

**Модели первого ряда семейства
суперкомпьютеров**

Общее техническое задание

1. Основание для выполнения работ

Работы по созданию моделей семейства суперкомпьютеров выполняются в соответствии с Совместной Программой "Разработка и освоение в серийном производстве семейства моделей высокопроизводительных вычислительных систем с параллельной архитектурой (суперкомпьютеров) и создание прикладных программно-аппаратных комплексов на их основе" (далее по тексту Совместная Программа), утвержденной

2. Основные цели и задачи создания семейства суперкомпьютеров

Главная цель Совместной Программы:

Возрождение суперкомпьютерной отрасли двух стран, *серийное промышленное производство ряда программно-совместимых моделей суперкомпьютеров с широким спектром производительности* — до 10 триллионов операций в секунду. Это позволит обеспечить достижение и *других целей Совместной Программы*:

- *Технологический прорыв в различных отраслях*—конструкторская деятельность, химическая промышленность, фармацевтика, медицина, перспективные наукоемкие разработки (термоядерный синтез, безопасные энергетические реакторы, ускорители), геологоразведка, национальная оборона, государственное и корпоративное управление.
- *Обеспечение национальной безопасности в суперкомпьютерной отрасли*— покрытие потребностей стран в такой технике за счет собственных изделий, ликвидация зависимости от поставок из западных стран.
- *Создание экспортных возможностей* для поставок суперкомпьютеров в третьи страны.
- *Создание основы* для дальнейшего развития отечественных перспективных разработок и наращивания отечественного потенциала в области производства высокопроизводительных средств вычислительной техники.

Важнейшие целевые показатели предлагаемой разработки:

- высокая производительность параллельных вычислительных комплексов;
- низкое соотношение цена/производительность;
- возможность развертывания серийного производства входящих в единый ряд моделей параллельных вычислительных комплексов на отечественных предприятиях;
- возможность реализации разнообразных прикладных систем, предназначенных для использования в различных областях деятельности на основе входящих в единый ряд моделей суперкомпьютеров;
- широкий набор базовых технических и программных средств.

3. Обобщенная система требований по созданию моделей семейства суперкомпьютеров

Система требований по созданию моделей семейства суперкомпьютеров базируется на следующих основных регламентирующих документах:

- Совместная белорусско-российская программа "Разработка и освоение в серийном производстве семейства моделей высокопроизводительных вычислительных систем с параллельной архитектурой (суперкомпьютеров) и создание прикладных программно-аппаратных комплексов на их основе".
- Концепция создания моделей семейства суперкомпьютеров (далее по тексту Концепция).
- Общее техническое задание (ОТЗ) на создание моделей семейства суперкомпьютеров.

- Общие технические требования (ОТТ) на создание моделей семейства суперкомпьютеров.
- Технические задания (ТЗ) на реализацию конкретных программных мероприятий (проектов) Совместной Программы.

Совместная Программа определяет:

- этапность выполнения работ;
- основных исполнителей программных мероприятий (в том числе, координаторов Совместной Программы, головного исполнителя и исполнителя Совместной Программы от РБ и РФ);
- перечень и наименование проектов Совместной Программы;
- ориентировочные сроки выполнения и ориентировочные объемы финансирования проектов.

Точные сроки выполнения и объемы финансирования определяются соответствующими договорными соглашениями по конкретным проектам.

Концепция определяет общие для всего семейства принципы создания суперкомпьютерных систем на всех этапах технологического процесса - от базовых концептуальных архитектурных принципов до процессов разработки, освоения в серийном производстве и поставки конечному пользователю конкретных прикладных суперкомпьютерных конфигураций.

В общем техническом задании на создание моделей семейства суперкомпьютеров в сжатой форме сформулированы общие технические принципы создания семейства суперкомпьютеров, базирующиеся на положениях Концепции.

ОТЗ разрабатывается для каждого этапа реализации Совместной Программы - ОТЗ на создание моделей первого ряда семейства суперкомпьютеров (I этап реализации Совместной Программы) и ОТЗ на создание моделей второго ряда семейства суперкомпьютеров (II этап реализации Совместной Программы).

В общих технических требованиях сформулированы общие технические требования к моделям семейства суперкомпьютеров, базирующиеся на положениях ОТЗ.

ОТТ также разрабатываются для каждого этапа реализации Совместной Программы - ОТТ к моделям первого ряда семейства суперкомпьютеров и ОТТ к моделям второго ряда семейства суперкомпьютеров.

Разделение общих технических принципов и требований на два документа (ОТЗ и ОТТ) отражает объективную целесообразность разработки ОТТ после согласования ОТЗ из-за разных (по очевидным причинам) сроков отработки концептуальных принципов и технических требований.

