

УДК 004.052.32

ПРОГРАММНАЯ СИСТЕМА ДИАГНОСТИКИ МНОГОПРОЦЕССОРНЫХ УЗЛОВ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

А. Б. Киселёв
(ФГУП "РФЯЦ-ВНИИЭФ", г. Саров Нижегородской области)

Описана программная система автоматизированной диагностики (проверки) неисправных многопроцессорных узлов вычислительной системы. Определены источники аппаратных и программных сбоев, указаны применяемые в системе методы восстановления работоспособности вычислительных узлов, находящихся под управлением системы пакетной обработки заданий JAM.

Ключевые слова: система пакетной обработки заданий, восстановление работоспособности многопроцессорного узла вычислительной системы.

Введение

С увеличением числа аппаратных и программных компонентов вычислительной системы снижается показатель ее надежности. В процессе эксплуатации аппаратные компоненты вычислительной системы проходят этапы приработки, нормального функционирования и старения. На этапах приработки и старения аппаратуры на первый план выходят вопросы оперативной локализации сбоев/отказов, уменьшения времени восстановления работоспособности¹ оборудования вычислительной системы и, в частности, вычислительных узлов (ВУ).

Поиск неисправного оборудования в ВУ — достаточно трудоемкая работа, выполняемая техническим персоналом. Чтобы выявить источник неисправности, техническому специалисту надо проанализировать информацию, предоставляемую средствами мониторинга аппаратуры вычислительной системы, находящуюся в системных журнальных файлах, запустить тестовые программы, принять решение о замене или перемещении аппаратных компонентов ВУ, записать информацию о ремонтных работах в электронный журнал эксплуатации вычислительной

системы и т. д. Восстановление работоспособности нескольких десятков ВУ может затянуться на часы.

Для помощи техническому персоналу и повышения эксплуатационных показателей вычислительной системы в рамках работ по расширению функциональных возможностей системы пакетной обработки заданий JAM (СПО JAM) [1] была создана автоматизированная программная система диагностики узлов (СДУ) [2]. СДУ стремится восстановить рабочее состояние неисправных ВУ², выявленных подсистемой мониторинга аппаратно-программных компонентов ВУ СПО JAM. В настоящий момент СДУ функционирует на всех вычислительных системах РФЯЦ-ВНИИЭФ, которые находятся под управлением СПО JAM.

СДУ была разработана с Б. П. Киселёвым (автор идеи) в 2010 году.

1. Источники неисправностей вычислительного узла

В табл. 1 перечислены основные источники неисправностей ВУ

Неисправности ВУ могут возникнуть как в процессе выполнения программы пользователя,

¹Работоспособность — состояние объекта, при котором значения всех параметров, характеризующих способность выполнять заданные функции, соответствуют требованиям нормативно-технической и конструкторской документации.

²Под неисправностью ВУ следует понимать состояние какого-либо его аппаратно-программного компонента, которое не позволяет программе пользователя запуститься или успешно выполняться.

Таблица 1

Источники неисправностей ВУ

Источники аппаратных неисправностей	Источники программных неисправностей
Оперативная память	Ядро операционной системы
Блоки питания	Системные программы
Адаптеры	Пользовательские программы
Процессор	
Материнская плата	
Разъемы	
Вентиляторы	

так и после перезагрузки ВУ. Например, после перезагрузки проявлениями неисправностей могут стать: невключение ВУ из-за отказа блока питания или разъема электропитания (контактная ошибка); отсутствие видеосигнала (отказ видеоадаптера); отсутствие работоспособности сетевого адаптера; потеря контакта между соединителем сетевого кабеля и гнездом; нераспознавание части оперативной памяти и т. д.

Очевидно, что дефект материнской платы, разъемов или кабелей связи/электропитания, изменение физических свойств термопасты может диагностировать и исправить только технический специалист. Тем не менее в процессе работы ВУ можно программным образом определить неисправность, если она связана с его оперативной памятью, процессором, сбоем сетевого адаптера, с системными или пользовательскими процессами.

