В данном случае может быть использована возможность защиты приложения от сбоев, путем построения отказоустойчивых конфигураций с помощью технологии VMware High Availability. Технология VMware VMotion позволяет перенести продуктивную нагрузку с одного аппаратного сервера на другой с минимальными простоями бизнес-приложений, а также практически в режиме реального времени балансировать нагрузку на аппаратных компонентах системы или даже между несколькими аппаратными комплексами (IBM BladeCenter или System X), перенося виртуальные машины с более нагруженных систем на менее нагруженные.

Но для реализации всех возможностей ПО VMware в таком решении необходимо разделяемое дисковое хранилище. Целесообразность применения дисковой системы того или иного уровня определяется специалистами компании-интегратора в сотрудничестве со специалистами заказчика на основании бизнес-требований к проектируемому комплексу (возможные варианты решения см. в приложении 1, стр 14).

При создании комплекса следует учитывать необходимость резервирования критичных данных, что подразумевает построение системы резервного копирования (СРК). В этой области специалисты компании «Инфосистемы Джет» обладают высокой экспертизой в построении таких систем с нуля, их миграции и построении централизованных систем резервного копирования для нескольких разнесенных площадок (консолидации). И основываясь на своем опыте, рекомендуют продукты компании IBM по резервированию данных и управлению их жизненным циклом, такие как IBM Tivoli Storage Manager, IBM Tivoli Storage Productivity Center и т.д., а также спектр аппаратного обеспечения для построения СРК, ленточные библиотеки начального, среднего и Hi-End уровня. Также компания IBM предлагает несколько законченных решений для виртуализации, управления хранением и дедупликацией резервируемых данных.

Но оптимальным комплексом оборудования с точки зрения соотношения цена/производительность – серверы, системы хранения и резервного копирования – специалисты нашей компании считают использование оборудования одного вендора, такого, например, как IBM, поскольку этот подход позволяет заказчику достичь целей в бизнесе и оптимизировать временные и финансовые затраты на построение и эксплуатацию ИТ-инфраструктуры.

Примером такого типового решения может служить построенный нашей компанией комплекс для одной финансовой организации. В разработанном комплексе был применен IBM BladeCenter E, обладающий неплохим соотношением цены и доступных возможностей по построению, наращиванию и управлению комплексом физических и виртуальных серверов. В составе комплекса применялись серверы IBM HS22, использующие новейшие процессоры Intel Xeon с архитектурой ядра Nehalem с производительностью, позволяющей разместить до нескольких десятков операционных систем на одном физическом сервере и исполнять несколько десятков или даже сотен задач в пределах кластера серверов VMware vSphere ESX. При этом в любое время клиент имеет возможность расширить построенный комплекс, не прерывая работы добавлением нового сервера или даже полного шасси. Все новейшие технологии виртуализации Intel® VT, Intel® VT-d в составе архитектуры процессоров Intel Xeon, эффективно используются компанией IBM при разработке серверов на их основе, что позволяет применить все достоинства ПО виртуализации от VMware, использующего эти техноло-

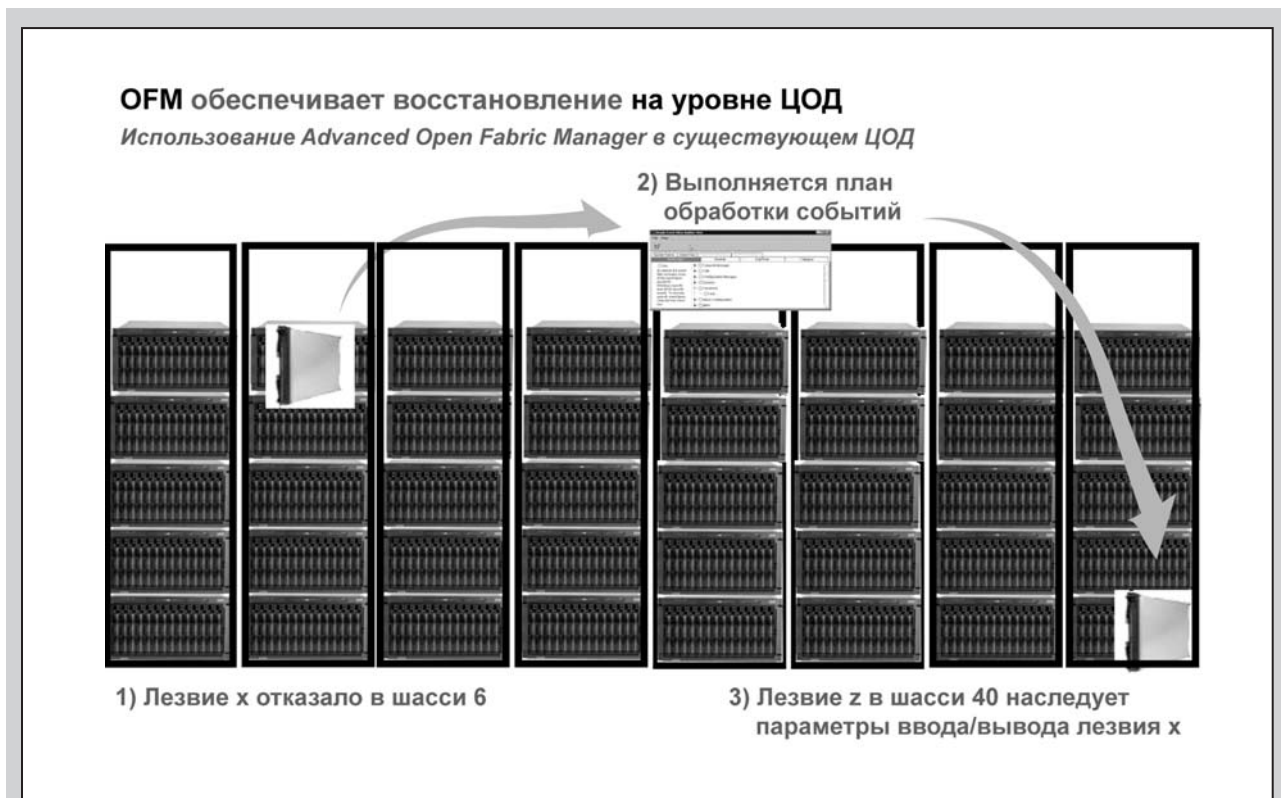


Рис. 1. Построение ЦОД с использованием IBM BladeCenter. Защита от сбоев с использованием технологий OFM

BladeCenter Open Fabric Manager

Программное обеспечение IBM BladeCenter Open Fabric Manager разработано, чтобы помочь в решении задач, связанных с ростом и усложнением среды. Оно облегчает управление системами ввода/вывода и сетевыми соединениями для максимум 100 шасси BladeCenter – то есть, для максимум 1400 blade-серверов. Программное решение Open Fabric Manager также сокращает время простоев за счет использования планов действий для автоматизированного восстановления после сбоев (с участием нескольких шасси или серверов в одном шасси). После установки эта утилита постоянно находится во встроенном модуле Advanced Management Module (АММ) и позволяет предварительно конфигурировать соединения с локальными сетями и сетями хранения данных. Соединения ввода/вывода создаются автоматически при подключении blade-сервера.

Основные характеристики ПО IBM BladeCenter Open Fabric Manager:

- Позволяет произвести первичную настройку и перенастройку LAN и SAN коммутаторов. Это можно сделать единожды для каждого блейд-сервера.
- Позволяет контролировать всю серверную структуру через один Advanced Management Module – до 100 шасси, до 1400 блейд-серверов.
- Работает со всеми Ethernet, FCoE, Fibre Channel и SAS-коммутационными модулями, со всеми шасси и с большинством Intel (включая HS22) и Power (включая JS23/43) блейд-серверов.
- Позволяет произвести автоматический ввод в эксплуатацию резервного сервера при отказе основного в соответствии с настроенными планами обработки событий (blade «rip and replace» – автоматическая подмена MAC, WWN, VLAN FCP Target настроек).
- Поддерживает все технологии высокой доступности VMware и XEN, позволяя на лету мигрировать настройки модулей ввода/вывода серверов.
- Обеспечивает простой доступ с использованием WEB-интерфейса, базирующегося на IBM Systems Director 6.1, позволяя легко осуществлять настройку, размещение и управление серверами и коммутаторами.

гии в комплексе с микрокодом BIOS серверов IBM и драйверами системных устройств.

Для хранения данных виртуальных машин и прикладных систем в построенном комплексе была применена дисковая система хранения среднего уровня IBM DS4700, подключающаяся к сети хранения данных (SAN) по высокоскоростным каналам FibreChannel и имеющая возможность использовать в своем составе как высокоскоростные диски с интерфейсом FC, так и диски большого объема с интерфейсом SATA.

Систему резервного копирования (СРК) данных решено было построить на основе ПО IBM Tivoli Storage Manager, развернутого на одном из серверов HS22 с использованием масштабируемой ленточной библиотеки IBM TS3310. Возможно произвести развертывание СРК в виртуальной машине Windows с среде VMware.

Для управления серверным комплексом было применено ПО настройки, управления и мониторинга, обладающее функционалом для повышения отказоустойчивости IBM BladeCenter Open Fabric Manager (OFM) Advanced (см.рис.1),

интегрированное с унифицированным средством управления серверами стандартной архитектуры x86 IBM Director, распространяемым в стандартной поставке серверов IBM System X.

