DOI: 10.15514/ISPRAS-2024-36(5)-5

Эффективность систем одинаково распределенных конкурирующих процессов при неограниченном и ограниченном параллелизме

П.А. Павлов, ORCID: 0009-0008-1233-7387 <pavlov.p@polessu.by>

Полесский государственный университет,
Республика Беларусь, 225710, г. Пинск, ул. Днепровской флотилии, д. 23.

Аннотация. В статье *с учетом ограниченного числа копий структурированного программного ресурса* проведен сравнительный анализ математических соотношений для вычисления общего времени выполнения множества одинаково распределенных конкурирующих процессов в асинхронном и двух синхронных режимах, в случае неограниченного и ограниченного параллелизма по числу процессоров многопроцессорной системы получено достаточное условие эффективности одинаково распределенной системы, доказано необходимое и достаточное условие существования эффективной системы одинаково распределенных конкурирующих процессов в зависимости от величины дополнительных системных расходов.

Ключевые слова: распределенный процесс; взаимодействующие процессы; программный ресурс; асинхронный (синхронный) режим; неограниченный (ограниченный) параллелизм; эффективность.

Для цитирования: Павлов П.А. Эффективность систем одинаково распределенных конкурирующих процессов при неограниченном и ограниченном параллелизме. Труды ИСП РАН, том 36, вып. 5, 2024 г., стр. 67–80. DOI: 10.15514/ISPRAS–2024–36(5)–5.

Efficiency of Systems of Identically Distributed Competing Processes with Unlimited and Limited Parallelism

P.A. Pavlov, ORCID: 0009-0008-1233-7387 <pavlov.p@polessu.by>

Polessky State University,
23, Dneprovskoy flotilii st., Pinsk, 225710, Belarus.

Abstract. In the article, taking into account the limited number of copies of a structured software resource, a comparative analysis of mathematical relationships for calculating the total execution time of a set of identically distributed competing processes in asynchronous and two synchronous modes was carried out; in the case of unlimited and limited parallelism by the number of processors of a multiprocessor system, a sufficient condition for the efficiency of an identically distributed system was obtained , a necessary and sufficient condition for the existence of an efficient system of identically distributed competing processes has been proven depending on the amount of additional system costs.

Keywords: distributed process; interacting processes; software resource; asynchronous (synchronous) mode; unlimited (limited) parallelism; efficiency.

For citation: Pavlov P.A. Efficiency of systems of identically distributed competing processes with unlimited and limited parallelism. Trudy ISP RAN/Proc. ISP RAS, vol. 36, issue 5, 2024. pp. 67-80 (in Russian). DOI: 10.15514/ISPRAS-2024-36(5)-5.

1. Введение

Быстрое развитие информационно-коммуникационных и сетевых технологий привело к интенсивному использованию географически распределенных вычислительных ресурсов и созданию на их основе динамически-масштабируемых высокопроизводительных *распределенных вычислительных систем* (РВС), одним из основных преимуществ которых является возможность параллельной обработки процессов [1-3]. При проектировании и создании РВС особую актуальность приобретают задачи построения и исследования математических моделей организации взаимодействия параллельных процессов, конкурирующих за программный ресурс (ПР). Данные задачи имеют как прямой, так и обратный характер. При постановке прямых задач условиями являются значения параметров распределенной вычислительной системы, а решением минимальное общее время реализации заданных объемов вычислений. Постановка обратных задач сводится к поиску критериев эффективности и оптимальности организации выполнения множества распределенных процессов. Случай, когда в общей памяти РВС имеется одна копия ПР, с различных точек зрения был изучен в работах [4-9]. В частности, были решены задачи нахождения минимального общего времени выполнения распределенных конкурирующих процессов, использующих структурированный на блоки ПР в различных режимах взаимодействия процессов, процессоров и блоков, проведен сравнительный анализ режимов взаимодействия процессов, процессоров и блоков, получены критерии эффективности и оптимальности структурирования программных ресурсов, решен ряд оптимизационных задач по расчету числа процессов, процессоров и др. Изучение задач, относящихся к оптимальной организации распределенных параллельных вычислений, приобретает особую актуальность в случае, когда в общей памяти РВС может быть одновременно размещено ограниченное число копий программного ресурса.

2. Математическая модель системы распределенных вычислений при ограниченном числе копий программного ресурса

Как и в работах [4-9] под *взаимодействующими процессами*, т. е. которые влияют на поведение друг друга путем обмена информацией, будем понимать выполнение последовательности наборов блоков $I_s = (1,\ 2,\ \ldots,\ s)$. Многократно выполняемую в многопроцессорной системе программу или ее часть будем называть *программным ресурсом* (ПР), а множество процессов его выполняющим – *конкурирующими.*

Математическая модель системы распределенной обработки взаимодействующих процессов, конкурирующих за использование ограниченного числа копий структурированного программного ресурса, включает в себя [10-12]: $p \geq 2$, процессоров многопроцессорной системы, которые имеют доступ к общей памяти; $n \geq 2$, распределенных взаимодействующих конкурирующих процессов; $s \geq 2$, блоков структурированного на блоки программного ресурса; матрицу $T = [t_{ij}]$, $i = \overline{1,n}$, $j = \overline{1,s}$, времен выполнения блоков программного ресурса распределенными взаимодействующими конкурирующими процессами; $2 \leq c \leq p$, число копий структурированного на блоки программного ресурса, которые могут одновременно находиться в оперативной памяти, доступной для всех p процессоров; $\theta > 0$ – параметр, характеризующий время дополнительных системных расходов, связанных с организацией конвейерного режима использования блоков структурированного программного ресурса множеством взаимодействующих конкурирующих процессов при распределенной обработке.