Технические задания на проведение работ по реализации конкретных проектов разрабатываются с учетом основных положений Совместной Программы, Концепции, ОТЗ и ОТТ и отражают специфику конкретных проектов.

4. Сроки выполнения работ

Работы по созданию моделей первого ряда семейства суперкомпьютеров выполняются в сроки первого этапа реализации Совместной Программы:

начало работ - I кв. 2000 г.

окончание работ - IV кв. 2001 г.

Сроки начала и окончания работ могут быть уточнены в зависимости от начала реального финансирования Совместной Программы.

5. Основные исполнители Совместной Программы

От Республики Беларусь:

- Национальная академия наук Беларуси (НАН Беларуси) - **координатор Совместной Программы;**
- Научно-исследовательское объединение (НИО) "Кибернетика" НАН Беларуси - **головной исполнитель Совместной Программы;**
- Научно-инженерное конструкторско-технологическое предприятие (НИКТП) "Белмикросистемы" Научно-производственного объединения (НПО) "Интеграл";
- Государственное предприятие "Научно-исследовательский институт электронных вычислительных машин" (ГП "НИИ ЭВМ");
- Государственное предприятие "Минское производственное объединение вычислительной техники" (ГП "МПОВТ").

От Российской Федерации:

- Федеральное государственное унитарное предприятие (ГУП) "Российские технологии" - **координатор Совместной Программы;**
- Институт программных систем Российской академии наук (ИПС РАН) - **исполнитель Совместной Программы;**
- Предприятие "Суперкомпьютерные системы" ("СКС");
- Институт высокопроизводительных вычислений и баз данных (ИВВ и БД) Министерства науки и технологий РФ.

Исполнители конкретных проектов по созданию моделей первого ряда семейства суперкомпьютеров и перечень этих проектов определены в Совместной Программе.

Состав работ по созданию моделей первого ряда семейства суперкомпьютеров может быть уточнен в зависимости от реальных объемов финансирования Совместной Программы.

6. Архитектура моделей семейства суперкомпьютеров

6.1. Архитектура всех моделей суперкомпьютеров базируется на основополагающем концептуальном архитектурном принципе - возможности использования для организации параллельного счета в рамках единой архитектуры высокопроизводительных вычислительных систем с параллельной архитектурой (ВВСПА) двух независимых (с возможностью взаимодействия) различных архитектурных компонент - сеть (кластер) вычислительных узлов или Асинхронная архитектурная Компонента (АК) и синхронная архитектурная компонента на базе Однородной Вычислительной Среды (ОВС):

1-й уровень (АК) - тесно связанная сеть (кластер) вычислительных узлов, работающих под управлением одного из клонов широко используемой многопользовательской универсальной операционной системы Unix и использующих для организации параллельного выполнения прикладных задач оригинальную распределенную операционную систему (Т-систему), реализующую автоматическое динамическое распараллеливание программ;

2-й уровень (ОВС) - однородная вычислительная среда - синхронная сеть большого числа (десятки тысяч) простых однопоровых процессоров, позволяющая эффективно реализовывать те задачи (фрагменты прикладных проблем), которые сложно распараллеливаются на первом уровне суперкомпьютера.

6.2. Реализация возможности организации параллельного счета на двух уровнях единой архитектуры обеспечивает:

- эффективную реализацию фрагментов прикладных проблем со сложной логикой вычисления, с крупноблочным скрытым (динамическим) параллелизмом с использованием Т-системы (на кластерном (асинхронном) архитектурном уровне суперкомпьютера);

- эффективную реализацию фрагментов прикладных проблем с простой логикой вычисления, с мелкозернистым явным параллелизмом, с большими потоками информации, требующими быстрой и несложной обработки “на пролете”, на синхронном (ОВС) архитектурном уровне суперкомпьютера.

6.3. Пропорции деления прикладной проблемы на фрагменты с крупноблочным и мелкозернистым параллелизмом определяют:

- пропорции объемов программного обеспечения в части Т-системы и ОВС;
- эффективную конфигурацию ВВСПА для решения конкретной прикладной проблемы (количество вычислительных узлов кластера, объемы оборудования ОВС и т.п.).

6.4. Архитектура моделей семейства суперкомпьютеров в зависимости от области применения может быть трех типов:

- кластерная (асинхронная) архитектура (1-й уровень единой двухуровневой архитектуры), реализуемая на кластере вычислительных узлов с Т-системой (кластер Т-узлов);
- синхронная архитектура (2-й уровень единой двухуровневой архитектуры), реализуемая на однородной вычислительной среде (синхронная сеть большого числа простых однобитовых процессоров);
- кластерная гибридная архитектура, включающая оба уровня единой архитектуры, реализуемая на кластере вычислительных узлов с Т-системой и ОВС (кластер Т-узлов и узлов ОВС).

