

УДК 004.75

Ю.В. Попов, А.В. Балута

Донецкий национальный технический университет
a.mag@front.ru

Алгоритм перераспределения вычислительной нагрузки в сети при распределённом логическом моделировании цифровых систем

Рассмотрена организация иерархической структуры вычислительной сети для распределённого логического моделирования цифровых систем. Определены характеристики, вычисляемые в процессе моделирования, которые используются в алгоритме перераспределения нагрузки на процессоры вычислительной сети

Распределённое логическое моделирование, цифровая схема, логический процессор, коммуникационный логический процессор (КЛП), сфера влияния, вычислительная сеть, перераспределение нагрузки, линия связи, архитектура вычислительной сети, поддерево КЛП, равновесие нагрузки, элемент схемы, узел схемы, внешнее событие, протокол синхронизации, оптимистическое моделирование

Введение

При повышении требований к производительности систем моделирования цифровых логических систем возникает необходимость использования распределённых вычислений [1,2,3]. При этом необходимо учитывать возможность возникновения неравномерной нагрузки на процессоры в сети. Кроме того, в условиях использования большого числа моделирующих процессоров в сети не исключена возможность внезапного отключения одного из компьютеров в сети либо потеря связи между некоторыми из них.

Вопросы, связанные с распределённым моделированием, исследовались, начиная с 70-х годов XX века. Важнейшим этапом в разработке систем распределённого моделирования является разработка алгоритмов синхронизации, которые используются при организации моделирования с одновременным использованием нескольких компьютеров в сети. Первые алгоритмы синхронизации были предложены в конце 70-х годов. Эти алгоритмы относятся к классу консервативных алгоритмов. Они описаны в работах Chandy K. M., Misra J. и Bryant R.E. [2]. В начале 80-х годов D. Jefferson предложил Time Warp алгоритмы (алгоритмы деформации времени). Эти алгоритмы были использованы при разработке оптимистических алгоритмов синхронизации [4,5,6].

Для систем распределённого моделирования имеется ряд исследований, связанных с организацией вычислительной сети и перераспределением моделируемой системы для ускорения моделирования [7]. Различные варианты решения задач, связанных с

использованием многоуровневой архитектуры моделирующей системы и более эффективного распараллеливания вычислений при распределённом моделировании рассматриваются в работах Ferscha и Tripathi [5]. Вопросы, связанные с организацией механизма управления вычислительной нагрузкой при распределённом моделировании цифровых схем рассматриваются в работах Schlagenhaft и Ruhwandl [8]. У Микова и Замятиной содержится описание методов динамической балансировки вычислительной нагрузки в системе распределённого моделирования Triad.Net [9].

Использование древовидной организации вычислительной сети при распределённом моделировании исследуется в работах Theodoropoulos и Logan, которые предлагают использовать данную архитектуру для улучшения декомпозиции моделируемой системы и упрощения решения задач, связанных с получением равномерной нагрузки на процессоры в сети [10].

Объект исследований данной статьи – организация вычислительной сети для распределённого логического моделирования цифровых систем с динамической топологией связей процессоров. Предмет исследования – параметры, характеризующие работу процессоров в сети при распределённом моделировании.

Цель исследований в статье – получение алгоритма, который на основании характеристик работы моделирующих процессоров выполняет перераспределение дерева процессоров и задания, выполняемого процессорами.

Задачи исследования в данной статье:

- рассмотреть иерархическую организацию вычислительной сети в системе распределённого

логического моделирования цифровых систем;

- выделить факторы, влияющие на нагрузку процессоров при моделировании с учетом архитектуры вычислительной сети;
- определить параметры, учитывающие наиболее существенные из этих факторов;
- разработать алгоритм перераспределения вычислительной нагрузки на процессоры в сети при использовании иерархической организации распределённой системы.

Архитектура вычислительной сети

Система распределённого логического моделирования используется для моделирования цифровых устройств, представленных их логической схемой. В задание моделирования входит структурно-функциональная модель схемы и входное воздействие в виде списка значения сигналов, подаваемых на входные узлы схемы в определенные моменты модельного времени.