IBM Director является очень удобным средством интеграции модулей управления всеми компонентами систем IBM, и специалисты компании «Инфосистемы Джет» в рассматриваемом решении с успехом применили его гибкость и функциональность, собрав в единую точку контроля и управления средства настройки дискового массива, средства настройки СРК, средства настройки, управления и мониторинга серверного комплекса.

Компонентная схема построенного комплекса приведена на рисунке 2.

Сеть передачи данных (СПД Ethernet) и сеть хранения данных (SAN FibreChannel) были построены на базе модульных коммутаторов, устанавливаемых в шасси BladeCenter. Сеть Ethernet была построена с применением стека Ethernet-коммутаторов Cisco по отказоустойчивой технологии, предусматривающей подключение к любому серверу или подсистеме комплекса

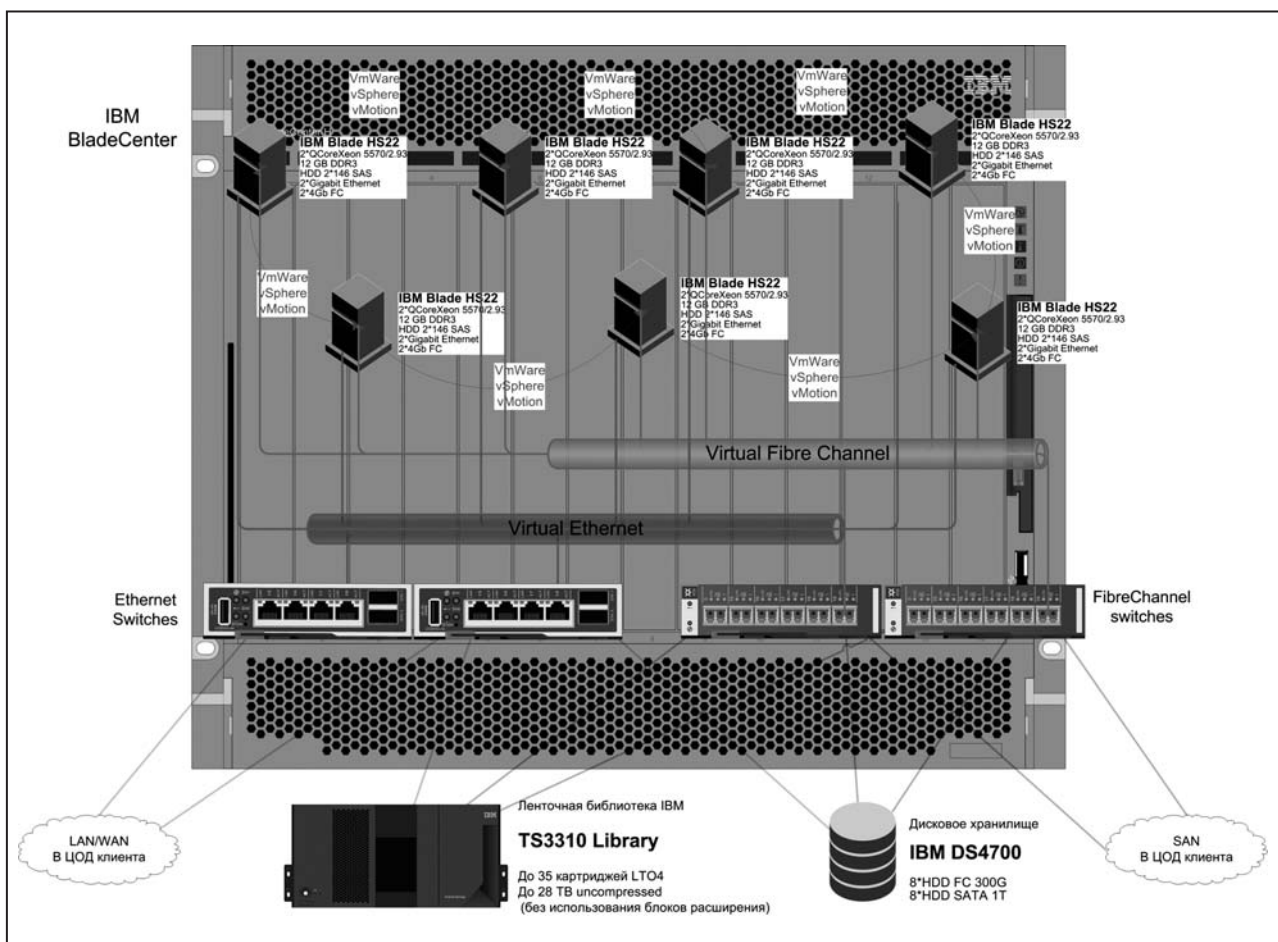


Рис. 2. Компонентная схема построенного комплекса

извне по двум физическим путям. SAN была построена с помощью FC-коммутаторов Brocade, также с использованием отказоустойчивой технологии Dual Fabric, обеспечивающей связь подсистем серверного комплекса с хранилищами данных по двум независимым путям.

IBM BladeCenter по сравнению с аналогичными продуктами конкурентов обладает несколькими выдающимися особенностями:

- виртуализация ввода\вывода при использовании широких возможностей IBM BladeCenter Open Fabric Manager Advanced для быстрого восстановления после отказа оборудования;
- встроенные средства управления (KVM-функционал), полный набор функций управления системами (IBM Light Path Diagnostics);
- наиболее высокая производительность инфраструктуры ввода/вывода (Gigabit и 10G Ethernet, 4X InfiniBand QDR, SAS, FC, FCoE);
- экономия на функционировании за счет энергосбережения (наиболее энергоэффективное решение); до 20-35% меньше затрат на электроэнергию по сравнению с конкурентами; до 50% при использовании IBM Active Power Management;
- позволяет осуществлять долгосрочную защиту инвестиций за счет совместимого форм-фактора шасси для любых серверов-лезвий.

Таким образом, преимущества предлагаемого компанией «Инфосистемы Джет» такого типового комплекса на основе продуктов IBM по сравнению с устаревшей схемой «одна задача — один сервер» очевидны и значительны:

- сокращение расходов на содержание ИТ-инфраструктуры (снижение TCO в расчете на 2-5 лет);
- видимый эффект экономии на эксплуатационных расходах (ROI) в течении 15-24 месяцев;
- внедрение инноваций, предлагаемых компанией IBM во многих своих решениях и продуктах, позволит справиться с большим количеством ответственных бизнес-нагрузок;
- те же инновации и «зеленые» технологии, обширно развиваемые западными производителями и в большом объеме компанией IBM, позволят клиентам сократить энергозатраты при эксплуатации и повысить экологичность своего ЦОД;
- широкая гамма средств управления позволит достаточно легко справляться с ростом сложности инфраструктуры и осуществлять контроль возникающих рисков.

Заключение

Таким образом, компания IBM предоставляет широкий спектр отличного оборудования и программного обеспечения, и благодаря компетенции специалистов компании «Инфосистемы Джет», заказчик комплексов и решений подобного типа получает современные конкурентноспособные технологические решения поставленных перед ним задач и имеет возможность эффективно решать вопросы последующей эксплуатации, модернизации и масштабирования своей ИТ-инфраструктуры.

Приложение 1. Линейка дисковых массивов IBM

Название	Системы начального уровня для малых предприятий и филиалов (небольшие комплексы)	Системы среднего уровня для средних и крупных предприятий (средние и крупные комплексы)			Системы корпоративного уровня для крупных предприятий (крупные разнородные серверные комплексы)	
	DS3200 / DS3300 / DS3400	DS3950 / DS5020	DS4700	DS5100 / DS5300	DS6800	DS8100 / DS8100/DS8700
Интерфейсы для серверов	3 Gbps SAS / 1 Gbps iSCSI / 4 Gbps FC	8 Gbps FC и 1 Gbps iSCSI	4 Gbps FC	8 Gbps FC и 1 Gbps iSCSI	4 Gbps FC, FICON	4 Gbps FC, FICON, ESCON
Количество каналов к серверам (при 2-х контроллерах)	От 2 до 6/ 4 / 4	До 4 FC+4 iSCSI/ до 8 FC или 4 FC+4 iSCSI	8	16	8	64/128/128
Метод соединения с сервером	Прямое подключение/ Прямое подключение или IP-сеть/ Прямое подключение или FC-сеть	Прямое подключение или FC-сеть или IP-сеть	Прямое подключение или FC-сеть	Прямое подключение или FC-сеть или IP-сеть	Прямое подключение или FC-сеть	Прямое подключение или FC-сеть
Форм-фактор	2U, 12 дисков	3U, 16 дисков	3U, 16 дисков	4U-контроллер, 4U-дисковые полки до 60 дисков	3U, 16 дисков	Стойка 42U
Максимальное число дисков	48	112	112	448	128	384/1024/1024
Тип дисков	SAS/SATA II	FC/FDE/SATA II	FC/SATA II	FC/FDE/SATA II	FC	FC/SATA II
Каналы дисков	2 по 3 Gbps	4 по 4 Gbps	4 по 4 Gbps	16 по 4 Gbps	8 по 4 Gbps	32/64/64
Объем кэш-памяти для 2-х контроллерной системы	1GB (288 MB for memory)	До 4 GB	До 4 GB	До 64 GB	4	128/256/384

Как построить эффективное решение с IBM System Power

Дмитрий Фомичев,
инженер-проектировщик в группе вычислительных систем,
компания «Инфосистемы Джет»

Виртуализация как концепция консолидации для отдельно взятых продуктов и элементов ИТ-инфраструктуры известна достаточно давно, еще со времен систем IBM S/360, и охватывает разные уровни — от платформ (включая микропроцессоры) до ПО и реализуется различными способами: программными, аппаратными и программно-аппаратными.

