

АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ ЦЕНТРА КОЛЛЕКТИВНОЙ ОБРАБОТКИ ИНФОРМАЦИИ

А. Н. Козлов (Санкт-Петербург)

Интеграция колоссальных вычислительных ресурсов в рамках одного процессора, огромные практические успехи при реализации архитектуры параллельных вычислений привели к очередному витку развития вычислительной техники – появлению современных, высокопроизводительных «мейнфреймов», позволяющих осуществлять эффективную коллективную обработку данных. И, как следствие этого, – пробуждению практического интереса крупных корпораций и банков к их внедрению. Централизация обработки данных, коллективное использование вычислительных ресурсов имеет множество преимуществ перед распределенной обработкой данных. Это и единый технологический процесс, единые программные средства и вычислительная среда, меньшая удельная стоимость владения вычислительными ресурсами, повышенная безопасность и конфиденциальность обработки и т. д. Особенно это актуально для банковских вычислительных систем.

В то же время создание мощных центров коллективной обработки данных остро ставит проблему оценки производительности центра, обеспечение его эффективной загрузки при безусловном обеспечении требуемой оперативности и качества вычислений.

Решению этой проблемы и посвящена данная статья. Исследовался крупный центр коллективной обработки данных банка, обрабатывающего запросы по переводу средств от кредитных организаций из других регионов. Т. е. в центр поступает случайный поток запросов различных типов. Их обработка производится поэтапно на вычислительных мощностях центра без прямого вмешательства персонала. Технология обработки достаточно сложна и включает множество этапов. При этом задействуется множество специализированных приложений и баз данных. В настоящее время перед центром, в связи с увеличивающейся рабочей нагрузкой, очень остро стоит проблема расширения конфигурации технических средств. Учитывая, с одной стороны, исключительно высокую их цену, а с другой – необходимость гарантированного обеспечения оперативности и качества обработки запросов, можно убедиться в том, что многократно возрастает ответственность за принятие таких решений. Чтобы обеспечить объективность и экономическое обоснование изменения конфигурации, было принято решение использовать моделирование.

Существует множество методов моделирования подобных систем.

Во-первых, это физическое моделирование. Чаще всего – создание испытательного полигона (центра), полностью соответствующего реальному центру по составу технических и программных средств. Для проверки полноты и надежности его функционирования, определения загрузки его компонент и выявления узких мест разрабатываются разнообразные, близкие к реальным, тестовые нагрузки.

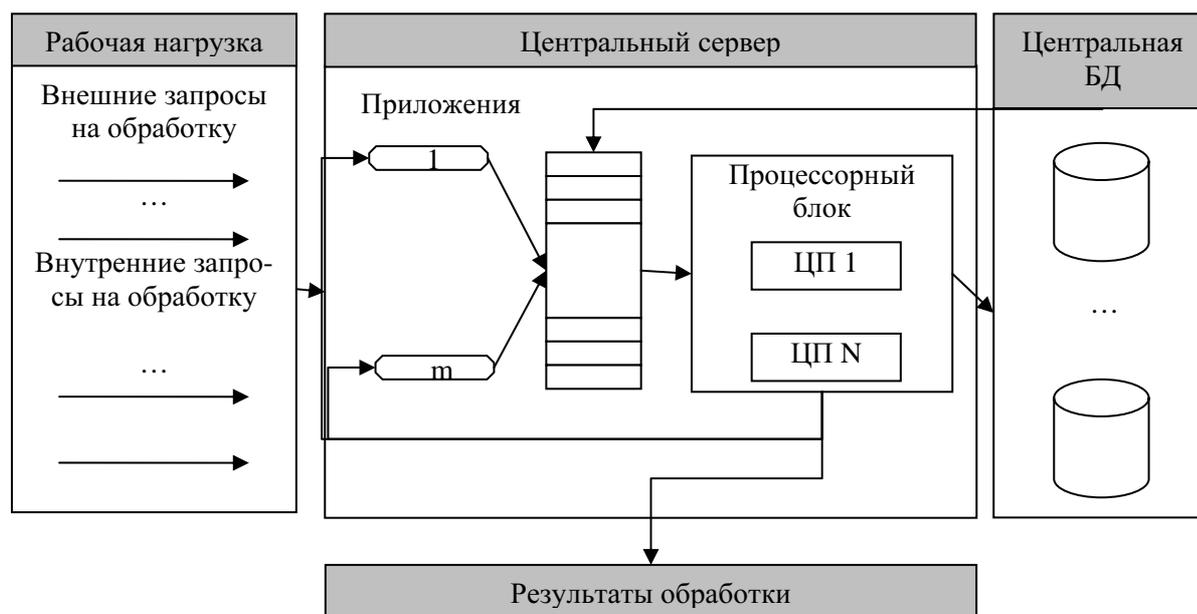
Во-вторых, математическое моделирование. Т. е. создание модели центра и проведение с ней экспериментов вместо экспериментов с реальной системой. Наиболее точные и устойчивые результаты можно получить, если модель построена аналитическими методами. Но чаще всего стремление исследователя представить систему аналитически приводит к неоправданно большим допущениям и ограничениям. И, вследствие этого, к малой практической пригодности таких моделей. В отличие от аналитических моделей *имитационные модели* позволяют создать более функционально полную, следовательно, и адекватную модель. Детализация в такой модели ограничивается лишь возможностями компьютера, наличием исходных данных и самим исследователем.

