

УДК 004.7

М.В. Мартынов

О построении вычислительного кластера для системы мониторинга гидрофизических полей Черного моря в квазиоперативном режиме

Для увеличения разрешающей способности численной модели диагноза гидрофизических параметров в системе спутникового мониторинга Черного моря, а также повышения оперативности получения результатов расчетов производительность имеющихся вычислительных средств должна быть увеличена в разы. Анализируются существующие на компьютерном рынке доступные готовые решения. Предлагается использовать вычислительный кластер, работающий как единая система и позволяющий наращивать вычислительную мощность по мере необходимости. Дана блок-схема кластера, в которой применяются имеющиеся в наличии вычислительные средства.

Постановка задачи. В настоящее время в Морском гидрофизическом институте НАН Украины существует функционирующая в близком к реальному режиму времени система непрерывного диагноза гидрофизических полей Черного моря [1]. Конечным продуктом работы системы являются карты трехмерных полей скорости течений, температуры, солености и других важных характеристик. Основная составляющая системы — численная гидродинамическая модель циркуляции моря, в которой ассимилируются данные спутниковых наблюдений (информация о температуре поверхности моря, данные альтиметрических и скаттерометрических измерений и др.). Для увеличения точности диагноза гидрофизических параметров в системе спутникового мониторинга Черного моря желательно уменьшить шаги сетки численной модели, на которой производятся расчеты. Кроме того, необходимо увеличить оперативность расчетов. На компьютере с процессором *AMD Athlon 64* с тактовой частотой 2 ГГц время расчета по этой модели может составлять несколько суток.

Отсюда можно сформулировать требование к производительности необходимой вычислительной системы. Она должна быть увеличена в разы по сравнению с имеющимися на настоящий момент вычислительными средствами. Потребуем, чтобы скорость расчетов была увеличена в восемь раз (такую цифру дает механическое уменьшение шага сетки в два раза для трехмерной задачи), и покажем, как эта задача может быть решена без существенных финансовых затрат.

Тактовая частота процессора компьютера (2 ГГц), на котором производятся расчеты в отмеченной выше системе мониторинга, близка к максимальной частоте доступных в настоящее время процессоров (у *AMD* максимальная достигнутая частота процессорного ядра 2,5 ГГц, у *Intel* — 3,6 ГГц). Поэтому единственная возможность многократного увеличения производительности вычислительной модели — создание и использование многопроцессорной системы. Данное направление развивается давно всеми производителями суперкомпьютеров. Однако суперкомпьютеры и поныне являются дорогими вычислительными средствами (счет идет на сотни тысяч и миллионы долларов).

Производители процессоров и процессорных систем для массового потребителя *Intel* и *AMD* предлагают процессоры и наборы системной логики, а также системные платформы, поддерживающие от 2 до 8 процессоров.

Многопроцессорные системы, предлагаемые фирмами *Intel* и *AMD*. В настоящее время компания *Intel* предлагает двухпроцессорные (*Intel Xeon*, максимальная частота процессора 3,6 ГГц.) и четырехпроцессорные решения (*Intel Xeon MP*, максимальная частота процессора 3 ГГц.) [2].

Компания *AMD* предлагает двухпроцессорные (*AMD Opteron* серии 200, максимальная частота процессора 2,5 ГГц) и восьмипроцессорные решения (*AMD Opteron* серии 800, максимальная частота процессора 2,5 ГГц.) [3]. Хотя системная логика и сами процессоры *AMD Opteron* серии 800 поддерживают одновременную работу до 8 процессоров, системных плат, реализующих это решение, нет. В настоящее время максимально можно установить на

одну материнскую плату 4 процессора *AMD Opteron* серии 800. Компания *AMD* для сравнения своих процессоров с процессорами *Intel* применяет рейтинговый коэффициент, который приблизительно равен 1,7 (для *Intel Xeon MP* он ниже). С этим коэффициентом в большинстве случаев можно согласиться. Таким образом, *Intel Xeon* (3,6 ГГц) по производительности приблизительно эквивалентен *AMD Opteron* (2,2 ГГц). Поэтому в дальнейшем мы не будем рассматривать варианты от *Intel*, поскольку они менее производительные и более дорогие.

