

УДК 004.75

СИСТЕМЫ ДУБЛИРОВАННЫХ ВЫЧИСЛИТЕЛЬНЫХ КОМПЛЕКСОВ С ПЕРЕРАСПРЕДЕЛЕНИЕМ ЗАПРОСОВ

И.Ю. Голубев

Исследованы подходы к организации структуры отказоустойчивой вычислительной сети на базе дублированных вычислительных комплексов, предусматривающих повышение надежности и производительности на основе динамического перераспределения потока запросов с учетом состояний работоспособности и загруженности компьютерных узлов. Представленные результаты демонстрируют существенную роль выбора кратности резервирования и алгоритмов управления запросами в эффективности системы. Выявлены зависимости эффективности системы от способа распределения вычислительных узлов по разным уровням сети и потока запросов между ними. Показана эффективность обеспечения надежности и производительности вычислительных систем, компонуемых из дублированных вычислительных комплексов на основе динамического распределения запросов.

Ключевые слова: дублированные вычислительные комплексы, распределение нагрузки, вычислительные сети.

Введение

В настоящее время широкое распространение получили вычислительные сети – распределенные вычислительные системы, включающие в себя вычислительные узлы и коммуникационную среду, обеспечивающую их взаимодействие [1–3]. Для систем ответственного целевого назначения актуальной является задача обеспечения их высокой производительности и надежности. Возрастание требований к надежности проектируемой системы обусловлено не только ответственностью целевого приложения, но

и коммерческими факторами и причинами, например, недопустимостью перерывов в работе системы, которые могут приводить к серьезным финансовым потерям.

Постановка задачи

Цель представленной работы – исследование подходов к организации структуры отказоустойчивой вычислительной сети на базе дублированных вычислительных комплексов (ДВК) [1], предусматривающих повышение надежности и производительности на основе динамического перераспределения потока запросов с учетом состояний работоспособности и загруженности компьютерных узлов [4–9].

Задачей исследования является выявление зависимостей эффективности вычислительной сети от кратности резервирования вычислительных элементов и от параметров распределения потока запросов между ними [4–9] с учетом использования в качестве компьютерных узлов ДВК.

На рис. 1 представлена структура рассматриваемой сети. ДВК объединены в кластеры: группу штатных кластеров и один – резервный. Мощности резервного кластера могут использоваться путем перераспределения диспетчерами запросов (ДЗ) потока запросов с целью повышения производительности и надежности вычислительной сети.

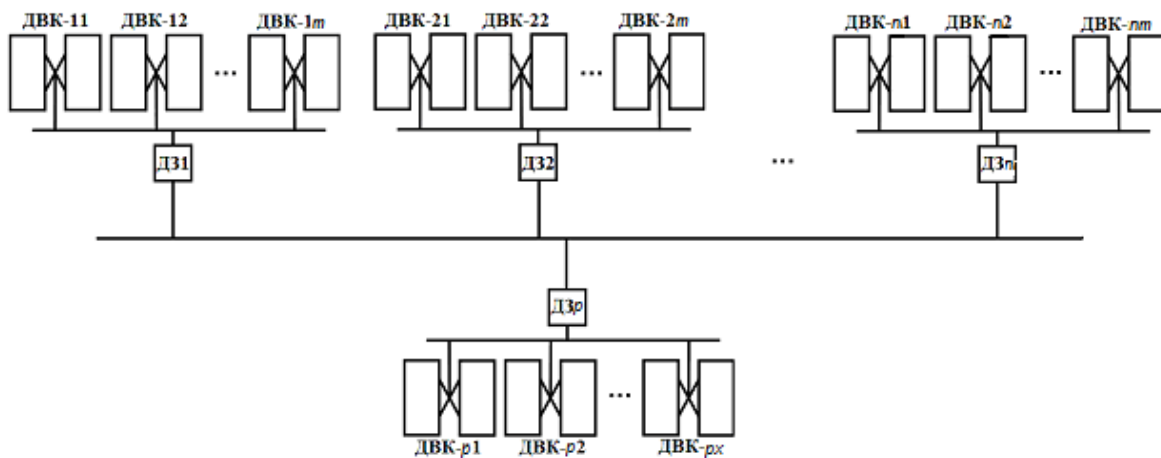


Рис. 1. Структура сети

Предполагается возможность отказов ДВК, а также каналов связи (КС) как внутри ДВК, так и между кластерами сети. ДЗ считаются абсолютно надежными.