Интерес к технологиям виртуализации обусловлен, прежде всего, как к действенному способу повышения эффективности использования ИТ-ресурсов. Одним из перспективных направлений развития ИТ-инфраструктуры предприятия является использование сервисной модели услуг и предоставления необходимых вычислительных ресурсов системы и информации по требованию (On Demand). Сделать это в рамках традиционной схемы (выделенные физические ресурсы для конкретного приложения) достаточно сложно. Если прежде проблема решалась путем приобретения дополнительного более мощного оборудования, то сегодня есть возможность более рационально использовать имеющиеся вычислительные ресурсы.

Положение дел

В качестве примера рассмотрим крупную компанию N. Как любая современная и динамичная компания для поддержки своего бизнеса наш заказчик опирается на информационные технологии и с первых дней своего существования старается обеспечить работоспособность большого количества сервисов, таких как электронная почта, телефония, доступ к ресурсам интернет, файловые серверы для хранения документов, серверы баз данных, а также различных бизнес-приложений.

Реалии современного бизнеса таковы, что заказчики стремятся сократить и оптимизировать эксплуатационные расходы, и на ИТ обычно выделяется очень ограниченный бюджет, но при этом заказчик хочет получить полнофункциональную и масштабируемую ИТ-инфраструктуру. Особенно это актуально для крупных компаний, которые уделяют большое внимание таким аспектам, как надежность и доступность сервисов, у которых затраты на владение ИТ-инфраструктурой велики. И рассматриваемый нами заказчик здесь не исключение. Периодически у него возникает необходимость решения таких задач, как ввод в действие новых ИТ-сервисов, например, новых функциональных модулей SAP R/3. Ведь оборудование информационных систем изнашивается физически, устаревает морально, и заказчик планирует замену устаревшего оборудования и миграцию ряда систем со старого на новое. При этом остро стоит проблема оптимизации расходов на лицензирование ПО (например, ORACLE), а также отсутствие инженерных ресурсов, например, подведение дополнительной электрической мощности для дальнейшего роста инфраструктуры.

При развертывании ИТ-сервисов наш заказчик (как best practice) использует трехуровневый ландшафт, состоящий из продуктивной системы, в состав которой, как правило, входят серверы доступа, серверы приложений SAP R/3 и СУБД Oracle, а также системы разработки и тестирования каждого модуля SAP R/3 и ряда вспомогательных систем. Используемый нашим заказчиком подход к построению информационных систем (принцип «1 сервис — 1 сервер») в современных условиях нельзя назвать оптимальным, поскольку сокращается соотношение между затратами на ИТ-инфраструктуру и экономической эффективностью автоматизации бизнес-процессов и прочих ИТ-сервисов предприятия.

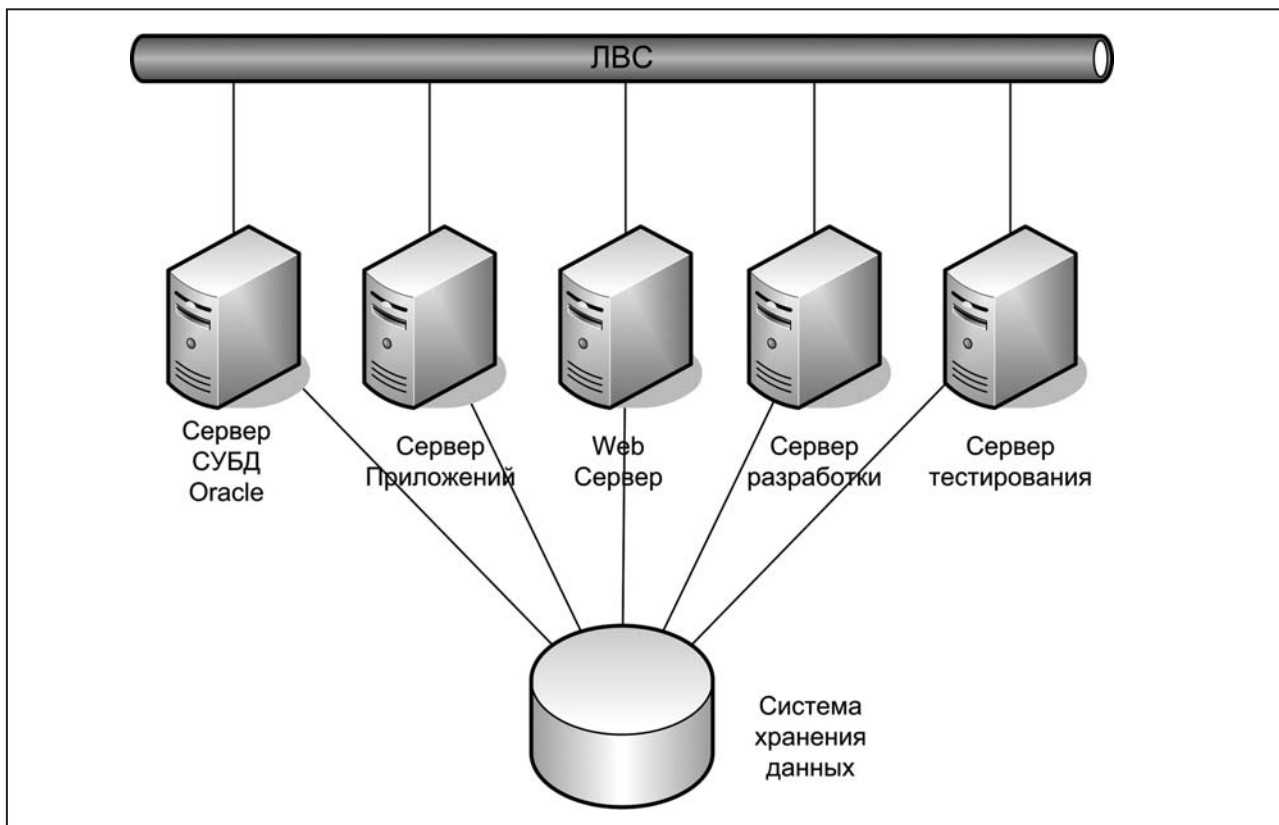


Рис. 1. Схема «1 сервис – 1 сервер»

Оптимальное решение

Как показал опыт специалистов компании «Инфосистемы Джет» наиболее эффективным с рассматриваемой точки зрения решением является консолидация серверов с использованием технологии виртуализации PowerVM.

В данном случае под виртуализацией следует понимать переход от физического представления ресурсов к логическому или виртуальному, что позволяет гибко менять набор исходных элементов ИТ-инфраструктуры, оставляя уровень представления сервисов неизменным. Виртуализировать можно не только серверы, но и системы хранения данных, сети передачи данных и т.д., что в совокупности составит законченную и логичную концепцию ИТ-инфраструктуры предприятия.

В настоящее время существует большое количество продуктов, позволяющих виртуализировать и консолидировать серверы, например, VmWare, Xen, Hyper-V, PowerVM, HP Integruty Virtual Mashines, Solaris containers и другие. В серверах серии IBM System Power используется технология PowerVM для виртуализации и консолидации серверов. Рассмотрим на примере нашей компании N, каким образом можно оптимизировать и организовать эффективную ИТ-инфра-

структуру, перейдя от традиционной модели построения вычислительного комплекса к консолидированной с использованием технологии виртуализации IBM PowerVM.

Переход к консолидированной модели построения вычислительного комплекса предусматривает широкий фронт работ, включая консалтинг, проектирование и внедрение предложенного решения.

Прежде всего проводится инвентаризация имеющихся у заказчика вычислительных ресурсов для определения состава оборудования, особенно морально устаревшего. Затем проводится мониторинг производительности вычислительных ресурсов для оценки текущей степени утилизации оборудования работающими на нем информационными системами и классификации их по признаку возможности консолидации. Информационные системы с малой степенью утилизации оборудования являются лучшими кандидатами для их консолидации в виртуальной среде. На основании полученных данных производится разработка стратегии развития консолидированной ИТ-инфраструктуры на основе технологии виртуализации PowerVM, плана консолидации приложений, включающего состав работ и оборудование, а также разработку проектной документации.

