

# СЕРВИСНАЯ ПОДСИСТЕМА УНИВЕРСАЛЬНОЙ КОМПАКТНОЙ СУПЕР-ЭВМ

*Н. А. Дмитриев, В. Н. Лашманов, В. Н. Стрюков, А. В. Шатохин, Е. Л. Шмаков, С. О. Черных*

ФГУП «РФЯЦ-ВНИИЭФ», г. Саров Нижегородской обл.

## Введение

Развитие различных областей науки и промышленности во многом определяется использованием и степенью совершенства вычислительной техники. Создание полноценных высокопроизводительных вычислительных систем (кластеров) – процесс трудоемкий и требующий значительных капиталовложений. В этой связи все большую популярность набирает направление компактных суперЭВМ – персональных супер-компьютеров (PSC, настольные суперЭВМ и т. д.), не требующих для эксплуатации специализированных инженерных систем и сооружений. Такие вычислительные машины ориентированы на использование в научно-исследовательских институтах, исследовательских лабораториях на производстве, в автомобильной, авиационной, аэрокосмической отраслях, атомной энергетике и коммерческих проектах. Компактные суперЭВМ позиционируются не только в качестве замены традиционных кластеров высокой производительности, но и как дополнение к ним. К этому классу вычислительных машин относится разработанная в РФЯЦ-ВНИИЭФ универсальная компактная суперЭВМ (далее КС-ЭВМ).

КС-ЭВМ оснащена базовым системным и прикладным программным обеспечением (в том числе и разработанным в РФЯЦ-ВНИИЭФ), ориентированным на решение конкретных задач 3-х мерного имитационного моделирования. Целевая аудитория – выделенные вычислительные центры и ведущие отраслевые предприятия РФ. Пиковая производительность КС-ЭВМ, составляющая 1.1 Тфлопс, достигается за счет использования трех вычислительных модулей (материнских плат), имеющих в своем составе по четыре универсальных процессора Magny-Cours и 32 модуля памяти. С аппаратной точки зрения КС-ЭВМ является комплексным оборудованием и состоит из ряда подсистем:

- вычислительной подсистемы;
- дисковой подсистемы;
- подсистемы энергообеспечения;
- подсистемы охлаждения;
- сервисной подсистемы.

## Вычислительная подсистема

Вычислительная подсистема КС-ЭВМ – это три вычислительных модуля, объединенных с использованием канала InfiniBand. Соединение осуществляется по бескоммутаторной схеме «треугольник» с использованием двухпортовых адаптеров QDR InfiniBand как показано на рис. 1.

## Дисковая подсистема

Дисковая подсистема КС-ЭВМ состоит из шести SATA-II дисков (по два диска на каждый модуль) и трех SSD-дисков (по одному на каждый модуль), как показано на рис. 2.

SATA-диски каждого вычислительного модуля объединены в RAID-массив (RAID-1) с целью повышения надежности и предназначены для хранения данных, SSD-диски хранят метаданные. Диски вычислительных модулей взаимодействуют посредством канала InfiniBand и образуют кластерную параллельную файловую систему Lustre.

## Подсистема энергообеспечения

Подсистема энергообеспечения КС-ЭВМ (рис. 3) построена на принципе отдельной и комбинированной подачи питания на энергопотребляющие компоненты. Питание вычислительных модулей построено на базе трех блоков питания с отдельной подачей питания на каждый модуль. В КС-ЭВМ предусмотрено два дополнительных маломощных блока питания для питания водяных помп (в системе жидкостного охлаждения), ethernet-коммутатора, сервисного модуля и 120-мм вентиляторов. Вся система подключается к энергосети посредством одной вилки стандарта 220V, суммарная мощность не превышает 2,5 кВт.

