

А. А. Московский, А. Ю. Первин, М. В. Стоцкий

## Прогнозирование аппаратных сбоев вычислительных кластеров с помощью систем мониторинга

Аннотация. В работе описан прототип системы прогнозирования аппаратных сбоев. Предполагается, что об отказе узла кластерной установки можно узнать заранее на основе показаний сенсоров [4]: температура процессора, скорость вращения вентилятора, напряжение. Описывается развернутая система сбора показаний сенсоров узла кластерной установки.

*Ключевые слова и фразы:* СПОКУ — система прогнозирования отказов кластерных установок, диагностика аппаратных сбоев, отказоустойчивость высокопроизводительных кластеров, мониторинг компонент кластера.

### 1. Введение

Основная тенденция в развитии высокопроизводительных параллельных компьютеров состоит в том, что среднее число вычислительных узлов, образующих кластер, растет очень быстро [3]. Так, например, лидер списка Top500 по итогам 2005 года, кластер BlueGene/L, состоит из 131072 вычислительных узлов. Кластеры сконструированы для того, чтобы быть эффективным и надежным инструментом для решения сложных научных и инженерных задач.

Однако, с ростом размера кластера возрастает и вероятность выхода из строя компонентов кластера. Для высокопроизводительного компьютера, включающего в себя тысячи процессоров, жестких дисков, модулей памяти и вентиляторов, вероятность отказа возрастает в десятки тысячи раз по сравнению с однопроцессорной установкой. В связи с этим особое значение приобретает задача обеспечения надежности вычислительных узлов кластера. Несмотря на выдающиеся достижения в микроэлектронике, отказы оборудования неизбежны даже в установке со средним количеством вычислительных узлов. Последствия таких отказов могут привести к ошибкам вычислений, потере данных и сбоям.

Существует несколько подходов, с помощью которых можно достигнуть необходимого уровня отказоустойчивости. Один из способов состоит в создании избыточных аппаратных средств с поддержкой «горячего» подключения. Например, существующая и апробированная технология RAID (redundant array of independent/inexpensive disks) [2] является эффективным решением для повышения надежности хранения информации. Но RAID-технологии дороги и, кроме того, могут снизить общую производительность системы из-за дополнительных требований к программному и аппаратному обеспечению.

Другой способ увеличить уровень отказоустойчивости — прогнозирование аппаратных сбоев. Это перспективная концепция, позволяющая снизить вероятность возникновения аппаратных сбоев, одновременно снизив эксплуатационные затраты. Замысел таких систем прогнозирования состоит в том, чтобы предсказать отказ оборудования прежде, чем он может произойти. Исследования, проводимые за рубежом, ясно показывают, что системы прогнозирования действительно способны защитить аппаратное обеспечение от потенциальных поломок.

Современные системы мониторинга, взаимодействуя с аппаратными датчиками (сенсорами [4]), предоставляют широкие возможности для сбора различной информации о состоянии «здоровья» компонент кластера. Анализ таких статистических данных может показать, через какое время тот или иной компонент кластера выйдет из строя. Так, например, аппаратному сбою конденсатора на материнской плате, будут предшествовать заметные броски напряжения на процессоре, что в свою очередь позволяет оценить оставшееся время наработки на отказ и принять упреждающие меры по замене вышедшей из строя детали.

Наше исследование направлено на разработку интеллектуальной Системы Прогнозирования Отказов Кластерных Установок (сокращенно СПОКУ), взаимодействующей с различными компонентами вычислительных кластеров (сенсоры температуры процессора, датчики напряжения на материнской плате, статистика использования сетевой карты и другие). Мы считаем, что анализируя динамику развития «здоровья» кластера, наше программное обеспечение будет способно прогнозировать аппаратные сбои, предоставлять информацию о состоянии кластера.

## 2. Требования к реализации СПОКУ

Анализ существующих систем прогнозирования аппаратных сбоев позволяет сформулировать набор необходимых требований к подобным решениям:

- Простота в настройке и установке.  
Программное обеспечение не должно быть дополнительной проблемой для системного администратора. Работа всех компонент системы должна быть абсолютно прозрачной.
- Своевременный прогноз.  
Ключевое свойство системы прогнозирования — своевременный прогноз аппаратных сбоев. Например, существующая технология S.M.A.R.T. [5] для жестких дисков сообщает о вероятной поломке за 48 часов до того, как отказ фактически происходит. Однако, даже такой период может показаться недостаточным с точки зрения пользователя.
- Рациональное использование ресурсов системы.  
«Агенты» СПОКУ не должны потреблять слишком много времени центрального процессора и других ресурсов, чтобы не помешать работе пользователей кластера, подписанного на услугу СПОКУ. Именно поэтому СПОКУ проектируется таким образом, чтобы решать ресурсоемкие задачи анализа на центральном сервере, а не на стороне клиента.
- Использование безопасных каналов связи и надежных баз данных.  
Для сокрытия конфиденциальной информации о состоянии системы подписчика необходимо передавать данные по каналам связи, использующим эффективные алгоритмы шифрования.