Потоки запросов с известной интенсивностью поступают на ДЗ каждого из n штатных кластеров, которые обмениваются между собой информацией об интенсивности входящих в них потоков запросов, а также о состоянии каждого из m включенных в них ДВК. Анализируя значения интенсивностей потоков и состояние ДВК, каждый ДЗ на основе заданного алгоритма принимает решение о необходимости перераспределения части запросов на резервный кластер. При обработке запросов на мощностях резервного кластера время их обслуживания включает в себя время транспортировки данных по КС.

Оценка надежности и задержек в системе

Кластер считается работоспособным, если исправен его КС и хотя бы один из ДВК. Вероятность работоспособного состояния каждого кластера $P_k(t)$ в течение времени t определяется следующим образом: $P_k(t) = p_{кск}(t)(1 - (1 - p_{вм}(t))^{2m})$, где $p_{кск}(t)$ и $p_{вм}(t)$ – вероятность работоспособного состояния канала связи кластера (КСК) и вычислительной машины (ВМ), составляющей ДВК; m – общее количество ДВК в кластере.

Система работоспособна, если либо исправны все кластеры верхнего уровня, либо исправны КС всех кластеров системы и хотя бы один из ДВК резервного кластера. Вероятность работоспособного состояния системы $P_c(t)$ в течение времени t оценивается так:

$$P_c(t) = p_{кск}(t)p_{кск}(t)^n \left[\begin{array}{l} (1 - (1 - p_{вм}(t))^{2d\gamma}) + \\ + (1 - (1 - p_{вм}(t))^{2\frac{d(1-\gamma)}{n-1}})^{n-1} - \\ - (1 - (1 - p_{вм}(t))^{2d\gamma})(1 - (1 - p_{вм}(t))^{2\frac{d(1-\gamma)}{n-1}})^{n-1} \end{array} \right], \quad (1)$$

где $p_{кск}$ – вероятность работоспособности КС между уровнями структуры сети; n – общее количество кластеров в системе; d – общее количество ДВК в системе; γ – доля ДВК в резервном кластере от общего числа ДВК системы.

В предположении экспоненциального распределения времени между отказами вероятность работоспособности компонент сети определяется как $p_{\text{КС}}(t) = \exp(-\lambda_{\text{КС}}t)$, $p_{\text{КСК}}(t) = \exp(-\lambda_{\text{КСК}}t)$, $p_{\text{ВМ}}(t) = \exp(-\lambda_{\text{ВМ}}t)$, где $\lambda_{\text{КС}}$, $\lambda_{\text{КСК}}$, $\lambda_{\text{ВМ}}$ – суммарные интенсивности отказов КС между уровнями структуры сети, КСК и ВМ соответственно. Предполагается, что поступающие в систему запросы образуют простейшие потоки, а длительности обслуживания запросов в ВМ и КС распределены по экспоненциальному закону. Сделанные предположения позволяют построить модель экспоненциальной сети массового обслуживания (СеМО), поддающейся исследованию аналитическими методами [10].

Среднее время пребывания запросов в каждом из кластеров верхнего и нижнего уровней (T и T_p соответственно), с учетом представления кластеров в виде СеМО, в которой ДВК – это система массового обслуживания (СМО) типа М/М/1, определяется формулами

$$T_p = \frac{2}{\mu_{\text{КСК}}} + \frac{\rho}{\mu(1-\rho)}, \quad T = \frac{2}{\mu_{\text{КС}}} + \frac{\rho_p}{\mu(1-\rho_p)},$$

где $\mu_{\text{КСК}}$, $\mu_{\text{КС}}$, μ – величины, обратные времени пребывания запроса соответственно в КСК, в КС между уровнями структуры сети и в каждой из СМО; ρ , ρ_p – коэффициенты использования каждой из СМО кластеров соответственно верхнего и нижнего уровней сети.

Предполагается, что система работает в устойчивом режиме, а нагрузка внутри кластеров верхнего и нижнего уровня распределяется равномерно, таким образом, что в любой момент времени

$$\rho = \frac{\lambda(1-\tau)}{r(1-\gamma)\mu} < 1, \quad \rho_p = \frac{\lambda\tau}{r\gamma\mu} < 1,$$

где λ – интенсивность поступления потока запросов в кластер; τ – доля запросов, перенаправляемых на обработку в резервном кластере; r – количество работоспособных ДВК в системе.

