



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G06F 13/00 (2020.02); G06F 13/10 (2020.02); G06F 9/54 (2020.02)

(21)(22) Заявка: 2019116927, 31.05.2019

(24) Дата начала отсчета срока действия патента:
31.05.2019

Дата регистрации:
11.03.2021

Приоритет(ы):

(22) Дата подачи заявки: 31.05.2019

(43) Дата публикации заявки: 30.11.2020 Бюл. № 34

(45) Опубликовано: 11.03.2021 Бюл. № 8

Адрес для переписки:

117997, Москва, ГСП-7, ул. Профсоюзная, 65,
ИПУ РАН, патентный отдел

(72) Автор(ы):

Стецюра Геннадий Георгиевич (RU)

(73) Патентообладатель(и):

Федеральное государственное бюджетное
учреждение науки Институт проблем
управления им. В.А. Трапезникова
Российской академии наук (RU)

(56) Список документов, цитированных в отчете

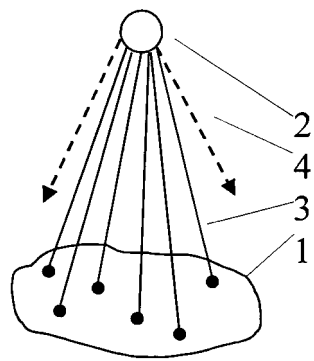
о поиске: US 9804901 B2, 31.10.2017. US
10169097 B2, 01.01.2019. US 7840833 B2,
23.11.2010. ЕА 200900001 А3, 26.02.2010. RU
2008138700 А, 10.04.2010.

(54) Способ и система управления связями компьютеров в многоуровневом составном компьютерном кластере

(57) Реферат:

Изобретение относится к вычислительной технике. Технический результат заключается в расширении арсенала средств того же назначения. Способ управления связями компьютеров в многоуровневом составном компьютерном кластере, содержащем контроллеры компьютеров - объекты kO_i и модули связи kMS , характеризующийся тем, что на каждом уровне иерархии, равном $k=0$ и более, лидер - один из объектов kO_i , непосредственно связанных с модулем связи kMS уровня иерархии k , посылает команду через kMS указанным в команде модулям связи следующего уровня иерархии ^{k+1}MS и

непосредственно связанным с ними объектам $^{k+1}O_i$, при этом модули ^{k+1}MS посылают в kMS сигнал начала барьерной синхронизации и прерывают передачу сигналов от kMS объектам $^{k+1}O_i$, взаимодействуя только со своим ^{k+1}MS под управлением одного из объектов $^{k+1}O_i$ - лидера уровня иерархии $k+1$, который после завершения заданных командой действий останавливает передачу от ^{k+1}MS к kMS сигнала барьерной синхронизации и восстанавливает связь между kMS и $^{k+1}O_i$. 2 н.п. ф-лы, 5 ил.



Фиг. 1

RU 2 7 4 4 5 9 1 C 2

RU 2 7 4 4 5 9 1 C 2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

G06F 13/00 (2020.02); G06F 13/10 (2020.02); G06F 9/54 (2020.02)(21)(22) Application: **2019116927, 31.05.2019**(24) Effective date for property rights:
31.05.2019Registration date:
11.03.2021

Priority:

(22) Date of filing: **31.05.2019**(43) Application published: **30.11.2020 Bull. № 34**(45) Date of publication: **11.03.2021 Bull. № 8**

Mail address:

**117997, Moskva, GSP-7, ul. Profsoyuznaya, 65,
IPU RAN, patentnyj otdel**

(72) Inventor(s):

Stetsyura Gennadij Georgievich (RU)

(73) Proprietor(s):

**Federalnoe gosudarstvennoe byudzhethoe
uchrezhdenie nauki Institut problem upravleniya
im. V.A. Trapeznikova Rossijskoj akademii nauk
(RU)**(54) **METHOD AND SYSTEM FOR CONTROLLING COMPUTER COMMUNICATIONS IN MULTILEVEL COMPOSITE COMPUTER CLUSTER**

(57) Abstract:

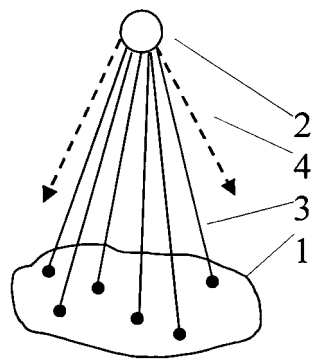
FIELD: computer equipment.

