УДК 681.3

В. А. Богатырев, А. В. Богатырев, С. В. Богатырев

ОЦЕНКА НАДЕЖНОСТИ ВЫПОЛНЕНИЯ КЛАСТЕРАМИ ЗАПРОСОВ РЕАЛЬНОГО ВРЕМЕНИ

Предложен подход к оценке надежности вычислительных систем кластерной архитектуры с учетом требования своевременного выполнения критичных запросов.

Ключевые слова: *надежность, реальное время, вычислительная система.*

Введение. Управляющие вычислительные системы, особенно функционирующие в реальном масштабе времени, должны характеризоваться высокой отказоустойчивостью и надежностью и прежде всего — при выполнении критичных запросов.

Надежность выполнения системами кластерной архитектуры критичных запросов (задач) должна оцениваться с учетом: готовности системы в момент поступления запроса, безотказности во время пребывания запроса в системе, своевременности получения результатов в режиме реального времени.

Первая составляющая оценивается по коэффициенту готовности, вторая — по вероятности безотказной работы, а вторая и первая совместно — по коэффициенту оперативной готовности, третья составляющая может оцениваться по вероятности выполнения запроса с задержкой, меньшей предельно допустимой величины. Рассмотрим вычислительную систему кластерной архитектуры, компонуемую из n одинаковых вычислительных узлов, на которые с интенсивностью Λ поступает общий поток.

Оценка коэффициента готовности. Готовность системы кластерной архитектуры определяется вероятностью застать ее в момент поступления запроса в одном из работоспособных состояний. Рассматривая каждый из узлов кластера как систему массового обслуживания типа M/M/1 [1], находящуюся в стационарном режиме, определим минимальное число s исправных узлов кластера, обеспечивающих его работоспособность. Значение s вычислим как ближайшее целое, не меньшее Λv (v — среднее время выполнения запросов).

Для оценки коэффициента готовности кластера модель его отказов и восстановлений представим моделью размножения и гибели [1, 2]. Будем считать, что при работоспособности i узлов кластера суммарная интенсивность отказов системы $\lambda_i = i\lambda$, а восстановлений $\mu_i = \mu$ (λ и μ — интенсивность отказов и восстановлений одного узла). Модель размножения и гибели с ограниченным восстановлением позволяет определить ее коэффициент готовности как сумму значений вероятности работоспособного состояния r_i [2, 3]:

$$K = \sum_{i=s}^{n} r_i .$$

В случае обслуживания кластера с неограниченным восстановлением, когда $\mu_i = (n-i)\mu$, его коэффициент готовности определим как

$$K = \sum_{i=s}^{n} C_n^i k^i (1-k)^{n-i} ,$$

где $k = t_1 / (t_1 + t_2)$, t_1=1/λ и t_2= 1/μ, $C_n^i = \dfrac{n!}{i!(n-i)!}$.

Полученная оценка значения коэффициента готовности может рассматриваться как верхнее приближение для систем с ограниченным восстановлением.

Оценка оперативной готовности. Будем считать, что каждый запрос распределяется в очередь одного из узлов, при отказе которого во время ожидания в очереди или его обслуживания запрос теряется, так как ограничение на время пребывания запросов в системе не позволяет повторно перераспределять их на выполнение в другой узел.

Рассматривая каждый из i работоспособных узлов кластера как систему массового обслуживания типа М/М/1, среднее время пребывания в нем запросов оценим как $T = v/(1-\Lambda v/i)$, а коэффициент оперативной готовности, т.е. вероятность поступления запроса при работоспособном состоянии кластера и безотказности выделенного узла, в течение времени пребывания запроса в системе, как

$$K = \sum_{i=s}^{n} r_i \exp(-\lambda T) = \sum_{i=s}^{n} r_i \exp\left(-\lambda\left(\frac{v}{1-\Lambda v/i}\right)\right).$$

Оценка вероятности своевременного выполнения запросов. Надежность функционирования системы P определим как вероятность $p(t_0)$ своевременного выполнения запроса с задержкой, меньшей предельно допустимой величины t_0 [4], при готовности системы в момент поступлении запроса (коэффициент готовности) и ее безотказности во время ожидания и обслуживания запроса (коэффициент оперативной готовности):

$$P = \sum_{i=s}^{n} r_i \exp\left(-\lambda\left(\frac{v}{1-\Lambda v/i}\right)\right) p(t_0),$$

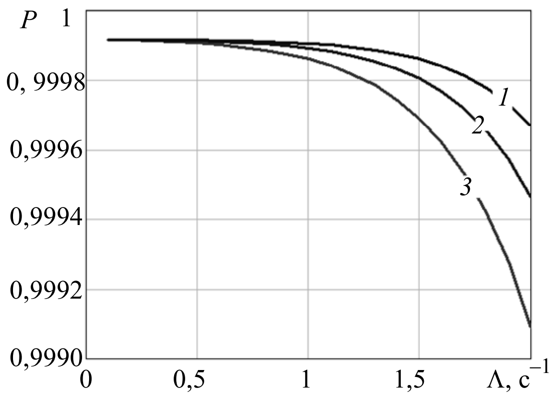
где $p(t_0) = 1 - \frac{\Lambda v}{i} \exp\left(-\left(\frac{1}{v} - \frac{\Lambda}{i}\right) t_0\right)$.