Среднее время пребывания запросов в системе вычисляется

$$T_{\text{ср}} = \sum_{r=1}^d [(1-\gamma)T + \gamma T_p] C_d^r p_{\text{ВМ}}^r(t) (1 - p_{\text{ВМ}}(t))^{d-r}. \quad (2)$$

В качестве оценки эффективности системы используется комплексный показатель, равный отношению полученных в формулах (1), (2) вероятности работоспособного состояния системы и среднего времени пребывания запросов в системе

$$I_{\text{эф}}(t) = P_c(t) / T_{\text{ср}}.$$

Определение кратности резервирования

Предполагается, что изначально нагрузка между штатными кластерами распределяется равномерно. Последующие рассуждения ведутся с учетом рационального в таком случае равномерного распределения ДВК между кластерами сети.

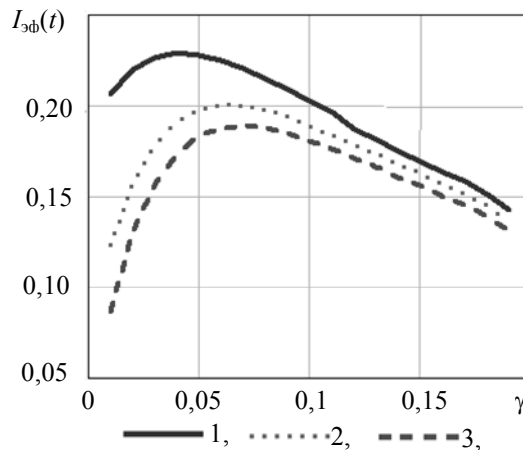


Рис. 2. Зависимость эффективности системы от способа распределения ДВК: кривая 1 соответствует 8 штатным кластерам сети; кривая 2 – 10; кривая 3 – 12

Для расчетов принимается $\lambda_{\text{КС}} = \lambda_{\text{КСК}} = 5 \times 10^{-6} \text{ ч}^{-1}$; $\lambda_{\text{ВМ}} = 5 \times 10^{-4} \text{ ч}^{-1}$; $d = 100$; $t = 5000 \text{ ч}$; $\mu_{\text{КС}} = 0,1 \text{ с}^{-1}$; $\mu_{\text{КСК}} = 2 \text{ с}^{-1}$; $\mu = 2 \text{ с}^{-1}$; интенсивность поступления запросов в систему $\lambda = 10 \text{ с}^{-1}$, доля запросов τ предполагается равной доле γ выделенных в резервный кластер ДВК.

На рис. 2 представлена зависимость эффективности системы по обработке запросов от количества кластеров верхнего уровня и от доли ДВК γ , предназначенных для работы на нижнем уровне сети. Кривая 1 соответствует 8 штатным кластерам сети, кривая 2 – 10, кривая 3 – 12. На рис. 2 видно, что оптимальным для структуры из 8 штатных кластеров является выделение в резервный кластер 4% от общего числа ДВК. Дальнейшие расчеты представлены для структуры из 8 штатных кластеров с найденной кратностью резервирования.

Определение параметров распределения запросов

Рассмотрен подход к управлению потоками запросов, основанный на перераспределении части потока запросов на резервный кластер в случае достижения порогового уровня загрузки. Для определения порогового значения загрузки в соответствии с формулой (2) и выбранной структурой вычислительной сети построена кривая зависимости эффективности системы по обработке запросов от доли потока запросов τ , направляемых на обработку в резервный кластер, представленная на рис. 3.

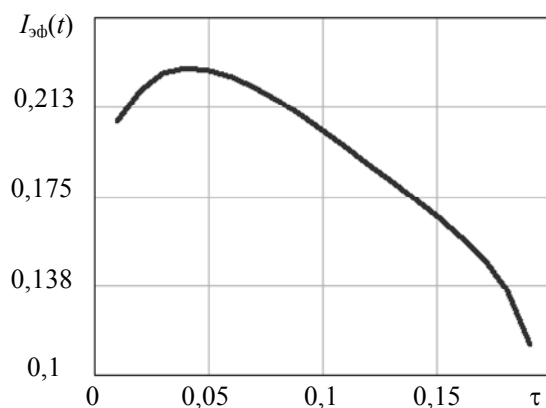


Рис. 3. Зависимость эффективности системы от доли распределения запросов

На рис. 3 видно, что оптимальное значение для доли потока запросов τ , направляемых на обработку в резервный кластер, равно 0,04.