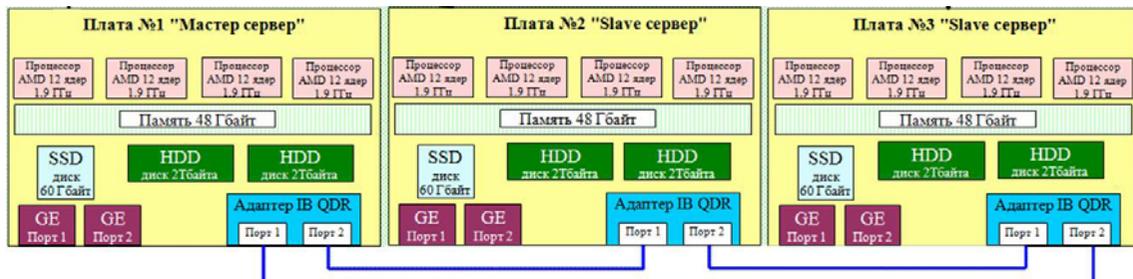


Рис. 1. Вычислительная подсистема КС-ЭВМ

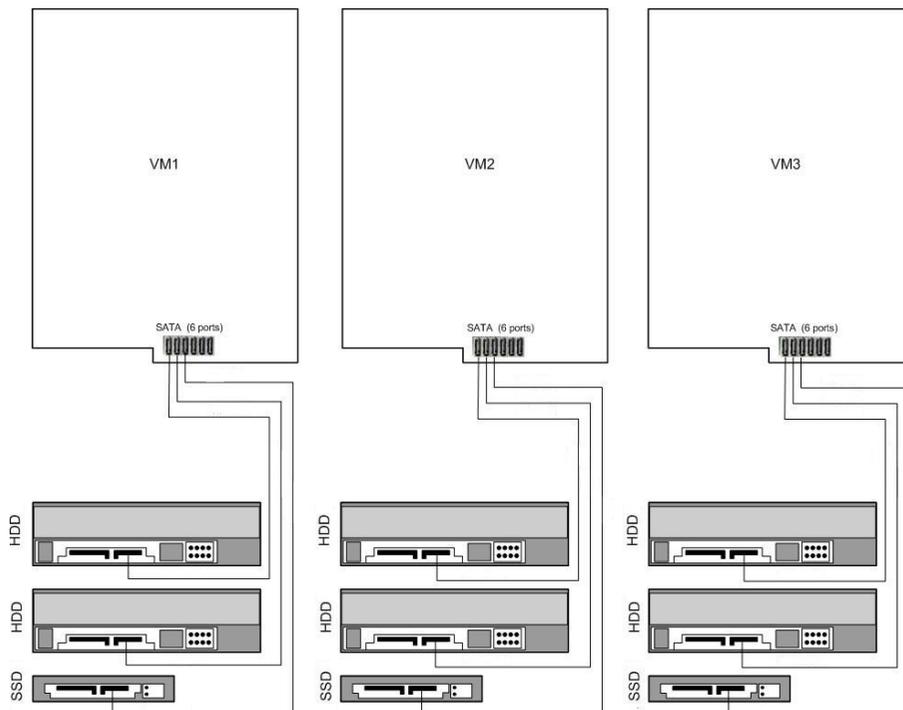


Рис. 2. Дисковая подсистема КС-ЭВМ

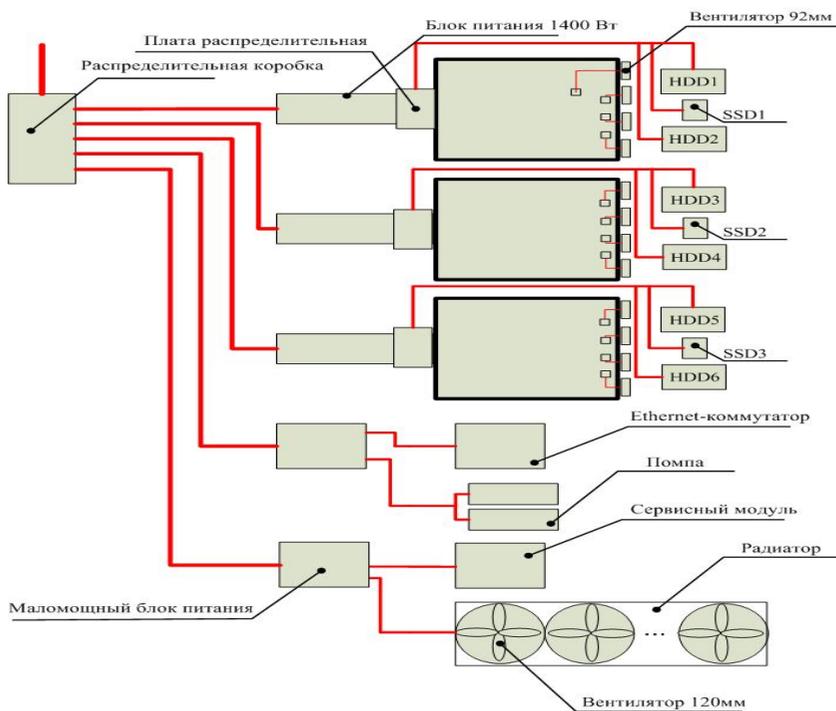


Рис. 3. Подсистема энергообеспечения КС-ЭВМ

## Подсистема охлаждения

Подсистема охлаждения КС-ЭВМ (рис. 4) является комбинированной жидкостно-воздушной системой. Система жидкостного охлаждения (далее СЖО) представляет собой совокупность элементов, образующих замкнутый контур, в котором циркулирует жидкость, отводящая тепло от процессоров. Воздушная система охлаждения используется для защиты от перегрева блоков питания, компонентов материнских плат, жестких дисков и для охлаждения жидкости, проходящей через радиаторы СЖО. Воздушная система охлаждения блоков питания построена на вентиляторах и является штатной для блоков питания и встроена непосредственно в корпус блоков питания.

## Сервисная подсистема

Контроль показателей работоспособности подсистем КС-ЭВМ возложен на сервисную подсистему (рис. 5), являющуюся, по сути, аппаратно-программным комплексом. Аппаратная часть представлена вычислительными модулями, объединенными сетью Ethernet с сервисной платой (сервисный модуль). Каждый вычислительный модуль имеет в своем составе плату ВМС для опроса аппаратных датчиков, характеризующих основные показатели состояния системы, такие, как температура, напряжения на энергопотребляющих компонентах, скорости вращения вентиляторов и т. д. Опрос осуществляется с определенной периодичностью, задаваемой программно.

