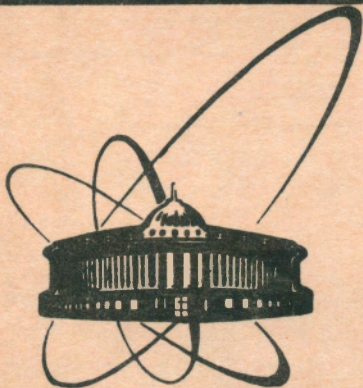


92-313



**СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА**

P11-92-313

В.В.Кореньков, Е.И.Ли, Е.А.Тихоненко

**INFSIO - СРЕДСТВО ДЛЯ ИССЛЕДОВАНИЯ
ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ
ДИСКОВОЙ ПОДСИСТЕМЫ
МНОГОМАШИННОГО КОМПЛЕКСА ЕС ЭВМ**

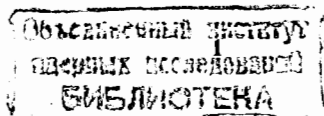
1992

Введение

Вопросы повышения эффективности функционирования как любой операционной системы, так и операционной системы СВМ (системы виртуальных машин)/1/ относятся к области весьма актуальных, но недостаточно изученных. Немаловажным в плане повышения эффективности функционирования операционной системы СВМ является рациональное использование дисковой памяти с точки зрения оптимального размещения системных областей (ядра системы, областей страничного обмена, накопления и временных наборов данных), системных сервисных средств, а также наборов данных пользователей.

На Центральном вычислительном комплексе (ЦВК) Объединенного института ядерных исследований (ОИЯИ) в Лаборатории вычислительной техники и автоматизации (ЛВТА) работает многомашинный комплекс ЕС ЭВМ (две ЭВМ ЕС-1037 и две ЭВМ ЕС-1066) на базе общей дисковой памяти 15.2 Гб на магнитных дисках MEMOREX (тип 3350, емкость 317 Мб каждый)/2/. Общее число зарегистрированных пользователей 765 человек, одновременно работают 20-80 человек (редактирование, счет в интерактивном или пакетном режиме). На всех ЭВМ комплекса работает единая операционная система СВМ.

В условиях работы многомашинного комплекса на базе общей дисковой памяти особенно остро встает проблема тщательного изучения нагрузок на отдельные тома (пакеты) дисковой подсистемы с целью устранения возможных перегрузочных ситуаций, связанных с нерациональным размещением дискового пространства.



Время обработки запросов ввода-вывода при обращении пользователей к накопителям на магнитных дисках T может быть представлено в виде суммы нескольких интервалов времени:

$$T = \sum_{i=1}^6 t_i, \quad (1)$$

где t_1 - время передачи управления от задачи пользователя к супервизору ввода-вывода,

t_2 - время нахождения в очереди на дисковый обмен,

t_3 - время перемещения магнитных головок на диске,

t_4 - время обмена с дисковым пакетом,

t_5 - время обработки прерывания ввода-вывода после завершения обмена,

t_6 - время передачи управления задаче пользователя.

При неудачном размещении наборов данных возможно появление больших очередей к отдельным пакетам, а также высокие интервалы времени перемещения магнитных головок на пакете. Следовательно, оптимизация размещения состоит в стремлении исключить возможности возникновения очередей и сократить время перемещения головок на пакете. Т.е. следует сокращать интервалы времени t_2 и t_3 . Это достигается, в частности, размещением часто используемых наборов данных на различных дисковых пакетах, и, если это невозможно, на максимально близком расстоянии друг от друга на одном и том же дисковом пакете.

1. Средство OMEGAMON

Существуют стандартные средства, такие как VPAC, DMON, OMEGAMON /з/, позволяющие оценить состояние вычислительной системы. С помощью многофункционального средства OMEGAMON возможно получать разнообразную информацию об использовании центрального процессора, состоянии системных ресурсов, количестве обращений к магнитным лентам и дискам, использовании страничной области, структуре основной памяти, работе виртуальных машин пользователей. В частности, OMEGAMON дает общую суммарную картину нагрузок на дисковые пакеты (имеется в виду количество обменов SIO с каждым отдельным дисковым пакетом) в любой данный момент времени.

Средство OMEGAMON ориентировано на работу на одной реальной ЭВМ. С целью использования его возможностей в рамках многомашинного комплекса был сделан ряд модификаций, расширяющих возможности этого средства.

В конкретной реализации на ЦВК ОИЯИ средство OMEGAMON настроено таким образом, что при вхождении в одноименную виртуальную машину OMEGAMON на любой реальной ЭВМ комплекса создаются файлы, в которых содержится информация о количестве обращений ко всем пакетам дисковой подсистемы на данной реальной ЭВМ с момента ее очередной загрузки.