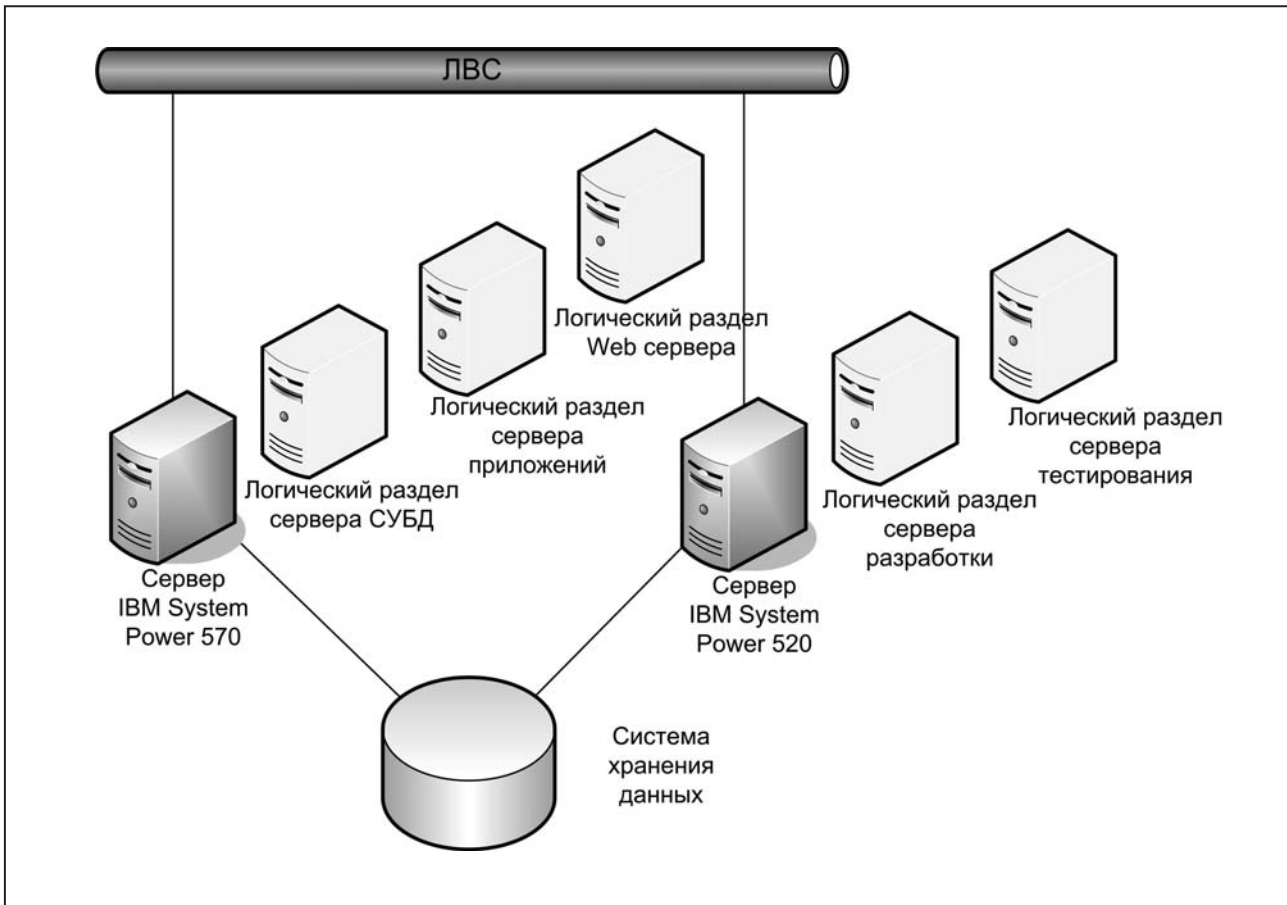


Рис. 2. ИТ-инфраструктура с консолидированными информационными системами

В соответствии с архитектурными требованиями на серверах IBM System Power создаются логические разделы с ОС AIX для установки приложений, каждому из которых выделены процессорные ресурсы и оперативная память, определенные на этапе сайзинга приложений, а также созданы виртуальные адаптеры ЛВС и СХД.

Суть предлагаемого заказчику решения, как вариант, в консолидации информационных систем продуктивной среды в логических разделах на сервере р570, а систем разработки и тестирования в логических разделах на сервере р520. Использование логических разделов позволяет проводить сравнительно простую миграцию унаследованных ИС с устаревшего и изношенного оборудования на новое. Кроме логических разделов с прикладными системами, на каждом из серверов создается дополнительно один или два служебных логических раздела (vios), предназначенных для организации ввода/вывода прикладных LPAR. С помощью vios обеспечивается сетевое взаимодействие логических разделов сервера как друг с другом, так и с внешними системами, а также взаимодействие с СХД.

На рисунке 2 приведен пример ИТ-инфраструктуры заказчика с консолидированными информационными системами.

Что получает заказчик

Построение консолидированной виртуальной ИТ-инфраструктуры позволяет значительно сократить как количество единиц оборудования, разместив информационные системы в логических разделах серверов, так и затраты на эксплуатацию и обслуживание вычислительной техники.

Таким образом, консолидация ИС в виртуальной среде позволяет сократить прямые финансовые издержки на выделение дополнительных площадей, кондиционирование и вентиляцию помещений, электроэнергию и приобретение нового оборудования для установки новых сервисов за счет более полного использования имеющихся мощностей и размещения на них новых сервисов.

Для резервирования избыточных на данный момент вычислительных мощностей хорошо подходит технология CoD (Capacity on Demand), при которой заказчик платит только за требуемую в данный момент мощность оборудования, а не за полную, которая может быть избыточной. Более эффективное использование вычислительных мощностей системы позволяет также оптимизировать затраты на лицензирование ПО — покупать лицензии требуется не на полную вычислительную мощность системы, а только на реально используемую.

Кроме сокращения прямых финансовых издержек, заказчик также получает технические и организационные преимущества, такие как унификация оборудования и ОС, обслуживание и техническая поддержка парка вычислительной

техники, снижение затрат на администрирование систем, широкие возможности модернизации ИС, гибкое управление серверной инфраструктурой и высокая надежность решения. Технология виртуализации PowerVM также позволяет сократить длительность цикла «разработка — тестирование — внедрение» программных продуктов в организации, упростить процедуру миграции унаследованных ИС.

В итоге экономическими преимуществами виртуализации являются снижение ТСО, быстрый возврат инвестиций, меньшее время внедрения. Еще одно важное достоинство виртуализации — упрощенный перенос унаследованных приложений на новые платформы, где можно эмулировать среду, необходимую для выполнения старого ПО.

Система виртуализации серверов IBM Power

Дмитрий Фомичев, Игорь Ушкало,
инженеры-проектировщики в группе вычислительных систем,
компания «Инфосистемы Джет»

Технологии виртуализации компании IBM охватывают широкий класс платформ — от серверов младшего уровня IBM System Power до мэйнфреймов System z10. Для пользователя логический раздел сервера ничем не отличается от обычного физического сервера, для ОС AIX виртуальные ресурсы также прозрачны и не отличаются от физических. Процессорные ресурсы и память всегда находятся под управлением гипервизора PowerVM. Гипервизор представляет собой абстрактный слой, реализованный на уровне микрокода сервера, лежащий между уровнем аппаратных ресурсов сервера и уровнем операционной системы, позволяет объединять в единый набор вычислительные ресурсы сервера, такие как процессоры, память, карты ввода/вывода,

формировать логические разделы с операционными системами и выделять логическим разделам сервера виртуальные вычислительные ресурсы.

Основной отличительной особенностью PowerVM является micro-partitioning — способность создавать разделы, кратные долям физического процессора (shared processing mode). С включением этой опции на Power Systems появляются следующие возможности по управлению ресурсами:

- динамически изменять параметры логического раздела, такие как количество выделенных долей процессора, объем выделенной памяти, конфигурацию устройства ввода/вывода;
- физические процессоры могут объединяться в вычислительные пулы (shared processor

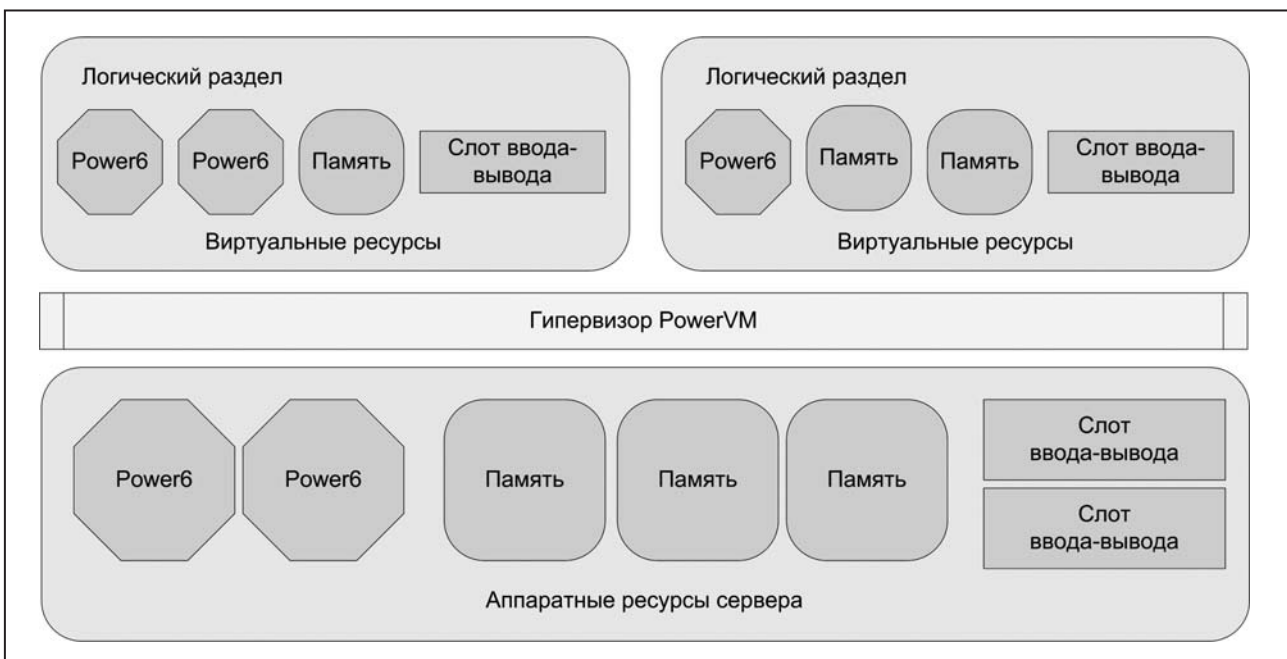


Рис. 1. Структура гипервизора PowerVM

pool); разные логические разделы могут быть привязаны к разным вычислительным пулам, чтобы гибко управлять выделением ресурсов и исключить конкуренцию;

- для микроразделов возможно установить режим Uncapped – раздел может использовать больше процессорной мощности, чем установлено в его конфигурации, в случае наличия в вычислительном пуле свободных ресурсов;
- разделы, созданные из целых физических ядер (dedicated), могут быть донорами процессорной мощности для микроразделов, созданных из долей ядер (shared), если ресурсы выделенных процессорных ядер не используются полностью – режим работы носит название shared dedicated capacity.

