



К.С. Жуманазаров, А.А. Азимов, Р.Д. Юлдашева

ИСПОЛЬЗОВАНИЕ ГРИД ТЕХНОЛОГИИ В ОБРАБОТКИ БОЛЬШИХ МАССИВОВ ДАННЫХ

(Каршинский филиал ТУИТ имени Мухаммад аль-Хоразми, Узбекистан)

Создание корпоративных информационных систем со сложной распределенной структурой предполагает постоянный обмен данными между рассредоточенными компьютерными системами на высокой скорости, которая обеспечивается средствами телекоммуникации. Поэтому на современном этапе развития науки об информатизации процессы обработки данных на мощных компьютерных системах немислима, рассматривать отдельно, без применения средств телекоммуникации и наоборот.

Бурное развитие информационно-коммуникационной технологии и на её основе распределенных систем привело к появлению новой технологии, получившей название Грид – технологии.

Идейной основой технологии Грид является создание на основе сетевых технологий компьютерной инфраструктуры нового типа, обеспечивающей глобальную интеграцию информационных и вычислительных ресурсов, а также ресурсов памяти.

Грид технологии с помощью специального программного обеспечения промежуточного уровня (между базовым и прикладным программным обеспечением), а также набора стандартизованных служб обеспечивают надежный совместный доступ к географически распределенным информационным и вычислительным ресурсам, а именно к отдельным компьютерам, кластерам, хранилищам информации и сетям[2].

Возможность применения Грид технологии в разных отраслях от бизнеса до науки делает данную технологию более удобной для вычисления более сложных корреляционных данных, также, по данным Институт системного анализа РАН технология Грид следующие популярные приложения:

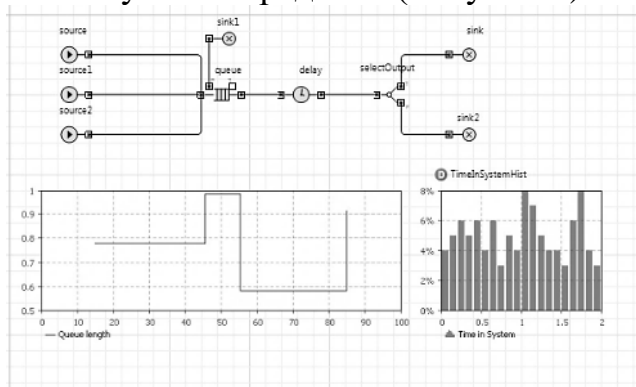
Высокопроизводительные вычисления

- Интерактивное моделирование (климат);
 - Имитационные эксперименты с высоким разрешением / большим числом объектов (формирование галактик, гравитационные волны, имитация боевых действий);
 - Проектирование (прогон вариантов, объединение моделей компонентов);
Работа с большими массивами данных;
 - Анализ экспериментальных данных (физика высоких энергий);
 - Анализ изображений и данных зондирования (астрономия, изучение климата, экология);
- Совместная работа в рамках распределенных коллективов;
- Удаленное управление приборами (микроскопы, рентгеноскопия и т.д.);

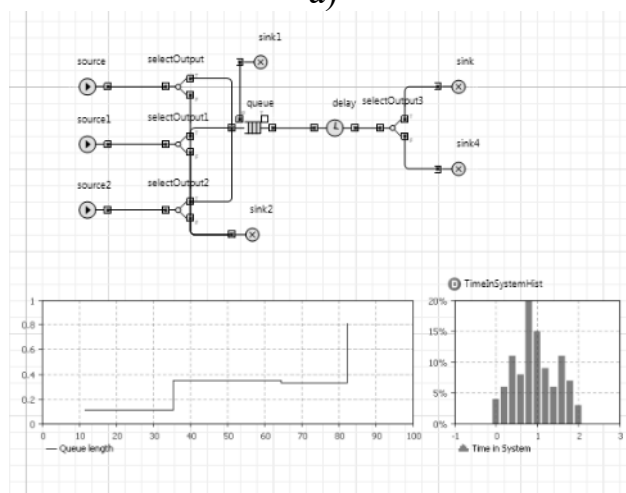


- Удаленная визуализация (изучение климата, биология);
- Инженерия (large-scale structural testing, chemical engineering, крупномасштабные структурные испытания, химическая инженерия)[1].

Исходя из вышеуказанных в особенности преимущество Грид технологии при работе с большими массивами цель данной работы является применения в банковском секторе Грид технологии, а именно в системе получения разных видов (краткосрочных, среднесрочных, долгосрочных) кредитов. Так как получения того или иного кредита подразумевает сбор нескольких видов документов, это требует от получателя затраты большого количество времени. Ко всему это также важна кредитная история получателя. Используя некоторые преимущество(высокопроизводительные вычисления, совместная работа в рамках распределенных коллективов, работа с большими массивами данных) Грид технологии можно в разы ускорить процесс обслуживания получателей, также исходя из больших массивов данных возможно предугадать крупно рискованных кредиторов что в свою очередь снижает финансовые потери для банков. При помощи симмуляционной программы AnyLogic можно экспериментально выстроит процесс обработки получения кредитов (Рисунок 1).