Моделируемая схема разрезается на части, каждая из которых моделируется на отдельном логическом процессоре (ЛП). При использовании распределённых вычислений в системе возникают события в узлах схемы на одном процессоре, которые влияют на узлы схемы, относящихся к другим процессорам, вследствие чего возникает необходимость передачи сообщений о таких событиях по сети [1,2,5]. Частота передачи сообщений по сети напрямую связана с тем, как именно выполнено разрезание схемы.

В разработанной системе используется модификация предложенного в [10] подхода к организации моделирующей системы на основании управления сферами влияния. Логические процессоры (ЛП) выполняют моделирование части схемы, входящей в их сферу влияния. Сфера влияния $SE(LP_i)$ представляет собой некоторое подмножество A множества B всех элементов схемы, которые LP_i должен пересчитывать (1):

$$SE(LP_i) = \{x : x \subset A\}, A \subseteq B \quad (1)$$

Если при моделировании возникает событие, которое говорит о том, что изменяется значение в некотором узле схемы и этот узел является входным для элемента в сфере влияния LP_i , то LP_i должен пересчитать значения на выходах этого элемента и сгенерировать соответствующие события.

При необходимости передачи сообщений между ЛП, данные передаются не напрямую, а через соответствующие коммуникационные логические процессоры (КЛП) [10]. Каждый из КЛП также имеет свою сферу влияния, то есть объединение множеств элементов, входящих в сферы влияния подчиненных ему ЛП и КЛП.

В вычислительной сети имеется администратор процесса моделирования (АПМ),

содержащий базу данных всех процессоров, доступных в сети (см. рис. 1а). Вначале моделирования используется один КЛП (корневой КЛП0), к которому присоединяется один ЛП (ЛП1) (см. рис. 1б). При этом остальные доступные в сети процессоры пока свободны.

Если нагрузка на ЛП1 при моделировании превышает некоторую границу (например значение, задаваемое пользователем), КЛП0 может запросить дополнительный ЛП, которому передать часть сферы влияния ЛП1. В процессе моделирования для уменьшения нагрузки на КЛП0 может потребоваться использование дополнительных КЛП, которые становятся подчиненными для КЛП0. Каждый из КЛП также может запрашивать в сети свободные ЛП и КЛП для устранения перегрузок и узких мест в системе [10]. Данная стратегия приводит к древовидной структуре вычислительной сети, где КЛП являются узлами дерева, а ЛП – листьями дерева [10] (см. рис. 1в).

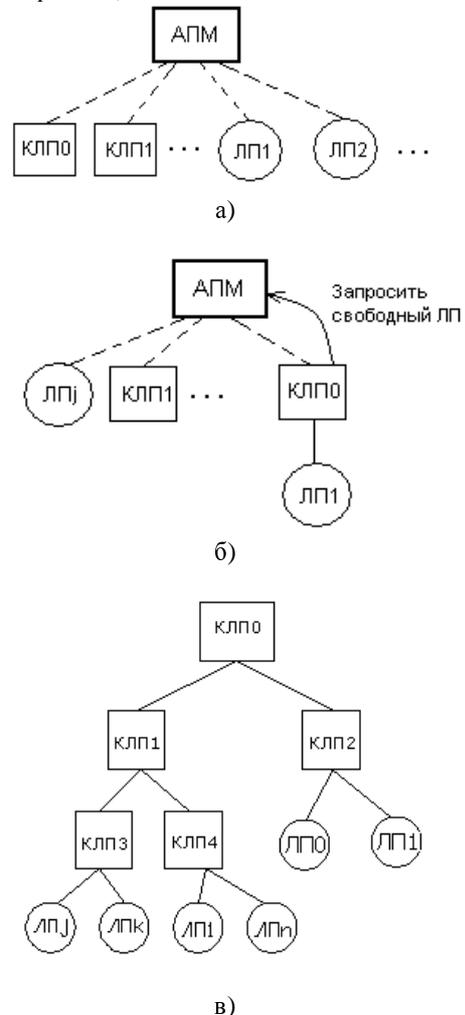


Рисунок 1 – Генерация дерева КЛП

а) АПМ и процессоры в сети

б) начало моделирования и запрос свободного ЛП

в) дерево КЛП после перераспределения

**Алгоритм перераспределения на
основании нагрузки на процессоры**

Особенности, связанные с древовидной архитектурой вычислительной сети, позволяют определить параметры нагрузки на процессоры в сети. На основании этих параметров можно принимать решение о необходимости перераспределения сфер влияния процессоров или изменений в структуре дерева процессоров.