Вычислительные модули имеют следующий список групп опрашиваемых сенсоров:

- температурные сенсоры температуры процессоров и чипсетов;
- подсистема питания – номиналы напряжений на компонентах;
- скорости вращения вентиляторов;
- состояние питания – статус, имеющий три значения (включено, выключено, отключено от питающей сети).

Сервисный модуль является компактной материнской платой VIA EPIA-N700-15L (форм-фактор NanoITX), оснащенной процессором с тактовой частотой 1.5 ГГц и 2 Гбайт DDR2 оперативной памяти, и предназначен для запуска и выполнения системы мониторинга. GE-порты (Gigabit Ethernet) сервисного и вычислительных модулей соединены с соответствующими портами Ethernet-коммутатора (switch), образуя сервисную сеть.

Программная составляющая сервисной подсистемы включает в себя системное ПО и систему управления и мониторинга (рис. 6). Система мониторинга предназначена для работы под управлением операционной системы Linux и представляет собой скриптовую оболочку для утилиты ipmitool. Это означает, что в основе функционирования системы лежит использование указанной утилиты для выполнения IPMI-опросов вычислительных узлов с последующей обработкой результатов опроса. Скриптовый язык, используемый в данной системе, – Perl (CGI).

На рис. 6 представлены основные функциональные блоки системы мониторинга. Центральным управляющим элементом системы является система сбора и обработки информации. В ее функции входит управление утилитой ipmitool и обработкой результатов опроса. Опрос плат ВМС вычислительных модулей осуществляется по протоколу IPMI.

