

Д.Ю. Кудинов

*Киевский национальный университет строительства и архитектуры, Киев***СРАВНЕНИЕ МОДЕЛЕЙ МАСШТАБИРУЕМЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ
ДЛЯ ИНТЕГРАЦИИ РЕСУРСОВ АРХИТЕКТУРНО-СТРОИТЕЛЬНЫХ САПР**

Рассмотрены существующие модели масштабируемых вычислительных систем и возможности их применения для интеграции ресурсов архитектурно-строительных САПР. Проведен сравнительный анализ моделей систем с точки зрения их стоимости, вычислительной мощности, скорости вычислений и скорости внедрения. Сделан вывод о целесообразности использования таких моделей систем для интеграции ресурсов архитектурно-строительных САПР.

Ключевые слова: масштабируемые вычислительные системы, интеграция ресурсов, САПР, анализ целесообразности использования

Розглянуто існуючі моделі масштабованих обчислювальних систем та можливості їх застосування для інтеграції ресурсів архітектурно-будівельних САПР. Проведено порівняльний аналіз моделей систем щодо їхньої вартості, обчислювальної потужності, швидкості обчислень та швидкості впровадження. Зроблено висновок щодо доцільності використання таких моделей систем для інтеграції ресурсів архітектурно-будівельних САПР.

Ключові слова: масштабовані обчислювальні системи, інтеграція ресурсів, САПР, аналіз доцільності використання

The paper examines existing models of scalable computing systems and their applicability to integrate resources of architectural CAD systems. It was held a comparative analysis of models of systems in terms of cost, processing power, computing speed and the speed of implementation. It was concluded regarding the feasibility of using such models of systems for the integration of resources of architectural CAD systems.

Keywords: scalable computing systems, integration of resources, CAD, analysis of the feasibility of using

Постановка проблемы

С развитием информационных систем все большее распространение получают технологии распределения вычислительных нагрузок. В сфере автоматизированного проектирования (САПР) для организации таких вычислений часто используются облачные технологии. Крупнейшие производители программного обеспечения предлагают пользователям свои собственные удаленные вычислительные ресурсы, которые находятся в инфраструктуре производителя и обслуживаются им же. Однако в нашей стране и странах СНГ существуют некоторые особенности в работе проектных организаций. В первую очередь они касаются обеспечения высокого уровня безопасности хранения и передачи проектных данных. Так же актуальной остается проблема скорости доступа и передачи данных. Учитывая эти

факторы, актуальной становится реализация масштабируемой вычислительной системы в инфраструктуре проектных организаций, обеспечение локализации системы вычислений и высокого уровня безопасности хранения и передачи данных.

**Анализ последних исследований
и публикаций**

В последнее время значительно возрос интерес к масштабируемым вычислительным системам. Большинство ведущих разработчиков программного обеспечения, особенно в сфере автоматизации проектных работ, предлагают свои облачные сервисы для хранения и обработки данных. Облачные вычисления (англ. cloud computing) – это модель обеспечения повсеместного и удобного сетевого доступа по требованию к общему пулу (англ. pool) конфигурируемых вычислительных

ресурсов (например, сетям передачи данных, серверам, устройствам хранения данных, приложениям и сервисам – как вместе, так и по отдельности), которые могут быть оперативно предоставлены и освобождены с минимальными эксплуатационными затратами и / или обращениями к провайдеру. Потребители облачных вычислений могут значительно уменьшить расходы на инфраструктуру информационных технологий (в краткосрочном и среднесрочном планах) и гибко реагировать на изменения вычислительных потребностей, используя свойства вычислительной эластичности (англ. *elastic computing*) облачных услуг.

В настоящее время облачные технологии активно внедряются в системы разного назначения, в том числе и в САПР. Провайдерами облачных решений обычно выступают крупные западные компании, которые предлагают очень гибкие условия приобретения сервиса и разнообразные возможности конфигурации услуг. Облачное хранение и обработка данных имеют ряд преимуществ (таких как скорость обработки данных, дешевизна, гибкость, обслуживание инфраструктуры провайдером), но в то же время не лишены весомых недостатков (таких как риски безопасности, риски доступности данных, низкий уровень контроля над процессами организации хранения данных и их обработки). Концепция использования облачных вычислений для решения задач, связанных с работой САПР, подробно описана в работе [1]. Там же проведен сравнительный анализ концепций распределения вычислений и хранения данных как в облаке, так и в локальной инфраструктуре предприятия с точки зрения целесообразности и эффективности их использования, приведены основные достоинства и недостатки данных подходов к архитектуре САПР, сформулирован вывод, относительно целесообразности использования гибридной облачной архитектуры.

