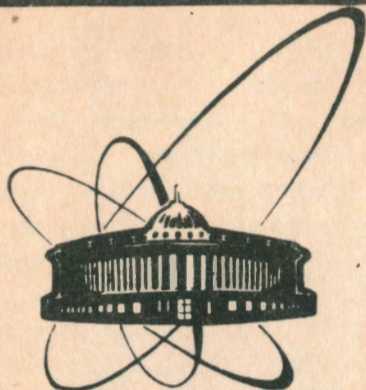


92-94



СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

P11-92-94

А.Д.Бавижев, В.В.Кореньков, Г.А.Коробова

СОВМЕСТНОЕ ИСПОЛЬЗОВАНИЕ
ОБЩЕЙ ДИСКОВОЙ ПОДСИСТЕМЫ
В МНОГОМАШИННОМ КОМПЛЕКСЕ ЕС ЭВМ

1992

Бавижев А.Д., Кореньков В.В.,
Коробова Г.А.

P11-92-94

Совместное использование общей
дисковой подсистемы
в многомашинном комплексе ЕС ЭВМ

Работа посвящена актуальной проблеме совместного использования общей дисковой памяти несколькими ЕС ЭВМ в среде операционной системы СВМ. Описан программный механизм резервирования мини-дисков.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1992

Перевод авторов

Bavizhev A.D., Koren'kov V.V.,
Korobova G.A.

P11-92-94

Shared Disk Storage Usage
for ES-Multicomputing System

The paper is devoted to the actual problem of shared disk storage usage on ES computers for VM operating system. The software providing mini-disks reserving is described.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

В последние годы интенсивно стали развиваться многомашинные комплексы ЭВМ, которые находят все более широкое применение в таких областях, как вычислительные системы с дистанционным доступом, сети ЭВМ, системы коллективного пользования, общие банки данных и т.п. Основными целями создания многомашинных комплексов могут быть повышение производительности, обеспечение надежности, создание общих баз данными, эффективное использование дисковой памяти и др.

Составляющие многомашинного комплекса должны быть связаны (скомплексированы) друг с другом для обмена информацией. Технические средства обеспечения комплексирования в ЕС ЭВМ представлены интерфейсом прямого управления (ИПУ), адаптером канал-канал (АКК) и устройствами НМД [1].

Работа посвящена решению проблемы совместного использования общей дисковой памяти несколькими ЕС ЭВМ в среде операционной системы СВМ.

Структура многомашинного комплекса ЕС ЭВМ в ОИЯИ

Наряду с другими типами ЭВМ центральный вычислительный комплекс включает машины серии ЕС, которым отводится значительная роль для решения стоящих перед ОИЯИ задач. В настоящее время серия ЕС ЭВМ представлена в ЦВК ОИЯИ двумя ЭВМ ЕС-1066 и двумя ЕС-1037. Все эти машины подключены к общей дисковой подсистеме типа "винчестер", имеющей четырехканальные переключатели. В качестве операционной системы верхнего уровня в многомашинном комплексе ЕС ЭВМ используется операционная система СВМ [2].

Отличительной особенностью системы СВМ является реализация концепции виртуальных машин, представляющей собой развитие идеи виртуализации отдельных технических средств ЭВМ. Виртуальная машина является логическим эквивалентом реальной ЭВМ, моделируемой при помощи технических средств и программы операционной системы СВМ. На одной реальной ЭВМ может одновременно функционировать несколько виртуальных

машин. Работа каждой виртуальной машины должна управляться своей операционной системой. Информация о каждой виртуальной машине, работающей под управлением СВМ, должна храниться в виде отдельного элемента в специальном файле, называемом оглавлением системы СВМ. Элементы оглавления для каждой виртуальной машины содержат идентификатор машины, пароль доступа, адреса и типы периферийных устройств, классы команд и др.