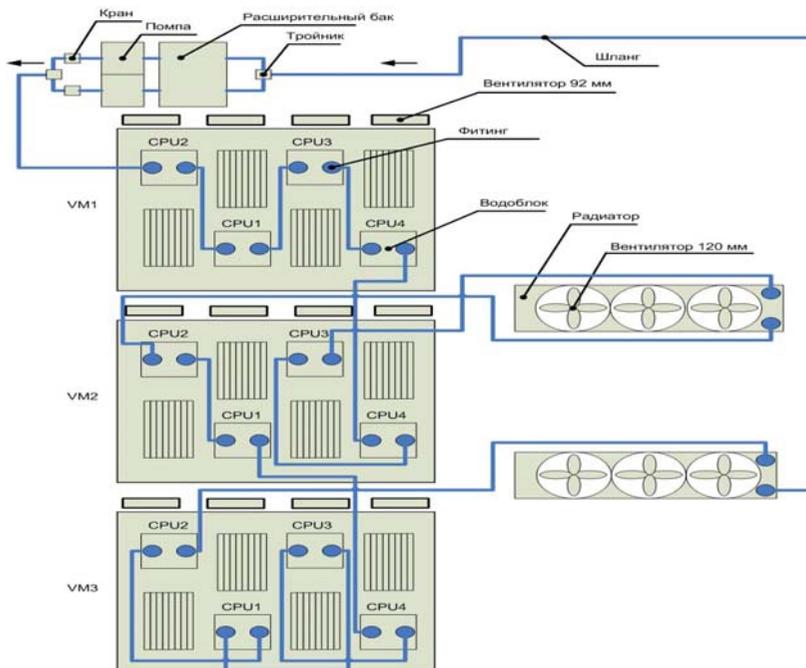


Рис. 4. Подсистема охлаждения КС-ЭВМ

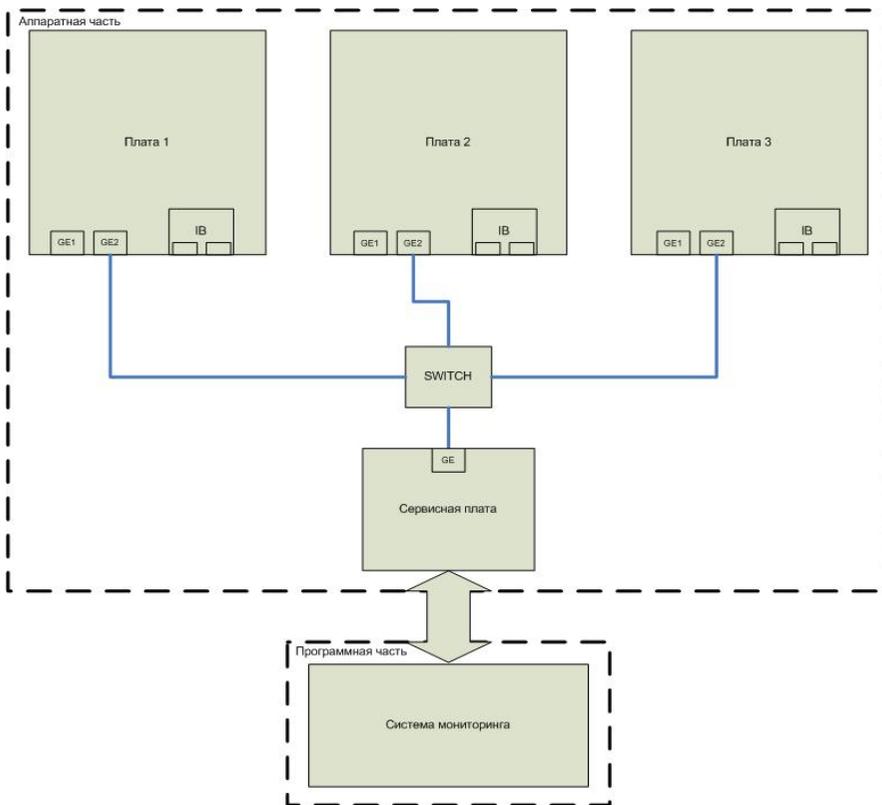


Рис. 5. Сервисная подсистема КС-ЭВМ

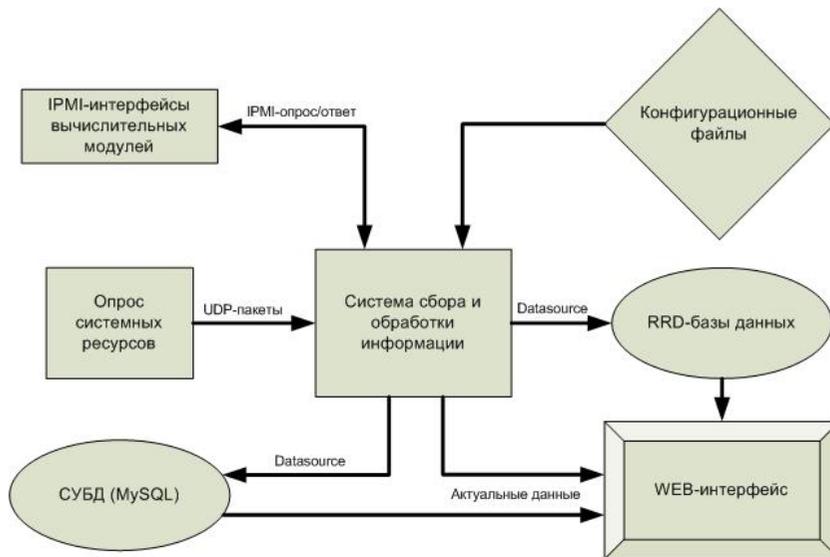


Рис. 6. Функциональная схема системы мониторинга

Данные, получаемые в результате опроса, фиксируются в СУБД (статистические данные), и предоставляется доступ к ним средствами WEB-интерфейса пользователя (актуальные данные). Конфигурационные файлы содержат необходимую для работы системы информацию: данные о подключении к СУБД и конфигурационные данные контролируемых устройств. При старте система сбора и обработки осуществляет чтение конфигурационных файлов, подключение к СУБД, в зависимости от заданных периодичностей опроса устройств выполняет балансировку нагрузки, после чего переходит

к процессу периодического опроса. RRD базы данных используются для накопления статистических сведений о показаниях датчиков за определенный промежуток времени. Исходные данные для формирования баз RRD предоставляются системой сбора и обработки. Кроме того, система мониторинга предусматривает возможность получения сведений о степени загруженности оборудования КС-ЭВМ. Эти функции возложены на модуль опроса системных ресурсов, который по протоколу UDP передает результаты опроса системе сбора и обработки информации.

WEB-страница системы мониторинга КС-ЭВМ спроектирована с учетом особенностей оборудования и имеет удобный, интуитивно понятный интерфейс. В функции WEB-интерфейса входит предоставление возможности:

- оперативного анализа состояния оборудования КС-ЭВМ;
- статистического анализа значений контролируемых сенсоров и полученных сообщений оборудования;
- управление питанием.

Карта WEB-интерфейса системы мониторинга представлена на рис. 7.

Далее будут рассмотрены разделы WEB-интерфейса и особенности их использования.

На рис. 8 представлено изображение главного окна WEB-интерфейса системы мониторинга КС-ЭВМ. Как видно из рисунка, окно состоит из трех основных областей: основного меню, заголовка, области данных.

Основное меню – это набор ссылок, предоставляющих доступ к основным разделам системы мониторинга. Механизм работы ссылок основан на обработке CGI-скриптов, определенным образом формирующих содержимое области данных.

Заголовок – статическая область, предназначенная для улучшения визуальных качеств интерфейса, реализованная с использованием изображения в формате .PNG.