Будем также предполагать, что число процессов n кратно числу копий c структурированного программного ресурса, т. е. $n = mc$, где $m = n/c \geq 2$, и что взаимодействие процессов, процессоров и блоков программного ресурса подчинено следующим условиям: 1) ни один из процессоров не может обрабатывать одновременно более одного блока; 2) процессы выполняются в параллельно–конвейерном режиме группами, т. е. осуществляется одновременное (параллельное) выполнение c копий каждого блока в сочетании с конвейеризацией группы из c копий Q_j–го блока, $j = \overline{1,s}$, по процессам и процессорам; 3) обработка каждого блока программного ресурса осуществляется без прерываний; 4) в случае, когда число блоков программного ресурса $s \leq \left[\dfrac{p}{c}\right]$, где $[x]$ – целая часть числа, для каждого i–го процесса, где $i = c(l-1) + q$, $l = \overline{1,m}$, $q = \overline{1,c}$, распределение блоков Q_j, $j = \overline{1,s}$, программного ресурса по процессорам осуществляется по правилу: блок с номером j распределяется на процессор с номером $c(j-1) + q$.

Далее, как и в случае одной копии программного ресурса [4-9], введем базовые режимы конвейерной реализации взаимодействия процессов, процессоров и блоков ПР, но с учетом наличия c копий программного ресурса, а также определение одинаково распределенной и стационарной систем.

Асинхронной режим взаимодействия процессов, процессоров и блоков структурированного программного ресурса предполагает, что начало выполнения копий очередного Q_j–го блока, $j = \overline{1,s}$, определяется наличием c процессоров и готовностью копий блока к выполнению, при этом программный блок считается готовым к выполнению, если он не выполняется ни на одном из процессоров.

Первый синхронный режим обеспечивает линейный порядок выполнения блоков программного ресурса внутри каждого из процессов без задержек, т. е. в случае, когда $2 \leq s \leq \left[\dfrac{p}{c}\right]$, момент завершения выполнения Q_j–го блока, $j = \overline{1,s-1}$, процессом с номером $i = (l-1)c + q$, $l = \overline{1,m}$, $q = \overline{1,c}$, на $((j-1)c + q)$–м процессоре совпадает с моментом начала выполнения следующего Q_{j+1}–го блока на процессоре с номером $(jc + q)$.

При **втором синхронном режиме** в случае, когда $2 \leq s \leq \left[\dfrac{p}{c}\right]$, момент завершения выполнения i–м процессом, где $i = (l-1)c + q$, $l = \overline{1,m-1}$, $q = \overline{1,c}$, j–го блока, $j = \overline{1,s}$, на $((j-1)c + q)$–м процессоре совпадает с моментом начала выполнения j–го блока процессом с номером $(i + c)$ на этом же процессоре, т.е. обеспечивается непрерывное выполнение каждого блока всеми процессами.

Определение 1. Систему распределенных конкурирующих процессов будем называть **одинаково распределенной**, если времена выполнения всех блоков ПР каждым из процессов совпадают и равны t_i^{θ}, т. е. справедлива цепочка равенств $t_{i1}^{\theta} = t_{i2}^{\theta} = \ldots = t_{is}^{\theta} = t_i^{\theta}$, для всех $i = \overline{1,n}$.

Будем рассматривать случаи *неограниченного* по числу процессоров распределенной вычислительной системы, т. е. когда $s \le \left[\dfrac{p}{c}\right]$, и *ограниченного*, когда $s > \left[\dfrac{p}{c}\right]$, параллелизма.

3. Анализ режимов функционирования одинаково распределенных систем конкурирующих процессов

В [10-12] подробно исследованы базовые асинхронный и два синхронных конвейерных режима взаимодействия распределенных процессов в условиях конкуренции за ограниченное число копий программного ресурса. Для вычисления общего времени выполнения множества распределенных процессов в рамках очерченных режимов с учетом $2 \le c \le p$ копий ПР получены различные математические соотношения. Определенный теоретический и практический интерес представляют задачи сравнительного анализа полученных соотношений. Проведем такой анализ для класса *одинаково распределенных* систем с учетом параметра $\theta > 0$, характеризующего время дополнительных системных расходов, связанных с организацией конвейерного режима использования блоков структурированного программного ресурса множеством конкурирующих процессов при распределенной обработке.

Для проведения сравнительного анализа множество из n процессов разобьем на c подмножеств по m процессов в каждом. В каждое q-е подмножество, $q = \overline{1,c}$, будут включены процессы с номерами $i = c(l-1)+q$, $l = \overline{1,m}$, блоки которых будут выполняться на $(c(j-1)+q)$-х процессорах, $j = \overline{1,s}$. Через $T_q^{\theta} = \sum_{i=1}^{m} t_{c(i-1)+q}^{\theta}$ обозначим суммарное время выполнения q-го подмножества процессов, а $t_{\max}^{q+} = \max\limits_{1\le i\le m} t_{c(i-1)+q}^{\theta}$ – максимальное время выполнения блока из этого подмножества, $q = \overline{1,c}$.

Определение 2. **Характеристическим** набором одинаково распределенной системы конкурирующих процессов будем называть набор параметров вида $(t_q^{\theta},\ t_{c+q}^{\theta},\ t_{2c+q}^{\theta},\dots,\ t_{c(m-1)+q}^{\theta},\ T_q^{\theta})$, $q = \overline{1,c}$.

