

В. А. БОГАТЫРЕВ, С. В. БОГАТЫРЕВ, А. В. БОГАТЫРЕВ

## ОЦЕНКА НАДЕЖНОСТИ ОТКАЗОУСТОЙЧИВЫХ КЛАСТЕРОВ С НЕПОСРЕДСТВЕННЫМ ПОДКЛЮЧЕНИЕМ УСТРОЙСТВ ХРАНЕНИЯ

Предложен подход к оценке надежности кластеров с прямым подключением двухходовых устройств хранения и серверов при ограничении предельно допустимого времени пребывания запросов в системе. Проанализировано влияние вариантов такого подключения на надежность системы с учетом вероятности выполнения запросов в заданные сроки.

**Ключевые слова:** отказоустойчивость, кластер, надежность, резервирование, устройство хранения, сервер.

**Введение.** Высокие отказоустойчивость, доступность, надежность и производительность распределенных вычислительных систем достигаются при объединении узлов обработки и хранения данных в кластеры [1, 2].

При построении небольших кластерных систем в ряде случаев целесообразно прямое подключение узлов разных уровней между собой, например, серверных узлов и устройств хранения (архитектура DAS, Directly Attached Storage [1]).

Поскольку надежность и производительность кластера зависят от вариантов объединения его узлов, представляет интерес исследование альтернатив, обеспечивающих большую отказоустойчивость, надежность и доступность системы при одинаковых затратах на ее реализацию [3—5].

**Двухуровневые структуры с непосредственным подключением резервированных узлов.** Рассмотрим варианты конфигурации двухуровневых кластерных систем. Будем считать, что на верхнем уровне (ВУ) выделяется  $m$  дублированных узлов. Таким образом, общее число узлов верхнего уровня равно  $2m$ , общее число узлов нижнего уровня (НУ) будем также считать равным  $2m$ . Узлы верхнего уровня объединяются в пары (дублируются) по функциональной принадлежности реализуемых приложений или по разделению обслуживаемого ими потока запросов [6—9].

Представляет определенный интерес исследование вариантов построения кластеров с прямым подключением узлов разного уровня, в частности серверов, к устройствам хранения [1, 2]. Цель таких исследований — выбор конфигурации, обеспечивающей при одинаковых затратах на реализацию системы ее большую отказоустойчивость, надежность и доступность.

В рамках указанного направления в работах [10, 11] исследованы варианты прямого подключения дублированных серверов к устройствам хранения без учета ограничений на время обслуживания запросов. Поскольку для систем, функционирующих в реальном времени, такие ограничения важны, в настоящей статье исследуется влияние конфигураций прямого подключения узлов двухуровневого кластера на выполнение функциональных запросов при ограничении времени их пребывания в системе.

**Варианты построения кластера.** Рассмотрим конфигурации кластерных систем с непосредственным подключением устройств хранения к серверам приложений  $m$  типов [10, 11]. На рис. 1 приведены варианты кластеров:  $a$  — S1,  $b$  — S2,  $c$  — S3. Будем считать, что запрос  $j$ -го потока (к  $i$ -му приложению) может быть обслужен хотя бы одним из выделенных для этого потока серверов (сервером  $i$ -го приложения) при взаимодействии с любым из подключенных к нему устройств хранения (узлов НУ). Кратность резервирования серверов будем считать  $2/1$ , в этом случае общее число узлов ВУ  $2m$ , общее число узлов НУ также будем считать равным  $2m$ .

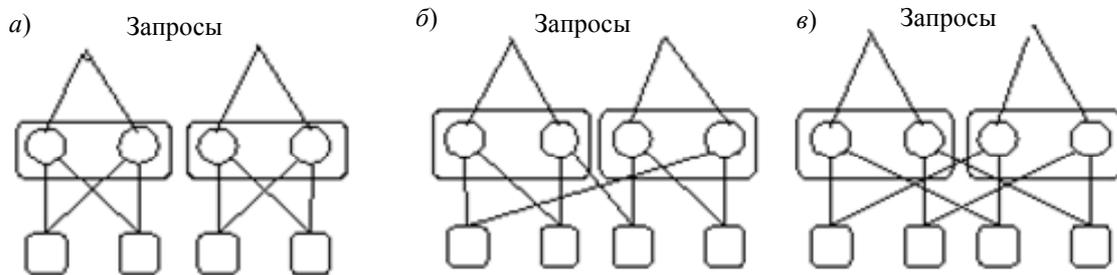


Рис. 1

Будем считать, что все серверы и устройства хранения (узлы ВУ и НУ) имеют по два порта (входа), что позволяет строить резервированные системы с отсутствием единой точки отказа без использования коммутаторов, которые в ряде случаев могут вносить дополнительную ненадежность в систему.

Проанализируем надежность представленных на рис. 1 вариантов структур [11] с учетом ограничений по предельно допустимому времени обслуживания запросов. Подчеркнем что эти варианты характеризуются одинаковыми затратами на их построение, но позволяют достичь различных уровней надежности и отказоустойчивости, что и обуславливает актуальность их исследования.