2. Подсистема мониторинга аппаратно-программных компонентов ВУ СПО ЖАМ

Подсистема мониторинга аппаратно-программных компонентов ВУ СПО ЖАМ обеспечивает автоматическую локализацию неисправных ВУ, а также ВУ, которые могут снижать или уже снижают эффективность выполнения программ пользователей. Главное отличие подсистемы мониторинга СПО ЖАМ от других систем мониторинга оборудования вычислительной системы — это доставка актуальной информации о состоянии оборудования ВУ планировщику заданий СПО ЖАМ.

Принцип работы этой подсистемы достаточно прост. На каждом узле периодически выполняется сценарий интерпретатора shell (сценарий

диагностики), который контролирует программные компоненты ВУ, а также сравнивает текущие метрики оборудования с эталонными значениями. С помощью сценария диагностики подсистема мониторинга контролирует оперативную память, сетевые адаптеры, файловые системы, а также системные и пользовательские процессы.

Если сценарий диагностики выявил неисправность, то ВУ в терминах СПО ЖАМ помечается как *неисправный*. Планировщик заданий СПО ЖАМ игнорирует такой ВУ до момента исчезновения неисправности, например завершения "зависшего" пользовательского процесса или восстановления дееспособности клиентского программного обеспечения параллельной файловой системы. Кроме того, сценарий диагностики вызывается сразу после завершения задания, поэтому подсистема мониторинга не позволяет следующему заданию использовать ВУ, если у него обнаружилась неисправность.

Если подсистема мониторинга не может выполнить сценарий диагностики из-за "зависания" ВУ, то СПО ЖАМ присваивает узлу статус *состояние неизвестно*. Это состояние также не позволяет планировщику заданий СПО ЖАМ распределять на ВУ задания пользователей.

3. Задачи СДУ

До реализации СДУ неисправные ВУ оставались недоступными для пользовательских заданий в ночное время, в выходные и праздничные дни, увеличивая показатель *время ремонта* на сотни часов. При этом опыт эксплуатации вычислительных систем РФЯЦ-ВНИИЭФ свидетельствует, что большая часть проблем на ВУ вызвана программными сбоями, которые исправляются тривиальной перезагрузкой.

Таким образом, цель, которую преследует СДУ, — это уменьшение времени восстановления рабочего состояния ВУ, повышение эксплуатационных показателей вычислительной системы посредством решения следующих задач:

- выполнение типовых операций по восстановлению работоспособности ВУ вместо технических специалистов в случаях, когда неисправность ВУ не требует выполнения каких-либо операций с оборудованием;
- оповещение технического персонала о том, что ВУ требует ремонта или детальной диагностики.

Кроме того, СДУ формирует историю ремонтов ВУ, записывая информацию в базу данных СПО JAM.

4. Алгоритм СДУ

С помощью утилиты командного интерфейса СПО JAM СДУ циклически осуществляет выборку названий ВУ со статусами *неисправный* и *состояние неизвестно*. Каждому статусу соответствует отдельный набор тестовых операций.

Например, если неисправный ВУ имеет статус *состояние неизвестно*, то СДУ в первую очередь выполняет поиск информации об аппаратных событиях, которые свидетельствуют о том, что узел необходимо ремонтировать. Если тестирование можно продолжить, то СДУ выполняет перезагрузку ВУ³. Таким способом проверяется функционирование аппаратных компонентов: блоков питания, видео- и сетевых адаптеров, материнской платы и разъемов.

Также устраняются неисправности, вызванные всеми перечисленными выше программными источниками (см. табл. 1). С помощью тестовой программы замеряется пропускная способность оперативной памяти ВУ. Посредством программы Linpack моделируется рабочая нагрузка на ВУ: проверяются оперативная память, блоки питания, процессор и материнская плата. В то же время проверка микропроцессора с помощью программы Linpack осуществляется формально, поскольку при современном техническом уровне отказ микропроцессора случается крайне редко. С учетом этого параметры программы Linpack подобраны так, чтобы она выполнялась не более десяти минут.