SUBSTANCE: invention relates to computer engineering. Method of controlling computer connections in a multilevel composite computer cluster comprising computer controllers - objects kO_i and communication modules kMS , characterized by that at each hierarchical level equal to $k = 0$ and more, leader - one of objects kO_i , directly connected to communication module kMS of hierarchy level k , sends command through kMS to communication modules of next level of hierarchy ^{k+1}MS and directly connected to them objects $^{k+1}O_i$, wherein ^{k+1}MS modules are sent

in kMS signal of beginning of barrier synchronization and interrupt transmission of signals from kMS to objects $^{k+1}O_i$, interacting only with their ^{k+1}MS under control of one of objects $^{k+1}O_i$ - leader of level hierarchy $k+1$, which, after completion of actions specified by command, stops transmission from ^{k+1}MS to kMS of barrier signal and restores communication between kMS and $^{k+1}O_i$.

EFFECT: technical result consists in wider range of the same purpose tools.

2 cl, 5 dwg



Фиг. 1

RU 2 7 4 4 5 9 1 C 2

RU 2 7 4 4 5 9 1 C 2

Изобретение относится к разделу вычислительной техники, в частности, к способам и техническим средствам, организующим взаимодействие компьютеров в компьютерных кластерах.

Под многоуровневым составным компьютерным кластером, выполняющим распределенные вычисления сетевыми средствами понимается кластер, состоящий из групп простых компьютерных кластеров. При этом в простом кластере его компьютеры обмениваются сообщениями с помощью средств, принадлежащих только этому кластеру. В составном кластере компьютеры, входящие в разные простые кластеры для обмена сообщениями должны привлекать внешние по отношению к этим кластерам средства. Многоуровневость составного кластера означает возможность объединить простые кластеры в иерархическую структуру так, что выше стоящие в иерархии кластеры могут управлять действиями ниже стоящих кластеров, и для любых двух простых кластеров существует иерархия кластеров, с помощью средств которых будет выполняться обмен сообщениями указанных кластеров.

В современных суперкомпьютерах и компьютерных кластерах используются группы компьютеров, которые передают по собственной инициативе сообщения и задания на выполнение распределенных вычислений другим компьютерам системы. Используется выполнение таких вычислений непосредственно сетевыми средствами (In-Network Computing), которые для этого содержат вычислительные средства - сетевые компьютеры. Возможно также проведение ряда вычислений сетевыми средствами, не имеющими сетевых компьютеров, (см. Компьютерная сеть с быстрой распределенной перестройкой своей структуры и обработкой данных в процессе их передачи, Проблемы управления №1. 2017. С. 47-56; Сетевая информационно-вычислительная поддержка взаимодействия подвижных роботов, Проблемы управления №5. 2018. С. 56-65.)

Вторая публикация взята в качестве прототипа для способа данного изобретения. Прототип ориентирован на имеющие компьютеры системы мобильных роботов, в которых компьютеры объединены беспроводными ненаправленными оптическими или радиоканалами и модулей связи.

Предлагаемые в патенте решения ориентированы на обеспечение возможностей прототипа, но для систем с отличающейся структурой. Поэтому более детально изложим основные возможности прототипа.

В прототипе каждый компьютер содержит сетевой контроллер управления обменом сообщениями с другими компьютерами системы; содержит модуль связи и направленные беспроводные оптические или радиоканалы связи компьютеров с модулем связи.

Каждый сетевой контроллер содержит источник и приемник оптических или радиосигналов. Сетевой контроллер содержит также устройство управления взаимодействием компьютера с источниками и приемниками сообщений в системе.

Цель предлагаемого изобретения - разработка способа и системы управления связями компьютеров в многоуровневом составном компьютерном кластере, выполняющем распределенные вычисления сетевыми средствами.

Для предлагаемого в патенте управления связями компьютеров в многоуровневом составном компьютерном кластере важны следующие компоненты структуры кластера прототипа и основные операции, в которых они участвуют. Эти компоненты - структура простого кластера, сетевые узлы, коммуникационный модуль. Операции - процедура запуска кластера, синхронизация, разрешение конфликтов доступа, асинхронная барьерная синхронизация, вычисления в сетевых средствах кластера.

Далее будут использованы следующие обозначения. Для модуля связи используется обозначение MS. Техническое средство компьютера - контроллер, выполняющий

взаимодействие компьютеров посредством MS, назовем объектом O_i .

А. Сетевая структура простого кластера, сетевые узлы - модуль связи, объекты. Каждый компьютер предлагаемого кластера содержит сетевой контроллер, который является первым из типов узлов для сети - объектом, представляющим этот компьютер в компьютерном кластере.

На фиг. 1 объекты O_i , обозначены как 1. Объект инициирует и обеспечивает выполнение всех необходимых распределенных операций в кластере.

Второй тип узла - это модуль связи MS, обозначенный здесь как 2. В сети имеется только один модуль связи. Объекты передают сигналы в модуль связи по каналам 3 и получают сигналы от модуля связи по каналам 4. Оба вида каналов могут быть направленными и ненаправленными, проводными и беспроводными, передающими оптические или радиосигналы. Взаимодействие между объектами через MS осуществляется следующим образом.

