

---

## ИМИТАЦИОННАЯ МОДЕЛЬ СИСТЕМЫ РАСПРЕДЕЛЕННОЙ ОБРАБОТКИ ДАННЫХ<sup>1</sup>

А.С. Родионов, С.В. Рудометов, В.В. Окольников (Новосибирск)

### Введение

Целью систем обработки большого потока данных является выделение значимой информации из массива начальных данных. Примером таких систем может быть метеорологический мониторинг с целью предсказания погодных аномалий.

Для решения таких задач применяются распределенные вычислительные сети с большим количеством узлов, каждый из которых представляет, как правило, сложный вычислительный центр с многоядерными вычислителями или даже суперкомпьютерами.

Для прогнозирования нагрузок в такой сети, а также для предсказания иных ее характеристик (например, пиковая производительность или оценка эффективности применяемых алгоритмов), могут применяться системы имитационного моделирования.

### Постановка задачи

Моделируемая система распределенной обработки данных (СРОД) представляет собой четырехуровневую распределенную сеть сбора и обработки больших объемов данных мониторинга окружающей среды.

Узлы сбора информации представляют собой разнородные устройства с различными протоколами работы и различной скоростью передачи данных (спутник, метеостанция). Каждый узел сбора информации подключен к одному, нижнему, вычислительному узлу в иерархии СРОД.

Все узлы в СРОД имеют полную связанность с различными задержками на передачу информации между узлами.

Узлы первого уровня принимают данные из различных источников мониторинга. Закон распределения входного потока каждого источника близок к нормальному с высокой интенсивностью.

Обработка входных данных в каждом узле СРОД осуществляется программными компонентами (фильтрами). Результаты обработки сохраняются на текущем узле обработки, а также передаются на узлы более верхнего уровня для дальнейшей обработки. Данные могут «спускаться» на нижние уровни, для повторной обработки. Конечным результатом обработки является сигнал возможной метеопасности, передаваемый на один из узлов верхнего уровня.

Имитационная модель должна решать следующие задачи:

1. Минимизация вероятности потери сигналов о метеопасной ситуации.
2. Минимизация времени обнаружения и обработки сигналов о метеопасной ситуации.
3. Минимизация ресурсов вычислительных узлов (память, вычислительные каналы).
4. Оценка эффективности используемых программ (фильтров).

---

<sup>1</sup> Работа поддержана Российским фондом фундаментальных исследований (код проекта 13-07-98023 р\_сибирь\_a).

### Описание имитационной модели СРОД

Имитационная модель СРОД была реализована с использованием механизма графической сборки в среде имитационного моделирования MTSS [1, 2, 3] с помощью *элементарных моделей* (ЭМ) этой системы.

Имитационная модель строится графически из предопределенных на этапе декомпозиции ЭМ, путем расположения на экране нужного количества ЭМ и соединения их между собою. Модель допускает произвольную модификацию набора ЭМ и связей между ними.

**Декомпозиция исходной системы.** Для моделирования СРОД, состоящей из  $N$  узлов обработки (вычислительных центров), соединенных между собой линиями связи, требуются следующие ЭМ:

**ЭМ «вычислительный центр» (ЭМ ВЦ)** имитирует работу вычислительного центра по сбору, хранению и обработке информации. Информация поступает в ЭМ ВЦ по входящим «линиям связи». Сама ЭМ ВЦ также может посылать информацию по исходящим «линиям связи».

**ЭМ «генератор данных» (ЭМ ГД)** требуется для имитации появления пакетов данных в системе. Пакеты данных появляются по заданному распределению (уникальному для каждой копии ЭМ ГД).

**ЭМ «линия связи» (ЭМ ЛС)** служит для задания топологии в модели СРОД. Наличие этой ЭМ между любыми двумя вычислительными узлами или между генератором данных и вычислительным узлом означает возможность посылки данных между такими объектами.

Графическая часть этих ЭМ служит для отображения состояния каждого экземпляра ЭМ. Также для каждой копии ЭМ задаются наборы **индикаторов** и **вычисляемых значений**. Эти значения могут собираться как статистические данные, а также отображаться в модели как графические индикаторы.

«Продуктом» в такой имитационной модели служит **пакет данных**. Пакет данных имеет тип, метку времени генерации, размер. Имитируется следующее поведение пакета (в процессе его прохождения по имитационной модели):

1. Порождение нового ПД реализуется в ЭМ ГД.
2. Передача пакета по линии связи реализуется в ЭМ ЛС. Возможная задержка определяется типом данных и их размером.
3. Обработка пакета реализуется в ЭМ ВЦ. При этом возможен как «вывод» пакета из системы (окончание его жизни), так и порождение нового пакета (после обработки исходного). Пакеты могут копироваться при передаче по нескольким линиям связи.

**ЭМ ВЦ** является подробной моделью стандартного вычислительного центра.

Пакет данных, поступивший в ЭМ ВЦ, проходит следующие этапы обработки:

1. Сохранение в подсистеме хранения данных (ПХД).
2. Определение вида обработки пакета (выбор «фильтра»).
3. Постановка в очередь на вычисление.
4. Вычисление.
5. Посылка результатов вычисления по выходным линиям связи.

**ЭМ ГД** позволяет имитировать входящий поток данных в виде пакетов данных для их последующей обработки. Пакеты данных имеют нормальное распределение по времени генерации и равномерное – по размеру данных.