Результаты прохождения всех этапов проверки анализируются. Если программа Linpack завершилась аварийно, или результат не совпал с ожидаемым, или тест "завис", то СДУ фиксирует это в электронном журнале эксплуатации вычислительной системы и завершает проверку ВУ. Если все этапы тестирования прошли успешно, то с помощью командного интерфейса на ВУ запускается программный демон СПО JAM, который выполняет сценарий диагностики (см. разд. 2). После успешного завершения данного сценария ВУ становится доступным для заданий.

³Перезагрузка ВУ выполняется с помощью BMC (Board Management Controller).

Если ВУ имеет статус *неисправный* и, например, источником неисправности является системная программа или пользовательские процессы, то СДУ перезагружает ВУ и запускает программный демон СПО JAM, который после успешного завершения сценария диагностики делает данный узел доступным для заданий. Если же неисправность устранить не удалось, то в электронный журнал эксплуатации вычислительной системы записывается сообщение, что узел должен ремонтировать технический специалист.

Описанные методы тестирования ВУ имеют исключения. Например, не проверяется ВУ с зависшим клиентским программным обеспечением параллельной файловой системы, если сетевые адаптеры ВУ работают с меньшей скоростью или нет связи, если количество распознанной оперативной памяти меньше, чем установлено. Перезагружать такой ВУ бесполезно, так как восстановить дееспособность параллельной файловой системы, устранить контактную ошибку разъема оперативной памяти или сетевого кабеля может только технический специалист.

Кроме того, если аппаратура ВУ является источником таких аппаратных событий, как *корректируемые/некорректируемые ошибки оперативной памяти, превышение температурных пределов, ошибки аппаратуры, обнаруженные ядром ОС*, то ВУ не тестируется, а с помощью командного интерфейса СПО JAM резервируется для ремонта: ВУ присваивается статус *ожидает ремонта*, который запрещает планировщику заданий СПО JAM распределять на него задания. Технические специалисты ремонтируют ВУ после завершения задания, которое его использует. Статус *ожидает ремонта* сбрасывается с помощью утилиты командного интерфейса СПО JAM вручную после окончания ремонта ВУ.

Набор тестовых операций, выполняемых с неисправными ВУ, можно легко изменить с помощью конфигурационного файла СДУ. Перезагружать систему не требуется, так как СДУ перечитывает свой конфигурационный файл автоматически в начале каждого цикла проверки.

5. Результаты эксплуатации СДУ

Положительный результат использования СДУ проявился в улучшении эксплуатационных показателей одной из вычислительных систем

РФЯЦ-ВНИИЭФ (далее ВС-1). В табл. 2 приведены результаты эксплуатации ВС-1 в 2013 (без СДУ) и 2014 (с СДУ) годах, взятые из статистического отчета, полученного с помощью применяемой в РФЯЦ-ВНИИЭФ программы сбора и обработки информации о счете задач и работе ЭВМ.

Из таблицы видно, что коэффициент работоспособности ВС-1 после внедрения СДУ повысился на 14,72 %, потери из-за ремонта уменьшились на 2,1 %. Кроме того, данные табл. 3 свидетельствуют, что непроизводительные потери, связанные с ремонтом узлов, также уменьшились в 4,7 раза. Это можно объяснить тем, что в 2013 году в число отбракованных ВУ попадали и узлы с программными сбоями, которые можно было легко устранить с помощью перезагрузки. После внедрения СДУ эксплуатационные показатели ВС-1 стали формироваться только на основе тех ВУ, которые действительно требовали ремонта оборудования.