Объекты должны отправлять сигналы частоты f_1 для передачи двоичной единицы и частоты f_0 для передачи двоичного нуля в MS. Модуль MS преобразует эти сигналы соответственно в сигналы *f_1 и *f_0 других частот и отправляет их всем объектам. Данные сигналы имеют одинаковую длительность для всей сети.

Для управления асинхронными процессами объекты используют сигналы f_2 и *f_2 . Сигналы f_2 отправляют объекты. Каждый f_2 имеет продолжительность, требуемую конкретным процессом. Модуль MS отправляет сигналы *f_2 объектам, если он принимает хотя бы один сигнал f_2 .

Компьютеры кластера являются стационарными. Следовательно, можно один раз определить время прохождения сигнала между каждым объектом кластера и MS. Это время необходимо для получения высокой производительности всех распределенных вычислений. Его можно не измерять, если не требуются быстрые распределенные операции.

В. Синхронизация. Задача синхронизации - доставка в MS группы сообщений объектов в виде одного общего сообщения, содержащего сообщения объектов, следующих одно за другим без временных пауз между ними, а также в виде одного сообщения с совпадением одинаковых битов сообщений объектов. Применяются три вида синхронизации: синхронизация с помощью специальной команды, синхронизация с устранением конфликтов доступа MS и барьерная синхронизация.

В первом варианте один из объектов отправляет команду синхронизации через MS всем объектам. После ее получения объекты отправляют сообщения в MS с задержкой, учитывающей удаленность объекта от MS. Объекту известен интервал времени $T_{\max}$ - интервал не менее времени распространения сигнала между самым удаленным объектом и MS. После получения команды объект отправляет сообщение с задержкой $D_i = T_{\max} - T_i$, где T_i - время распространения сигнала между O_i и MS.

В этом случае биты сообщений с одинаковыми именами поступают в MS одновременно от всех объектов, и все объекты получают общее сообщение от MS - результат наложения сигналов от объектов сообщений.

Во втором варианте, при отсутствии сигналов от MS, объекты начинают передавать сообщения, обнаруживают конфликт в MS и отправляют двоичную шкалу в MS. Шкала имеет количество битов, равное количеству объектов, и имеет цифру один в битах, соответствующих объектам, которые должны отправить сообщение. Получив эту шкалу

от MS, объекты упорядочение отправляют сообщения. Этим конфликт устраняется для всех конфликтующих объектов.

При асинхронной барьерной синхронизации участвующие в ней объекты посылают сигналы в MS на частоте f_2 . Модуль MS возвращает объекту сигнал с частотой f_2^* . По завершении работы объект удаляет сигнал f_2 . Исчезновение сигнала f_2^* в объектах является моментом синхронизации.

С. Представление данных в кластере. Распределенные данные обрабатываются либо полностью в модуле связи, либо с выполнением в нем значительной части таких вычислений.

Для этого используются парафазное кодирование двоичных цифр и представление цифр чисел в виде логических шкал.

При парафазном кодировании цифра «один» представлена последовательностью сигналов $f_1 f_0$, цифра «ноль» - последовательностью $f_0 f_1$.

При использовании логической шкалы каждая цифра числа, заданного в системе счисления с произвольным основанием p , представлена двоичной шкалой, число битов в которой равно p . В шкале только один бит, соответствующий значению цифры, равен 1, а остальные равны 0. Например, для $p=10$ и цифры 7 шкала имеет вид 001000000.

Примеры вычислений.

Д. Логические операции И, ИЛИ. Объекты синхронно передают биты данных в MS, используя парафазное представление двоичных цифр 1 и 0. Если выполняется операция И, комбинация сигналов $f_1 f_0$, которая возникает в результате наложения сигналов в MS, дает результат операции, равный 1, в противном случае получается результат равный 0. Для операции ИЛИ комбинация $f_0 f_1$ дает результат операции 0, в противном случае 1.

Эти операции выполняются в MS без задержки сигнала; Время расчета не зависит от количества участников операции. Логические элементы не используются для расчетов в MS.

Е. Поиск MAX и MIN. Для вычисления MAX, объекты передают в MS наибольшую цифру своего числа, представленную в виде логической шкалы, и при получении от MS шкалы с несколькими единицами определяют наибольшее значение этой цифры для всех объектов одновременно. Следующая цифра передается объектами, которые передали наибольшую предшествующую цифру. MAX будет получен после передачи всех цифр номеров. Для расчета MIN объекты определяют минимальные значения цифр. Увеличение базы p ускоряет операцию, поскольку уменьшается количество обменов сообщениями.

Ф. Аналого-цифровые операции. В прототипе показано, как при наличии в модуле связи аналого-цифрового преобразователя уровней сигналов непосредственно в сетевом модуле выполняются быстрые сетевые операции суммирования. Время суммирования не зависит от количества объектов, участвующих в операции.