## 3. Архитектура СПОКУ

Разрабатываемая система прогнозирования сбоев построена на принципе веб-сервисов. Для того чтобы подписаться на услугу прогнозирования сбоев, необходимо установить пакет программ «Агент» на целевой кластер и сообщить организаторам СПОКУ учетную информацию о кластере. Концептуально, работа СПОКУ описывается следующими пунктами:

- Первый шаг — извлечение первичной информации. На кластере, подписанном на услугу прогнозирования сбоев, запускается специальная программа «Сборщик», входящая в состав «Агента», которая отвечает за получение информации от сенсоров. Сбор данных производится периодически через равные интервалы времени, которые при необходимости можно изменять. Кроме того «Сборщик» имеет возможность делать серию замеров, для того чтобы обнаружить случайные отклонения измеряемых величин от их средних значений.
- На следующем шаге накопленная статистическая информация от сенсоров пересылается во временное хранилище на головном вычислительном узле кластера. При этом подразумевается, что головной вычислительный узел имеет выход в Интернет. Далее данные в сжатом виде передаются в центральное хранилище для последующей обработки.
- Третий шаг — хранение и анализ. На этом шаге накопленная информация аккумулируется в общее хранилище. Данные о подписанных кластерах хранятся в различных документах единой базы данных. Во избежание нежелательных потерь данных регулярно производится резервное копирование базы данных с последующим архивированием. Таким образом в центральном хранилище доступна информация о состоянии кластера за весь период подписки.

В нашей реализации используется системный планировщик задач для того, чтобы периодически запускать «Сборщик» на вычислительных узлах кластера и передавать информацию от сенсоров в центральное хранилище. Текущий прототип СПОКУ способен получать от сенсоров информацию о температуре процессора, напряжении, скорости вентиляторов и количестве ошибок на сетевых интерфейсах.

Во время установки статическая информация, с описанием вычислительных узлов, сохраняется в базе данных. Эта информация содержит описания моделей процессоров, их тактовых частот, моделей материнских плат. База данных организована с использованием XML СУБД Sedna [7]. Статическая информация может быть использована при анализе данных для составления более точных прогнозов.