Пусть $\beta = \left\{(t_q^{\theta},\ t_{c+q}^{\theta},\dots,\ t_{c(m-1)+q}^{\theta},\ T_q^{\theta})\ \middle|\ T_q^{\theta} = \sum_{i=1}^{m} t_{c(i-1)+q}^{\theta},\ q=\overline{1,c}\right\}$ – множество всех допустимых характеристических наборов системы одинаково распределенных конкурирующих процессов. Выделим из множества β подмножество характеристических наборов вида:

$$S(\beta) = \left\{\begin{matrix} (t_q^{\theta},\ t_{c+q}^{\theta},\ t_{2c+q}^{\theta},\dots,\ t_{c(m-1)+q}^{\theta},\ T_q^{\theta}) \in \beta,\ \textit{где} \\ t_q^{\theta} \le t_{c+q}^{\theta} \le t_{2c+q}^{\theta} \le \dots \le t_{c(k-1)+q}^{\theta} \ge t_{ck+q}^{\theta} \ge \dots \ge t_{c(m-1)+q}^{\theta},\ k=\overline{1,m},\ q=\overline{1,c} \end{matrix}\right\}.$$

Для выделенного подмножества характеристических наборов $S(\beta)$ справедлива

Теорема 1. *В случае неограниченного параллелизма по числу процессоров РВС минимальные общие времена выполнения множества одинаково распределенных конкурирующих процессов в многопроцессорной системе с параметрами* $p \ge 2$, $n \ge 2$, $s \ge 2$, $2 \le c \le p$, $\theta > 0$ *в асинхронном и двух синхронных режимах совпадают, т. е.*

$$T_{ac}^{op}(p,n,s,c,\theta) = T_{1c}^{op}(p,n,s,c,\theta) = T_{2c}^{op}(p,n,s,c,\theta).$$

Доказательство. Для любого характеристического набора одинаково распределенной системы в случае неограниченного параллелизма для асинхронного режима и второго синхронного режима, обеспечивающего непрерывное выполнение каждого блока всеми процессами, минимальное общее время выполнения n одинаково распределенных процессов, конкурирующих за использование c копий структурированного на s блоков программного ресурса в многопроцессорной системе с p процессорами с учетом параметра $\theta > 0$, в том числе и для любого характеристического набора $\mu \in S(\beta)$, вычисляется по формуле [9,10]:

$$T_{ac}^{op}(p,n,s,c,\theta) = T_{2c}^{op}(p,n,s,c,\theta) = \max_{1 \le q \le c}(T_q^{\theta} + (s-1)t_{\max}^{q+}).$$

Если же взаимодействие процессов, процессоров и блоков структурированного ПР осуществляется в первом синхронном режиме, при котором обеспечивается выполнение блоков программного ресурса внутри каждого из процессов без задержек, то для любого характеристического набора из β при $2 \le s \le \left[\frac{p}{c}\right]$ выполняется равенство [11]:

$$T_{1c}^{op}(p,n,s,c,\theta) = \max_{1 \le q \le c}\left(T_q^{\theta} + (s-1)\left[t_{c(m-1)+q}^{\theta} + \sum_{i=2}^{m}\max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0)\right]\right).$$

Покажем, что $t_{c(m-1)+q}^{\theta} + \sum_{i=2}^{m}\max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) = t_{\max}^{q+}$. Так как $t_{\max}^{q+} = \max_{1 \le i \le m} t_{c(i-1)+q}^{\theta}$, $q = \overline{1,c}$, то для всех номеров $1 \le i \le k \le m$ имеем $\sum_{i=2}^{k}\max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) = 0$, а для $1 \le k \le i \le m$ имеет место равенство $\sum_{i=k+1}^{m}\max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) = t_{c(k-1)+q}^{\theta} - t_{c(m-1)+q}^{\theta}$.

Следовательно, $t_{c(m-1)+q}^{\theta} + \sum_{i=2}^{m}\max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) = t_{c(k-1)+q}^{\theta} = t_{\max}^{q+}$. Теорема доказана.

Теорема 2. *В случае неограниченного параллелизма по числу процессоров многопроцессорной системы если допустимый характеристический набор* μ *одинаково распределенной системы с параметрами* $p \ge 2$, $n \ge 2$, $s \ge 2$, $2 \le c \le p$, $\theta > 0$, *не принадлежит подмножеству* $S(\beta)$, *то*

$$T_{1c}^{op}(p,n,s,c,\theta) > T_{ac}^{op}(p,n,s,c,\theta) = T_{2c}^{op}(p,n,s,c,\theta).$$

Доказательство. Так как для асинхронного и второго синхронного режимов минимальные общие времена выполнения одинаково распределенных конкурирующих процессов с учетом ограниченного числа копий программного ресурса равны и определяются по формуле

$$T_{ac}^{op}(p,n,s,c,\theta) = T_{2c}^{op}(p,n,s,c,\theta) = \max_{1 \le q \le c}(T_q^{\theta} + (s-1)t_{\max}^{q+}),$$

а для первого синхронного режима по формуле

$$T_{1c}^{op}(p,n,s,c,\theta) = \max_{1 \le q \le c}\left(T_q^{\theta} + (s-1)\left[t_{c(m-1)+q}^{\theta} + \sum_{i=2}^{m}\max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0)\right]\right),$$

то доказательство неравенства $T_{1c}^{op}(p,n,s,c,\theta) > T_{ac}^{op}(p,n,s,c,\theta)$ и будет доказательством теоремы.

Доказательство неравенства

$$t_{c(m-1)+q}^{\theta} + \sum_{i=2}^{m} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) - t_{\max}^{q+} > 0, \ q = \overline{1,c}, \tag{1}$$

проведем индукцией по числу процессов $m \geq 2$.

Пусть $m = 2$. В этом случае множество всех допустимых характеристических наборов системы одинаково распределенных конкурирующих процессов $\beta = (t_q^{\theta},\ t_{c+q}^{\theta}) \in S(\beta)$, $q = \overline{1,c}$.

Пусть неравенство (1) выполняется при $m = k$, покажем, что оно справедливо и при $m = k + 1$. При $m = k + 1$ получим:

$$t_{ck+q}^{\theta} + \sum_{i=2}^{k+1} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) - t_{\max}^{q+} > 0, \text{ где } t_{\max}^{q+} = \max_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta}, \ q = \overline{1,c}.$$

Последнее неравенство для всех $q = \overline{1,c}$ равносильно неравенству вида:

$$t_{ck+q}^{\theta} + \sum_{i=2}^{k} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) + \max(t_{c(k-1)+q}^{\theta} - t_{ck+q}^{\theta}, 0) - \max_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta} > 0.$$

Если $\max\limits_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta}$ достигается при $i = k+1$, т. е. $\max\limits_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta} = t_{ck+q}^{\theta}$, то

$$t_{ck+q}^{\theta} + \sum_{i=2}^{k} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) + \max(t_{c(k-1)+q}^{\theta} - t_{ck+q}^{\theta}, 0) - t_{ck+q}^{\theta} =$$
$$= \sum_{i=2}^{k} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) + \max(t_{c(k-1)+q}^{\theta} - t_{ck+q}^{\theta}, 0) > 0.$$

Здесь первое слагаемое больше нуля, ибо в противном случае $\mu \in S(\beta)$, что противоречит условию теоремы, а второе слагаемое равно нулю, так как $t_{ck+q}^{\theta} \geq t_{c(k-1)+q}^{\theta}$.