С целью эффективного использования дисков в СВМ используется концепция мини-дисков [3]. Мини-диск является частью реального диска. Его минимальный размер составляет один цилиндр, максимальный - размер полного пакета. Адресация цилиндров начинается с нуля независимо от места расположения мини-диска на реальном диске. Мини-диски закрепляются за виртуальными машинами (описываются в оглавлении) и трактуются ими как обычные дисковые пакеты, с той лишь разницей, что имеют меньший размер.

Для обеспечения корректного использования мини-дисков в условиях возможной одновременной работы с ними нескольких виртуальных машин в запросе на организацию доступа пользователь может оговорить получение доступа некоторыми условиями. Например: необходим доступ только для чтения информации с мини-диска, но при условии, что ни одна виртуальная машина еще не получила доступ к этому мини-диску для записи; или необходим доступ для записи, если еще ни одна машина не получила никакого доступа к мини-диску.

Благодаря наличию общей дисковой памяти стала возможна работа всех ЭВМ с общим оглавлением. Это означает, что любая виртуальная машина, описанная на одной из ЭВМ, существует и на других ЭВМ. А так как все используемые в системе мини-диски описываются в элементах оглавления виртуальных машин, то и мини-диски могут совместно использоваться с разных машин.

Такая организация программного обеспечения устраняет дублирование данных, а также избавляет от необходимости приведения их в соответствие при использовании на разных ЭВМ. Это касается как системных мини-дисков, так и мини-дисков пользователей. Но при этом возникает проблема корректного

использования мини-дисков при возможной одновременной работе с ними нескольких виртуальных машин. Как подчеркивалось выше, если виртуальные машины функционируют на одной реальной ЭВМ, пользователь может оговорить получение доступа к мини-диску некоторыми условиями. Иначе говоря, корректное использование мини-диска несколькими виртуальными машинами на одной реальной ЭВМ обеспечивается средствами операционной системы СВМ.

По-другому обстоит дело при совместном использовании системных и пользовательских мини-дисков с различных ЭВМ. Операционная система СВМ не обеспечивает корректность такого использования мини-дисков.

Решение проблемы совместного использования общих мини-дисков

Каждая ЭВМ комплекса управляется своей операционной системой СВМ и поэтому использование (чтение или модификация файлов) какого-либо мини-диска одной из ЭВМ не будет "замечено" другими машинами. А это означает, что могут возникать такие ситуации, когда две или более ЭВМ будут параллельно осуществлять доступ к одному и тому же мини-диску. Параллельное использование мини-диска несколькими ЭВМ в режиме чтения не вызовет нарушений в структуре мини-диска. Однако одновременный доступ на запись к одному и тому же мини-диску несколькими ЭВМ приведет к разрушению структуры мини-диска, а следовательно к порче файлов. Поэтому для обеспечения корректного использования общей дисковой памяти несколькими машинами комплекса необходимо исключить параллельный доступ на запись к мини-диску с нескольких ЭВМ. Для решения этой проблемы разработан механизм резервирования и освобождения мини-дисков, используемых в режиме записи с разных машин. Это позволяет последовательно осуществлять доступ на запись к мини-дискам несколькими ЭВМ, использующими общую дисковую подсистему.

Каждой ЭВМ комплекса во время открытия доступа к

определенному мини-диску должна быть известна информация о состоянии этого мини-диска с точки зрения использования его в режиме записи другими ЭВМ. Информацию о состоянии мини-диска удобнее всего хранить на самом мини-диске. Для этой цели используется один байт в метке каждого мини-диска, который будем называть байтом состояния.

Пусть A_i - значение i -го разряда байта состояния мини-диска. Тогда $A_i=1$ означает, что i -тая ЭВМ многомашинного комплекса зарезервировала данный мини-диск; в этом случае другими ЭВМ не может быть открыт доступ на запись к этому мини-диску. При $A_i=0$ данный мини-диск не используется на запись i -той ЭВМ.