6.5. Двухуровневая архитектура является **масштабируемой реконфигурируемой надплатформенной**, т.е. может взаимодействовать с кластерными узлами, работающими под управлением разных операционных систем и реализованными на разных программно-аппаратных платформах.

6.6. Двухуровневая архитектура позволяет создавать прикладные суперкомпьютерные системы в широком диапазоне производительности - от изделий класса высокопроизводительных серверов (до 50-100 миллионов операций в секунду) и мультипроцессорных рабочих станций (миллиарды операций в секунду) до вычислительных систем с массовым параллелизмом сверхвысокой производительностью (триллионы операций в секунду).

6.7. основополагающие архитектурные принципы должны быть адекватно отображены на программно-аппаратном уровне реализации моделей семейства суперкомпьютеров.

7. Требования к базовому (общесистемному) программному обеспечению моделей суперкомпьютеров

7.1. Универсальная сетевая операционная система

7.1.1. В качестве сетевой операционной системы (ОС) в моделях суперкомпьютеров с кластерной (асинхронной) и кластерной гибридной архитектурой должна использоваться ОС Linux - клан универсальной ОС Unix.

ОС Linux распространяется свободно, т.е. бесплатно и с исходными текстами. Распространение ОС Linux не подвержено каким-либо ограничениям каких-либо стран или фирм.

ОС Linux является открытой, т.е. реализована не только для платформ класса IBM PC, но и для многих других аппаратных платформ.

ОС Linux —обеспечивает наличие готовой поддержки в кластерном уровне суперкомпьютера всех существующих (и возможных в будущем) требований в части:

- периферийных устройств (включая поддержку подключения к LAN/WAN/Internet);
- различных режимов функционирования (многопользовательский, удаленный с криптографическими протоколами доступа, оконный интерфейс и т.п.);

- сервисных программных средств.

7.1.2. Для реализации дополнительных свойств в суперкомпьютерных кластерных конфигурациях наряду с базовой сетевой ОС Linux могут использоваться и другие операционные системы, например, ОС MS Windows NT Server.

7.2. Операционная система динамического распараллеливания (Т-система)

7.2.1. Управление работой моделей суперкомпьютеров с кластерной (асинхронной) архитектурой и кластерной гибридной архитектурой должно осуществляться распределенной операционной системой динамического распараллеливания (Т-системой), являющейся оригинальным программным обеспечением для мультипроцессорных систем.

Задачей Т-системы является назначение процессов на кластерные узлы системы. Совокупность кластерных узлов, используемых для решения задачи определяется набором процессов, готовых к исполнению, и изменяется в ходе решения задачи. Т-система должна анализировать занятость кластерных узлов и выбирать те из них, которые позволяют наиболее эффективно реализовать процесс.

7.2.2. Т-система должна обеспечивать:

- автоматическое разбиение прикладных проблем (задач) на:
 - фрагменты со сложной логикой вычисления, с крупноблочным скрытым (динамическим) параллелизмом - такие фрагменты эффективно реализовывать с использованием Т-системы, на кластерном уровне суперкомпьютера;
 - фрагменты с простой логикой вычисления, с мелкозернистым явным параллелизмом, с большими потоками информации, требующими быстрой и несложной обработки "на пролете" - такие фрагменты эффективно реализовывать в ОВС.
- автоматическую передачу фрагментов задач со скрытым и явным параллелизмом для выполнения соответственно на асинхронном и синхронном уровнях параллелизма.

7.2.3. При выполнении фрагментов задач с крупноблочным (скрытым) параллелизмом на кластерном уровне суперкомпьютера Т-система должна обеспечивать автоматическое динамическое распараллеливание программ, реализуя следующие основные функции:

- обнаружение готовых к исполнению фрагментов задачи (процессов);
- их распределение по процессорам;
- их синхронизацию по данным.

7.2.4. Т-система должна обеспечивать глубокий уровень параллелизма за счет поддержки корректного выполнения некоторых операций над неготовыми значениями (реализация понятия "неготовое значение", т.е. поддержка возможности выполнения счета в некотором процессе-потребителе в условиях, когда часть из обрабатываемых им значений еще не готова - не вычислена в соответствующем процессе-поставщике).

7.2.5. Для снижения трафика межпроцессорных передач, а, следовательно, и требований к пропускной способности аппаратуры объединения процессорных элементов в кластер при организации параллельного счета, алгоритм динамического автоматического распределения процессов по процессорам должен учитывать специфику неоднородных распределенных вычислительных сетей.