Современные облачные сервисы обычно имеют собственные системы безопасности, что позволяет обеспечить высокий уровень надежности хранения данных. Вопросы обеспечения безопасности хранения данных в облачных сервисах и пути минимизации выхода данных за пределы САПР подробно описаны в работе [2].

В практике сложных вычислений (особенно в сфере научной деятельности) все шире применяются системы на основе технологии Грид (англ. – GRID). Грид представляет собой компьютерную сеть, в которой вычислительные ресурсы каждого компьютера объединены с ресурсами других компьютеров системы и являются общими и открытыми для использования. Таким образом, любой пользователь, являющийся участником

данной системы, может использовать те или иные её ресурсы для решения различного рода научных, математических, медицинских, экологических, коммерческих и прочих задач. Узлы сети при этом могут иметь различные операционные системы и работать на разных платформах аппаратного обеспечения. Технология Грид имеет ряд преимуществ по сравнению с суперкомпьютерами, в первую очередь, – это стоимость развертывания и обслуживания, а так же простота использования [3].

Одной из распространенных архитектур крупных вычислительных систем является кластер. Кластер – группа компьютеров, объединённых высокоскоростными каналами связи и представляющая, с точки зрения пользователя, единый аппаратный ресурс. Решение о внедрении данного типа архитектуры может быть принято только на основе планируемых нагрузок и ожидаемой производительности вычислений. Данная архитектура отличается высокой стоимостью внедрения и обслуживания. Кроме того, для администрирования и поддержания работоспособности системы необходим штат высококвалифицированного персонала [4].

Цель статьи

Целью статьи является сравнительный анализ возможностей применения наиболее распространенных моделей масштабируемых вычислительных систем для решения вычислительных задач систем автоматизации проектировочных работ.

Изложение основного материала

Масштабируемые параллельные системы можно разделить на такие основные типы:

- мультикомпьютеры;
- кластеры;
- симметричные мультипроцессоры;
- архитектуры с распределенной памятью;
- массово-параллельные системы.

Основным критерием выбора той или иной системной архитектуры является увеличение стоимости, приходящееся на единицу увеличения производительности. По этой причине системы на серийных процессорах и модулях памяти (массово параллельные системы) стали вытеснять векторно-конвейерные суперкомпьютеры.

Мультикомпьютер – совокупность объединенных сетью отдельных вычислительных модулей, каждый из которых управляется своей операционной системой (ОС). Узлы мультикомпьютера, как правило, не имеют общих структур кроме сети и характеризуются высоким уровнем автономности. Мультикомпьютерные системы могут состоять из отдельных компьютеров или представлять собой различные комбинации кластеров, SMP-, DSM- и

MPP-систем [5]. Мультикомпьютер рассматривается пользователем как общий ресурс, который выполняет его процессы.

Кластер – набор компьютеров, который рассматривается операционной системой, программным обеспечением и пользователями как единая система. Кластеры имеют высокий уровень готовности благодаря наличию собственной оперативной памяти и копии операционной системы в каждом узле. Кластеры могут иметь один из типов архитектуры: с разделяемыми дисками и без разделения дисков. В первом случае узлы кластера имеют общие диски, на которых может располагаться общая база данных. Второй тип предполагает наличие собственной оперативной памяти и дисков, а общей является только коммуникационная система. Кластеры являются масштабируемой архитектурой и подходят для использования в высокоструктурированных приложениях, в том числе САПР. Скорость передачи данных по протоколу TCP/IP сильно уступает скорости внутрикластерных протоколов. Кластерные системы имеют неограниченный потенциал наращивания производительности, но при этом растут расходы на администрирование.

Симметричные мультипроцессоры (SMP) – это системы, состоящие из десятков процессоров с общей памятью, объединенных коммуникационной системой. Каждый из процессоров имеет доступ к общей памяти и может прерывать другие процессоры. Скорость обмена данными между процессорами намного превышает скорость передачи между узлами мультикомпьютера или кластера. Однако при наращивании числа процессоров время записи данных увеличивается из-за наличия синхронизации. Таким образом, наращивание числа процессоров приводит к снижению скорости работы системы.

Системы с распределенной разделяемой памятью (DSM) – системы, в каждом процессорном узле которых, помимо кэша, есть локальная память. Узел может иметь архитектуру SMP и состоять из нескольких процессоров. Система работает с единым адресным пространством памяти, однако сама память является распределенной по узлам. Следовательно, время доступа к данным зависит от их места расположения. DSM системы можно разделить на две группы: NUMO (Non-Uniform Memory Access) и ccNUMO (cache coherent NUMO). В системах типа NUMO отсутствует проблема когерентности, т. к. данные не заносятся в кэш каждого процессора. В данных системах аппаратно не поддерживается когерентность, а управление осуществляется специальным контроллером. Целостность данных поддерживается только программными методами, возможности которых

существенно ограничены. Однако программирование таких систем сложнее и требует больших затрат. В системах типа ccNUMO обеспечивается согласованность данных в кэш-памяти. Состояние кэшей отслеживается с помощью протокола на основе справочника (directory). Справочник распределен по узлам ccNUMO системы и хранит состояние каждого блока данных, его копию, а также сведения об изменении данных.