Если значение $\max\limits_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta}$ находится в промежутке $1 \leq i \leq k$, то имеем:

$$t_{ck+q}^{\theta} + \sum_{i=2}^{k} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) + \max(t_{c(k-1)+q}^{\theta} - t_{ck+q}^{\theta}, 0) - \max_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta} =$$
$$= t_{c(k-1)+q}^{\theta} + \sum_{i=2}^{k} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) - \max_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta} +$$
$$+ t_{ck+q}^{\theta} - t_{c(k-1)+q}^{\theta} + \max(t_{c(k-1)+q}^{\theta} - t_{ck+q}^{\theta}, 0).$$

Здесь $t_{c(k-1)+q}^{\theta} + \sum\limits_{i=2}^{k} \max(t_{c(i-2)+q}^{\theta} - t_{c(i-1)+q}^{\theta}, 0) - \max\limits_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta} > 0$ по индукционному предположению и в силу того, что $\max\limits_{1 \leq i \leq k+1} t_{c(i-1)+q}^{\theta} = \max\limits_{1 \leq i \leq k} t_{c(i-1)+q}^{\theta}$.

Покажем далее, что $t_{ck+q}^{\theta} - t_{c(k-1)+q}^{\theta} + \max(t_{c(k-1)+q}^{\theta} - t_{ck+q}^{\theta}, 0) \geq 0$:

– если $t^{\theta}_{c(k-1)+q} = t^{\theta}_{ck+q}$ равенство нулю очевидно;

– если $t^{\theta}_{c(k-1)+q} > t^{\theta}_{ck+q}$ получим, что

$$t^{\theta}_{ck+q} - t^{\theta}_{c(k-1)+q} + \max(t^{\theta}_{c(k-1)+q} - t^{\theta}_{ck+q}, 0) = t^{\theta}_{ck+q} - t^{\theta}_{c(k-1)+q} + t^{\theta}_{c(k-1)+q} - t^{\theta}_{ck+q} = 0;$$

– если $t^{\theta}_{c(k-1)+q} < t^{\theta}_{ck+q}$, имеем, что $\max(t^{\theta}_{c(k-1)+q} - t^{\theta}_{ck+q}, 0) = 0$, тогда

$$t^{\theta}_{ck+q} - t^{\theta}_{c(k-1)+q} + \max(t^{\theta}_{c(k-1)+q} - t^{\theta}_{ck+q}, 0) = t^{\theta}_{ck+q} - t^{\theta}_{c(k-1)+q} > 0,$$

что и требовалось доказать.

4. Эффективность одинаково распределенных систем конкурирующих процессов при неограниченном параллелизме

В п.3 доказано, что в случае неограниченного параллелизма по числу процессоров распределенной вычислительной системы минимальные общие времена выполнения n *одинаково распределенных* процессов, конкурирующих за использование c копий структурированного на s блоков программного ресурса в многопроцессорной системе с p процессорами с учетом параметра $\theta > 0$ в асинхронном и двух синхронных режимах совпадают и вычисляются по формуле (теорема 1):

$$T^{op}(p,n,s,c,\theta) = \max_{1\le q\le c}(T^{\theta}_q + (s-1)t^{q+}_{\max}), \tag{2}$$

где $T^{\theta}_q = \sum_{i=1}^{n/c} t^{\theta}_{c(i-1)+q}$ – время выполнения q–го подмножества процессов, $t^{q+}_{\max} = \max_{1\le i\le n/c} t^{\theta}_{c(i-1)+q}$ – максимальное время выполнения блока из этого подмножества с учетом системных расходов θ, $q = \overline{1,c}$.

Определение 3. Одинаково распределенная система конкурирующих процессов называется **стационарной**, если $t^{\theta}_1 = t^{\theta}_2 = \ldots = t^{\theta}_n = t^{\theta}$.

В случае *стационарной* одинаково распределенной системы конкурирующих процессов минимальные общие времена выполнения определяются равенством [9-11]:

$$T^{cc}(p,n,s,c,\theta) = \left(\frac{n}{c} + s - 1\right)t^{\theta}, \text{ где } t^{\theta} = t + \theta.$$

Определение 4. Одинаково распределенную систему конкурирующих процессов будем называть **эффективной** при фиксированных $p \ge 2$, $s \ge 2$, $c \ge 2$, $\theta > 0$, если выполняется соотношение $\Delta_{op} = \max_{1\le q\le c} sT_q - T^{op}(p,n,s,c,\theta) \ge 0$, где sT_q – время выполнения s блоков программного ресурса всеми процессами q -го подмножества в последовательном режиме, а $T_q = \sum_{i=1}^{n/c} t_{c(i-1)+q}$.

При наличии двух эффективных одинаково распределенных систем конкурирующих процессов будем считать, что первая не менее эффективная, чем вторая, если $\Delta^1_{op} \ge \Delta^2_{op}$.

Для введенного подмножества одинаково распределенных систем справедливо следующее утверждение.