Исследуемую систему можно представить как дискретную массового обслуживания. В нее по определенной технологии и алгоритмам поступает поток запросов, и на ряде вычислительных ресурсов осуществляется их обработка. Время между поступлениями запросов и время обработки на ресурсах случайно. В процессе обработки могут возникать очереди к тем или иным ресурсам. Запросы могут иметь различный приоритет при обслуживании.

Учитывая сложность данной системы и необходимость учета разнообразных, сложно формализуемых факторов, в качестве метода исследования в данной работе было выбрано **имитационное моделирование**.

При разработке имитационной модели автором был детально проанализирован технологический процесс организации вычислений в центре коллективной обработки данных. В процессе анализа исследовались алгоритмы работы приложений в режимах пакетной обработки и «Online», архитектура используемых аппаратных и программных средств и т. д. В результате было получено достаточно детализированное формализованное описание функционирования вычислительного центра. И в то же время удалось избежать отражения в модели множества второстепенных деталей, которые могли сделать модель слишком большой и затруднить исследование.

В самом общем виде исследуемую систему представим следующим образом:



На основе статистических данных была разработана статистическая модель рабочей нагрузки. Обработка статистических данных производилась с применением ППП Statistica. А по результатам формализации технологического процесса – имитационная модель центра. Имитационная модель была написана на языке GPSS World. В качестве варьируемого параметра может выступать практически любой элемент модели. Но в силу поставленных задач в процессе исследования варьировались характеристики рабочей нагрузки (W) – интенсивность, объем требуемых вычислений, приоритет запроса и т. д. и конфигурация аппаратных средств (S) – количество центральных процессоров, объем оперативной памяти, число каналов ввода-вывода и т.д. При этом технология обработки запросов, перечень и функции приложений оставались неизменными.

На основе опыта эксплуатации системы, возникавших ранее проблем и узких мест результирующими показателями работы системы были выбраны коэффициенты

загрузки центральных процессоров $U = F_1(W, S)$ и среднее время обработки запросов $T = F_2(W, S)$ где $W = \{w_1, w_2, \dots, w_n\}$, $S = \{s_1, s_2, \dots, s_m\}$.

Для облегчения использования модели в процессе исследования был разработан специальный комплекс программ для ввода данных, запуска модели и анализа результатов. В качестве компонент визуально-графического ввода данных о структуре и логике системы была использована система VSM.

В результате практического использования данного программного комплекса был решен ряд практических задач, стоящих перед вычислительным центром. В частности, была проведена серия экспериментов с моделью, которая позволила разработать методику определению числа центральных процессоров в зависимости от параметров рабочей нагрузки.

Таким образом, можно с уверенностью сказать, что применение метода имитационного моделирования для оценки производительности вычислительного центра в процессе его эксплуатации позволяет принимать технически обоснованные и экономически просчитанные решения о проведении модернизации аппаратных и программных средств центра.