Для ЛП используется параметр загрузки $Ld(LPi)$, вычисляемый для промежутка времени $[t1, t2]$ и соответствующий количеству операций по пересчету значений на выходах элементов схемы в сфере влияния $SE(LPi)$. $Ld(LPi)$ зависит от количества элементов, которые находятся в сфере влияния ЛПи.

Нагрузка в поддереве КЛП на линию связи между ЛПи и ЛПj $Z_{i-j}[t1, t2]$ представляет собой количество сообщений о событиях, передаваемых от ЛПи к ЛПj через этот КЛП в интервале времени $[t1, t2]$. На КЛП можно вычислить значение нагрузки на линии связи Z_{i-j} между всеми возможными группами по два ЛП в пределах поддерева КЛП (см. рис. 2).

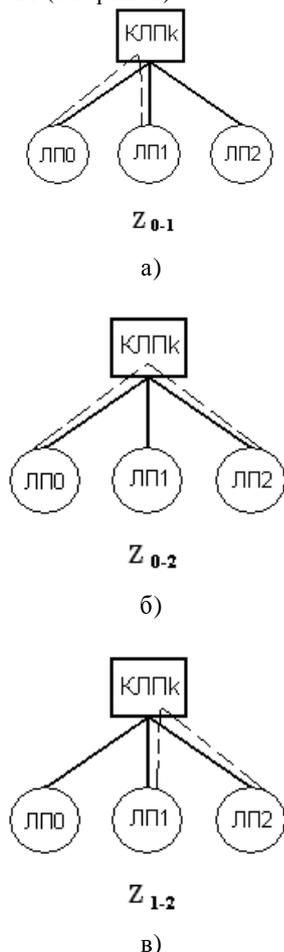


Рисунок 2 – Нагрузка на линии связи в поддереве КЛПк а) при передаче сообщений между ЛПО и ЛП1 б) между ЛПО и ЛП2 в) между ЛП1 и ЛП2

Нагрузка на линии связи Z_{i-j} вычисляется для интервала времени $[t1, t2]$ и используется для определения равновесия нагрузки на линии связи в поддереве КЛП (2) [11]:

$$\text{Равновесие (КЛП}_k) = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (z_i - \bar{z})^2}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (z_i)^2}} = \frac{\sqrt{\sum_{i=1}^n (z_i - \bar{z})^2}}{\sqrt{\sum_{i=1}^n (z_i)^2}}, \quad (2)$$

где z_i – нагрузка на i -ю линию связи в поддереве КЛП, n – количество линий связи между всеми возможными группами по два ЛП в поддереве КЛП;

\bar{z} – среднее значение нагрузки,

$\sqrt{\frac{1}{n-1} \sum_{i=1}^n (z_i - \bar{z})^2}$ – это значение стандартного отклонения нагрузки на линии связи от среднего значения,

$\frac{1}{n-1} \sum_{i=1}^n (z_i)^2$ – это наибольшее значение дисперсии нагрузки на линии связей в поддереве.

Значение равновесия КЛП может изменяться от 0 в наилучшем случае (когда загрузка всех линий связи одинакова) до 1, когда одна из линий связи перегружена, в то время как остальные линии связи не используются [11].

Если равновесие КЛП превышает некоторое пороговое значение, то можно попытаться уменьшить нагрузку на КЛП путем передачи группы ЛПи–ЛПj с наибольшим значением Z_{i-j} другому КЛП, доступному в сети. Также для уменьшения нагрузки на КЛП в этом случае можно выполнить объединение сфер влияния группы процессоров ЛПи–ЛПj на одном из них (ЛПи или ЛПj) и как следствие – освободить один из этих двух ЛП.

Алгоритм перераспределения задания моделирования, использующий параметры нагрузки на ЛП и линии связей на КЛП, представлен на рисунке 3. Этот алгоритм позволяет генерировать дерево КЛП и перераспределять сферы влияния ЛП во время выполнения моделирования. Также в алгоритме используются параметры, которые могут устанавливаться непосредственно перед началом моделирования и выбираться в соответствии с особенностями моделируемой схемы и входного воздействия в задании моделирования. При использовании алгоритма необходимо ограничивать количество выполняемых в системе изменений.