Для проверки полученных аналитическим путем результатов в системе моделирования общецелевого назначения GPSS World была разработана имитационная модель вычислительной сети выбранной структуры. Реализованный в модели алгоритм распределения нагрузки учитывает суммарную длину очередей к каждому из кластеров верхнего уровня: в случае достижения порогового значения длины очереди запросы перенаправляются в резервный кластер.

№	Доля запросов τ , перенаправляемых для обработки в резервном кластере	Среднее время пребывания в системе запроса, обслуживаемого в штатном кластере, с	Среднее время пребывания в системе запроса, обслуживаемого в резервном кластере, с	Эффективность системы по обработке запросов
1	0,01	19,2549	26,72	0,0517
2	0,02	12,7366	22,435	0,0773
3	0,03	9,6016	35,661	0,0964
4	0,04	5,1485	51,791	0,1644
5	0,045	1,9616	55,271	0,229
6	0,05	1,7689	57,399	0,2181
7	0,06	1,6855	72,525	0,1669
8	0,07	1,6654	88,921	0,1331
9	0,08	1,6468	113,572	0,0979
10	0,09	1,6326	154,646	0,0639

Таблица. Результаты имитационного эксперимента

Результаты серии имитационных экспериментов представлены в таблице. Расхождение значения оптимальной доли потока запросов τ , направляемых на обработку в резервный кластер, с полученными ранее аналитическим путем данными не превосходит 10%, что является допустимой для модели погрешностью.

Заключение

Проведенные исследования демонстрируют существенную роль выбора кратности резервирования в эффективности системы дублированных вычислительных комплексов. Показана эффективность обеспечения надежности и производительности вычислительных систем, компонуемых из дублированных вычислительных комплексов на основе динамического распределения запросов.

Исследование выполнено при финансовой поддержке Правительства Санкт-Петербурга.

Литература

1. Голубев И.Ю., Богатырев В.А., Беззубов В.Ф. Сравнительный анализ структур отказоустойчивых дублированных вычислительных комплексов // Информационно-измерительные и управляющие системы. – М.: Радиотехника. – 2011. – Т. 9. – № 2. – С. 8–12.
2. Половко А.М., Гуров С.В. Основы теории надежности: Учебное пособие. – 2-е изд., перераб. и доп. – СПб: БВХ–Петербург, 2008. – 704 с.
3. Takefusa A., Matsuoka S., Aida K. [et al.] Overview of a performance evaluation system for global computing scheduling algorithms // Proceedings of the Eighth IEEE International Symposium on High Performance Distributed Computing (HPDC'99). – 1999. – P. 97–104.
4. Богатырев В.А. Протоколы динамического распределения запросов и отображения функциональных ресурсов в отказоустойчивых вычислительных системах // Электронное моделирование. – 1999. – № 6. – С. 87–96.
5. Богатырев В.А. Комбинаторный метод оценки отказоустойчивости функционально-распределенных вычислительных систем // Электронное моделирование. – 2000. – № 4. – С. 84–92.
6. Богатырев В.А. Оценка надежности функционально избыточных многомашинных вычислительных систем с реконфигурацией на основе перераспределения функций // Электронное моделирование. – 1994. – № 2. – С. 88–90.
7. Богатырев В.А. Распределение заданий в многомашинных вычислительных системах // Изв. вузов. Приборостроение. – 1986. – № 5. – С. 43–47.
8. Богатырев В.А. Безотказность адаптивно-перестраиваемой системы с реконфигурацией на основе перераспределения функций // Изв. вузов. Приборостроение. – 1993. – № 4. – С. 84–87.
9. Богатырев В.А. Динамическое отображение конфигурации в локальных сетях магистральной топологии // Изв. вузов. Приборостроение. – 1993. – № 9–10. – С. 30–35.
10. Алиев Т.И. Основы моделирования дискретных систем: Учебное пособие. – СПб: СПбГУ ИТМО, 2009. – 363 с.

Голубев Иван Юрьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, www.golubev@mail.ru