Недостаток способа в прототипе в том, что в нем не предусмотрена возможность разделить все компьютеры системы на иерархически организованные группы - кластеры так, чтобы в каждом кластере одновременно выполнялась независимая (автономная) обработка данных, с обменом сообщениями в пределах кластера. Это снижает производительность такой системы.

В качестве прототипа системы, взята система, приведенная в указанной выше статье - Сетевая информационно-вычислительная поддержка взаимодействия подвижных

роботов, Проблемы управления №5. 2018. С. 56-65.

Недостаток системы в прототипе состоит в том, что она не позволяет использовать одновременно несколько иерархически связанных модулями связи кластеров для одновременного автономного выполнения работ в нескольких кластерах с независимыми обменами сообщениями в кластере. Это замедляет процесс обработки данных.

Задача настоящего изобретения для способа состоит в устранении недостатка прототипа - невозможности обработки данных одновременно несколькими группами компьютеров, взаимодействующих в пределах группы и разработка для этого способа управления иерархическими связями компьютеров в многоуровневом составном компьютерном кластере, использующем группу модулей связи для выполнения распределенных вычислений сетевыми средствами.

Техническим результатом для способа является включение в него способов взаимодействия компьютеров с группой модулей связи и способов управления иерархическими связями компьютеров в многоуровневом составном компьютерном кластере, использующем группу модулей связи для выполнения распределенных вычислений сетевыми средствами.

Технический результат способа достигается тем, что на каждом уровне иерархии, равном $k=0$ и более, лидер - один из объектов kO_i , непосредственно связанных с модулем связи kMS уровня иерархии k , посылает команду через kMS указанным в команде модулям связи следующего уровня иерархии ${}^{k+1}MS$ и непосредственно связанным с ними объектам ${}^{k+1}O_i$, при этом модули ${}^{k+1}MS$ посылают в kMS сигнал начала барьерной синхронизации и прерывают передачу сигналов от kMS объектам ${}^{k+1}O_i$, связанным с этими ${}^{k+1}MS$, которая впоследствии восстанавливается по команде от kO_i , и эти ${}^{k+1}O_i$ продолжают работать автономно, взаимодействуя только со своим ${}^{k+1}MS$ под управлением одного из объектов ${}^{k+1}O_i$ - лидера уровня иерархии $k+1$, который после завершения заданных командой действий останавливает передачу от ${}^{k+1}MS$ к kMS сигнала барьерной синхронизации и восстанавливает связь между kMS и ${}^{k+1}O_i$ в результате чего kMS , ${}^{k+1}MS$, kO_i и ${}^{k+1}O_i$ начинают посылать сигналы связанным с ними непосредственно модулям связи и объектам и принимать от них сигналы.

Задача настоящего изобретения для системы устройств состоит в устранении указанного недостатка системы устройств прототипа - невозможности использовать одновременно несколько модулей связи и создания организованного многоуровневого составного компьютерного кластера, использующего эту группу модулей связи для выполнения распределенных вычислений сетевыми средствами.

Технический результат для системы устройств состоит в том, что она решает указанную задачу создания иерархическими организованного многоуровневого составного компьютерного кластера, использующего группу модулей связи для выполнения распределенных вычислений сетевыми средствами.

Технический результат для системы достигается тем, что на каждом уровне иерархии $k \geq 0$, она содержит модули связи kMS и контроллеры в компьютерах - объекты kO_i , причем беспроводными или проводными линиями связи выполнены соединения

соответственно между ${}^k\text{MS}$ и ${}^{k+1}\text{MS}$, между ${}^k\text{MS}$ и ${}^k\text{O}_i$, между ${}^{k+1}\text{MS}$ и ${}^{k+1}\text{O}_i$, и в каждом из указанных устройств есть приемники посланных ему сигналов и источники посылаемых им сигналов приемникам этих устройств, а в модулях связи ${}^{k+1}\text{MS}$ содержатся блоки, разрывающие и восстанавливающие связь между устройствами уровней иерархии k и $k+1$, и блок управления ими, причем ${}^k\text{O}_i$ содержат также блок управления для обмена командами и сигналами с компьютером.

Техническая сущность и принцип действия предложенной системы устройств поясняются чертежами.

Фиг. 1. Простой кластер.

Фиг. 2. Составной кластер.

Фиг. 3. Структура модуля связи.

Фиг. 4. Блок изоляции объектов соседних уровней иерархии.

Фиг. 5. Структура объекта.

Приведенные на чертежах устройства следует рассматривать как примеры реализации предлагаемых в патенте устройств.

Краткое описание предложенного способа.

Основная особенность способа управления связями компьютеров в многоуровневом составном компьютерном кластере следующая. На каждом уровне k иерархии составного кластера, отличном от верхнего нулевого уровня, один из объектов уровня k через модуль ${}^k\text{MS}$ посылает команду в модули связи уровней больше k , которая на время своего действия изолирует в указанных командой модулях все непосредственно связанные с ними объекты от сигналов предшествующих уровней. Это позволяет выделенным таким способом компьютерам действовать взаимосвязано и изолированно от других устройств, увеличивая общее быстродействие составного кластера.