Из сказанного выше следует, что даже теоретически на многопроцессорных системах, предлагаемых фирмами *Intel* и *AMD*, поставленная задача не может быть решена. Суммарная частота четырех процессоров составляет всего $2,5 \text{ ГГц} \cdot 4 = 10 \text{ ГГц}$, в то время как для увеличения производительности вычислительной системы по сравнению с имеющимся компьютером в восемь раз необходим гипотетический процессор с тактовой частотой 16 ГГц. Существует перспектива того, что такая система появится в течение одного-двух лет. Фирма *AMD* анонсировала в середине 2004 г. двухъядерный процессор, т.е. процессор, в одном корпусе которого располагаются два. В широкой продаже такие процессоры появились в середине 2005 г. Их стоимость выше, чем стоимость двух одноядерных процессоров аналогичной частоты, при этом максимальные частоты ниже, чем у одноядерных. Компания *Intel* массовые продажи двухъядерных процессоров планирует начать в 2006 г.

Резюмируя сказанное, отметим, что многопроцессорная система не позволяет наращивать вычислительные мощности по мере необходимости. Повысить производительность в два раза путем добавления второго процессора обычно не удастся, если решается одна задача. Рост производительности меньше и зависит от того, насколько удастся разделить реализуемую задачу и равномерно загрузить процессоры.

Кластерное решение. В соответствии с определением компании *DEC*, кластер — это группа вычислительных машин, которые связаны между собой и функционируют как один узел обработки информации [4]. Вычислительные кластеры появились сравнительно недавно. Первый крупный вычислительный кластер был создан в 1994 г. в *NASA*. Это был 16-процессорный кластер, получивший название *Beowulf* [5]. Проект существует до настоящего времени. На его сайте [5] расположен учебник, призванный облегчить создание вычислительных кластеров, в которых используются недорогие стандартные коммутационные устройства и бесплатное программное обеспечение.

Первый кластер (*Avalon* [6]), превзошедший суперкомпьютеры по показателю цена/производительность, был создан в 1998 г. в Лос-Аламосской национальной лаборатории. Первоначально он состоял из 68 процессоров. Кластер был собран на стандартном оборудовании *Fast Ethernet* и *Gigabit Ethernet* по обычной топологии многолучевой звезды с коммутатором в центре.

Список крупных вычислительных кластеров, работающих на территории СНГ, и применяемого на них оборудования можно найти на сайте Лаборатории параллельных информационных технологий НИВЦ МГУ (кластерные установки России и СНГ) [7]. Согласно [7], на территории Украины функционируют три кластера.

На кластерах в качестве операционной системы обычно устанавливается система на основе ОС *Linux*, но возможна установка и *Windows*. В современную *Windows 2003 Server* входит поддержка кластера. В качестве коммуникационной библиотеки, отвечающей за параллельное выполнение задачи на различных процессорах, наиболее распространенной является *MPI (Message Passing Interface)*. Существует коммерческая версия *MPI* для *Windows* — *WMPI*. Получить программу можно по адресу <http://www.criticalsoftware.com/wmpi>. При количестве процессоров меньше восьми *WMPI* распространяется бесплатно. Рекомендуется конфигурировать *WMPI* так, чтобы на одном процессоре выполнялся один процесс. Однако на одном процессоре возможно выполнение и нескольких процессов.

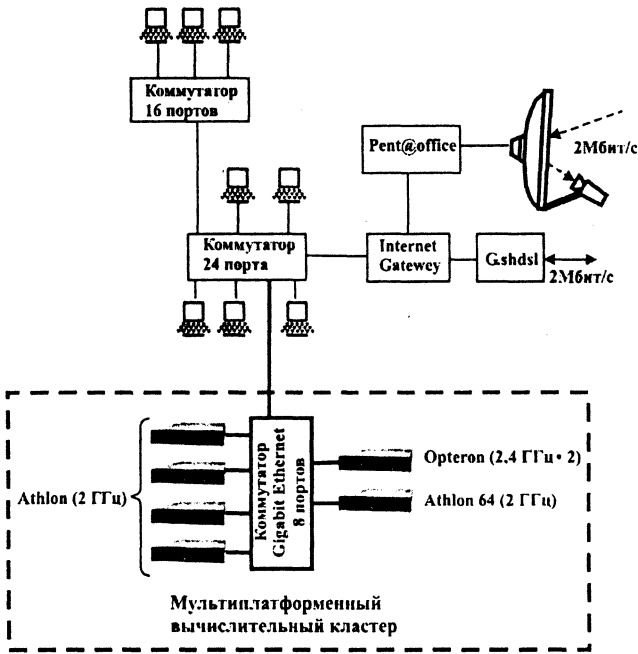
Основные преимущества кластера перед многопроцессорными суперкомпьютерами — меньшая стоимость и возможность наращивания кластера по мере увеличения требований к производительности вычислительной системы. Основной недостаток — меньшая скорость передачи данных между узлами кластера, чем между процессорами, входящими в многопроцессорную систему.