На рис. 1 приведена структура гипервизора PowerVM.

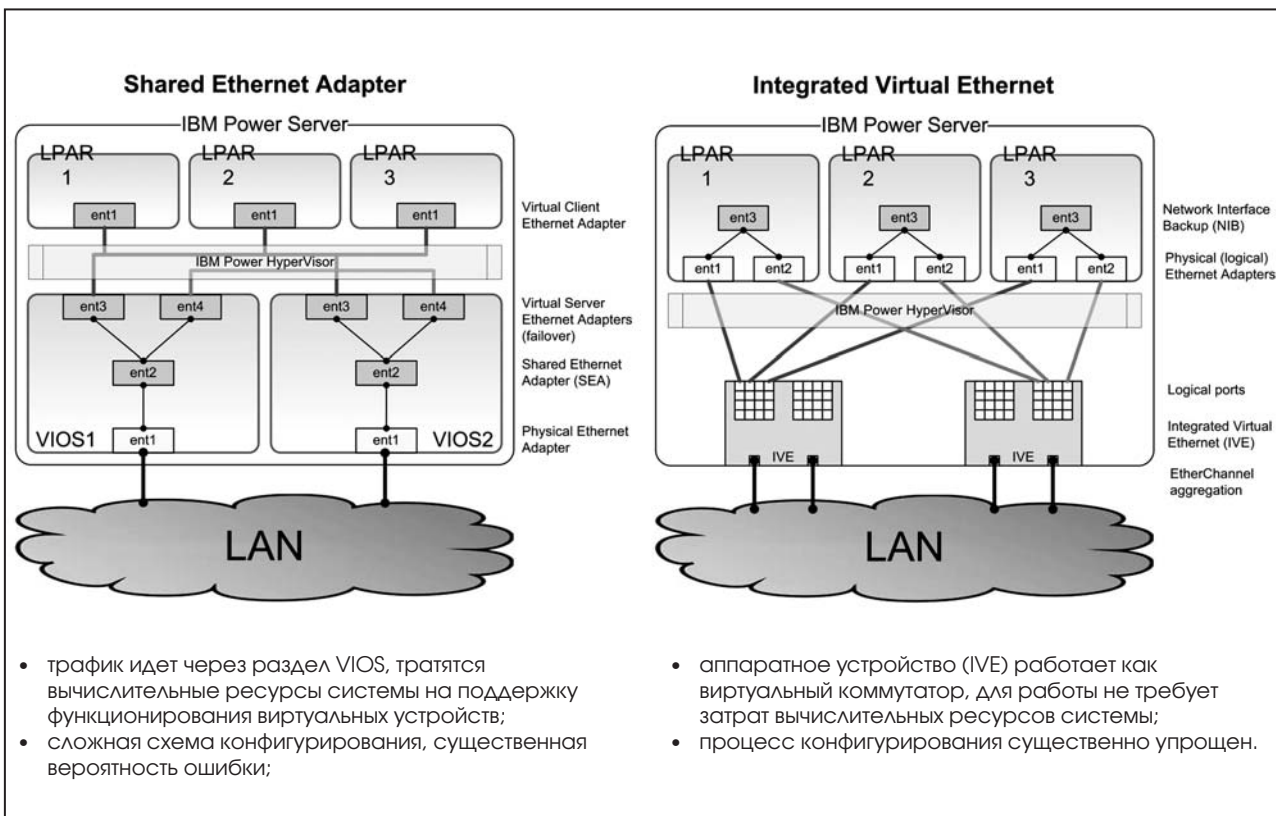
Кроме самих логических разделов сервера (LPAR) гипервизор также формирует виртуальные устройства ввода/вывода для этих разделов. В любом логическом разделе System Power всегда создаются два виртуальных устройства – адаптеры последовательных портов, используемые для виртуальной консоли LPAR. Прочие виртуаль-

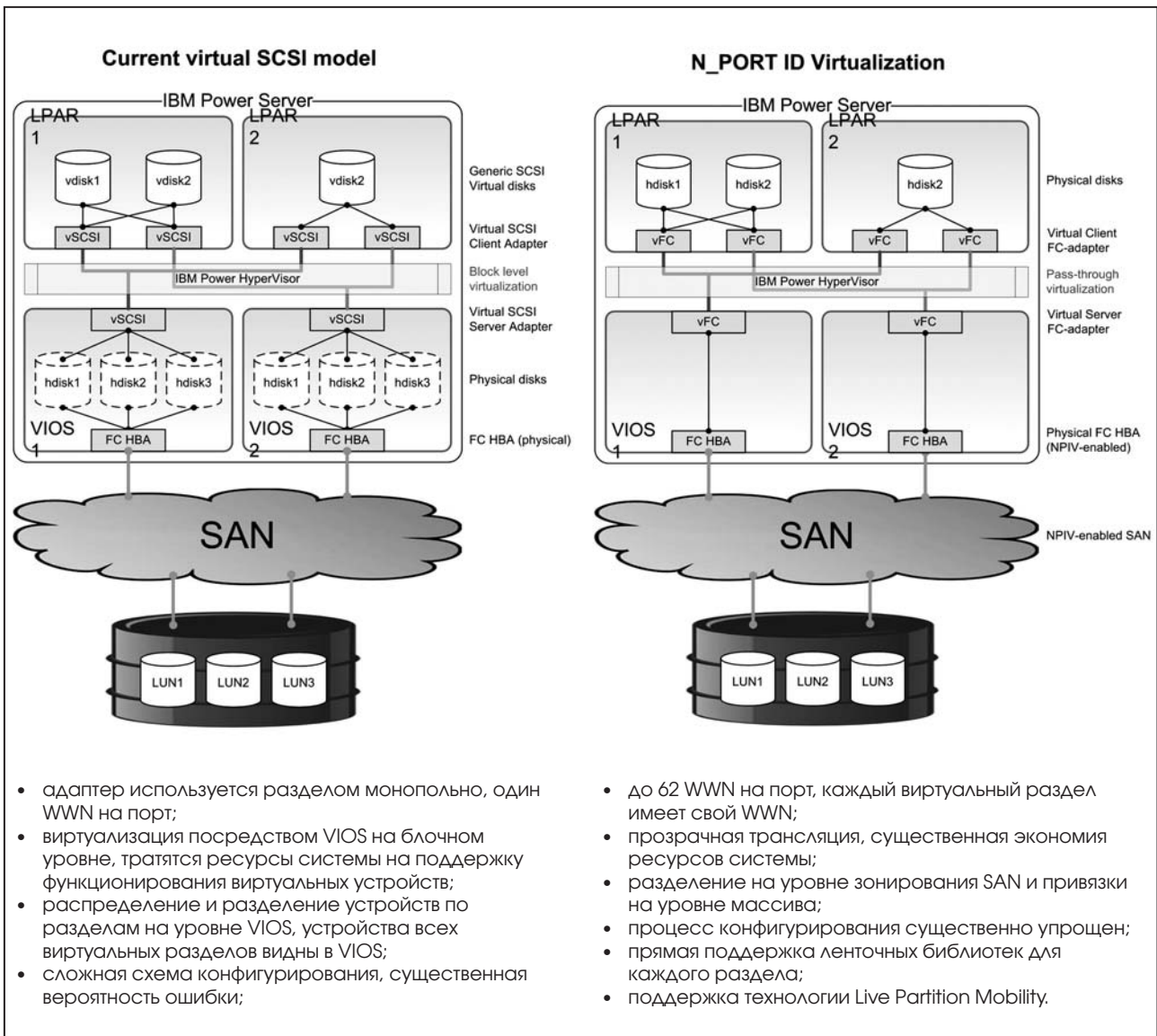
ные устройства могут понадобиться, когда на одной системе создаются несколько разделов.

Для гибкой конфигурации устройств ввода/вывода в IBM System Power используется Virtual IO Server (VIOS), который представляет собой отдельный логический раздел со специализированным ПО на основе ОС. Основная задача VIOS – создание и поддержка функционирования виртуальных устройств, таких как виртуальные диски, адаптеры LAN или SCSI/FC из физических устройств и предоставление к ним доступа другим логическим разделам. PowerVM позволяет создавать в логических разделах виртуальные устройства следующих типов:

- виртуальный последовательный порт;
- виртуальные сетевые адаптеры:
- разделяемый виртуальный сетевой адаптер (Shared Ethernet Adapter);
- логический сетевой адаптер (на базе аппаратного устройства Integrated Virtual Ethernet);
- виртуальный SCSI HBA;
- виртуальный FC HBA.

На схемах ниже приведены работы различных виртуальных устройств и их основные особенности:





Таким образом, используя возможности системы виртуализации PowerVM, мы можем на одном сервере IBM System Power создать несколько виртуальных серверов (логических разделов) со своими виртуальными устройствами, работающими под управлением ОС AIX, Linux или IBM i (на текущий момент теоретический предел составляет 254 раздела). Возможна миграция лицензий PowerVM, например, с Express Edition до Standard Edition или до Enterprise Edition. В приложении 1 (см. стр. 22) перечислены версии Power VM и их возможности.