Краткое описание предложенной системы устройств.

Основная особенность системы управления связями компьютеров в многоуровневом составном компьютерном кластере состоит в том, что на каждом уровне иерархии $k \geq 0$, она содержит модули связи ${}^k\text{MS}$, а в компьютерах содержит контроллеры - объекты ${}^k\text{O}_i$. Соединения между ${}^k\text{MS}$ и ${}^{k+1}\text{MS}$, между ${}^k\text{MS}$ и ${}^k\text{O}_i$, между ${}^{k+1}\text{MS}$ и ${}^{k+1}\text{O}_i$ выполнены беспроводными или проводными линиями связи. В каждом из указанных устройств есть приемники посланных ему сигналов и источники посылаемых им сигналов. Модули связи ${}^{k+1}\text{MS}$ содержат блоки, разрывающие и восстанавливающие связь между устройствами уровней иерархии k и $k+1$, и блок управления ими. Объекты ${}^k\text{O}_i$ содержат блок управления для обмена командами и сигналами с компьютером.

Подробное описание предложенного способа.

Способ применяется к многоуровневому составному кластеру, содержащему центральный кластер и подчиненные кластеры. Он ориентирован на развитие структуры фиг. 1. В составном кластере рассматривается способ, состоящий из двух взаимодействующих способов - синхронного и асинхронного (автономного).

Синхронный способ функционирования составного кластера.

Организация взаимодействия узлов составного кластера, состоящего из центрального и подчиненных кластеров, показана на фиг. 2. Приведенным ниже компонентам составного кластера на фиг. 2 присвоены числа, указанные ниже в тексте в скобках.

Здесь ${}^0\text{MS}$ (5) - основной модуль связи нулевого уровня, принадлежащий

центральному кластеру. Объекты ${}^0O_1, \dots, {}^0O_i$ (6) напрямую связаны с 0MS . Подчиненные кластеры первого уровня 1O_i (8) с номерами $i=1, \dots, k$ также подключены к 0MS через свои модули связи первого уровня ${}^1MS_1, \dots, {}^1MS_k$ (7). Волоконно-оптические каналы (9) подключают объекты к модулям первого уровня: к модулю 1MS_1 - объекты ${}^1O_{1.1}, \dots, {}^1O_{1.l}$, к модулю 1MS_k -объекты ${}^1O_{k.1}, \dots, {}^1O_{k.m}$. Модули 1MS подключены к 0MS также каналами (9).

Для всех объектов составного кластера известны их удаленность от 0MS и наибольшая удаленность всех объектов от 0MS .

Модуль 0MS работает, как описанный выше модуль MS простого кластера. Модули 1MS_1 принимают сигналы f_0, f_1 и f_2 подключенных к ним объектов и, не изменяя частоту сигналов, передают их в 0MS . Поэтому 0MS воспринимает модули 1MS как объекты, соединенные с каналами (9).

Модуль 0MS посылает сигналы ${}^*f_1, {}^*f_0$ и *f_2 не только объектам, но и модулям 1MS .

Последние, как и 0MS , посылают эти сигналы своим объектам.

Таким образом, модули 1MS выполняют только функции повторителей (ретрансляторов) сигналов, принимаемых в 0MS от объектов. Система с модулями 1MS действует как система, содержащая только 0MS . Отличие только в том, что сигналы объектов любого кластера непосредственно не поступают к объектам других кластеров.

Процесс синхронизации в составном кластере должен быть уточнен. Пусть для упрощения кластеры для синхронизации используют шкалы синхронизации одинакового числа двоичных цифр. Количество подчиненных кластеров известно. Процесс синхронизации включает в себя все кластеры, которые передают свои шкалы в общем сообщении. От 0MS общая шкала вернется ко всем объектам. Это позволит объектам кластера синхронно отправлять свои сообщения одинаковой продолжительности.

В этом процессе должно быть выполнено дополнительное условие. Сообщения объекта должны поступать без конфликтов не только на 0MS , но и на 1MS . Поскольку все расстояния между объектами 0MS и 1MS известны, для этого каждый объект должен отправлять любое из своих сообщений с дополнительной задержкой, учитывающей расстояние между 0MS и его 1MS , а также порядок передачи сообщений в 0MS от других 1MS , предшествующих в шкале синхронизации 1MS этого объекта.

В таком составном кластере все методы обмена данными и цифровые операции аналогичны действиям кластера, который не имеет подчиненных кластеров. Также легко организовать работу со шкалами и сообщениями различной, но заранее известной длины. Можно работать с динамически изменяемыми длинами сообщений, но для этого необходимо указать эти длины в шкале синхронизации. Вместо этого можно отправить обычную шкалу синхронизации, за которой следует более короткая шкала, отправляемая только объектами, которые передают сообщения. Во второй шкале сообщается длина сообщений.