Стандартная информация дополняется датой последней загрузки операционной системы, что придает этой информации полную наглядность, давая представление, за какой интервал времени эти обращения к дисковым пакетам были совершены. Создана также процедура, позволяющая получать сводную таблицу нагрузок на дисковую подсистему для всего комплекса ЕС ЭВМ в целом.

2. Средство INFSIO

К сожалению, обладая определенной информативностью, картина распределения обменов по всем дисковым пакетам не дает представления о нагрузках внутри каждого отдельного дискового пакета. Чтобы оценить эту нагрузку внутри каждого дискового пакета, было создано средство INFSIO. Работая на любой реальной ЭВМ, виртуальная машина INFSIO дает возможность получить картину распределения обменов на любом заданном пакете по цилиндрам и, далее, по минидискам или системным областям данного пакета. Информация о цилиндре, на котором в данный момент находятся магнитные головки, и о числе обменов со всем пакетом на данный момент извлекается из соответствующих разделов системного управляющего блока RDEVBL0K для данного пакета. Опрос производится через каждые Δt секунд. Заметим, что мы выбрали Δt равным 1 секунде, хотя для получения более точной картины нагрузок на дисковые пакеты следовало бы значительно уменьшить периодичность сбора информации, однако уменьшение этого периода сбора приводит к существенному увеличению затрат времени центрального процессора, что нарушает нормальную работу пользователей.

После окончания работы средства INFSIO в непрерывном цикле сбора информации образуются файлы, имеющие вид таблиц, в которых указаны интервал времени сбора информации, число обменов с данным пакетом за этот интервал, а для каждого цилиндра указано, сколько раз за это время данный цилиндр был "актуальным" цилиндром пакета (т.е. над этим цилиндром располагались магнитные головки).

Имея эту информацию, можно получить файл, содержащий сведения об усреднённом числе обменов для каждого минидиска или системной области. Понятно, что расположение магнитных головок над тем или иным цилиндром не свидетельствует о том, идет ли в этот момент обмен с пакетом. Поэтому необходимо оценить реальное количество обменов за период сбора информации. Предполагая, что на больших выборках усредненное число обменов для всех цилиндров пакета можно рассматривать как математическое ожидание с равными весовыми коэффициентами, усреднённое число обменов с любой группой цилиндров с номерами от i_1 до i_2 данного дискового пакета можно определить как

$$K_{SIO}^{(i_1-i_2)} = N_{SIO} * \frac{\Delta t * \sum_{i=i_1}^{i_2} n_i}{t}, \quad (2)$$

где t - время сбора информации,
 Δt - периодичность опроса состояния пакета,
 N_{SIO} - реальное число обменов с данным пакетом за время t ,
 n_i - число раз, когда i -й цилиндр был "актуальным".

Получив на нескольких (или всех четырех) реальных ЭВМ комплекса таблицы распределения обменов по минидискам или системным областям для одного и того же пакета, можно получить сводную таблицу с информацией о суммарном усредненном числе обменов для данного пакета по всем ЭВМ комплекса.

Программы средства INFSIO реализованы на языке ASSEMBLER /4/ и языке процедур REXX /5/.

3. Оценка нагрузок на дисковую подсистему многомашинного комплекса с помощью средств OMEGAMON и INFSIO

Совместное последовательное использование описанных в п.п.1 и 2 средств OMEGAMON и INFSIO позволяет производить оценку нагрузок на дисковую подсистему в целом и на отдельные тома дисковой подсистемы. Таблицы, получаемые посредством возможностей OMEGAMON на всех реальных ЭВМ комплекса, дают наглядное представление о нагрузочной ситуации на дисковую подсистему с момента очередной загрузки операционной системы. Типичный пример приведен в табл. I, из которой видно, что на ЭВМ 1066A произведено наибольшее число обменов с собственными томами системы VMSRES, VMSEXP и VMPACK, а также с томами пользователей JINR01, JINR02, JINR11 и JINR16 (DISK- имя пакета(тома), DADR-реальный адрес пакета, SIO- число обменов).

Таблица 1.

Э В М JINR66A ЗАГРУЖЕНА 12/05/92 16:48:00							
DISK	JINR12	JINR13	JINR14	JINR15	COPCOP	JINR17	VMSP57
DADR	100	101	102	103	104	105	110
SIO	102	50799	24	594	2	901	13785
DISK	JINR19	HELPME	COP103	VMSYS1	COP111	BAZA90	COP113
DADR	112	113	114	115	121	122	123
SIO	4799	2	2	25719	2	1	2
DISK	COP115	COP100	COP101	COP102	OSUSE1	LIB619	VMSEXP
DADR	125	130	131	132	133	134	140
SIO	2	2	2	2	90	2	73945
DISK	JINR01	JINR02	JINR03	JINR04	JINR05	JINR06	JINR07
DADR	141	142	143	144	145	150	151
SIO	82930	95916	7226	13534	2	6190	548
DISK	JINR09	JINR10	JINR11	COP150	COP151	COP152	COP153
DADR	153	154	155	160	161	162	163
SIO	38164	74	156436	2	2	2	2
DISK	JINR16	COP140	VMSRES	VMPACK			
DADR	165	170	8C0	8C2			
SIO	85260	2	52651	78047			

Далее, исходя из полученной информации, выбирается пакет, нагрузку на который следует оценить более детально, после чего запускаются программы средства INFSIO.