Важной особенностью древовидной организации распределенной системы является то, что перераспределение сфер влияния группы ЛП во время моделирования требует приостановки выполнения моделирования только на этих ЛП и не влияет на работу остальных ЛП и КЛП в сети.

Установить параметры:

```
ConstLoad; //максимальная загрузка ЛП
ConstBalance; //предельное значение
//равновесия в поддереве на КЛП
если КЛПi
(БД_процессоров.количество_ЛП=1) то
если Загрузка (ЛП ∈ КЛПi) > ConstLoad то
запросить_свободный_ЛП;
если КЛПi
(БД_процессоров.количество_ЛП = 2) то
если Загрузка (ЛП1 ∈ КЛПi) > ConstLoad
или Загрузка (ЛП2 ∈ КЛПi) > ConstLoad
то запросить_свободный_ЛП;
если КЛПi
(БД_процессоров.количество_ЛП > 2) то
Вычислить_равновесие;
если равновесие > ConstBalance то
определить_максимально
загруженную_линию (ЛПi, ЛПj);
если ∃ в сети (КЛПk) то
передать (ЛПi, ЛПj) на КЛПk
иначе
объединить_сферы_влияния
(ЛПi, ЛПj) на ЛПi;
освободить_от_задания (ЛПj);
end;
end;
end;
```

Рисунок 3 – Алгоритм перераспределения на основании нагрузки на ЛП и КЛП

При выполнении перераспределения сфер влияния в поддереве некоторого КЛП на этот КЛП могут приходить события от других ЛП, которые временно не могут быть отправлены тем ЛП, которые должны их обработать. Поэтому для КЛП должен быть предусмотрен буфер для хранения таких событий, которые после завершения перераспределения в поддереве КЛП должны быть разосланы согласно изменившимся сферам влияния подчиненных ЛП.

Перераспределение сфер влияния ЛП на основании топологии моделируемой схемы

В некоторых случаях добавление новых ЛП не приводит к ускорению выполнения моделирования, а перераспределение сфер влияния ЛП может уменьшить нагрузку на линии связи соответствующих КЛП, поскольку по сети будет передаваться меньшее число событий.

Например, перераспределение сфер влияния двух ЛП с учетом топологии

моделируемой схемы позволит уменьшить количество событий, передаваемых по сети от одного ЛП другому (см. рис. 4).

Пусть в поддереве КЛП имеется ЛП1 и ЛП2 и сферы влияния ЛП1 и ЛП2 распределены так, как показано на рисунке 4а. При моделировании ЛП1 будет генерировать три события, которые должны будут передаваться по сети на ЛП2. В случае, если элемент Dk переместить в сферу влияния ЛП2, то количество событий, передаваемых от ЛП1 к ЛП2 очевидно уменьшится (см. рис. 4б).

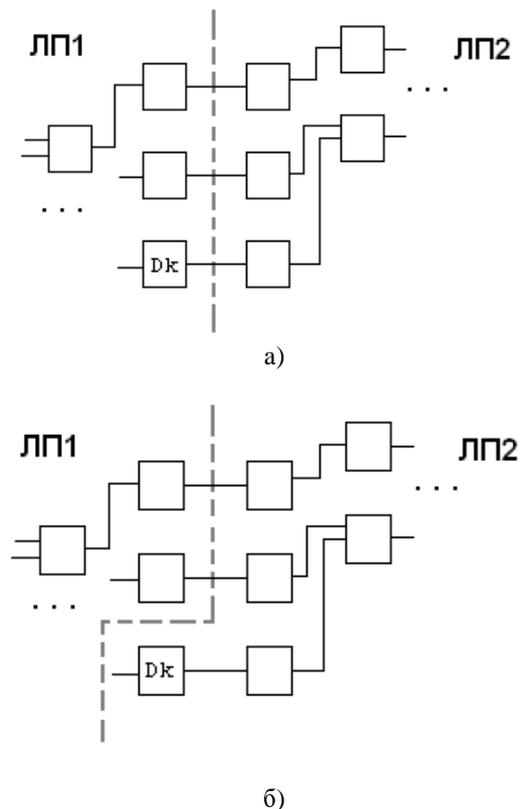


Рисунок 4 – Изменение количества событий, передаваемых по сети а) по сети передаётся три события б) по сети передаётся два события