Синхронный способ может применяться независимо, но он также входит в состав следующего автономного способа.

Автономный способ функционирования составного кластера.

В отличие от предыдущего раздела, объекты, подключенные к модулям 1MS , работают в этом кластере автономно, не задерживая работу объектов в других кластерах. Вернемся сначала к простому кластеру. Пусть один из объектов U в кластере отправляет задание другим объектам. Каждый из этих объектов запускает асинхронную операцию, отправляет сигнал f_2 барьерной синхронизации в 0MS и завершает передачу f_2 после завершения задачи U . Модуль 0MS в ответ на f_2 отправляет сигнал *f_2 объектам и завершает передачу *f_2 при прекращении прихода к нему сигнала f_2 . При отсутствии *f_2 объект U может отправлять новое задание.

Для упрощения изложения действия модуля 1MS далее для упрощения изложения представлены в виде взаимодействия двух его частей 1MS_a и 1MS_b . Хотя эти части могут быть физически реализованы на полностью или частично совпадающих технических средствах, их логическое функционирование различно. Перевод центральным кластером модуля 1MS в автономный режим состоит из следующих этапов.

Этап 1. Произвольный объект центрального кластера - U отправляет команду $C1$.

Модуль 1MS , получив команду $C1$, сообщает подключенным к нему объектам о предстоящей автономной работе. Модули 1MS и их объекты, которые получают команду $C1$, ждут команду $C2$ от объекта U . В $C2$ перечислены имена объектов - получателей $C2$. Для каждого объекта указывается действие или последовательность действий (программа), которые должен выполнить объект. Модуль 1MS выполняет этап 2.

Этап 2. После получения $C2$ модуль 1MS запрещает прохождение сигналов от 0MS к объектам данного модуля 1MS и сигналов этих объектов к 0MS . Часть 1MS_a отправляет сигнал f_2 - признак барьерной синхронизации в 0MS . От модуля 0MS поступает сигнал *f_2 . Объекты модуля 1MS будут асинхронно выполнять задачи, указанные в команде $C2$, выполняя этап 3.

Этап 3. Задание, полученное от 0MS , выделяет произвольный объект *U среди объектов 1MS . Объект *U должен управлять действиями остальных объектов модуля 1MS . Объект *U действует подобно объекту U центрального кластера. Объект *U также может быть выбран совместными действиями объектов 1MS .

Каждый объект, получивший команду от *U , отправляет части 1MS_b сигнал f_2 , а 1MS_b превращает сигнал f_2 в сигнал *f_2 , отправляемый объектам модуля 1MS .

После завершения работы объект отключает свой сигнал f_2 . При отсутствии сигналов f_2 модуль 1MS_b отключает сигнал *f_2 . В ответ объект *U либо продолжает работать с объектами 1MS , либо завершает работу и отправляет команду в часть 1MS_a для завершения отправки сигнала f_2 в модуль 0MS . Объект *U также отменяет изоляцию своих объектов от центрального кластера.

Этап 4. Когда модуль 0MS прекратит прием сигналов f_2 , он прекратит передачу

сигнала *f_2 объектам и модулям 1MS . После выполнения этапов 1-4 результаты вычислений в модулях 1MS остаются в объектах, подключенных к этим модулям, и недоступны в объектах, подключенных к 0MS . Для доступа к ним выполняется этап 5.

Этап 5. Предполагается, что в команде C2 объект *U получил задачу собрать данные из объектов, подключенных к 1MS , и сформировать из них данные, необходимые для объектов, подключенных к 0MS . Объект *U выполняет эти действия. Аналогично объект U собирает через модуль 0MS все данные от объектов.

Таким образом, разделение модуля 1MS на части 1MS_a и 1MS_b позволяет взаимодействовать двум асинхронным процессам. Часть 1MS_a обеспечивает подключение модуля 1MS как объекта к 0MS . Часть 1MS_b позволяет объектам, подключенным к модулю 1MS , работать автономно от всех объектов, подключенных к 0MS .

Подчиненные модули 1MS по сравнению с 0MS должны иметь дополнительное оборудование для выполнения вышеуказанных действий.

Этапы 1-5 могут быть изменены, чтобы включать количество уровней иерархии k , превышающее рассмотренное выше $k=0, 1$. Введем модули kMS с $k>1$ и подключим их в систему подобно подключению модулей 1MS к модулю 0MS . Если любой модуль связи ^{k+1}MS получает команду, разрывающую его соединение с вышестоящим в иерархии модулем связи kMS , тогда этот ^{k+1}MS по отношению к нижестоящим модулям связи и объектам действует подобно 1MS по отношению к 0MS . Это создает многоуровневую структуру.