Приложение 1. Основные характеристики линейки IBM BladeCenter

Опция	Описание	Express Edition	Standard Edition	Enterprise Edition	Power 5	Power 6
Virtual I/O Server (VIOS)	Специализированный серверный виртуальный раздел, обеспечивающий работу виртуальных устройств ввода/вывода для клиентских разделов	V ¹	V	V	V	V
Micro-Partitioning	Возможность разделять ресурсы физических процессоров между виртуальными разделами и распределять их с точностью до 0,1 процессора	V	V	V	V	V
Capped/Uncapped	Возможность задавать приоритет и ограничивать использование свободных процессорных ресурсов конкретными виртуальными разделами	V	V	V	V	V
PowerVM Lx86	Возможность устанавливать на виртуальный раздел OS Linux on Power	V	V	V	V	V
NPIV, IVE	Поддержка аппаратных возможностей виртуализации устройств	V	V	V	X	V
Multiple Shared Processor Pools	Возможность создания разделяемых пулов процессоров и распределения процессорных ресурсов между логическими разделами	X	V	V	X	V
Active Memory Sharing	Возможность создания разделяемых пулов оперативной памяти и распределения ресурсов между логическими разделами	X	X	V	X	V
Live Partition Mobility	Возможность переноса виртуального раздела с одного физического сервера на другой без остановки работы ОС и приложений	X	X	V	X	V
Management	Управление инфраструктурой с помощью Integrated Virtualization Manager (IVM) либо Hardware Management Console (HMC)	IVM	IVM, HMC	IVM, HMC		
Max LPARs	Максимальное количество виртуальных разделов	3 / Server	10 / Core	10 / Core		

1 В версии Express Edition можно создать только один раздел VIOS, в остальных версиях — без ограничений.

Вычислительные комплексы масштаба предприятия на платформе IBM Power

Игорь Ушкало,
инженер-проектировщик в группе вычислительных систем,
компания «Инфосистемы Джет»

Информационные технологии в настоящее время являются одной из основных движущих сил прогресса. Прикладные информационные системы, без которых уже никто не может себе представить существование любых организаций, постоянно совершенствуются, разработчики реагируют на быстро меняющиеся потребности заказчиков и реализуют новую функциональность — следовательно, растут требования к производительности оборудования и уже само собой разумеющейся считается высокая доступность всех информационных сервисов в режиме 24 x 7.

В любой организации рано или поздно принимается решение о построении новой серверной платформы, на которой будет происходить развертывание, промышленная эксплуатация и дальнейшее развитие больших корпоративных информационных систем. Иногда это связано с глобальным переходом на новый программный продукт (например, внедрение новой АБС или ERP-системы), иногда же просто приходит понимание, что существующая платформа перестала удовлетворять возросшим требованиям. Попробуем обозначить эти требования и наметить пути, следуя которым, можно эти требования удовлетворить.

Доступность. Информационные сервисы, обеспечивающие поддержку критичных бизнес-процессов организации, должны работать всегда, чтобы ни случилось с компонентами поддерживающей инфраструктуры. Обеспечить доступность на уровне 99,9% и выше имеют возможность только кластерные архитектуры, способные автоматически изолировать сбойные участки и перераспределить нагрузку. Во многих случаях нужно также обеспечить катастрофоустойчивость решения, разделив инфраструктуру на несколько площадок, территориально удаленных друг от друга.

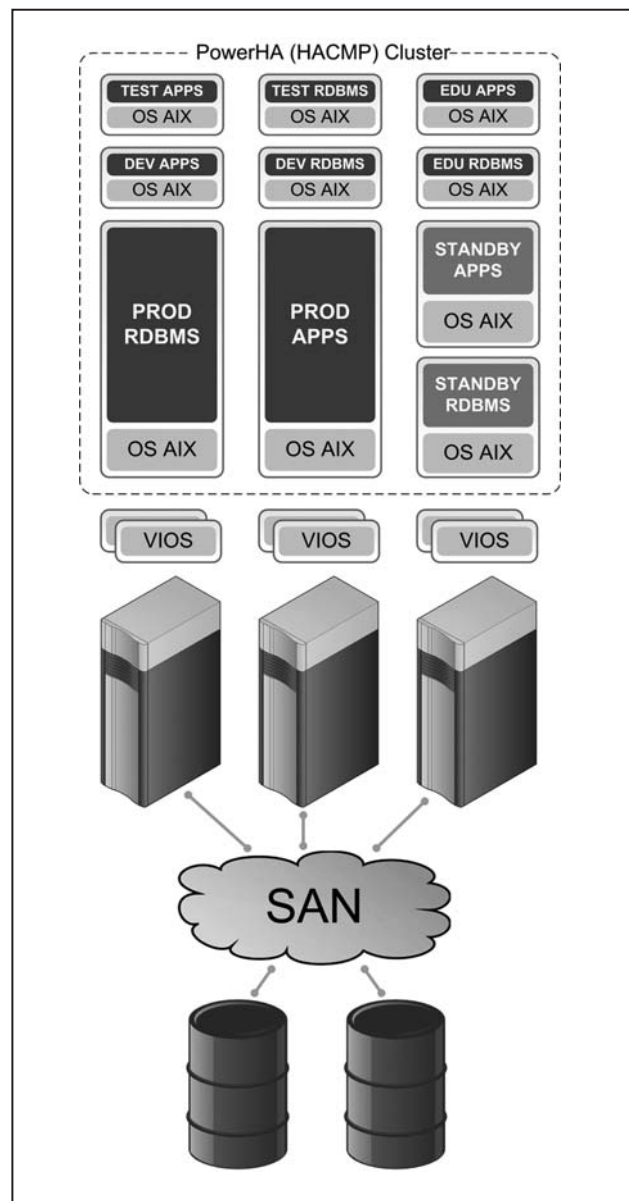


Рис. 1. Современная схема ВК

Производительность. В большинстве случаев нагрузка на информационные системы распределяется неравномерно в течение времени. Чтобы справиться с пиками, требуется запас по производительности. Оптимальным решением является не закупка оборудования с запасом, которое будет простаивать большую часть времени, а перераспределение ресурсов с целью использования временно свободных — технологии виртуализации и консолидации подходят для этого как нельзя лучше.

Масштабируемость. Нагрузка на оборудование обычно существенно растет в процессе развития системы. Необходимо предусмотреть оптимальную стратегию будущей модернизации (либо замены) серверов, систем хранения данных и подсистемы резервного копирования, обеспечив защиту первоначальных инвестиций и минимизировав последующие.

Поддержка развития. Серверная платформа должна поддерживать все стадии жизненного цикла информационной системы. Необходимо обеспечить возможность разработчикам развивать систему, гарантировать качественное тестирование вносимых в систему изменений и возможность их безопасного применения к промышленным окружениям. Следовательно, является весьма востребованной функциональность создания и управления «снимками» файловых и дисковых ресурсов, используемая для создания клонов окружений и быстрого восстановления промышленного окружения, например, в случае внесения изменений, вызвавших серьезный сбой, либо устранения ошибки пользователя, повлекшей за собой потерю целостности данных.

Эффективность. В конечном счете, как всегда, речь идет только о двух ресурсах — это время и деньги. Иногда важнее одно, иногда другое, однако современные технологии позволяют избежать компромисса — минимальное время реакции на потребность увеличить производительность системы или экономия бюджета — программы CoD (Capacity on Demand) позволяют иметь в наличии аппаратный резерв мощности и мгновенно активировать его по требованию на постоянной либо временной основе. Весьма полезной является и технология Thin Provisioning, позволяющая оптимизировать требования к дисковой подсистеме.

Персонал. По сути это еще один аспект защиты инвестиций. Помимо совершенно очевидных инвестиций в оборудование и программное обеспечение следует помнить об инвестициях в технологии и персонал, поддерживающий эксплуатацию комплекса. Технологии не работа-

ют сами по себе, их реализуют конкретные специалисты, от компетентности которых зависит большая часть успеха. При смене аппаратной платформы, существенная часть знаний и опыта будет утеряна, потребуется дополнительное обучение сотрудников. Одной из кардинальных мер, позволяющих обойти эту проблему, является принятие решения о передаче поддержки на аутсорсинг, но следует понимать, что это решение должно быть осмысленным и сопровождаться комплексом мер по организации процесса управления качеством сервиса.

Чтобы построить перспективную программно-аппаратную платформу, удовлетворяющую всем вышеописанным требованиям, недостаточно просто провести закупку по спецификации — в этом случае получится всего лишь набор ящиков с оборудованием на складе и коробки с ПО на полках. В итоговый комплекс должны войти оригинальные проектные решения (учитывающие специфику конкретного проекта), а также методики развертывания и настройки, основывающиеся на опыте квалифицированных специалистов. Результат в итоге должен получиться существенно больше, чем сумма составляющих его компонент — это и будет надежная и гибкая платформа для развития. Комплексный подход позволит избежать проблем в будущем и обеспечит все потребности заказчика сейчас и в долгосрочной перспективе.

Из каких частей состоит современный вычислительный комплекс? Попробуем обозначить основные компоненты:

- серверная платформа (серверы, операционные системы и кластерное ПО);
- системы и сети хранения данных (СХД);
- подсистема резервного копирования и восстановления (ПРК);
- СПО управления комплексом и мониторинга его состояния.

Переходя от теории к практике, рассмотрим оптимальную, с нашей точки зрения, реализацию вычислительного комплекса на оборудовании IBM.

