

УДК 681.3.06
ББК 32.973

Б.П. Ильин

ИСПОЛЬЗОВАНИЕ СВОБОДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОРГАНИЗАЦИИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ*

Дан обзор истории развития вычислительной техники в области высокопроизводительных вычислений. Приведены определения суперкомпьютеров и кластеров. Описаны способы использования свободного программного обеспечения для организации параллельных вычислений.

Ключевые слова: параллельные вычисления; кластер; суперкомпьютер; свободное программное обеспечение.

B.P. Ilyin

USING FREE SOFTWARE IN PARALLEL COMPUTING

The article provides an overview of history of high-performance computing development, gives definitions to 'supercomputer' and 'cluster', and describes the ways of using free software in parallel computing.

Keywords: parallel computing, cluster, supercomputer, free software.

Одним из ярчайших изобретений человечества по праву считается компьютер. Примерно с середины прошлого века в научных институтах ведущих держав начали появляться ЭВМ первого поколения, они, а также компьютеры некоторых других поколений представляли собой массивные вычислительные устройства, занимающие большие площади, требовавшие огромных энергозатрат и непростые в эксплуатации. Однако, с течением времени, а также с уменьшением элементной базы, на которой строились устройства, размер этих устройств уменьшался, что привело к появлению к концу 70-х гг. прошлого века персональных компьютеров.

Известно, что первопричиной создания компьютеров была настоятельная необходимость быстрого проведения вычислительных работ, связанных с решением больших научно-технических задач в атомной физике, авиастроении, климатологии и т.п. Решение таких задач и сейчас остается главным стимулом совершенствования компьютеров. Однако, в настоящее время, основная сфера их применения связана совсем с иной деятельностью, в которой вычислительная составляющая либо отсутствует, либо занимает небольшую часть [1].

Исходя из задач, которые ставились перед компьютерами, и технологической базой, на которых они строились, со временем выделилось два подхода к организации сложных вычислений — суперкомпьютеры и совокупность персональных компьютеров, объединенных в вычислительную сеть.

Началом эры суперкомпьютеров с полным правом может считаться 1976 г. — год появления первой векторной системы Cray 1. Результаты, показанные этой системой, пусть и на ограниченном в то время наборе приложений, были столь впечатляющими, в сравнении с остальными, что

* Работа выполнена при финансовой поддержке проекта ФБ-40 «Повышение эффективности информационно-телекоммуникационных систем на основе свободного программного обеспечения (СПО) для бюджетных организаций».

система заслуженно получила название «суперкомпьютер» и, в течение длительного времени, определяла развитие всей индустрии высокопроизводительных вычислений.

Попытки дать определение термину «суперкомпьютер», опираясь только на производительность, неизбежно приводят к необходимости постоянно поднимать планку, отделяющую его от рабочей станции или даже обычного настольного компьютера. Так, по определению Оксфордского словаря вычислительной техники 1986 г., для того чтобы получить это гордое название, нужно было иметь производительность в 10 MFlops). Сегодня, как известно, производительность настольных систем на два порядка выше.

Из альтернативных определений наиболее интересны два: экономическое и философское. Первое из них гласит, что суперкомпьютер — это система, цена которой выше 1–2 млн дол. Второе — что суперкомпьютер — это компьютер, мощность которого всего на порядок меньше необходимой для решения современных задач.

Кластер — группа компьютеров, объединенных в локальную вычислительную сеть (ЛВС) и способных работать в качестве единого вычислительного ресурса. Дополнительно предполагается, что для кластера обеспечивается более высокая надежность и эффективность, нежели для ЛВС, и существенно более низкая стоимость в сравнении с другими типами параллельных вычислительных систем (за счет использования типовых аппаратных и программных решений) [2].

Сейчас в научных организациях и университетах, как правило, имеются энтузиасты свободно распространяемого ПО и специалисты по ОС Linux. В то же время парк более-менее современных персональных компьютеров в этих организациях также имеется. Закономерно появилась идея создания параллельных вычислительных систем из общедоступных компьютеров на базе процессоров Intel и недорогих Ethernet-сетей, установив на эти компьютеры Linux и объединив их с помощью одной из бесплатно распространяемых коммуникационных библиотек (PVM или MPI) в кластер. Оказалось, что на многих классах задач и при достаточном числе узлов такие системы дают производительность, сравнимую с той, что можно получить, используя дорогие суперкомпьютеры.