Если модули и объекты в текущий момент времени не используют автономный способ взаимодействия, то их действия не отличаются от действий при синхронном способе функционирования составного кластера.

В составном кластере любой компьютер может быть заменен составным кластером.

Подробное описание предлагаемой системы управления связями компьютеров в многоуровневом составном компьютерном кластере.

Предлагаемая система представлена для упрощения только нулевым и первым уровнями иерархии. Она показана на фиг. 2 и содержит модуль связи нулевого уровня 0MS (5), модули связи первого уровня $^1MS_1 \dots ^1MS_k$ (7), объекты нулевого уровня $O_1 - O_i$ (6), объекты первого уровня $^1O_1 - O_m$ (8). Перечисленные устройства соединены каналами связи (9).

На фиг. 3 приведено устройство модулей kMS 7, первого и более высоких уровней иерархии.

Устройство модулей связи 0MS и MS не рассматривается, так как оно непосредственно следует из устройства модулей 7. На фиг. 3 слева общий канал 9 представлен каналами 3 и 4, соединяющими модуль связи 7 с предшествующим уровнем иерархии, аналогично справа каналы 3 и 4 соединяют модуль связи со следующим уровнем иерархии.

Модуль 7 связи содержит блоки 13 и 11 приемников сигналов f_0, f_1, f_2 и $^*f_0, ^*f_1, ^*f_2$ соответственно, переводящих их в электрические сигналы; блоки 12 и 14 источников

сигналов f_0, f_1, f_2 и $^*f_0, ^*f_1, ^*f_2$ соответственно, формирующих их в ответ на приход электрических сигналов; содержит блок 10 изоляции объектов соседних уровней иерархии, организация которого представлена на фиг. 4; содержит блок управления

В блоке 15 его входы 16, 20 подключены к выходам блоков 11, 13 соответственно, выходы 17, 19 подключены к входам блоков 12, 14, выход 18 подключен управляющему входу блока 10, который на фиг. 4 обозначен как 21. Здесь блок 21 содержит ключи 22, 23, 25. Для выполнения действий способа по изоляции объектов соседних уровней иерархии блок 15 управляет ключами 22, 23. Для разрыва связей посланный на вход 24 (фиг. 4) сигнал размыкает 22, 23 и замыкает 25.

В прототипе используются аналого-цифровые операции, применяющие АЦП. Чтобы иметь такую возможность в предлагаемой системе, введен узел 26, включающий АЦП. Если с выхода 13 на узел 26 поступает требование выполнения аналого-цифровой операции, то на время выполнения этой операции и передачи ее результата в систему узел размыкает ключ 25.

Конструкция объекта показана на фиг. 5. Здесь 27 - объект, т.е. сетевой контроллер компьютера, входящего в один из кластеров; 28 - блок управления; 29 - блок источников сигналов f_0, f_1, f_2 , посылаемых в модуль связи по каналу 3; 30 - блок приемников

сигналов $^*f_0, ^*f_1, ^*f_2$, принимаемых от модуля связи по каналу 4. Блок управления 28 соединен с 29 и 30, а также соединен с компьютером кластера каналами 31 и 32, по которым передаются команды и сигналы, обеспечивающие действия предлагаемого способа.

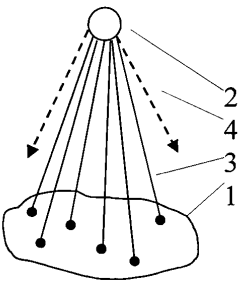
(57) Формула изобретения

1. Способ управления связями компьютеров в многоуровневом составном компьютерном кластере, содержащем контроллеры компьютеров - объекты kO_i и модули связи kMS , характеризующийся тем, что на каждом уровне иерархии, равном $k=0$ и более, лидер - один из объектов kO_i , непосредственно связанных с модулем связи kMS уровня иерархии k , посылает команду через kMS указанным в команде модулям связи следующего уровня иерархии ^{k+1}MS и непосредственно связанным с ними объектам $^{k+1}O_i$, при этом модули ^{k+1}MS посылают в kMS сигнал начала барьерной синхронизации и прерывают передачу сигналов от kMS объектам $^{k+1}O_i$, связанным с этими ^{k+1}MS , которая впоследствии восстанавливается по команде от kO_i , и эти $^{k+1}O_i$ продолжают работать автономно, взаимодействуя только со своим ^{k+1}MS под управлением одного из объектов $^{k+1}O_i$ - лидера уровня иерархии $k+1$, который после завершения заданных командой действий останавливает передачу от ^{k+1}MS к kMS сигнала барьерной синхронизации и восстанавливает связь между kMS и $^{k+1}O_i$, в результате чего kMS , ^{k+1}MS , kO_i и $^{k+1}O_i$ начинают посылать сигналы связанным с ними непосредственно модулям связи и объектам и принимать от них сигналы.

