

Суперкомпьютеры в промышленном секторе – вызовы и задачи

Многие предприятия сегодня в попытках снизить затраты на разработку и производство продукции прибегают к различным способам оптимизации отдельных звеньев своей деятельности, например цепочки поставок, при этом пытаясь сохранить высокий темп выпуска и высокое качество новых продуктов. Между тем, успех предприятия в высококонкурентной среде зависит от множества факторов, включающих увеличение производительности всех операций, ускорение цикла разработки продуктов и соответствие нормативным требованиям. Существуют разные способы увеличить операционную производительность и сократить издержки, но независимо от выбранной стратегии использование высокопроизводительных вычислительных сред является критичным для достижения успеха.

Для того, чтобы быстро реагировать на изменение спроса со стороны потребителей продукции и создавать продукты, обладающие конкурентными преимуществами, предприятию необходим непрерывный цикл разработки, позволяющий постоянно сокращать время, требуемое для выпуска новых продуктов на рынок.

Использование на каждом этапе разработки или испытаний компьютерного моделирования дает возможность значительно ускорить выход новой, качественной и конкурентной продукции, обеспечив при этом сокращение издержек. Программные решения, позволяющие обеспечить максимально точную симуляцию поведения продукта, требуют огромных вычислительных ресурсов. Существуют разные подходы к тому, каким именно образом следует организовать эти ресурсы. С одной стороны, использование одного мейнфрейма позволяет консолидировать множество производственных задач и обеспечивает высокую производительность и отказоустойчивость вычислительной системы. Но при этом предприятие может оказаться заложником такого подхода, поскольку мейнфреймы используют проприетарные архитектуры и технологии, что означает привязку предприятия к тому или иному производителю. Являясь закрытой инфраструктурой, мейнфрейм требует существенных инвестиций в квалификацию обслуживающего персонала, услуги по введению в эксплуатацию и дальнейшей модернизации решения.

Более целесообразным для решения задач CAE (Computer Aided Engineering) представляется использование высокопроизводительных вычислительных кластеров (High Performance Computing Cluster, HPCC). Подобное решение, построенное на основе открытых архитектур среды вычислений (x86), передачи (например Infiniband) и хранения данных (CIFS/NFS/FC/iSCSI), а

также среды управления позволяет радикально сократить совокупную стоимость владения и максимально быстро вернуть сделанные инвестиции.

Следует отметить, что в последнее время наметился явный тренд на использование GPU (графических ускорителей) в решениях HPCC. Несмотря на существенные отличия в программировании таких решений, данный подход обладает серьезными преимуществами, обеспечивая:

- ▶ невероятный потенциал производительности:
 - CPU – 96 GFLOPS на сокет (~ 1500 \$),
 - GPU – 1000+ GFLOPS (~500 – 1500 \$);
- ▶ лучшее соотношение цена/производительность.

Одним из факторов, сдерживающих рост популярности подобных решений, является достаточно вялая поддержка со стороны производителей ПО.

Компания Dell уже сейчас предлагает решения по построению

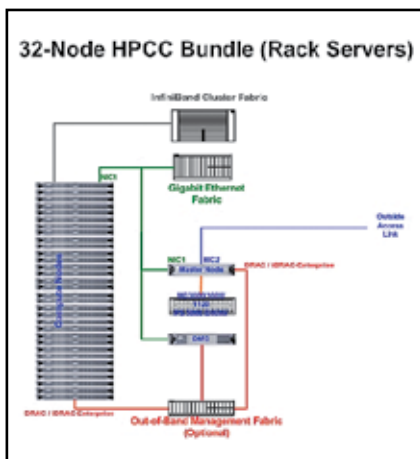


Рис. 1

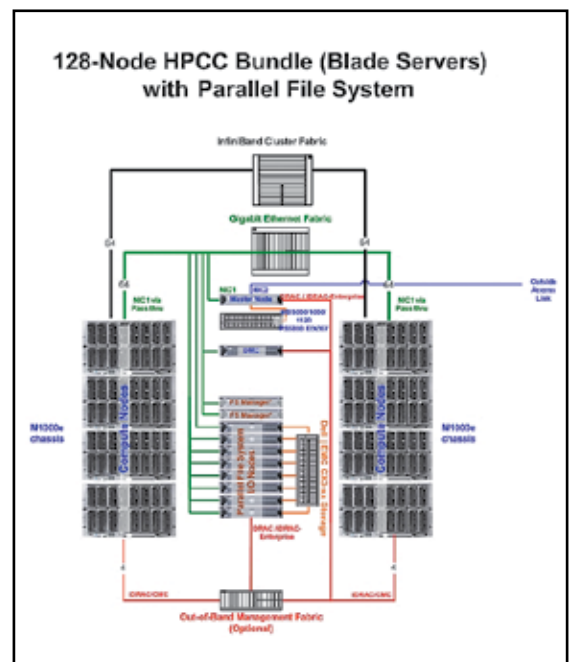


Рис. 2

НРСС с использованием GPU, но в рамках данной статьи мы остановимся все же на решениях, основанных на CPU x86.