Серверная платформа. Вычислительное ядро комплекса составляют высокопроизводительные серверы IBM Power. Система виртуализации IBM PowerVM, в целях сведения к минимуму накладных расходов, использует встроенные аппаратные возможности сетевой подсистемы (IVE — Integrated Virtual Ethernet) и технологию NPIV (N_Port ID Virtualization). Создается два системных виртуальных раздела под управлением VIOS, которые дублируют друг друга и обеспечивают

работу виртуальных устройств ввода/вывода. Каждой программной компоненте корпоративной ИС (СУБД либо серверу приложений) выделяется свой собственный виртуальный раздел. Продвинутое возможности системы виртуализации по управлению утилизацией аппаратных ресурсов оборудования (Micro-Partitioning, Active Memory Sharing) позволяют оптимизировать требования к процессорам и оперативной памяти. В целях сокращения запланированного времени простоя системы инфраструктура подготавливается к использованию уникальной функциональности Live Partition Mobility, позволяющей переносить виртуальный раздел с одного физического сервера на другой без прекращения работы пользователей системы.

В случае возникновения незапланированной потребности в дополнительных вычислительных ресурсах, решение использует возможности программы CoD (Capacity on Demand). Функциональность CoD позволяет задействовать установленные, но неактивные процессоры и оперативную память на постоянной (активировать ресурсы) либо временной основе (активировать ресур-

сы по мере необходимости и платить по факту использования).

Виртуальные разделы объединены в кластеры IBM PowerHA, обеспечивающие высокую доступность информационных сервисов. Кластерное решение использует возможности тесной интеграции с удаленной репликацией дисковых систем хранения данных. Особое внимание уделено дублированию линий связи между узлами кластеров, находящихся на разных площадках — вероятность сегментирования кластера сведена к теоретическому минимуму.

СХД. В плане организации систем и сетей хранения данных, в решении могут быть использованы самые прогрессивные технологии и компоненты: дисковые массивы класса midrange IBM DS5000 либо high-end IBM DS8000, производительные SAN-коммутаторы и система виртуализации IBM SAN Volume Controller (SVC). Решение готово к построению катастрофоустойчивой конфигурации, предусмотрена возможность разнесения оборудования по основной и резервной площадкам. Распределенная сеть хранения данных позволяет обеспечить работу системы уда-

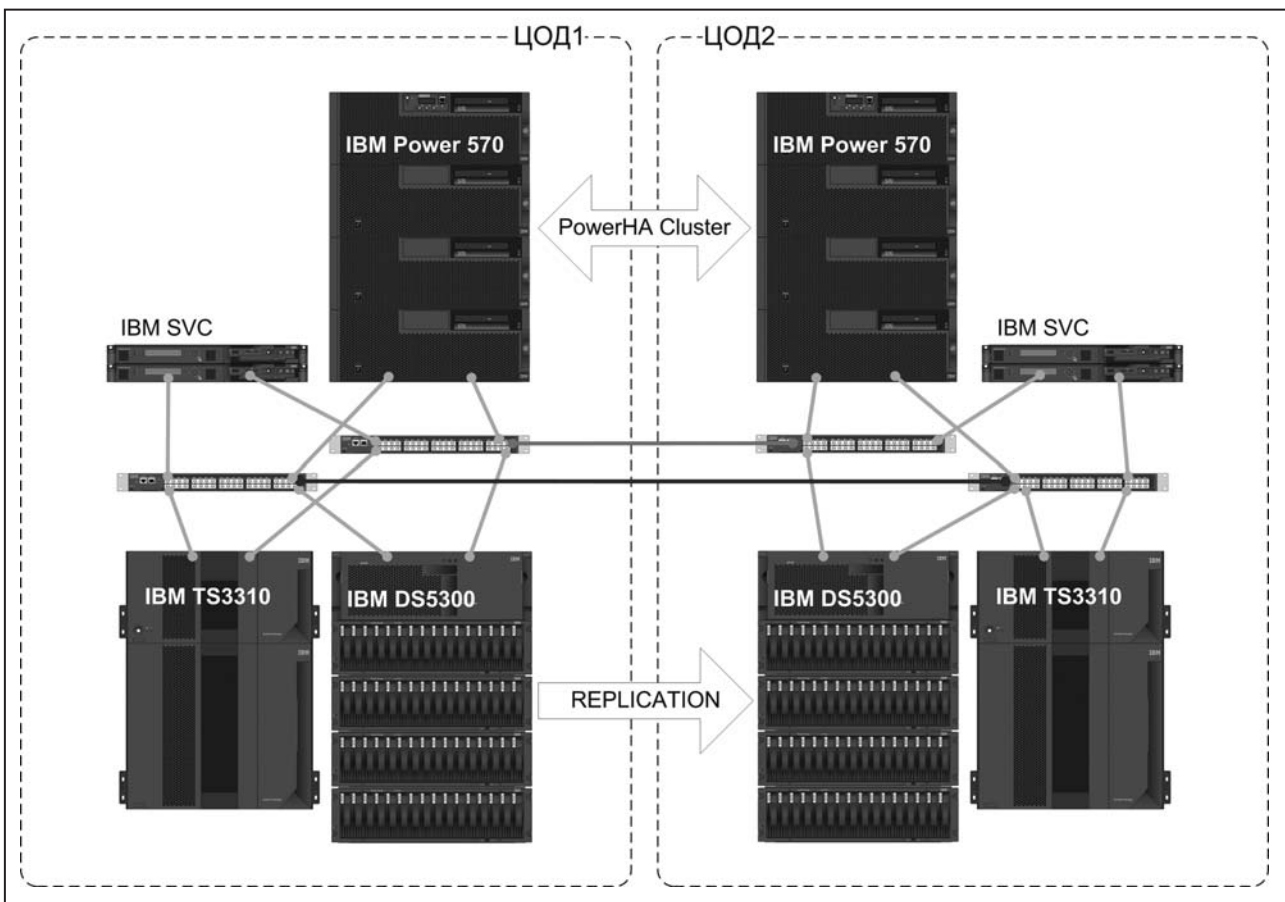


Рис. 2. Аппаратная схема распределенного ВК

ленной репликации практически без ограничений по расстоянию между основным и резервным ЦОД. Выделение дискового пространства оптимизируется посредством технологии Thin Provisioning.

Тщательно продумана инфраструктура для дальнейшего развития информационных систем. Помимо продуктивных окружений для каждой подсистемы предусмотрены окружения для разработки и тестирования, включая тестирование качества и производительности после внесения изменений. Эффективная реализация всей этой инфраструктуры основывается на использовании функций удаленной репликации и снимков, реализованных с помощью встроенных возможностей и системы виртуализации СХД.

СРК. В качестве подсистемы резервного копирования и восстановления используются ленточные библиотеки семейства IBM TS3000 Tape Library, управляемые ПО IBM Tivoli Storage Manager (TSM). Подсистема управляет режимами и расписанием процессов резервного копирования, обеспечивает полную автоматизацию резервного копирования СУБД и файловых ресурсов. TSM поддерживает все основные виды системного ПО (СУБД, серверы приложений) и отлично интегрируется с ними, обеспечивая проведение резервного копирования в режиме реального времени без остановки системы.

СПО управления и мониторинга представляет собой программный комплекс, построенный на базе семейства продуктов IBM Tivoli и обеспечивающий эффективный, согласованный и централизованный мониторинг ключевых распределенных ресурсов с помощью единого интерфейса управления. Система позволяет контролировать как актуальные данные, практически в режиме реального времени, так и ретроспективные данные. Элементы проактивного мониторинга существенно повышают общую надежность всего решения, позволяя персоналу реагировать на инциденты, предупреждая прерывание сервиса пользователям. Опционально можно настроить интеграцию с системой управления инцидентами.

В процессе проектирования, на этапе уточнения требований, происходит адаптация решения и выбор оборудования под конкретные запросы заказчика. Решение обладает отличной вариативностью; если не требуется исключительная производительность, то можно использовать модели начального уровня (серверы Power 550 и дисковые массивы DS5020), однако можно масштабировать решение до уровня hi-end и построить его на серверах Power 595 и системах хранения данных DS8700.

Большие ИС требуют длительной доработки и тестирования перед запуском, поэтому достаточно часто используется принцип двухэтапного построения серверного комплекса.

На этап доработки и тестирования системы достаточно оборудования начального уровня (которое потом будет использовано для поддержки сред разработки и тестирования), а перед запуском в промышленную эксплуатацию настроенная система переносится на тяжелое решение, обеспечивающее требуемые производительность и надежность. При выборе стратегии модернизации оборудования следует рассматривать и сравнивать между собой несколько вариантов:

- старт на серверах начального уровня, затем закупка тяжелых серверов;
- использование возможностей модульного масштабирования серверов старших линеек (закупка в минимальной конфигурации, затем модернизация);
- использование программы CoD (закупка сразу целевой конфигурации комплекса с неактивными ресурсами и последующая активация перед запуском);

Выбор конкретного варианта производится исходя из экономической эффективности и реальной оценки рисков (например, если весьма вероятен сдвиг сроков запуска проекта в промышленную эксплуатацию, более привлекательным выглядит первый вариант).