а)



б)

Рис. 1. Симмуляционная модель обработки заявок для получения кредитов. а) без использования Грид технологии, б) с использованием Грид технологии



В симмуляционном процессе используются следующие элементы:

- Первым объектом в процессе симмуляции является объект типа Source. Объект Source генерирует заявки определенного типа. Заявки представляют собой объекты, которые производятся, обрабатываются, обслуживаются, или еще каким-нибудь образом подвергаются действию моделируемого в этом модели это клиенты в банковской системы;

- Следующий объект — Queue. Он моделирует очередь заявок, ожидающих приема объектами, следующими за данным объектом в процессе симмуляции. В нашем случае он будет моделировать очередь запросов, ждущих освобождения предыдущего заявителя (кредитный отдел).

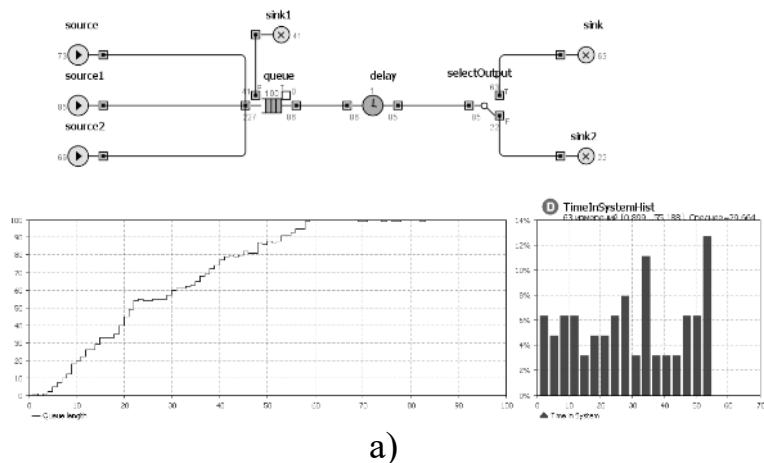
- Объект select Output направляет входящие заявки в один из двух выходных портов T и F в зависимости от выполнения заданного условия. Условие может зависеть как от заявки, так и от каких-то внешних факторов. Поступившая заявка покидает объект в тот же момент времени. Может использоваться для сортировки заявок согласно заданному критерию, для случайного разделения потока заявок (заданием вероятностей) на части и т.д.

- Объект Delay. Он задерживает заявки на заданный период времени, представляя в нашей модели непосредственно банк, на котором обрабатываются запросы.

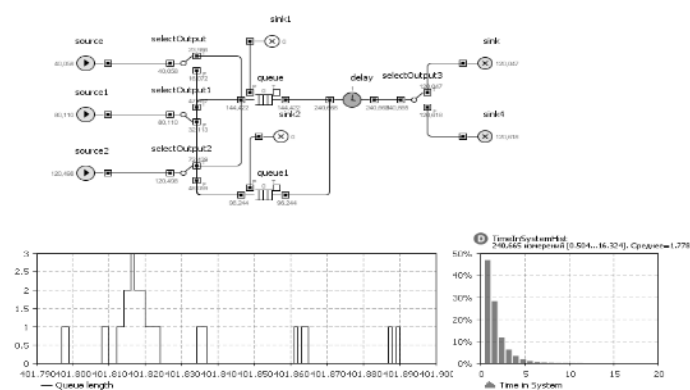
- Последним в диаграмме нашей дискретно-событийной модели находится объект Sink. Этот объект уничтожает поступившие заявки. Обычно он используется в качестве конечной точки потока заявок. В нашем случае он выводит из модели обработанные банком заявки клиентов[3].

В данном симмуляционном процессе показывает как обработка данных без использования Грид технологии снижает количество обработанных заявок в то время как обработка заявок с использованием Грид технологии в разы увеличивает количество обработанных данных с уменьшением затраты времени на каждую заявку (Рисунок 2).

Как видно из результатов симмуляционной модели с увеличением количество подаваемых заявок, время затрачиваемой времени на обработку одной из заявок без использования Грид технологии увеличивается в разы, также исходя из времени ожидания сброс из 227 заявок, обработано всего 85 заявок, в то время, как технология с использованием Грид технологии уменьшает время затраты. У него в корни отсутствуют необработанные заявки. Из 240,666 заявок обработано 240,666 заявок, тем самым в разы увеличивает доход банка от кредитных операций снижением рисков факторов.



a)



b)

Рис. 2. Результат симмуляционной модели обработки заявок для получения кредитов. а) без использования Грид технологии, б) с использованием Грид технологии

Литература

1. Институт системного анализа РАН. Курс "Grid-технологии" Лекция 1. Введение. Грид технологии. Москва-2016. Стр. 34.
2. Ian Foster. What is the Grid? A Three Point Checklist. Argonne National Laboratory & University of Chicago. July 20, 2002
3. Боев В.Д., Кирик Д.И., Сыпченко Р.П. Компьютерное моделирование пособие для курсового и дипломного проектирования военная академия связи. Санкт-Петербург-2011. Стр. 125. УДК 681.142.1.001.57 681.142.33 62-506/507