Теорема 3. *Для любой эффективной одинаково распределенной системы конкурирующих процессов при* $s \le \left[\dfrac{p}{c}\right]$ *и* $\theta > 0$ *существует более эффективная стационарная одинаково распределенная система.*

Доказательство. Рассмотрим любую эффективную одинаково распределенную систему. Согласно определению 4 условие ее эффективности с учетом (2) можно записать в виде следующего неравенства:

$$\Delta_{op} = \max_{1\le q\le c} sT_q - T^{op}(p,n,s,c,\theta) = \max_{1\le q\le c}\left(sT_q - T_q^{\theta} - (s-1)t_{\max}^{q+}\right) =$$

$$= (s-1)\max_{1\le q\le c}(T_q - t_{\max}^q) - \left(\frac{n}{c} + s - 1\right)\theta \ge 0 \text{, где } t_{\max}^q = \max_{1\le i\le n/c} t_{c(i-1)+q} . \quad (3)$$

Для любой *стационарной* одинаково распределенной системы условие эффективности запишется следующим образом:

$$\Delta_{cc} = sT^m - T^{cc}(p,n,s,c,\theta) = sT^m - \left(\frac{n}{c} + s - 1\right)t^{\theta} =$$

$$= (s-1)(T^m - t) - \left(\frac{n}{c} + s - 1\right)\theta \ge 0 \text{, где } T^m = \frac{n}{c}t . \quad (4)$$

Для доказательства теоремы достаточно доказать выполнение неравенства $\Delta_{op} \le \Delta_{cc}$.

Подставив вместо Δ_{op} и Δ_{cc} выражения из (3) и (4), получим:

$$\max_{1\le q\le c}(T_q - t_{\max}^q) \le \left(\frac{n}{c} - 1\right)t .$$

Докажем справедливость последнего неравенства. Рассмотрим стационарную одинаково распределенную систему конкурирующих процессов, в которой $t = \max_{1\le i\le n} t_i = t_{\max}^q$. Пусть для определенности $t_{\max}^q = t_k$, тогда справедлива цепочка соотношений:

$$\max_{1\le q\le c}(T_q - t_{\max}^q) = \max_{1\le q\le c}\left(\sum_{i=1}^{k-1} t_{c(i-1)+q} + \sum_{i=k+1}^{n/c} t_{c(i-1)+q}\right) \le \left(\frac{n}{c} - 1\right)t_{\max}^q = \left(\frac{n}{c} - 1\right)t .$$

Теорема доказана.

Следующее утверждение в случае неограниченного параллелизма по числу процессоров распределенной вычислительной системы с учетом ограниченного числа копий структурированного программного ресурса определяет достаточное условие эффективности одинаково распределенной системы.

Теорема 4. *Система одинаково распределенных конкурирующих процессов с параметрами* p, n, s, c, θ, *удовлетворяющими соотношениям*

$$3 \le s \le \left[\frac{p}{c}\right],\ s = \frac{n}{c} \ne 3,\ ns \ge 2[n + c(s-1)],\ 0 < \theta \le t_{\min} = \min_{1\le i\le n} t_i ,$$

является эффективной.

Доказательство. Согласно формуле (3), условие эффективности равносильно неравенству

$$\max_{1\le q\le c}\frac{T_q - t_{\max}^q}{\theta} \ge \frac{n + c(s-1)}{c(s-1)} . \quad (5)$$

Следовательно, для доказательства теоремы достаточно убедиться в справедливости неравенства (5). В силу выбора $0<\theta\le t_{\min}$ будет выполняться неравенство $t_{\min}/\theta\ge 1$, тогда справедлива цепочка

$$\max_{1\le q\le c}\frac{T_q-t^q_{\max}}{\theta}\ge\left(\frac{n}{c}-1\right)\frac{t_{\min}}{\theta}\ge\frac{n}{c}-1\,. \tag{6}$$

Далее, из $ns\ge 2[n+c(s-1)]$ следует справедливость неравенства

$$\frac{n}{c}-1\ge\frac{n+c(s-1)}{c(s-1)}\,. \tag{7}$$

Следствием неравенств (6) и (7) является неравенство (5). Таким образом, теорема доказана.

Ниже формулируется и доказывается необходимое и достаточное условие существования эффективной системы одинаково распределенных конкурирующих процессов при неограниченном параллелизме с учетом c копий программного ресурса в зависимости от величины дополнительных системных расходов θ.

Теорема 5. *Для существования эффективной одинаково распределенной системы конкурирующих процессов с заданными параметрами* $\left[\frac{p}{c}\right]\ge 3,\ \ 2\le s\le\left[\frac{p}{c}\right],\ \ 2\le c\le p$ *и* $\theta>0$ *необходимо и достаточно выполнение следующих условий:*

$$\theta\le\begin{cases}\varphi[c(1+\sqrt{s}\,)], & \textit{если } \sqrt{s} \textit{ целое,}\\ \max\{\varphi[c(1+\sqrt{s}\,)],\ \varphi[c(2+\sqrt{s}\,)]\}, & \textit{если } \sqrt{s} \textit{ нецелое,}\end{cases} \tag{8}$$

где $\varphi(x)=\dfrac{c(s-1)T^m(x-c)}{x^2+xc(s-1)}$, *а* $[x]$ *– наибольшее целое, не превосходящее* x.

Доказательство. Согласно (4), условие эффективности любой стационарной одинаково распределенной системы конкурирующих процессов определяется соотношением

$$\Delta_{cc}=(s-1)(T^m-t)-\left(\frac{n}{c}+s-1\right)\theta\ge 0\text{, где } T^m=\frac{n}{c}t\,,$$

которое равносильно выполнению неравенства

$$\theta\le\frac{c(s-1)T^m(n-c)}{n^2+nc(s-1)}\,. \tag{9}$$

Введем в рассмотрение функцию $\varphi(x)=\dfrac{c(s-1)T^m(x-c)}{x^2+xc(s-1)}$. В силу того, что $\varphi'(x)=\dfrac{c(s-1)T^m[-x^2+2cx+c^2(s-1)]}{[x^2+xc(s-1)]^2}$, то функция φ при $x>0$ достигает своего максимального значения в точке $x=c(1+\sqrt{s}\,)$. Положим

$$n_э=\begin{cases}c(1+\sqrt{s}\,), & \textit{если } \sqrt{s} \textit{ — целое,}\\ \max\{c(1+[\sqrt{s}\,]),\ c(2+[\sqrt{s}\,])\}, & \textit{если } \sqrt{s} \textit{ — нецелое.}\end{cases} \tag{10}$$

Необходимость условий (8) будет доказана, если будет установлена невозможность существования эффективной одинаково распределенной системы n конкурирующих процессов, для которой выполнялось бы неравенство вида:

$$\theta > \frac{c(s-1)T^m(n-c)}{n^2 + nc(s-1)} . \quad (11)$$

Очевидно, такой системы нет при $n = n_э$, так как в силу определения функции φ, для такого n выполняется неравенство (11) и, следовательно, такая система не может быть эффективной.