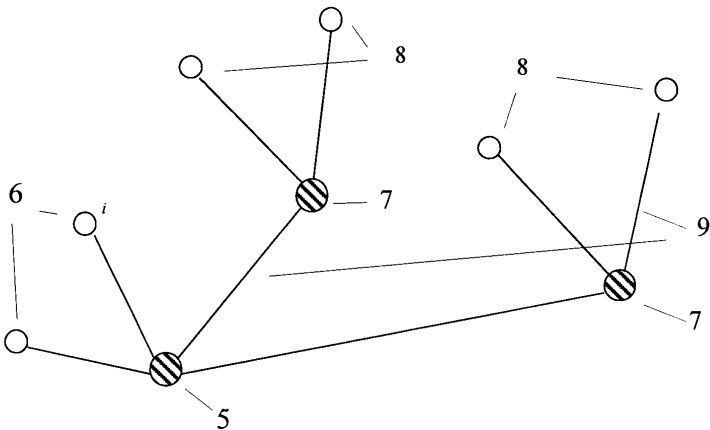
2. Система управления связями компьютеров в многоуровневом составном компьютерном кластере, характеризующаяся тем, что на каждом уровне иерархии $k \geq 0$,

она содержит модули связи ${}^k\text{MS}$ и контроллеры в компьютерах - объекты ${}^k\text{O}_i$, причем беспроводными или проводными линиями связи выполнены соединения соответственно между ${}^k\text{MS}$ и ${}^{k+1}\text{MS}$, между ${}^k\text{MS}$ и ${}^k\text{O}_i$, между ${}^{k+1}\text{MS}$ и ${}^{k+1}\text{O}_i$, и в каждом из указанных устройств есть приемники посланных ему сигналов и источники посылаемых им сигналов приемникам этих устройств, а в модулях связи ${}^{k+1}\text{MS}$ содержатся блоки, разрывающие и восстанавливающие связь между устройствами уровней иерархии k и $k+1$, и блок управления ими, причем ${}^k\text{O}_i$ содержат также блок управления для обмена командами и сигналами с компьютером.

1

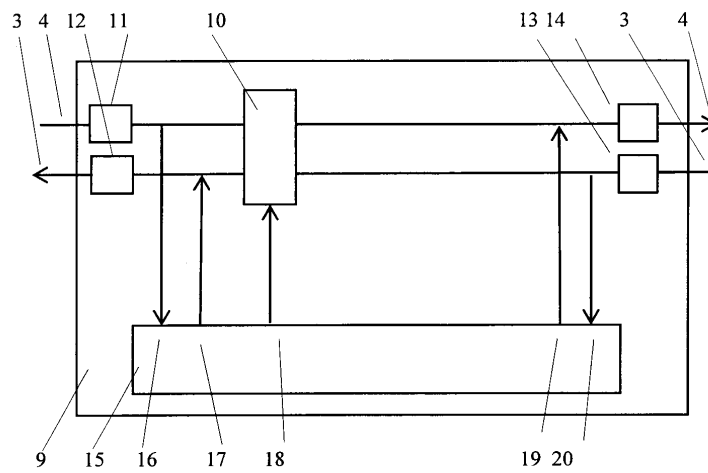


Фиг.1

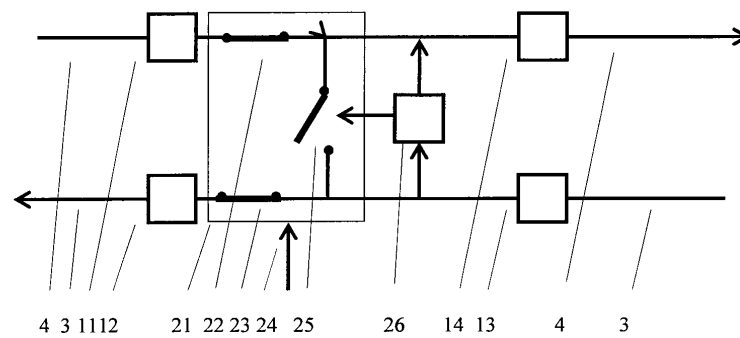


Фиг. 2

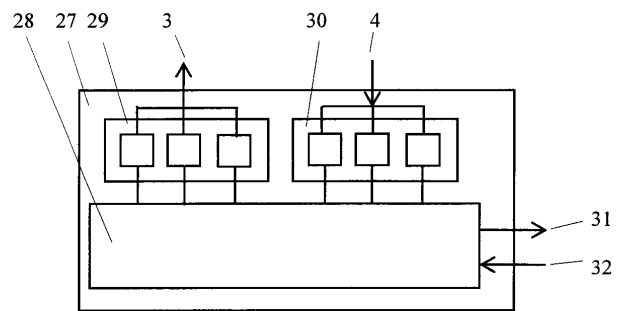
2



Фиг. 3



Фиг. 4



Фиг. 5