Высокопроизводительные вычислительные кластеры относятся к разряду сложных решений в силу большого числа входящих в них как аппаратных, так и программных компонентов. Компания Dell считает необходимым переосмысление существующего в мире подхода к построению высокопроизводительных систем и видит основную задачу здесь в максимальной стандартизации и упрощении НРСС. Важнейшим элементом новой стратегии компании является активное участие Dell в программе Intel Cluster Ready. Данная инициатива Intel призвана максимально упростить подбор правильного решения за счет тестирования и валидации его аппаратных и программных составляющих для того, чтобы гарантировать совместимость и корректную работу всего комплекса в целом.

С этой целью Dell предлагает ряд простых и удобных инструментов, в том числе и референсные архитектуры комплексов. Рассмотрим НРСС как начального уровня (32-Node, стоечные серверы), так и средние по производительности (128-Node, блейд-серверы). Решение начального уровня представлено на рис. 1.

Данная конфигурация НРСС включает в себя 32 сервера PowerEdge R410 (compute nodes), каждый из которых оснащен двумя процессорами Intel Xeon 55xx, оперативной памятью DDR3, интегрированным двухпортовым контроллером Gigabit Ethernet и адаптером HCA InfiniBand. Управление комплексом (master node), а также функции шлюза NFS обеспечивает сервер PowerEdge R710/R510 с подключенной системой хранения данных PowerVault (MD3000/1000/1200) или Equallogic. Для мониторинга со-

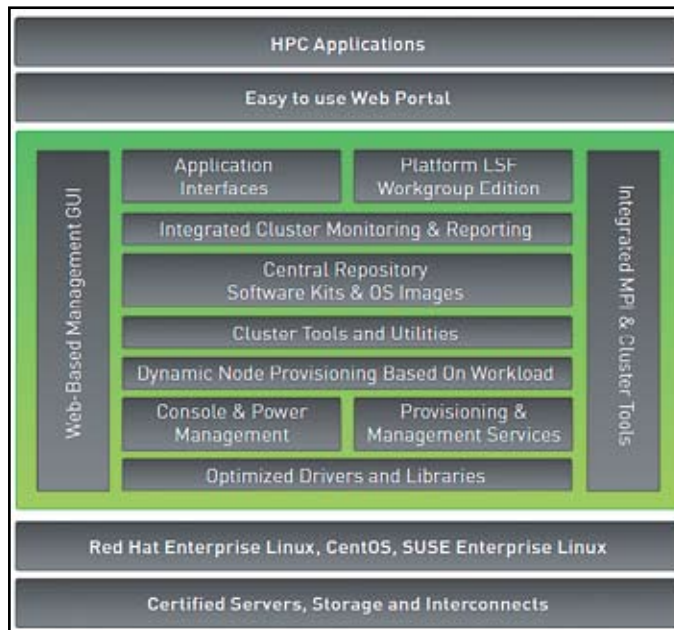


Рис. 3

стояния аппаратных составляющих решения используется ПО управления Dell Management Console (DMC), которое дополняет функционал встроенных в серверы контроллеров Lifecycle Controller и iDRAC6, обеспечивающих существенное упрощение развертывания комплекса, управление жизненным циклом и полный контроль серверов. Среда коммутации состоит из трех фабрик: Infiniband для межнодовой коммуникации, Gigabit Ethernet для in-band-управления и (опционально) фабрики для ООВ-управления (out-of-band) аппаратными платформами.

Решение среднего уровня, представленное на рис. 2, использует в качестве вычислительных узлов 128 блейд-серверов (например PowerEdge M610).

Использование блейд-серверов позволяет существенно увеличить плотность комплекса. Так, для 128 узлов достаточно 4-х шасси PowerEdge M1000e (суммарно 80 U в стойках). Для обеспечения достаточной производительности и масштабируемости среды хранения данных используются мощные системы блочного доступа Dell|EMC, подключенные через фабрику Fibre Channel к узлам, на которых работает кластерная файловая система.

Отдельно стоит упомянуть платформу управления вычислительным комплексом. Наиболее популярными в настоящий момент

являются решения на основе Linux. Несмотря на то, что такие решения обеспечивают хорошую производительность и гибкость, не все организации обладают квалификацией, позволяющей осуществлять управление и обслуживание подобными решениями. Именно поэтому Dell предлагает интегрированную, простую в работе платформу управления Platform HPC Workgroup Manager (HPC-WM). Весь стек комплекса представлен на рис. 3, цветом

выделен сегмент стека, перекрываемый HPC-WM.

HPC-WM включает в себя простую в работе web-консоль, средства управления нагрузками и выделения ресурсов, тулkit для организации работы кластера (в том числе различные варианты MPI), средства мониторинга и многое другое.

Большое внимание при выборе НРСС-решения необходимо уделять вопросам энергоснабжения и охлаждения. Серверы Dell отличаются энергоэффективностью, которая достигается за счет использования инновационного подхода к дизайну и управлению системой охлаждения серверов, блоков питания с КПД преобразования более 90 % и возможностью выбора различных номиналов мощности, а также благодаря многим другим решениям, воплощенным в "железе" и встроенном ПО. Применение блейд-серверов позволяет добиться еще большей энергоэффективности, но в силу высокой плотности решения вопрос организации охлаждения требует аккуратного и вдумчивого подхода. Хорошим вариантом в этом отношении является использование стандартизованных решений, предлагаемых компаниями APC и Schneider Electric, использующих чиллеры (водяное охлаждение) и изолированные коридоры.

Артём Гениев, компания Dell