Если все же предположить существование такой системы с n процессами, то должно выполняться соотношение $n \neq n_э$.

Выше установлено, что одинаково распределенная система с $n_э$ процессами эффективна, следовательно, для нее имеет место неравенство $\theta \le \frac{c(s-1)T^m(n_э - c)}{n_э^2 + n_э c(s-1)}$. В то время как для гипотетической системы с n процессами в силу предположения должно выполняться неравенство $\theta > \frac{c(s-1)T^m(n-c)}{n^2 + nc(s-1)}$. Очевидным следствием полученных неравенств, является неравенство $\theta > \theta$.

В случае $n < n_э$ в силу (11) должна выполняться цепочка неравенств

$$\frac{c(s-1)T^m(n_э - c)}{n_э^2 + n_э c(s-1)} < \frac{c(s-1)T^m(n-c)}{n^2 + nc(s-1)} < \theta ,$$

из которой в силу (9) следует неэффективность одинаково распределенной системы с $n_э$ процессами.

Наконец, если $n > n_э$, то $\frac{c(s-1)T^m(n_э - c)}{n_э^2 + n_э c(s-1)} > \frac{c(s-1)T^m(n-c)}{n^2 + nc(s-1)} \ge \theta$, что указывает на эффективность гипотетической одинаково распределенной системы n конкурирующих процессов. Полученные противоречия во всех возможных случаях доказывают необходимость условий (8).

Достаточность условий (8) непосредственно следует из наличия функции φ со свойством (9). Действительно, в этом случае требуемой эффективной одинаково распределенной системы является система с $n = n_э$ конкурирующими процессами, где $n_э$ определяется формулой (10). Теорема доказана.

З а м е ч а н и е. При $\left[\frac{p}{c}\right] = s = 2$ одинаково распределенная система конкурирующих процессов будет эффективной, если выполняется неравенство $\frac{\theta}{T_m} \le \frac{c(n-c)}{n(n+c)}$.

5. Эффективность одинаково распределенных систем в условиях ограниченного параллелизма

В случае ограниченного параллелизма для вычисления минимального общего времени выполнения одинаково распределенных конкурирующих процессов в асинхронном и втором синхронном режимах имеют место формулы [10-12]:

$$T_{ac}^{op}\left(p,n,s,c,\theta\right) = T_{2c}^{op}\left(p,n,s,c,\theta\right) =$$

$$= \max_{1\le q\le c}\begin{cases} kT_q^{\theta}+\left(\left[\frac{p}{c}\right]-1\right)t_{\max}^{q+}, \ \textit{при}\ s=k\left[\frac{p}{c}\right],\ k>1,\ T_q^{\theta}>\left[\frac{p}{c}\right]t_{\max}^{q+}, \\ (k+1)T_q^{\theta}+(r-1)t_{\max}^{q+}, \ \textit{при}\ s=k\left[\frac{p}{c}\right]+r,\ k\ge 1,\ 1\le r<\left[\frac{p}{c}\right],\ T_q^{\theta}>\left[\frac{p}{c}\right]t_{\max}^{q+}. \end{cases}$$

Здесь $T_q^{\theta}=\sum_{i=1}^{n/c}t_{c(i-1)+q}^{\theta}$, $t_{\max}^{q+}=\max_{1\le i\le n/c}t_{c(i-1)+q}^{\theta}$, $q=\overline{1,c}$.

Теорема 6. *Если параметры системы* $\frac{n}{c}\ge 3$ *одинаково распределенных конкурирующих процессов в многопроцессорной системе с* p *процессорами удовлетворяют соотношениям* $s\ge 3$, $s=\frac{n}{c}\ne 3$, $2\le c\le p$, $0<\theta\le t_{\min}=\min_{1\le i\le n}t_i$ *и* $T_q^{\theta}>\left[\frac{p}{c}\right]t_{\max}^{q+}$, *то в случае ограниченного параллелизма рассматриваемая система будет эффективной, если выполняются условия:*

$$k\frac{n}{c}\left[\frac{p}{c}\right]\ge\begin{cases} 2\left(k\frac{n}{c}+\left[\frac{p}{c}\right]-1\right), & s=k\left[\frac{p}{c}\right],\ k>1, \\ \frac{n}{c}[2(k+1)-r]+2(r-1), & s=k\left[\frac{p}{c}\right]+r,\ k\ge 1,\ 1\le r<\left[\frac{p}{c}\right]. \end{cases}$$

Доказательство. Для случая $s=k\left[\frac{p}{c}\right]$, $k>1$, условие эффективности одинаково распределенной системы конкурирующих процессов запишется в виде:

$$\Delta_{op}^{ac,2c}\left(s=k\left[\frac{p}{c}\right]\right)=\max_{1\le q\le c}k\left[\frac{p}{c}\right]T_q-T_{ac,2c}^{op}\left(p,n,k\left[\frac{p}{c}\right],c,\theta\right)=$$

$$=\left(\left[\frac{p}{c}\right]-1\right)\max_{1\le q\le c}(kT_q-t_{\max}^{q})-\left(k\frac{n}{c}+\left[\frac{p}{c}\right]-1\right)\theta\ge 0,\ \text{где}\ T_q=\sum_{i=1}^{n/c}t_{c(i-1)+q},\ t_{\max}^{q}=\max_{1\le i\le n/c}t_{c(i-1)+q}$$

Или

$$\left(\left[\frac{p}{c}\right]-1\right)\max_{1\le q\le c}\frac{kT_q-t_{\max}^{q}}{\theta}\ge k\frac{n}{c}+\left[\frac{p}{c}\right]-1. \tag{12}$$

В силу того, что $0<\theta\le t_{\min}$, получим:

$$\max_{1\le q\le c}\frac{kT_q-t_{\max}^{m}}{\theta}\ge\left(k\frac{n}{c}-1\right)\frac{t_{\min}}{\theta}\ge k\frac{n}{c}-1. \tag{13}$$

Из (13) и (12) следует первая формула теоремы.