Топология моделируемой схемы может использоваться при перераспределении сфер влияния между ЛП. Задача уменьшения нагрузки на сеть решается на КЛП и поэтому не требует использования вычислительных ресурсов ЛП. Каждый из КЛП содержит сведения о том, какие элементы моделируются на каждом из его подчиненных ЛП. КЛП может для каждого выходного j -го узла для каждого из элементов $\{ek\}$ на ЛП_i определить, сколько внешних событий сгенерируется вследствие изменения значения сигнала на j -м узле. Тогда можно определить количество N внешних событий от ЛП_i, генерируемых при заданном подмножестве

элементов {ek}, входящих в сферу влияния ЛП_i. Параметр N может быть вычислен для всех ЛП, входящих в поддереву КЛП.

Уменьшение нагрузки на линии связи в поддереве КЛП может быть выполнено следующим образом. Можно определить процессор ЛП_i, который генерирует наибольшее количество внешних событий при данном распределении сфер влияния. Далее необходимо определить подмножество элементов {em} в сфере влияния ЛП_i, которые участвуют в формировании этих событий, то есть изменения сигналов на выходных узлах этих элементов передаются другим процессорам в виде событий. Далее на КЛП необходимо проверить, на какие ЛП должны приходить эти события от ЛП_i и определить процессор ЛП_j, который принимает наибольшее количество этих событий от ЛП_i. Выделенная подгруппа элементов {em} может быть частично или полностью передана в сферу влияния ЛП_j.

Решение задачи перераспределения может корректироваться на КЛП путем проверки изменившегося значения параметра N на всех ЛП в рамках поддереву КЛП после выполнения изменения сфер влияния согласно этому способу.

Необходимость перераспределения сфер влияния при моделировании может определяться не только на КЛП, но и при сборе статистики моделирования на ЛП. Поскольку моделирование на ЛП происходит по комбинированному протоколу, то ЛП обрабатывает события с временной меткой $ts: GVT < ts < GVT + ConsVT$, где GVT – глобальное виртуальное время, ConsVT – параметр, определяющей, на сколько времени вперед может продвигаться локальное виртуальное время (LVT) на ЛП. Когда на процессор ЛП_i поступает событие, время выполнения которого меньше LVT(ЛП_i), то на ЛП_i выполняется откат, то есть ЛП_i, используя историю изменений моделируемой части схемы восстанавливает значения в узлах схемы и генерирует события в списке событий вплоть до момента виртуального времени, непосредственно предшествующего моменту, до которого должен выполняться откат.

Пока на одном из ЛП в сети выполняется откат, другие ЛП при этом переходят в режим ожидания появления новых событий, после того как они продвинули LVT вперед и не имеют событий для обработки. Частые переходы ЛП в режим ожидания могут говорить о том, что распределение сфер влияния на ЛП должно быть откорректировано. ЛП при переходе в режим ожидания может посылать соответствующее уведомление родительскому КЛП, содержащее также количество единиц времени, потраченного на ожидание. КЛП после получения определённого количества этих уведомлений принимает решение о необходимости

перераспределения сфер влияния ЛП, которые наиболее часто переходят в режим ожидания. В этом случае при выполнении перераспределения сфер влияния должна учитываться топология моделируемой схемы.

Оценка стоимости доступа процессоров друг к другу в вычислительной сети

Одним из параметров, который позволяет оценивать перераспределение задания на вычислительную сеть является стоимость доступа ЛП друг к другу [10]. Пусть на процессоре ЛП_i возникло событие $[n=v@t]$, которое должно быть передано процессору ЛП_j, поскольку узел n влияет на элемент в сфере влияния ЛП_j (см. рис. 5). Поскольку ЛП_i и ЛП_j входят в сферы влияния различных КЛП, то передача события $[n=v@t]$ будет происходить по дереву КЛП от ЛП_i к ЛП_j, через соответствующие уровни дерева КЛП.