Результаты, отраженные в табл. 3, получены экстраполяцией эксплуатационных показателей работы СДУ ВС-1 на вычислительную систему,

Таблица 2

Эксплуатационные параметры ВС-1 без СДУ (2013 год) и с СДУ (2014 год)

Показатель	Без СДУ, %	С СДУ, %
Коэффициент работоспособности ⁴	82,54	97,26
Коэффициент потерь из-за ремонта ⁵	2,93	0,83

Таблица 3

Экстраполированные эксплуатационные показатели ВС-1

Год	Количество отказов ВУ	Время ремонта ВУ, ч
2013	560	5 684
2014	119	392

⁴Коэффициент работоспособности — это календарное время работы вычислительной системы за вычетом потерь времени на ремонт систем жизнеобеспечения, затрат времени на выполнение профилактических работ с оборудованием и его ремонт, деленное на календарное время работы вычислительной системы.

⁵Коэффициент потерь из-за ремонта — это затраты времени на ремонт оборудования, деленные на календарное время работы вычислительной системы.

содержащую 1 000 ВУ⁶, с учетом того, что число ее сбоев и отказов пропорционально числу ВУ⁷. Из табл. 3 следует, что на такой вычислительной системе СДУ уменьшит общее время ремонта ВУ на 5 292 часа. Разделив это число на количество часов в году, получим эквивалентное количество узлов, которые не были доступны для заданий пользователей в течение года, т. е. 1 ВУ. Кроме того, если за норму времени, в течение которого технический специалист устраняет простую неисправность ВУ, условно принять 5 минут, то с учетом разницы в количестве отказов ВУ за год СДУ экономит техническому персоналу не менее 36 часов рабочего времени. Конечно, на конкретной вычислительной системе данный показатель будет зависеть от стадии эксплуатации оборудования (приработка, нормальная эксплуатация или старение) и от запускаемых на ней прикладных программ (наличия в них ошибок, их корректной настройки на аппаратно-программную инфраструктуру и т. д.).

Еще одна важная характеристика СДУ — это круглосуточная работа. На разных вычислительных системах показатель срабатывания СДУ в интервале времени от 17 до 8 часов находится в пределах 20–69 %.

Заключение

СДУ является важным программным компонентом, способствующим повышению готовности, эксплуатационной надежности, пропускной способности вычислительной системы и уменьшению затрат времени на ее обслуживание техническими специалистами.

Такие характеристики СДУ, как круглосуточная работа, оперативное взаимодействие с техническим персоналом, являются основой эффективной работы системы. Так, в результате внедрения СДУ в 4,7 раза уменьшились непроизводительные потери машинного времени, связанные с ремонтом ВУ. Работоспособность большинства из них СДУ восстановила без ремонта, и, как следствие, значительно повысился коэффициент работоспособности вычислительной системы.

⁶Производительность 1 000 современных ВУ примерно равна 1 Пфлопс.

⁷Пропорциональность имеет место в случае использования топологий *nD-решетка*, *nD-тор* (при фиксированной размерности $n = 2, 3, \dots$), *Fat Tree* (при фиксированном числе уровней 2, 3, 4).

Новизна технического решения СДУ по восстановлению работоспособности ВУ в составе вычислительной системы под управлением СПО JAM подтверждена свидетельством ФИПС [2].

Список литературы

1. Киселёв А. Б., Киселёв С. Н. Система пакетной обработки заданий JAM // Вопросы атомной науки и техники. Сер. Математическое моделирование физических процессов. 2009. Вып. 4. С. 60–66.

2. Богданов А. А., Киселёв А. Б. Программный комплекс контроля аппаратных метрик и тестирования вычислительных узлов кластера (КАМиТ). Свидетельство о государственной регистрации программы для ЭВМ № 20116182 // Бюллетень ФИПС. 2012. № 78(1). С. 38.

Статья поступила в редакцию 13.11.17.

SOFTWARE SYSTEM FOR DIAGNOSING MULTIPROCESSOR NODES OF A COMPUTING SYSTEM / A. B. Kiselev (FSUE RFNC-VNIIEF, Sarov, N. Novgorod region).

The paper describes the software system for the automatized diagnosing (checking) of faulty multiprocessor nodes in a computing system. Sources of hardware and software failures have been identified, the methods used in the system to restore serviceability of computing nodes controlled by the JAM batched job processing system are described.

Keywords: batched job processing system, restoration of serviceability of a computing system's multiprocessor node.