В случае, когда $s=k\left[\frac{p}{c}\right]+r$, $k\ge 1$, $1\le r<\left[\frac{p}{c}\right]$, условие эффективности будет иметь вид:

$$k\left(\left[\frac{p}{c}\right]-1\right)\max_{1\le q\le c}\frac{T_q}{\theta}+(r-1)\max_{1\le q\le c}\frac{T_q-t_{\max}^{q}}{\theta}\ge(k+1)\frac{n}{c}+r-1. \tag{14}$$

Так как, $t_{\min}/\theta\ge 1$, то

$$k\left(\left[\frac{p}{c}\right]-1\right)\max_{1\le q\le c}\frac{T_q}{\theta}+(r-1)\max_{1\le q\le c}\frac{T_q-t_{\max}^q}{\theta}\ge k\frac{n}{c}\left(\left[\frac{p}{c}\right]-1\right)+\left(\frac{n}{c}-1\right)(r-1)\,. \qquad (15)$$

Из (15) и (14) следует вторая формула теоремы. Теорема доказана.

Сформулируем и докажем необходимое и достаточное условие существования эффективной системы одинаково распределенных конкурирующих процессов в зависимости от величины накладных расходов θ в случае ограниченного параллелизма по числу процессоров РВС.

Теорема 7. *Для существования эффективной одинаково распределенной системы конкурирующих процессов с заданными параметрами* $\left[\frac{p}{c}\right]\ge 3$, $T^m=\frac{n}{c}t$, $2\le c\le p$, $\theta>0$ *необходимо и достаточно, чтобы выполнялись следующие условия:*

1) *при* $s=k\left[\frac{p}{c}\right]$, $k>1$,

$$\theta\le\begin{cases}\varphi_1\left[\frac{c}{k}\left(1+\sqrt{\left[\frac{p}{c}\right]}\right)\right], & \textit{если}\ \frac{c}{k}\left(1+\sqrt{\left[\frac{p}{c}\right]}\right)\ -\ \textit{целое},\\ \max\left\{\varphi_1\left(\left[\frac{c}{k}\left(1+\sqrt{\left[\frac{p}{c}\right]}\right)\right]\right),\ \varphi_1\left(\left[\frac{c}{k}\left(1+\sqrt{\left[\frac{p}{c}\right]}\right)\right]+1\right)\right\}, & \textit{в противном случае},\end{cases}$$

где $\varphi_1(x)=\dfrac{aT^m c(kx-c)}{kx^2+acx}$, $a=\left[\frac{p}{c}\right]-1$;

2) *при* $s=k\left[\frac{p}{c}\right]+r$, $k\ge 1$, $1\le r<\left[\frac{p}{c}\right]$,

$$\theta\le\begin{cases}\varphi_2(x), & \textit{если}\ x\ -\ \textit{целое},\\ \max\{\varphi_2([x]),\ \varphi_2([x]+1)\}, & \textit{если}\ x\ -\ \textit{нецелое},\end{cases}$$

где $\varphi_2(x)=\dfrac{cT^m[akx+b(x-c)]}{(k+1)x^2+bcx}$, $x=\dfrac{bc}{ak+b}\left(1+\sqrt{1+\dfrac{ak+b}{k+1}}\right)$, $b=r-1$.

Доказательство. В случае *стационарной* одинаково распределенной системы конкурирующих процессов во всех трех режимах минимальное общее время с учетом параметра $\theta>0$ определяется по формулам:

$$T^{cc}(p,n,s,c,\theta)=$$

$$=\begin{cases}\left(k\frac{n}{c}+\left[\frac{p}{c}\right]-1\right)t^{\theta}, & \textit{если}\ \min(m,s)>\left[\frac{p}{c}\right]\ \textit{и}\ s=k\left[\frac{p}{c}\right],\ k>1,\\ \left((k+1)\frac{n}{c}+r-1\right)t^{\theta} & \textit{если}\ \min(m,s)>\left[\frac{p}{c}\right]\ \textit{и}\ s=k\left[\frac{p}{c}\right]+r,\ k\ge 1,\ 1\le r<\left[\frac{p}{c}\right].\end{cases}$$

Условие эффективности стационарной одинаково распределенной системы конкурирующих процессов в случае $s=k\left[\frac{p}{c}\right]$, $k>1$, определяется соотношением

$$\Delta_{cc}\left(s=k\left[\frac{p}{c}\right]\right)=k\left[\frac{p}{c}\right]T^m-T^{cc}\left(p,n,k\left[\frac{p}{c}\right],c,\theta\right)=$$

$$=\left(\left[\frac{p}{c}\right]-1\right)(kT^m-t)-\left(k\frac{n}{c}+\left[\frac{p}{c}\right]-1\right)\theta\geq 0,$$

которое равносильно выполнению неравенства $\theta\leq\dfrac{aT^m c(kn-c)}{kn^2+acn}$, где $a=\left[\dfrac{p}{c}\right]-1$.

Введем в рассмотрение функцию $\varphi_1(x)=\dfrac{aT^m c(kx-c)}{kx^2+acx}$, которая при $x>0$ достигает своего максимума в точке $x=\dfrac{c}{k}\left(1+\sqrt{\left[\dfrac{p}{c}\right]}\right)$.