Стоимость доступа ЛП_i к ЛП_j $S(ЛП_i, ЛП_j)$ – это количество КЛП, которые должны быть пройдены, прежде чем событие достигнет ЛП_j. Например, стоимость $S(ЛП_i, ЛП_j)$ на рисунке 5 равна 3.

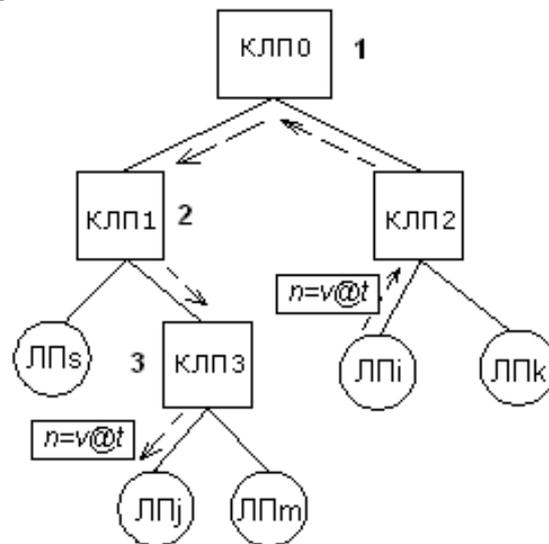


Рисунок 5 – Передача по сети события от ЛП_i к ЛП_j

Количество пройденных КЛП при передаче события соответствует количеству уровней дерева КЛП, через которые было передано это событие.

Загрузка всех КЛП в сети может характеризоваться таким параметром, как общая стоимость S_{Σ} доступа всех ЛП друг к другу в некотором интервале времени $[t1, t2]$ либо при обработке некоторого количества событий (3):

$$S_{\Sigma} = \sum_{i=1}^n \sum_{j=i+1}^n S(ЛП_i, ЛП_j) \quad (3)$$

Стоимость доступа ЛП друг к другу в рамках поддеревя одного КЛП равна 0. Наилучшим распределением задания на процессоры в сети будет то, при котором общая стоимость S_{Σ} будет наименьшей [10].

Уменьшение общей стоимости S_{Σ} может быть получено, например, при использовании алгоритма перераспределения сфер влияния на основании загрузки ЛП и КЛП (см. рис. 3). Однако в случае, если количество ЛП и КЛП, доступных в сети, ограничено и при моделировании возникает необходимость уменьшения нагрузки на ЛП, то одним из решений в этом случае является перераспределение сфер влияния на некоторых ЛП, учитывающее топологию моделируемой схемы.

Выводы

Предложена иерархическая организация вычислительной сети для распределённого логического моделирования цифровых систем. Выделены параметры, которые позволяют оценивать нагрузку на процессоры в сети. Рассмотрены особенности работы системы распределенного моделирования, при которых возникает необходимость уменьшения нагрузки на ЛП и КЛП.

Разработан алгоритм перераспределения сфер влияния ЛП и генерации дерева КЛП, учитывающий рассмотренные параметры. Данный алгоритм позволяет получить вариант перераспределения задания моделирования и выполнить генерацию дерева КЛП с целью улучшения таких характеристик распределенной системы, как время, необходимое для выполнения

моделирования и общая стоимость доступа процессоров друг к другу в сети.

Выполнение перераспределения сфер влияния между процессорами в сети необходимо также для решения возникающих при моделировании проблем, связанных с ненадежными процессорами, которые могут отключаться во время моделирования. При внезапном отключении компьютера в сети во время моделирования не требуется начинать моделирование задания заново благодаря возможности использования специально сохраняемых на КЛП предыдущих состояний моделируемой системы. При оптимистическом моделировании на ЛП допускается выполнение откатов. Поэтому сохранённые на КЛП сведения о предыдущих состояниях ЛП позволяют восстановить одно из состояний моделируемой системы, предшествующих моменту отключения процессора.

При использовании древовидной структуры вычислительной сети с относительно невысокой общей стоимостью доступа процессоров друг другу потребуется задействовать меньшее число процессоров при выполнении откатов к состоянию, предшествовавшему моменту отключения процессора. Важно учесть также, что любое перераспределение сфер влияния процессоров во время выполнения моделирования требует остановки процесса моделирования на некоторых из процессоров и передачи необработанных событий от одних процессоров на другие, что негативно влияет на скорость моделирования задания.