Иерархичность доступа к памяти в NUMO-архитектурах ограничивает увеличение количества процессорных узлов. Обычно их количество составляет 64, а количество процессоров – 128. Большое количество трафика в процессе работы системы является еще одним ограничивающим фактором, что делает архитектуру малоэффективной для обработки процессов с большим количеством обращений процессора к памяти.

Массово-параллельные системы – характеризуются большим количеством вычислительных узлов. Узлы, как правило, состоят из процессора, локальной памяти и устройств ввода/вывода.

Таким образом, все рассмотренные архитектуры имеют общие черты, это – распределённость вычислений и распределённость данных. С одной стороны в каждой из моделей рассмотренных вычислительных систем заложена возможность масштабирования, однако вследствие увеличения интенсивности обмена данными, производительность работы может сильно снижаться. Кроме того, существует максимальный порог роста производительности, который определяется долей последовательных процессов.

Наиболее важными параметрами, имеющими наибольший вес при выборе архитектуры распределенной вычислительной системы, являются следующие: скорость работы (скорость передачи данных внутри системы), готовность (быстрота внедрения), возможность масштабирования (возможность увеличения производительности), скорость после масштабирования (скорость передачи данных внутри системы при наращивании количества вычислительных узлов), уровень затрат на внедрение (стоимость внедрения системы), уровень затрат на сопровождение (стоимость поддержки, обслуживания и администрирования системы), стоимость за единицу увеличения производительности. Для сравнения типов архитектур по параметрам эффективности предлагается использовать следующую градацию: «низкий» – минимальное значение параметра, «средний» – промежуточное значение параметра, «высокий» – максимальное значение параметра. Сравнение описанных систем по ключевым параметрам эффективности представлено в таблице.

Сравнение масштабируемых систем по основным параметрам эффективности

Параметр	Тип системы			
	Мультикомпьютер	Кластер	SMP, DSM	MPP
Скорость работы	Средняя	Средняя	Высокая	Средняя
Готовность	Высокая	Высокая	Низкая	Высокая
Возможность масштабирования	Высокая	Высокая	Низкая	Высокая
Скорость после масштабирования	Средняя	Средняя	Низкая	Средняя
Уровень затрат на внедрение	Низкий	Средний	Высокий	Низкий
Уровень затрат на сопровождение	Низкий	Средний	Высокий	Низкий
Стоимость увеличения производительности	Низкая	Средняя	Высокая	Низкая

Выводы

Проведен сравнительный анализ технологий построения распределенных систем с рассмотрением их возможностей для интеграции вычислительных ресурсов архитектурно-строительных САПР. Результаты исследования показали, что все исследуемые архитектуры в той или иной степени подходят для решения вычислительных задач в процессе работы САПР. На сегодня наиболее предпочтительной является архитектура мультикомпьютер. Она обеспечивает высокую производительность с довольно умеренным уровнем затрат.

Показано, что выбор вычислительной системы для каждой конкретной организации, зависит от целей и ожидаемого уровня производительности при прогнозируемом характере нагрузок, а также является предметом дальнейших исследований в этом направлении. Результаты сравнения масштабируемых вычислительных систем приведены в таблице.

Список литературы

1. Айрапетов Д.А. Исследование и выбор концепции распределения и использования ресурсов проектирования для САПР систем управления сложными технологическими объектами // Известия ВолгГТУ. – 2012. – № 7(94). – С.196-199.
2. Некрасова К.А. Средства защиты образовательных облачных сервисов // Сб. научн. трудов SWorld. Материалы международной научно-практической конференции «Современные проблемы и пути их решения в науке, транспорте, производстве и образовании 2012». – Одесса: Куприенко, 2012. – Вып. 4. – 116 с.
3. Палюх Б.В., Борисов А.Л., Борисов С.Ю. Основные проблемы и перспективы развития систем распределенных вычислений // Вестник ТвГТУ. – 2012. – Вып. 20 (180). – С. 6–11.
4. Воеводин В. В., Жуматий С. А., Вычислительное дело и кластерные системы. – М.: МГУ, 2007. – 150 с.
5. Топорков В. В. Модели распределенных вычислений. – М.: Физматлит, 2004. – 320 с.

Статья поступила в редколлегию 30.07.2013

Рецензент: д-р техн. наук, проф. В.М. Михайленко, Киевский национальный университет строительства и архитектуры, Киев.