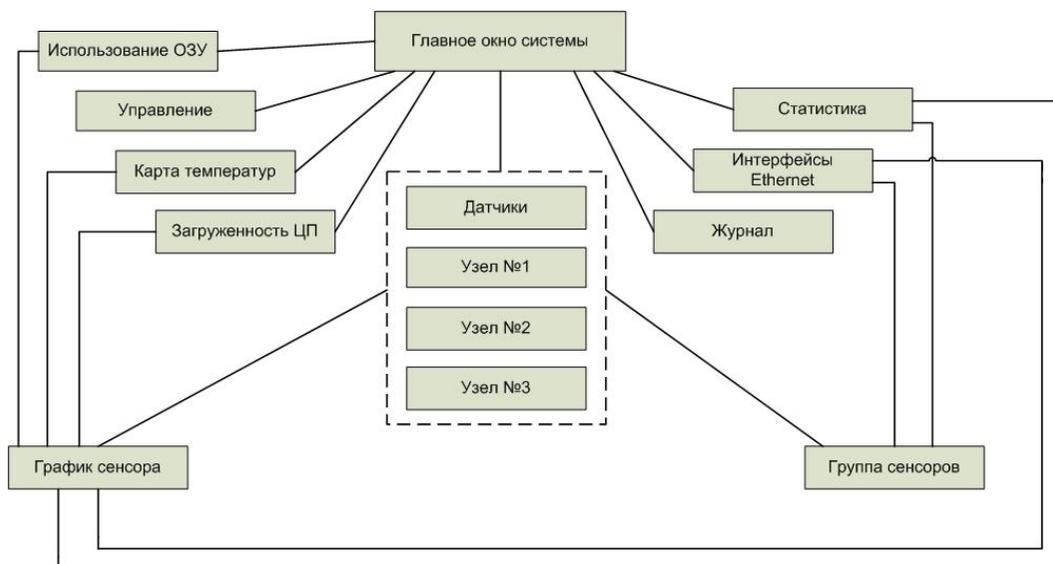


Рис. 7. Карта WEB-интерфейса системы мониторинга КС-ЭВМ

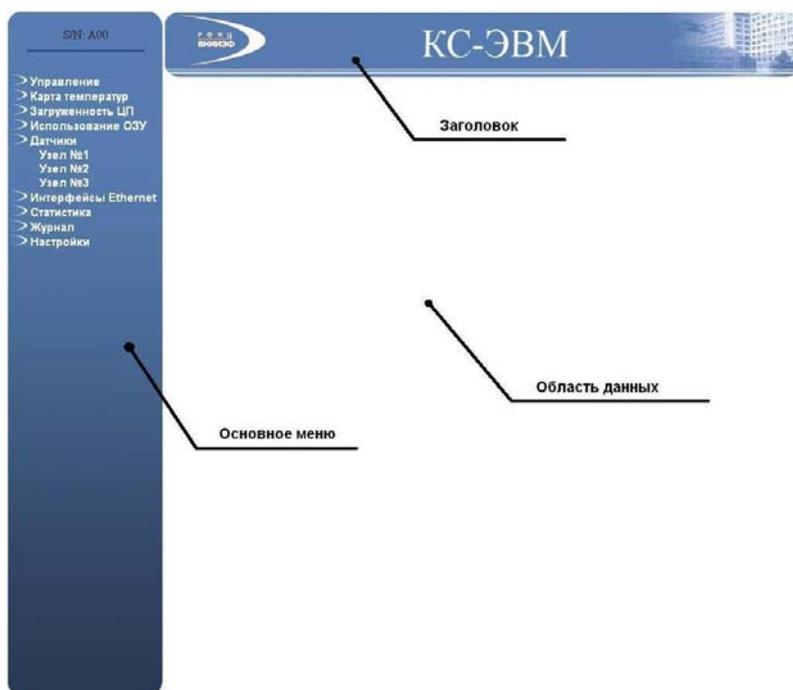


Рис. 8. Главное окно WEB-интерфейса системы мониторинга КС-ЭВМ

Область данных – динамически изменяющееся содержимое окна WEB-интерфейса. Вывод информации осуществляется в табличной форме.

В разделе «Управление» доступны базовые операции над питанием каждого вычислительного модуля и возможность просмотра статуса питания (рис. 9).

На рис. 10 представлена карта температур. Идея данного режима отображения информации заключается в том, чтобы предоставить пользователю удобную возможность визуальной оценки степени нагрева компонент КС-ЭВМ.

Карта загрузки центральных процессоров (рис. 11) построена по аналогии с температурной картой и предназначена для предоставления возможности оценки степени нагрузки на процессоры КС-ЭВМ.

Использование ОЗУ (рис. 12) – раздел системы мониторинг, построенный по принципу температурной

карты и предназначен для наблюдения за степенью загруженности оперативной памяти каждого вычислительного модуля КС-ЭВМ.

Датчики (рис. 13) – это раздел системы мониторинга, предоставляющий сведения об аппаратных сенсорах. Информация выводится в таблицу в четыре колонки: ссылка на график сенсора, наименование, показание, статус. Для удобства просмотра сенсоры сгруппированы по группам. Название группы является ссылкой, позволяющей в отдельном окне просматривать показания сенсоров только выбранной группы.