Анализ получаемой картины распределения обменов внутри конкретного пакета позволяет делать выводы о рациональности существующего распределения дискового пространства. Так, например, в табл. 2 представлена типичная картина обменов с дисковым пакетом VMSEXP (VMSEXP – собственный том операционной системы, на котором содержатся многие системные средства, в том числе системы CMS и PTS, различные библиотеки массового использования и т.п.).

Таблица 2.

 TOMNAME = VMSEXP

СВОДНАЯ ТАБЛИЦА РАСПРЕДЕЛЕНИЯ ОБМЕНОВ ПО ВСЕМ ЭВМ КОМПЛЕКСА
 ЗА ПЕРИОД С 13/05/92 12:31:24 по 13/05/92 15:50:32

КОЛИЧЕСТВО ОБМЕНОВ SIO ЗА ЭТОТ ПЕРИОД = 23817

Имя минидиска	Номер минидиска	Начальный цилиндр	Конечный цилиндр	Усредненное кол-во обменов
⇔ALLOC⇔	124	0	0	0
⇔DIRECT⇔	102	1	2	1
	⋮			
CBM	19E	54	93	513
MAINT	325	94	180	1226
VPAC	191	181	190	0
	⋮			
TCPMAINT	608	236	240	0
MAINT	310	241	310	9259
⇔SAVSYS⇔	105	311	330	215
	⋮			
DIRMON	193	421	424	74
MAINT	500	425	429	8201
MAINT	19F	430	435	1
ISPVM	191	436	445	0
	⋮			
CSYS009	192	473	476	0
MAINT	190	477	516	2337
MAINT	19E	517	526	2702
MAINT	19D	527	553	223
DIRADMIN	191	554	554	17

Как видно из табл.2, наиболее часто идут обращения к

минидискам с номерами 325,310,190 и 19E виртуальной системной машины MAINT, которые расположены соответственно с 94 по 180, с 241 по 310, с 477 по 516 и с 517 по 526 цилиндры. Вполне очевидно, что данное расположение минидисков не может быть признано оптимальным: обращения идут как к начальным, так и к конечным цилиндрам пакета. Следовательно, необходимо произвести перераспределение пространства дискового пакета VMSEXP.

Заключение

Предлагаемая методика учета и анализа нагрузочной ситуации на дисковую подсистему с помощью средств OMEGAMON и INFSIO дает возможность прийти к наиболее рациональному распределению дискового пространства, что существенно повышает эффективность эксплуатации операционной системы CBM на многомашинном комплексе ЦВК ОИЯИ.

ЛИТЕРАТУРА

1. Булко И.М. и др. Система виртуальных машин для ЕС ЭВМ. М.: Финансы и статистика, 1985.
2. Бавижев А.Д. и др. Совместное использование общей дисковой подсистемы в многомашинном комплексе ЕС ЭВМ. ОИЯИ, P11-92-94, Дубна, 1992.
3. Система виртуальных машин. Руководство системного программиста. Средства анализа производительности. E1.00005-03 32 01-7, 1987.
4. Система виртуальных машин. Подсистема диалоговой обработки. Описание макрокоманд. E1.00005-03 92 01, 1987.
5. Пекер Ф.Л., Морозов Б.А. Прикладное программирование в системе виртуальных машин ЕС ЭВМ. Минск: Высшая школа, 1989.

Рукопись поступила в издательский отдел
 17 июля 1992 года.

Кореньков В.В., Ли Е.И., Тихоненко Е.А. P11-92-313
INFSIO - средство для исследования
эффективности использования дисковой подсистемы
многомашинного комплекса ЕС ЭВМ

Работа посвящена актуальной проблеме повышения эффективности эксплуатации операционной системы СВМ на многомашинном комплексе ЕС ЭВМ. Предлагается методика оценки и анализа нагрузок на дисковую подсистему с помощью разработанного средства INFSIO.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1992

Перевод авторов

Koren'kov V.V., Li E.I., Tickonenko E.A. P11-92-313
INFSIO - Software Package for
Investigation of Shared Disk Usage Efficiency at
ES-Multicomputing System

The paper is devoted to the actual problem of VM operating system usage efficiency improvement at ES-multicomputing system. The method of accounting and analysis of Shared Disk Storage usage with the help of developed package INFSIO is proposed.

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

Communication of the Joint Institute for Nuclear Research. Dubna 1992