Таким образом, возможность совместного использования общих мини-дисков несколькими ЭВМ достигается за счет того, что каждая машина при попытке открытия доступа на запись к определенному мини-диску проверяет (используя байт состояния), не зарезервирован ли он любой другой ЭВМ. Если мини-диск свободен, то данная ЭВМ выставляет соответствующий разряд в байте состояния и открывает к нему доступ на запись. В противном случае запрос на предоставление доступа на запись к мини-диску отвергается и об этом сообщается пользователю. При освобождении мини-диска, открытого для записи, соответствующий разряд байта состояния выставляется в ноль.

Открытие доступа к мини-диску в системе СВМ осуществляет команда ACCESS, а закрытие - команда RELEASE. Поэтому были разработаны преобразователи этих команд, которые выполняют соответственно резервирование и освобождение мини-диска.

Таким образом, предложенная схема позволяет корректно использовать общие мини-диски несколькими ЭВМ многомашинного комплекса. Однако, если произошел сбой той ЭВМ, которая осуществляла доступ на запись к мини-диску (или нескольким мини-дискам), то ни одна из остальных ЭВМ не сможет осуществить доступ на запись к этому мини-диску (мини-дискам). Для обработки такой аварийной ситуации предусмотрена команда FREEDISK, позволяющая оператору по просьбе пользователя освободить все мини-диски определенной виртуальной машины от любой ЭВМ комплекса. Кроме того,

пользователь сам может выполнить команду ACCESS с параметром OVER. Этот параметр позволяет открыть доступ к мини-диску, если он даже заблокирован со стороны другой ЭВМ. Например, если мини-диск 193 зарезервирован машиной JINR66A, то после ввода на любой другой ЕС ЭВМ команды

ACC 193 E

на экране терминала появится сообщение:

MINI-DISK 193 WRITE ACCESS TO JINR66A

Тогда открытие доступа к мини-диску обеспечит команда:

ACC 193 E (OVER)

В качестве идентификатора ЭВМ может быть JINR66A, JINR66B, JINR37A или JINR37B.

При выполнении действий по разблокированию мини-дисков через оператора или командой ACCESS с использованием параметра OVER необходимо убедиться, что мини-диск действительно не используется на той ЭВМ, которая указана в сообщении. Иначе содержимое мини-диска может быть разрушено.

Необходимо отметить, что окончание сеанса работы виртуальных машин производится по команде LOGOFF. Выполнение этой команды вызывает неявное исполнение команды RELEASE для всех использовавшихся в сеансе мини-дисков, что, в свою очередь, обеспечивает их разблокирование. Однако иногда возникает необходимость выхода из сеанса с помощью команд CP LOGOFF и #CP LOGOFF. В этом случае не происходит выполнения команды RELEASE. При таком варианте завершения сеанса реальное разблокирование мини-дисков осуществляется в виртуальной машине оператора с помощью программы DISKOFF.

Факт завершения сеанса работы какой-либо виртуальной машины по командам CP LOGOFF или #CP LOGOFF передается машине оператора с помощью системного средства СВМ "Программируемый оператор" [4].

В заключение отметим, что рассмотренный в работе программный продукт успешно используется в многомащинном комплексе ЕС ЭВМ в ЦВК ОИЯИ. С его внедрением число разрушений структур мини-дисков значительно уменьшилось.

Литература

1. Королев Л.Н. Структуры ЭВМ и их математическое обеспечение. Второе издание. М., Наука, Главная редакция физ.-мат. литературы, 1978.-351.
2. Ларионов К.А., Наумов В.В. Функциональная структура операционной системы ОС 7 ЕС // ВТ соц. стран, М., финансы и статистика, 1989, Вып. 25 - с. 56-61.
3. Система виртуальных машин для ЕС ЭВМ. Под редакцией Ковалевича Э.В., М., финансы и статистика, 1985.-360.
4. Virtual Machine/Extended Architecture. System Product. CP Programming Services, Release 2, SC23-0370-1, 1988.-499.

Рукопись поступила в издательский отдел
4 марта 1992 года.