Разделы «Узел № 1», «Узел № 2», «Узел № 3» являются подразделами раздела «Датчики». Данные выводятся только по одному (выбранному) вычислительному модулю.



Рис. 9. Управление



Рис. 10. Карта температур



Рис. 11. Загруженность ЦП



Рис. 12. Загруженность ОЗУ

Сетевой трафик и количество ошибок в Ethernet-сетях можно посмотреть в разделе «Интерфейсы Ethernet» (рис. 14).

Раздел статистики позволяет просматривать показания групп датчиков за определенный (выбираемый пользователем) промежуток времени в графическом виде (рис. 15).

Настройку сетевого интерфейса, а также имени пользователя и пароля предоставляет раздел системы мониторинга «Настройки» (рис. 16).

### Заключение

Поскольку КС-ЭВМ является уникальной разработкой, требующей использования специализированного оборудования и программного обеспечения, создание сервисной подсистемы позволило решить ряд задач, связанных с управлением и мониторингом ЭВМ такого класса. Накопленный опыт и наработки будут использованы для создания компактных вычислительных машин следующего поколения. Пла-

нируется расширение функциональных возможностей WEB-интерфейса. На данный момент сервисная подсистема в составе КС-ЭВМ сдана в опытную эксплуатацию, успешно функционирует, а также ис-

пользуется для тестирования и отладки небольших групп серверов и активно модернизируется.

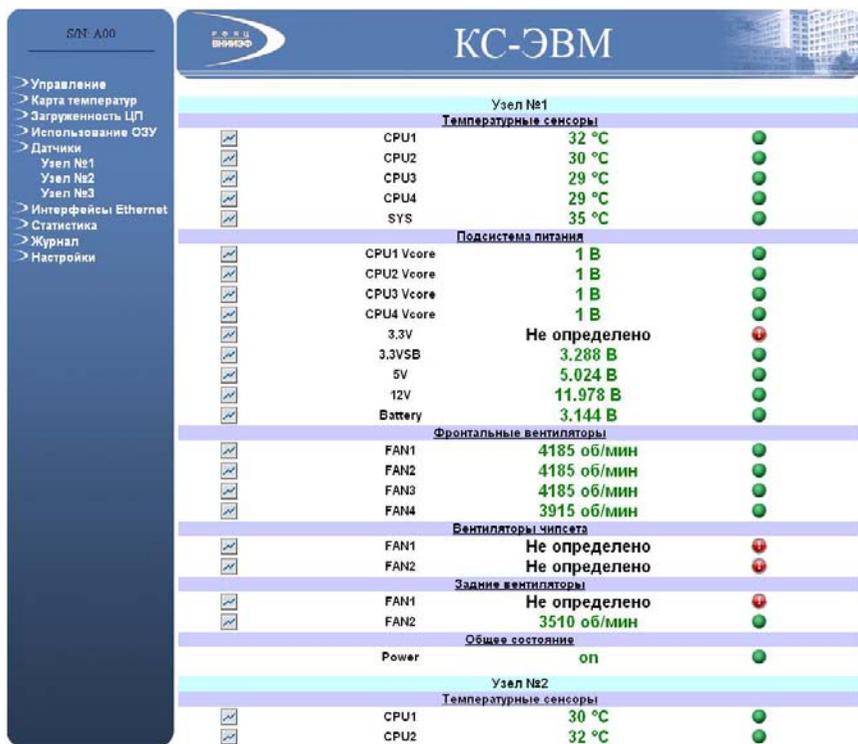


Рис. 13. Датчики



Рис. 14. Интерфейсы Ethernet



Рис. 15. Статистика

The figure shows the configuration page for the network interface. The 'Настройка сетевого интерфейса' section includes: IP-адрес: 172.19.38.51, Шлюз: 172.19.38.1, Маска подсети: 255.255.255.0. The 'Настройка имени пользователя и пароля' section has input fields for 'Пользователь' and 'Пароль', and buttons for 'Подтвердить', 'Добавить', 'Изменить', and 'Удалить'. A table lists the following users: Shatoichin\_av, Chernykh\_go, Lashmanov\_vn, Dembnev\_na, and Strelkov\_el.

Рис. 16. Настройки