Список литературы

1. Ладыженский Ю.В. Исследование комбинированного протокола синхронизации логических процессоров при распределенном событийном моделировании цифровых устройств / Ю.В. Ладыженский, Ю.В. Попов // Моделирование и компьютерная графика: материалы 1-й международной научно-технической конференции, г. Донецк, 04-07 октября 2005. – Донецк, ДонНТУ, 2005. – С. 165–169.
2. Chandy K.M., Misra J. Asynchronous distributed simulation via a sequence of parallel computations // Communications of The ACM – CACM, vol. 24, no. 4, P. 198–206, 1981.
3. Fujimoto R.M. Parallel and distributed simulation systems // Winter Simulation Conference 2001.– Arlington, VA, USA: ACM, 2001.– P.147–157.
4. Jefferson D.R. Virtual time // USC Tech. rep. TR-83-213.– Los Angeles: Univ. of Southern California.– 1983.
5. Fersha A., Tripathi S.K. Parallel and Distributed Simulation of Discrete Event Systems // Hardbound of Parallel and Distributed Computing.– McGraw-Hill, 1995.
6. Mattern F. Efficient Algorithms for Distributed Snapshots and Global Virtual Time Approximation // Journal of Parallel and Distributed Computing Vol. 18, No 4, 1993. – P. 423-434.
7. Maruringsith W., Roland N. Specification-based Verification in a Distributed Shared Memory Simulation Model // SIMULATION vol. 86, No 4, 2010.- P. 229-245.
8. Schlagenhaft R., Ruhwandl M., Sporrer C., Bauer H. Dynamic load balancing of a multi-cluster simulator on a network of workstations // 9th Workshop on Parallel and Distributed Simulation.– New York: Parallel and Distributed Simulation, Workshop on, 1995.– P. 175–180.

9. Миков А.И. Метод динамической балансировки процессов имитационного моделирования / А.И. Миков, Е.Б. Замятина, К.А. Осмехин // Материалы Всероссийской научно-технической конференции. Методы и средства обработки информации МСО-2005. – М.: МГУ, 2005. – С.472-478.

10. Theodoropoulos G. An approach to interest management and dynamic load balancing in distributed simulation / G. Theodoropoulos, B. Logan // 2001 European Simulation Interoperability Workshop. – London: University of Westminster, 2001. – P.566–571.

11. Попов Ю.В. Організація обчислювальної мережі для розподіленого логічного моделювання цифрових систем: автореф. дис. канд. техн. наук / Ю.В. Попов. – Донецьк: Донецький національний технічний університет, 2007. – 21 с.

Надійшла до редколегії 20.09.2011

Ю.В. ПОПОВ, А.В. БАЛУТА

Донецький національний технічний університет

Алгоритм перерозподілу обчислювального навантаження у мережі при розподіленому логічному моделюванні цифрових систем

Розглянуто організацію ієрархічної структури обчислювальної мережі для розподіленого логічного моделювання цифрових систем. Визначено характеристики, що обчислюються у процесі моделювання, які використовуються в алгоритмі перерозподілу навантаження на процесори обчислювальної мережі

Розподілене логічне моделювання, цифрова схема, логічний процесор, комунікаційний логічний процесор (КЛП), сфера впливу, обчислювальна мережа, перерозподіл навантаження, лінія зв'язку, архітектура обчислювальної мережі, піддерево КЛП, рівновага навантаження, елемент схеми, вузол схеми, зовнішня подія, протокол синхронізації, оптимістичне моделювання

Y. POPOFF, A. BALUTA

Donetsk national technical University

An Algorithm to Redistribute Workload in a Network for Distributed Logic Simulation of Digital Systems

An organization of a hierarchical structure of computer network for distributed logic simulation of digital systems is reviewed. Characteristics evaluated at the simulation processors are defined. These characteristics are used in the algorithm to redistribute workload between processors in a computer network.

Distributed logic simulation, digital circuit, logic processor, communication logic processor (CLP), sphere of influence, computer network, redistribute workload, communication line, computer network architecture, subtree of CLP, workload equilibrium, circuit element, circuit node, external event, synchronization protocol, optimistic simulation