Достаточно часто строящаяся платформа должна обеспечить функционирование не только новых корпоративных информационных систем, но и уже работающих, причем их платформа может быть самой разной. Миграция старых систем на новую платформу является одной из самых сложных задач в процессе реализации комплексного проекта, как из-за технических сложностей самого процесса миграции, так и по причине очень узких окон запланированного даунтайма у работающих промышленных систем. Однако наши квалифицированные специалисты владеют самыми современными технологиями, и, обладая большим опытом проведения миграций различных корпоративных систем, способны решить самые сложные задачи заказчиков. При миграции на платформу Power, компания IBM совместно с системным интегратором, ведущим проект, может предоставляться сервис миграции — Migration Factory. Продвинутое программное решение обеспечивают надежную миграцию с минимальным временем простоя информационных сервисов (см. статью «XenoBridge с точки зрения «ораклиста»).

После завершения проекта заказчик получает работоспособное и хорошо документированное комплексное решение, где все подсистемы взаимосвязаны между собой, что существенно облегчает задачу последующей поддержки инфраструктуры. Решение такого уровня будет удовлетворять всем требованиям заказчика на длительный период времени и способно адаптироваться к любым изменениям, обеспечивая управляемую, надежную и гибкую платформу для развития. Построение таких решений невозможно без синергетического взаимодействия между персоналом заказчика и квалифицированными специалистами компании интегратора, имеющими большой опыт реализации проектов самого высокого уровня сложности.


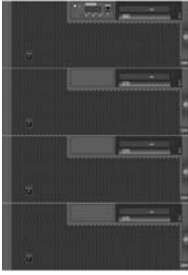
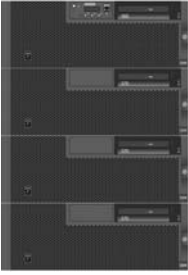


Заключение

Наверное, вы уже захотели провести виртуализацию и консолидацию всего, что только возможно, если еще этого не сделали. А возможно вы уже хорошо знакомы с этими концепциями и давно их используете? Тогда не стоит останавливаться на достигнутом. Ведь несмотря на кризис, в области ИТ постоянно что-то происходит, на рынке возникают новые продукты и услуги, которые помогают клиентам тратить меньше, а получать больше. В качестве примера можно привести технологии, которые позволяют превращать разрозненные мелкие дисковые массивы в оборудование класса High-End, а систему резервного копирования наделять функциями защиты данных и учить саму писать планы аварийного восстановления. Но это уже совсем другая история.

Приложение 1. Модельный ряд серверов IBM

Параметры	System p5 510/510Q	Power 520 Express	Power 550 Express
Сервер			
Параметры			
Тип-модель	9111-51A	8203-E4A	8204-E8A
Конструктив	2U 19"	4U 19" / Tower	4U 19" ck / Tower
Процессор	64-bit POWER5+	64-bit POWER6+	64-bit POWER6+
Кол-во ядер (частота)	1,2 (2.1 GHz) 4 (1,65 GHz)	1,2,4 (4,2 GHz, 4,7 GHz)	1,4,6,8 (3,5 GHz, 4,2 GHz, 5,0 GHz)
Память (min/max)	1GB – 32GB DDR2 SDRAM	4GB – 256GB DDR2 SDRAM	4GB – 256GB DDR2 SDRAM
Кол-во HDD	4 x 3,5" tswap	4 x 3,5" / 8 x 2,5" hotswap	4 x 3,5" / 8 x 2,5" hotswap
Контроллер HDD	Ultra320 SCSI	3G SAS	3G SAS
Встроенные сетевые порты	2 x GBE	2 / 4 x GBE 2 x 10GBE	2 / 4 x GBE 2 x 10GBE
Разъемы PCI-X	3	2 hotswap	2 hotswap
Разъемы PCIe	–	3 hotswap	3 hotswap
Память Chipkill	✓	✓	✓
Сервисный процессор	✓	✓	✓
Динамическое освобождение проц.	✓	✓	✓
Повторение процессорных инструкций	✓	✓	✓
Автоматическое замещение процессора	X	✓	✓
«Горячее» добавление проц. блока	X	X	X
«Горячий» ремонт проц. блока	X	X	X
Ресурсы по требованию (CoD)	X	X	X
PowerVM Express	X	✓	✓
PowerVM Standard	✓	✓	✓
PowerVM Enterprise	X	✓	✓
rPerf для AIX (кол-во ядер)	2.10 GHz: 12,46 (2)	4,2 GHz: 31,48 (4), 4,7 GHz: 39,73 (4)	3,5 GHz: 58,80 (8), 4,2 GHz: 68,20 (8), 5,0 GHz: 78,60 (8)

Параметры	DS3200	DS3300	DS3400	DS3950	DS4700
Кэш-память (на контроллер)	512Mb / 1Gb	512Mb / 1Gb	512Mb / 1Gb	2 / 4 Gb	2 / 4 Gb
Внешние интерфейсы (на контроллер)	3 x SAS 3Gbps	2 x iSCSI 1Gbps	2 x FC 4Gbps	4 x FC 8Gbps 1x iSCSI 1Gbps	8x FC 4Gbps
Тип контроллера	Dual active	Dual active	Dual active	Dual active	Dual active
Типы RAID	0,1,3,5,6,10	0,1,3,5,6,10	0,1,3,5,6,10	0,1,3,5,6,10	0,1,3,5,6,10
Кол-во дисков (макс)	48	48	48	112	112
Тип дисков	SAS / SATA	SAS / SATA	SAS / SATA	FC / SATA	FC / SATA
Сырая емкость (макс)	21,6Tb SAS 48Tb SATA	21,6Tb SAS 48Tb SATA	21,6Tb SAS 48Tb SATA	67,2Tb FC 112Tb SATA	33,6Tb FC 112Tb SATA

Power 560 Express	Power 570	Power 570/32	Power 575	Power 595
				
8234-EMA	9117-MMA	9117-MMA	9125-F2A	9119-FHA
4U/CEC 19" Rack	4U/CEC 19" Rack	4U/CEC 19" Rack	2U 24"	24" Rack
64-bit POWER6+	64-bit POWER6+	64-bit POWER6+	64-bit POWER6+	64-bit POWER6+
4,8,16 (3,6 GHz)	2,4,8,12,16 (3,5 4,4 и 5,0 GHz)	4,8,16,24,32 (4,2 GHz)	32 (4,7 GHz)	8 – 64 (4,2 GHz, 5,0 GHz)
8GB – 384GB DDR2 SDRAM	4GB – 768GB DDR2 SDRAM	4GB – 768GB DDR2 SDRAM	4GB – 256GB DDR2 SDRAM	16GB – 4096GB DDR2 SDRAM
6 (12) x 3,5" hotswap	6 (24) x 3,5" hotswap	6 (24) x 3,5" hotswap	2 x 2,5" tswap	832 x 2,5" hotswap
3G SAS	3G SAS	3G SAS	3G SAS	3G SAS
2 / 4 (4 / 8) x GBE 2 (4) x 10GBE	2 / 4 (8 / 16) x GBE 2 (8) x 10GBE	2 / 4 (8 / 16) x GBE 2 (8) x 10GBE	4 x GBE	–
2 (4) hotswap	2 (8) hotswap	2 (8) hotswap	–	–
4 (8) hotswap	4 (16) hotswap	4 (16) hotswap	4 hotswap	640 hotswap
V	V	V	V	V
V	V	V	V	V
V	V	V	V	V
V	V	V	V	V
V	V	V	V	V
X	V	V	V	V
X	V	V	X	V
X	V	V	X	V
X	X	X	X	X
V	V	V	V	V
V	V	V	V	V
3,6 GHz: 100,30 (16)	3,5 GHz: 105,75 (16), 4,4 GHz: 127,32 (16), 5,0 GHz: 141,21 (16)	4,2 GHz: 193,25 (32)	–	4,2 GHz: 553,01 (64), 5,0 GHz: 479,89 (64)

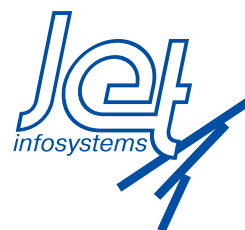
DS4800	DS5020	DS5100/5300	DS6000	DS8000	XIV
4 / 8 / 16 Gb	2 / 4 Gb	до 64 Gb	2 Gb	256 Gb	120 Gb
8xFC 4Gbps	4xFC 4Gbps	16xFC 8Gbps 2xiSCSI 1Gbps	8 ports total FC 4Gbps, FICON	128 ports total FC 4Gbps, FICON	24xFC 8Gbps 6xiSCSI 1Gbps
Dual active	Dual active	Dual active	Dual active	Multi active	Multi active
0,1,3,5,6,10	0,1,3,5,6,10	0,1,3,5,6,10	5,10	5,6,10	RAID-X
224	112	224/448	128	1024	180
FC / SATA	FC / SATA	FC / SATA	FC	FC / SATA	SATA
100,8Tb FC 224Tb SATA	67,2Tb FC 112Tb SATA	268,8Tb FC 448Tb SATA	57,6Tb	460,8Tb FC 1024Tb SATA	180Tb

Jet Info

ИНФОРМАЦИОННЫЙ БЮЛЛЕТЕНЬ

Издается с 1995 года

Главный редактор: Дмитриев В.Ю. (vlad@jet.msk.su)
Редактор: Слободчикова Т.А. (slobodchikova@jet.msk.su)
Россия, 127015, Москва, Б. Новодмитровская, 14/1
тел. (495) 411 76 01
факс (495) 411 76 02
[email: JetInfo@jet.msk.su](mailto:JetInfo@jet.msk.su) <http://www.jetinfo.ru>



Издатель: компания «Инфосистемы Джет»

Подписной индекс по каталогу Роспечати

32555

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