Варианты построения вычислительного кластера на базе ОС Linux:

1. На машины, которые будут узлами кластера, устанавливается Linux. Для установки используется дистрибутив общего назначения, например ALT Linux или Ubuntu. После установки системы и настройки сети, на узлы будущего кластера устанавливается UNIX-вариант системы MPI. Этот вариант хорош еще тем, что избавляет от необходимости решать проблему, связанную с переходом на незнакомую операционную систему. Например, если кластер организуется на базе компьютерного класса, используемого в учебном процессе вуза, нет необходимости перестраивать учебные планы с учетом использования Linux в качестве базовой операционной системы. Установка ОС Linux позволяет применять мультизагрузку. То есть имеется возможность не уничтожать ранее установленную на компьютере операционную систему, а использовать обе системы совместно.

2. Одной из ведущих мировых Linux-компаний Mandrake выпускается специализированный дистрибутив CLIC. Этот дистрибутив предназначен для создания высокоэффективного кластерного суперкомпьютера, основанного на сети рабочих станций. После установки дистрибутива все необходимое матобеспечение для работы и мониторинга кластера уже будет установлено в системе.

3. Для быстрого развертывания виртуального кластера потребуется один компакт-диск с последней версией дистрибутива PelicanHPC GNU Linux. С этого диска загружается операционная система кластера (без установки ее на винчестер) на компьютере, который будет играть роль консоли кластера, т.е. того компьютера, непосредственно на котором будут компилироваться и запускаться ваши параллельные программы.

Остальные узлы кластера будут загружаться по сети. Для загрузки ОС вычислительных узлов кластера по сети необходимо, чтобы сетевые карты этих компьютеров умели выполнять загрузку по сети. Большинство современных карт, в том числе встроенных, это делать умеют.

4. Как и в предыдущем варианте, в данном случае используется специализированный дистрибутив: Rocks Cluster Distribution. Для установки кластера потребуются два диска: загрузочный Rocks Base и дополнительный HPC Roll. Эти два диска позволяют установить кластер в минимально необходимой конфигурации. В дополнение к минимальной конфигурации, дистрибутив кластера содержит компакт-диск Intel Roll, позволяющий установить MPI, интегрированный с оптимизирующим компилятором языка FORTRAN 95 от компании Intel. Также полезным может оказаться дополнение, находящееся на компакт-диске PBS/Maui Roll. Пакет PBS является программным средством для управления очередью заданий, запускаемых на кластере. Управление очередью заданий может оказаться необходимым компонентом кластера в случае, когда он интенсивно используется большим количеством пользователей.

5. Проект ThinCluster. Для организации кластера с помощью этого проекта необходимо на имеющуюся Linux-систему определенным образом установить необходимые пакеты (NFS, DHCP, MPICH, TFTP, gcc) [3].

Итак, из вышеизложенного можно сделать вывод о том, что в настоящее время параллельные вычисления являются одним из перспективнейших средств организации научных вычислений и могут осуществляться на базе свободно распространяемого программного обеспечения.

Список использованной литературы

1. Воеводин В.В. Параллельные вычисления / В.В. Воеводин, Вл.В. Воеводин. — СПб.: БХВ-Петербург, 2004. — 608 с.
2. Гергель В.П. Теория и практика параллельных вычислений: учеб. пособие / В.П. Гергель. — М.: Интернет-Университет информационных технологий; БИНОМ. Лаборатория знаний, 2007. — 423 с.
3. Сбитнев Ю. Практическое руководство по параллельным вычислениям / Linux. Кластер: сайт. — URL: <http://cluster.linux-ekb.info/>.

References

1. Voevodin V.V. Parallel'nye vychisleniya / V.V. Voevodin, Vl.V. Voevodin. — SPb.: BKhV-Peterburg, 2004. — 608 s.
2. Gergel' V.P. Teoriya i praktika parallel'nykh vychislenii: ucheb. posobie / V.P. Gergel'. — M.: Internet-Universitet informatsionnykh tekhnologii; BINOM. Laboratoriya znani, 2007. — 423 s.
3. Sbitnev Yu. Prakticheskoe rukovodstvo po parallel'nym vychisleniyam / Linux. Klaster: sait. — URL: <http://cluster.linux-ekb.info/>.

Информация об авторе

Ильин Борис Петрович — старший преподаватель, кафедра информатики и кибернетики, Байкальский государственный университет экономики и права, г. Иркутск, e-mail: molotoff_2001@mail.ru.

Author

Ilyin Boris Petrovich — Senior Instructor, Chair of Computer Science and Cybernetics, Baikal State University of Economics and Law, Irkutsk, e-mail: molotoff_2001@mail.ru.