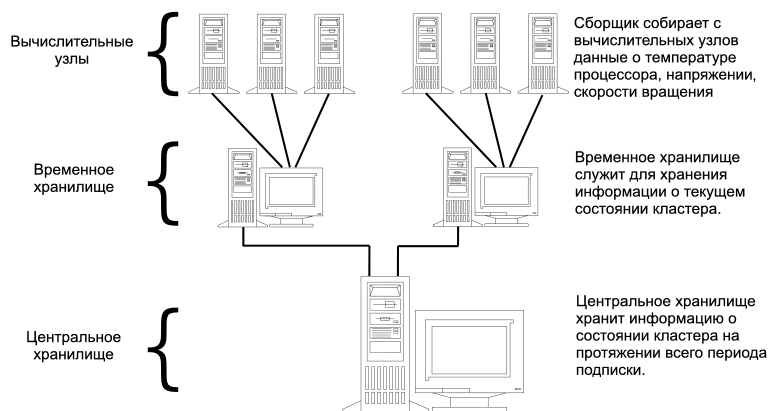


Рис. 1. Диаграмма транспорта данных

Выше перечисленная информация о «здоровье» кластера передается и хранится в формате XML [6]. Несмотря на очевидную избыточность XML для хранения сенсорных данных, мы думаем, что XML СУБД существенно упростит систему в целом. Мы надеемся, что, в конечном счете, накладные расходы окажутся невелики. Для каждого кластера выделен отдельный XML-документ в центральном хранилище, который содержит информацию о состоянии кластера с момента подписки на услугу прогнозирования сбоев. Этот XML-документ имеет следующую структуру:

```
<cluster>
  <entries>                                <!-- Блок динамической -->
    <entry>                                  <!-- информации -->
      <date>...</date>                       <!-- Дата сбора -->
      <node>
        <name>...</name>                     <!-- Имя вычислительного узла -->
        <cpus>                                <!-- Информация о процессорах -->
          <cpu>
            <temp>...</temp>                 <!-- Температура процессора -->
            <fan>                             <!-- Скорость вращения -->
            <speed>...</speed>              <!-- вентилятора -->
            </fan>
            <voltage>...</voltage>          <!-- Напряжение на процессоре -->
            <usage>                           <!-- Коэффициент загрузки -->
            </usage>                          <!-- процессора -->
          </cpu>
        ...
      ...
    ...
  ...
</cluster>
```

```
</cpus>
  <voltages>...</voltages>                 <!-- Напряжения на -->
</node>                                     <!-- материнской плате -->
  ...
</entry>
...
</entries>
<description>                               <!-- Блок статической -->
  <!-- информации -->
  <name>...</name>                           <!-- Имя кластера -->
  <location>...</location>                  <!-- Местоположение кластера -->
  <organization>...</organization>        <!-- Организация-владелец -->
  <!-- кластера -->
  <nodes>                                     <!-- Статическое описание -->
    <!-- узлов кластера -->
    <node>
      <name>...</name>                       <!-- Имя вычислительного узла -->
      <motherboard>...</motherboard>        <!-- Марка и модель -->
      <!-- материнской платы -->
      <cpus>                                  <!-- Информация о процессорах -->
        <cpu>
          <name>...</name>                   <!-- Марка процессора -->
          <vendor_id>...</vendor_id>         <!-- Модель процессора -->
          <model_name>...</model_name>      <!-- Тактовая частота -->
          <mhz>...</mhz>                    <!-- процессора -->
        </cpu>
      ...
    </cpus>
  ...
</nodes>
</description>
</cluster>
```

Мы исследуем различные подходы к статистическому анализу накопленной информации о «здоровье» кластера для прогнозирования аппаратных сбоев. Мы полагаем, что даже простой анализ (например, экстраполяция характера изменения температуры процессора) даст возможность составления адекватных прогнозов и, тем самым, позволит увеличить надежность кластеров.

#### 4. Выводы

В ходе исследовательской работы разработан прототип системы прогнозирования аппаратных сбоев. Система должна удовлетворять двум целям: повысить надежность кластерной установки для пользователей и в то же самое время уменьшить общую стоимость владения

кластером. Прототип установлен и тестируется на кластерах «Первенец-М» в Институте Программных Систем РАН и СКИФ-К1000 в Объединенном Институте Проблем Информатики НАН Белоруси. В центральном хранилище накоплена статистическая информация об этих кластерах за тестовый период.

Несмотря на то, что наша разработка ориентирована прежде всего на работу в высокопроизводительных кластерах, однако, мы полагаем, что сервис прогнозирования отказов может быть востребован для других систем, таких как банки данных, фермы серверов и даже персональные компьютеры.

### Список литературы

- [1] Лаборатория Параллельных информационных технологий НИВЦ МГУ. PARALLEL.RU — Информационно-аналитический центр по параллельным вычислениям. — <http://parallel.ru>.
- [2] RAID — Wikipedia, the free encyclopedia. — [http://en.wikipedia.org/wiki/Redundant\\_array\\_of\\_independent\\_disks](http://en.wikipedia.org/wiki/Redundant_array_of_independent_disks).
- [3] Jack Dongarra Supercomputers and Clusters and Grids, Oh My!. — <http://www.hpc2n.umn.se/para06/index.php?content=ip1>: University of Tennessee, Oak Ridge National Laboratory.
- [4] Lm\_sensors — Linux hardware monitoring. — <http://secure.netroedge.com/~lm78/>.
- [5] IBM Self-Monitoring, Analysis and Reporting Technology (SMART). — <http://smartlinux.sourceforge.net/smart>.
- [6] O'Reilly Media I. XML.com: XML From the Inside Out — XML development, XML resources, XML specifications. — <http://www.xml.com>.
- [7] ISPRAS MODIS MODIS group. — <http://www.modis.ispras.ru/index.htm>.
- [8] Воеводин В. В., Воеводин В. В. Параллельные вычисления.: БХВ-Петербург, 2002, с. 600.
- [9] Жуматий С. А. Программная среда поддержки эффективного выполнения задач на параллельных вычислительных системах. — Москва, 2005.

ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР МУЛЬТИПРОЦЕССОРНЫХ СИСТЕМ

A. A. Moskovsky, A. Y. Pervin, M. V. Stotsky. *Cluster Hardware Monitoring for Failure Predictive Analysis.* (in Russian.)

ABSTRACT. This paper is devoted to the study of the failure predictive analysis systems. We consider several ways to achieve necessary level of reliability of hardware components of compute clusters. We introduce the prototype of the cluster hardware monitoring system for failure predictive analysis and its applications in real conditions.