7.2.6. Программы, разработанные для Т-системы, должны без переработки и перекомпиляции исполняться на кластерном суперкомпьютере с любым числом процессоров (вычислительных узлов).

Это позволяет:

- Без переработки системного и прикладного программного обеспечения расширять конфигурации систем (добавлять число процессоров в систему) и тем самым достигать ускорения счета прикладных программ.
- Без переработки системного и прикладного программного обеспечения исполнять задачи на частично неисправной системе (с вышедшими из строя компонентами) без перезапуска задачи после выхода из строя части оборудования.

7.2.7. В качестве языка программирования Т-системы должно использоваться расширение языка с новыми Т-конструкциями.

7.2.8. Т-система должна быть надплатформенной, т.е. должна взаимодействовать с кластерными узлами, работающими под управлением разных операционных систем и реализованными на разных программно-аппаратных платформах.

7.2.9. Конкретные технические требования к Т-системе должны быть отражены в технических заданиях (ТЗ) на проведение работ по реализации соответствующих проектов Совместной Программы.

7.3. Программное обеспечение ОВС (синхронного архитектурного уровня)

7.3.1. Программное обеспечение (ПО) ОВС должно включать набор программных инструментов и функций (таких как диспетчер задач, компоновщик программ для матрицы ОВС и драйверы поддержки соответствующих внешних устройств, подключаемых к матрице ОВС).

7.3.2. ПО ОВС должно обеспечивать:

- организацию виртуальных спецпроцессоров на матрице ОВС (за счет программного управления), реализующих решаемую функцию или задачу;
- одновременное решение нескольких задач или функций;
- реализацию механизма перезагрузки сегментов ОВС, позволяющего перезагружать часть матрицы без остановки выполнения еще незавершенных задач;
- полную аппаратную и программную масштабируемость (для обеспечения гибкости и перестраиваемости вычислительных систем на базе ОВС);
- использование естественного параллелизма решаемой задачи вплоть до битового уровня, то есть уровня структуры обрабатываемых данных;
- организацию конвейеров произвольной глубины;
- возможность одновременной обработки множества независимых некогерентных потоков;
- гибкую организацию ввода/вывода данных на вычислительную матрицу ОВС;
- рост производительности решения задач на ОВС (близкий к линейному) с увеличением рабочей частоты поля и площади вычислительной матрицы.

7.3.3. В состав ПО ОВС должны входить необходимые инструментальные средства разработчика программ для ОВС, соответствующие различным уровням абстракции, и библиотеки стандартных функций для ОВС.

7.3.4. Конкретные требования к ПО ОВС должны быть отражены в ТЗ на проведение работ по реализации соответствующих проектов Совместной Программы.

7.4. Программные средства сопряжения асинхронного и синхронного архитектурных уровней

7.4.1. Программные средства сопряжения *в части асинхронного (кластерного) уровня* должны включать:

- набор входящих в состав ядра ОС Linux широкой номенклатуры стандартных драйверов;
- базовую библиотеку стандартных примитивов обмена информацией и управления ОВС;
- библиотеку прикладных задач и подпрограмм, реализуемых с использованием ОВС;
- структуры данных и программные механизмы, обеспечивающие:
 - передачу Т-процесса, из которого осуществляется взаимодействие с ОВС, в один из вычислительных узлов кластерного уровня, имеющих физический интерфейс с ОВС;
 - осуществление удаленного вызова функции/прикладной задачи из вычислительного узла кластерного уровня, не имеющего интерфейса с ОВС с использованием механизмов, предназначенных для распределенной работы с файлами.

7.4.2. *В части ОВС* программные средства сопряжения включают в себя реализованный в виде специализированной библиотеки набор предназначенных для загрузки из кластерной компоненты в ОВС фрагментов программного кода, каждый из которых

непосредственно реализует в ОВС ту или иную прикладную задачу или фрагмент вычислений, в частности:

- получает из кластерного уровня наборы входных данных;
- организует и осуществляет выполнение в ОВС вычислений в соответствии с алгоритмом решения соответствующей прикладной задачи;
- передает из ОВС в кластерный уровень наборы данных, содержащие результаты вычислений.

7.4.3. Программные средства сопряжения асинхронного и синхронного архитектурных уровней должны обеспечивать:

- передачу фрагмента решаемой задачи из Т-программы в ОВС;
- передачу фрагмента решаемой задачи из выполняемого в ОВС кода в кластерную компоненту.

7.4.4. Конкретные требования к программным средствам сопряжения асинхронного и синхронного архитектурных уровней должны быть отражены в ТЗ на проведение работ по реализации соответствующих проектов Совместной Программы.