Рассмотрим решение задачи, требующей своевременного выполнения всех запросов, поступающих с интенсивностью Λ.

Надежность функционирования систем при своевременном выполнении запросов, поступающих во время решения задачи (с выделением на каждую задачу одного узла кластера), оценим как

$$P = \sum_{i=s}^{n} r_i \exp(-\lambda t) \left[1 - \Lambda v \exp\left(-\left(\frac{1}{v} - \Lambda\right) t_0\right)\right]^{\Lambda t}.$$

Пример расчета. Зависимость вероятности своевременного выполнения критических запросов от интенсивности их поступления Λ представлена на рисунке кривыми *1—3* соответственно для $t_0 = 10$, 11, 12 с. Расчеты произведены при n=5 $v = 1$ с, $\lambda = 10^{-4}$ ч$^{-1}$, $\mu = 1$ ч$^{-1}$.



Математические модели надежности распределенных систем реального времени [5] могут быть детализированы с учетом влияния временных издержек на диспетчеризацию и множественный доступ к сетевым ресурсам [6—12].

Заключение. Таким образом, для кластерных систем предложен комплексный показатель надежности, учитывающий готовность вычислительной системы, ее безотказность и вероятность своевременного выполнения критичных запросов.

Предложенная модель может быть использована при оценке надежности вычислительных и управляющих систем в случае выполнении ими критичных запросов в реальном времени.

СПИСОК ЛИТЕРАТУРЫ

1. *Алиев Т. И.* Основы моделирования дискретных систем. СПб: СПбГУ ИТМО, 2009. 363 с.

2. *Гуров С. В., Половко А. М.* Основы теории надежности. СПб: БХВ-Петербург, 2006. 704 с.

3. *Черкесов Г. Н.* Надежность аппаратно-программных комплексов. СПб: Питер, 2005. 479 с.

4. *Хинчин А. Я.* Работы по математической теории массового обслуживания. М.: Гл. изд-во физ.-мат. лит-ры, 1963. 236 с.

5. *Богатырев В. А., Богатырев А. В.* Функциональная надежность систем реального времени // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 4. С. 150—151.

6. *Богатырев В. А.* Счетно-эстафетный метод множественного доступа с динамическим отображением конфигурации локальных сетей магистральной топологии // Автоматика и вычислительная техника. 1992. № 2. С. 50—54.

7. *Богатырев В. А.* Счетно-интервальный метод множественного доступа к общей магистрали // Электронное моделирование. 1991. № 2. С. 96—98.

8. *Богатырев В. А.* Счетно-эстафетный метод множественного доступа к магистрали // Электронное моделирование. 1991. № 3. С. 32—36.

9. *Богатырев В. А.* Бесприоритетный счетно-импульсный метод множественного доступа // Изв. вузов СССР. Приборостроение. 1990. № 12. С. 25—29.

10. *Богатырев В. А.* Повышение отказоустойчивости многомагистрального канала с помощью межмагистральной ретрансляции пакетов // Автоматика и вычислительная техника. 1999. № 2. С. 81—88.

11. *Богатырев В. А.* Отказоустойчивость и сохранение эффективности функционирования многомагистральных распределенных вычислительных систем // Информационные технологии. 1999. № 9. С. 44—48.

12. *Bogatyrev V. A.* Exchange of Duplicated Computing Complexes in Fault tolerant Systems // Automatic Control and Computer Sciences. 2011. Vol. 46, N 5. P. 268—276.

Сведения об авторах

Владимир Анатольевич Богатырев — д-р техн. наук, профессор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: Vladimir.bogatyrev@gmail.com

Анатолий Владимирович Богатырев — аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; E-mail: Vladimir.bogatyrev@gmail.com

Станислав Владимирович Богатырев — Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра безопасности информационных систем; младший научный сотрудник; E-mail: Vladimir.bogatyrev@gmail.com

Рекомендована кафедрой вычислительной техники

Поступила в редакцию 23.12.13 г.