**ЭМ ЛС** позволяет визуально соединить 2 объекта в модели: ЭМ ГД и ЭМ ВЦ или ЭМ ВЦ и ЭМ ВЦ. ЭМ ЛС имитирует однонаправленную связь.

С использованием построенной имитационной модели были решены следующие задачи:

1. Контроль загрузки всей вычислительной сети и каждого узла в отдельности.
2. Контроль количества обработанных пакетов данных.
3. Имитация отказа части сети.
4. Модификация линий связи и контроль скорости обработки информации после такой модификации.
5. Количество и качество применяемых вычислительных «фильтров» на каждом из узлов.
6. Контроль расхода памяти в ПХД каждого узла.

Фрагмент имитационной модели СРОД, реализованной в системе MTSS, представлен на рис. 1.

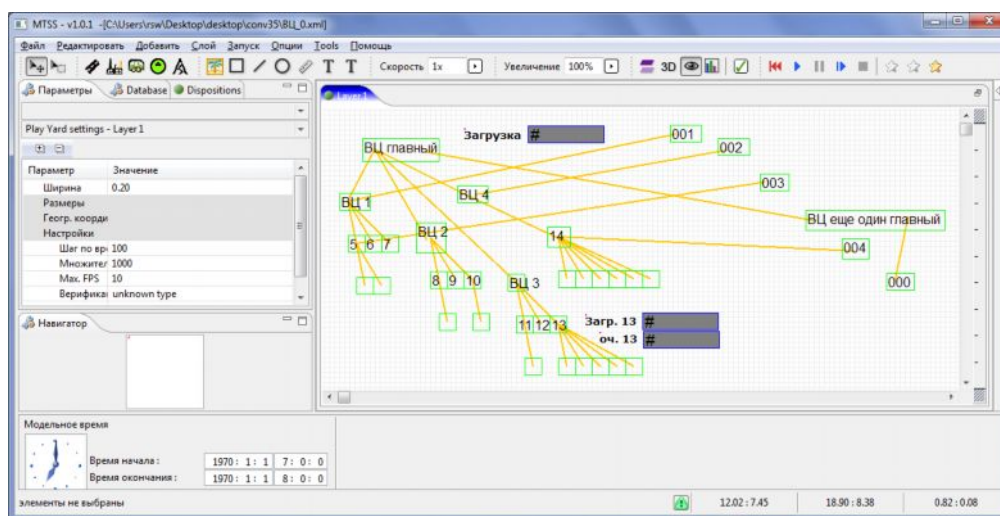


Рис. 1. Имитационная модель СРОД в системе MTSS

В процессе работы имитационной модели выполняется анализ и вывод статистических данных. Эти статистические данные используются для быстрого определения глобальных параметров модели, таких как:

1. Средняя загрузка вычислительных узлов.
2. Количество обработанных пакетов данных.
3. Расход электроэнергии и других ресурсов и т.д.

#### Валидация имитационной модели

Валидация имитационной модели СРОД осуществлялась с использованием результатов аналитического решения для упрощенной сети распределенной обработки данных. Аналитическое решение выполнено другим коллективом авторов и является самостоятельным научным результатом. Для валидации потребовалось также некоторое упрощение элементарных моделей. Результаты, полученные с помощью имитационной модели, разработанной в системе MTSS, совпадают с аналитическим решением упрощенной сети распределенной обработки данных, содержащей 5 узлов.

Результаты, полученные с помощью имитационной модели СРОД в MTSS, аналогичны результатам, полученным при моделировании СРОД с помощью среды моделирования AGNES [4, 5].

#### Выводы

Подход и архитектурно-программные решения, реализованные в системе MTSS, позволяют моделировать широкий класс сложных вычислительных сетей обработки дан-

---

ных. Библиотека ЭМ СРОД в системе MTSS может быть переиспользована или использована в качестве прототипа для создания сложных имитационных моделей распределенных систем обработки данных.

### Литература

1. **Рудометов С.В.** Создание системы имитационного моделирования технологических систем // Тр. Пятой всероссийской научно-практической конференции по имитационному моделированию и его применению в науке и промышленности. Имитационное моделирование, Теория и практика (ИММОД-2011). – Санкт-Петербург. – 2011. – Т.1. – С. 383–387.
2. **Рудометов С.В.** Визуально-интерактивная система имитационного моделирования технологических систем // Вестник СибГУТИ. – 2011. – № 3. – С. 14–27.
3. **Рудометов С.В.** Система имитационного моделирования MTSS. Новосибирск, ФАП СО РАН, 2011, URL: <http://fap.sbras.ru/node/2325>.
4. **Глинский Б.М., Родионов А.С., Марченко М.А., Подкорытов Д.И., Винс Д.В.** Агентно-ориентированный подход к имитационному моделированию суперЭВМ экзафлопсной производительности в приложении к распределенному статистическому моделированию // Вестник ЮУрГУ. – 2012. – № 18 (277). – Вып. 12. – С. 93–106.
5. **Podkorytov D., Rodionov A., Choo H.**, Agent-based simulation system AGNES (AGent NEtwork Simulator) for networks modeling // Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication, ICUIMC'12, 2012. Article No. 115 ACM New York, NY, USA ©2012 table of contents ISBN: 978-1-4503-1172-4 doi>10.1145/2184751.2184883 – 8 p.