Согласно исследованиям Лаборатории параллельных информационных технологий НИВЦ МГУ существует десять наиболее распространенных коммуникационных технологий, позволяющих объединять отдельные компьютеры в кластер. Скоростные характеристики наиболее распространенных коммуникационных технологий, применяемых при построении кластеров, приведены в таблице [8].

Время задержки и скорость передачи, обеспечиваемые наиболее распространенными коммуникационными технологиями

Тип интерконнекта	Время задержки передачи, мкс	Пиковая производительность, Мбайт/с
Gigabit Ethernet	120	125
Myrinet	10	200
SCI	4	400
InfiniBand	5-7	800

Для сравнения, шина *Hyper Transport*, служащая для передачи данных между процессорами *AMD*, обеспечивает скорость передачи данных 1000 Мбайт/с. У *Intel Xeon* скорость межпроцессорного обмена 537 Мбайт/с. Следовательно, лучшие современные коммуникационные технологии обеспечивают скорости, близкие или даже превосходящие скорость межпроцессорного обмена у *Intel Xeon* [9]. Тем самым стирается, на аппаратном уровне, различие между многопроцессорными системами и вычислительными кластерами. К сожалению, цены на коммутационное оборудование от *Myrinet*, *SCI*, *InfiniBand* очень высоки [7].



Блок-схема компьютерной сети

Планируется организовать кластер на основе *Gigabit Ethernet*. Как видно из таблицы, это самая медленная из коммуникационных технологий, применяемых для организации вычислительных кластеров. Такой выбор объясняется низкой стоимостью коммутационного оборудования. Адаптеры *Gigabit Ethernet* бесплатные, поскольку они установлены на материнскую плату.

Первоначально кластер будет создан на основе имеющихся компьютеров — компьютера с двумя процессорами *AMD Opteron* (2,4 ГГц), компьютера с процессором *AMD Athlon 64* (2 ГГц) и четырех компьютеров с процессорами *AMD Athlon 32* (около 2 ГГц). Эти компьютеры могут использоваться и как рабочие станции, и в составе вычислительного кластера.

Блок-схема компьютерной сети приведена на рисунке. При полной равномерной загрузке процессоров вычислительный кластер должен быть эквивалентен гипотетическому компьютеру с частотой процессора 14,8 ГГц, что позволяет вплотную приблизиться к решению поставленной задачи.

СПИСОК ЛИТЕРАТУРЫ

1. *Дорофеев В.Л., Кортаев Г.К., Ратнер Ю.Б., Мартынов М.В.* Система мониторинга гидрофизических полей Черного моря в квазиоперативном режиме // Экологическая безопасность прибрежной и шельфовой зон и комплексное использование ресурсов шельфа.– Севастополь: МГИ НАН Украины, 2004.– Вып.11.– С. 9–23.
2. *Сайт* компании Intel, подраздел «Семейство процессоров Intel Xeon».– <http://www.intel.com/cd/products/services/emea/rus/browse/processors/164854.html>.
3. *Сайт* компании AMD, подраздел «Объяснение нумерации моделей процессора AMD Opteron».– http://www.amd.com/ru-ru/Processors/ProductInformation/0,,30_118_8796_9240,00.html.
4. *Савяк В.* Эффективные кластерные решения.– www.ixbt.com/cpu/clustering.shtml.
5. *Официальный сайт* проекта Beowulf.– <http://www.beowulf.org/>.
6. *Сайт* проекта Avalon.– <http://cnls.lanl.gov/avalon/>.
7. *Сайт* Лаборатории параллельных информационных технологий НИВЦ МГУ (кластерные установки России и СНГ).– <http://www.parallel.ru/about/lab.html>.
8. *Сайт* «Дата Технологии» – крупнейшего производителя кластеров в России.– <http://www.datatec.ru/index.php>.
9. *Сайт* компании AMD, подраздел «Сравнение с конкурентами».– http://www.amd.com/ru-ru/Processors/ProductInformation/0,,30_118_8826_9014,00.html.

Морской гидрофизический институт НАН Украины,
Севастополь

Материал поступил
в редакцию 31.03.05
После доработки 06.04.05

ABSTRACT To increase resolution of a numerical model of the diagnosis of hydrophysical parameters in the system of the Black Sea satellite monitoring and efficiency of obtaining calculation results, capacity of the existing computing means should be multiply raised. Software available on the computer market is analyzed. A computer cluster operating as a uniform system and permitting to increase computing capacity is proposed for application. The scheme of a cluster including the available computer means is given.