В случае $s=k\left[\dfrac{p}{c}\right]+r$, $k\geq 1$, $1\leq r<\left[\dfrac{p}{c}\right]$, условие эффективности стационарной одинаково распределенной системы конкурирующих процессов определяется неравенством:

$$\Delta_{cc}\left(s=k\left[\frac{p}{c}\right]+r\right)=akT^m+b(T^m-t)-\left((k+1)\frac{n}{c}+b\right)\theta\geq 0, \text{ где } b=r-1.$$

С учетом того, что $t=\dfrac{c}{n}T^m$, это равносильно $\theta\leq\dfrac{cT^m[akn+b(n-c)]}{(k+1)n^2+bcn}$.

Рассмотрим функцию $\varphi_2(x)=\dfrac{cT^m[akx+b(x-c)]}{(k+1)x^2+bcx}$. В силу того, что $\varphi_2'(x)=\dfrac{cT^m[(-ak-b)(k+1)x^2+2bc(k+1)x+b^2c^2]}{[(k+1)x^2+bcx]^2}=0$, функция $\varphi_2(x)$ при $x>0$ достигает своего максимума в точке $x=\dfrac{bc}{ak+b}\left(1+\sqrt{1+\dfrac{ak+b}{k+1}}\right)$. Теорема доказана.

6. Заключение

В данной работе для класса *одинаково распределенных систем конкурирующих процессов с учетом ограниченного числа копий программного ресурса* проведен сравнительный анализ математических соотношений времен выполнения множества процессов в асинхронном и двух синхронных режимах, в случаях неограниченного и ограниченного параллелизма по числу процессоров РВС получены достаточные условия эффективности одинаково распределенных систем, доказываются необходимые и достаточные условия существования эффективных систем одинаково распределенных конкурирующих процессов в зависимости от величины дополнительных системных расходов. Полученные условия эффективности имеют многочисленные области применения, в частности, они могут быть использованы при проектировании системного и прикладного программного обеспечения, ориентированного на многопроцессорные вычислительные системы и комплексы, а также при решении проблем *оптимального* использования вычислительных ресурсов.

Список литературы / References

[1]. Andrew S. Tanenbaum, Maarten Van Steen. Distributed Systems. Amazon Digital Services LLC, 2023. 684 p.

[2]. Robey R., Zamora Y. Parallel and High Performance Computing. Manning, 2021. 800 p.

[3]. Бабичев С.Л., Коньков К.А. Распределенные системы. М.: Юрайт, 2019. 507 с. / Babichev S., Konkov K. Distributed Systems. Moscow, Juright, 2019, 507 p. (in Russian).

[4]. Павлов П.А. Оптимальность систем одинаково распределенных конкурирующих процессов. Вестник БГУ. Серия 1: Физика. Математика. Информатика. 2009, №3. С. 114-118. / Pavlov P. Optimality of systems of identically distributed competing processes. BSU Bulletin. Series 1: Physics. Mathematics. Computer Science. 2009, №3. pp. 114-118 (in Russian).

[5]. Павлов П.А., Коваленко Н.С. Математическое моделирование паралллельных процессов. Germany: Lambert Academic Publishing, 2011. 246 с. / Pavlov P., Kovalenko N. Mathematical modeling of parallel processes. Germany, Lambert Academic Publishing, 2011, 246 p. (in Russian).

[6]. Павлов П.А. Оптимальность структурирования программных ресурсов при конвейерной распределенной обработке. Программные продукты и системы. 2010, №3. С. 79-85. / Pavlov P. Optimal structuring of software resources during pipeline distributed processing. Software products and systems. 2010, №3. pp. 79-85 (in Russian).

[7]. Павлов П.А. Задача оптимизации числа процессоров при распределенной обработке. Вестник государственного Самарского аэрокосмического университета имени академика С.П. Королева. 2011, №4. С. 230-240. / Pavlov P. The problem of optimizing the number of processors in distributed processing. Bulletin of the State Samara Aerospace University named after Academician S.P. Koroleva. 2011, №4. pp. 230-240 (in Russian).

[8]. Pavlov P.A. The optimality of software resources structuring through the pipeline distributed processing of competitive cooperative processes. International Journal of Multimedia Technology (IJMT). 2012, Vol.2, №1. pp. 5–10.

[9]. Kovalenko N.S., Pavlov P.A. Optimal Grouping Algorithm of Identically Distributed Systems. Programming and Computer Software. 2012, №3. pp. 143-150.

[10]. Павлов П.А., Коваленко Н.С. Распределенные вычисления при ограниченном числе копий программного ресурса. Программные продукты и системы. 2011, №4. С. 155-163. / Pavlov P., Kovalenko N. Distributed computing with a limited number of copies of a software resource. Software products and systems. 2011, №4. pp. 155-163 (in Russian).

[11]. Kovalenko N.S., Pavlov P.A., Ovseec M.I. Asynchronous distributed computations with a limited number of copies of a structured program resource. Cybernetics and systems analysis. 2012, №1. pp. 86-98.

[12]. Павлов П.А., Коваленко Н.С. Синхронный режим распределенных вычислений при непрерывном выполнении блоков ограниченного числа копий программного ресурса. Программные продукты и системы. 2024, №1. С. 43-53. / Pavlov P., Kovalenko N. Synchronous mode of distributed computing with continuous execution of blocks of a limited number of copies of a software resource. Software products and systems. 2024, №1. pp. 43-53 (in Russian).

Информация об авторах / Information about authors

Павел Александрович ПАВЛОВ – кандидат физико-математических наук, доцент, доцент кафедры информационных технологий и интеллектуальных систем Полесского государственного университета. Сфера научных интересов: математическое моделирование распределенных вычислительных систем конкурирующих процессов, исследование операций, математическое программирование.

Pavel Alexsandrovich PAVLOV – Cand. Sci. (Phys.-Math.), Associate Professor of the department of information technologies and intelligent systems (Polessky State University). Research interests: mathematical modeling of distributed computing systems of competing processes, operations research, mathematical programming.