При исправности двух резервированных серверов некоторого приложения конфигурации S1, S2 и S3 сохраняют функциональность, если действует хотя бы одно из двух, трех либо четырех связанных с рассматриваемыми серверами устройств хранения соответственно. При дееспособности одного сервера приложения для всех вариантов конфигураций (см. рис. 1) функциональность сохраняется, если исправно хотя бы одно из двух связанных с сервером устройств хранения. Таким образом, отказоустойчивость конфигурации S3 выше, оценим надежность ее и базовой конфигурации S1 [10].

Вычислим вероятность безотказной работы (ВБР) конфигурации S1 (рис. 1,  $a$ ) при условии исправности хотя бы одного сервера и одного устройства хранения в каждой из  $m$  дублированных групп [11]:

$$P_1(t) = \left\{ \left[ 1 - (1 - p_1(t))^2 \right] \left[ 1 - (1 - p_2(t))^2 \right] \right\}^m,$$

где  $p_1(t) = \exp(-\lambda_1 t)$ ,  $p_2(t) = \exp(-\lambda_2 t)$ ,  $\lambda_1$ ,  $\lambda_2$  — интенсивность отказов сервера и устройства хранения.

В этой формуле не введены ограничения на время выполнения запросов, характерные для систем реального времени. С учетом этих ограничений считается, что система может выполнить некоторое приложение (поток запросов), если запрос может быть обслужен хотя бы одним узлом ВУ и хотя бы одним узлом НУ, а время пребывания запросов не превышает  $t_0$ .

Будем считать, что среднее время обслуживания запросов в серверах включает их взаимодействие с устройствами хранения, тогда

$$P_1(t) = \left[ 1 - (1 - p_2(t))^2 \right] \left[ (p_1(t))^2 g_2 + 2p_1(t)(1 - p_1(t))g_1 \right].$$

Вычислим вероятность того, что время пребывания запросов в  $i$  параллельно работающих серверах меньше  $t_0$  [12]:

$$g_i = \begin{cases} 0, & \text{if } (\Lambda v / i) \geq 1, \\ 1 - \left( \frac{\Lambda v}{i} \right) \exp\left(-\left(\frac{1}{v} - \frac{\Lambda}{i}\right)t_0\right), & \text{if } (\Lambda v / i) < 1, \end{cases}$$

Оценим ВБР системы для конфигураций S2 и S3 при условии неперевышения  $t_0$  (ограничения по времени пребывания запросов не учитываются, т.е.  $g_1=1, g_2=1$ ):

$$P_2(t) = \left[ 1 - (1 - p_2(t))^3 \right] (p_1(t))^2 g_2 + 2p_1(t)(1 - p_1(t)) \left[ 1 - (1 - p_2(t))^2 \right] g_1,$$

$$P_3(t) = \left[ 1 - (1 - p_2(t))^4 \right] (p_1(t))^2 g_2 + 2p_1(t)(1 - p_1(t)) \left[ 1 - (1 - p_2(t))^2 \right] g_1.$$

Кривые 1—3 на рис. 2 иллюстрируют результаты расчета вероятности безотказной работы структур S1—S3 при  $\lambda_1=10^{-4}$ ,  $\lambda_2=2 \cdot 10^{-4}$  ч<sup>-1</sup>. Из графиков видна зависимость ВБР исследуемых структур от варианта соединения двухвходовых серверов и устройств хранения, причем затраты на реализацию сравниваемых структур одинаковы.

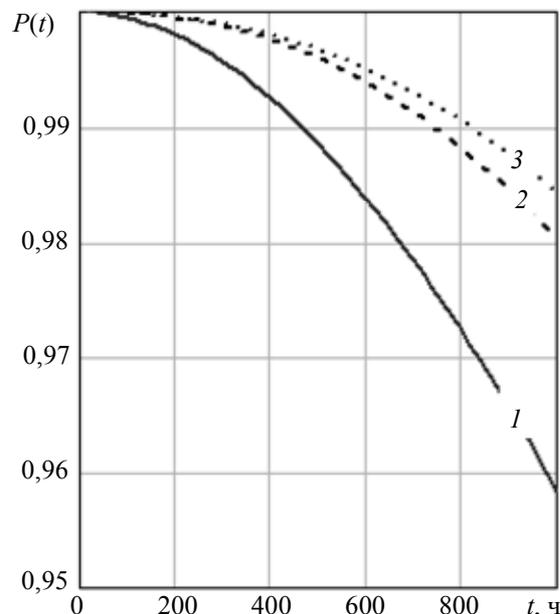


Рис. 2

На рис. 3 представлены результаты расчета вероятности выполнения запросов за время, меньшее  $t_0$  ( $\lambda_1=10^{-4}$ ,  $\lambda_2=2 \cdot 10^{-4}$  ч<sup>-1</sup>, среднее время обслуживания запроса сервером, с учетом обращений к системе хранения — 1 с). Кривые 1, 2 соответствуют ВБР конфигураций S3 и S1;

кривые 3, 4, 5 для конфигурации S3 (6, 7, 8 — для S1) соответствуют вероятности выполнения запросов при  $t_0=10, 5, 2$  с соответственно. Интенсивность потока рассматриваемых функциональных запросов  $1 \text{ с}^{-1}$  (рис. 3, а) и  $1,5 \text{ с}^{-1}$  (рис. 3, б).

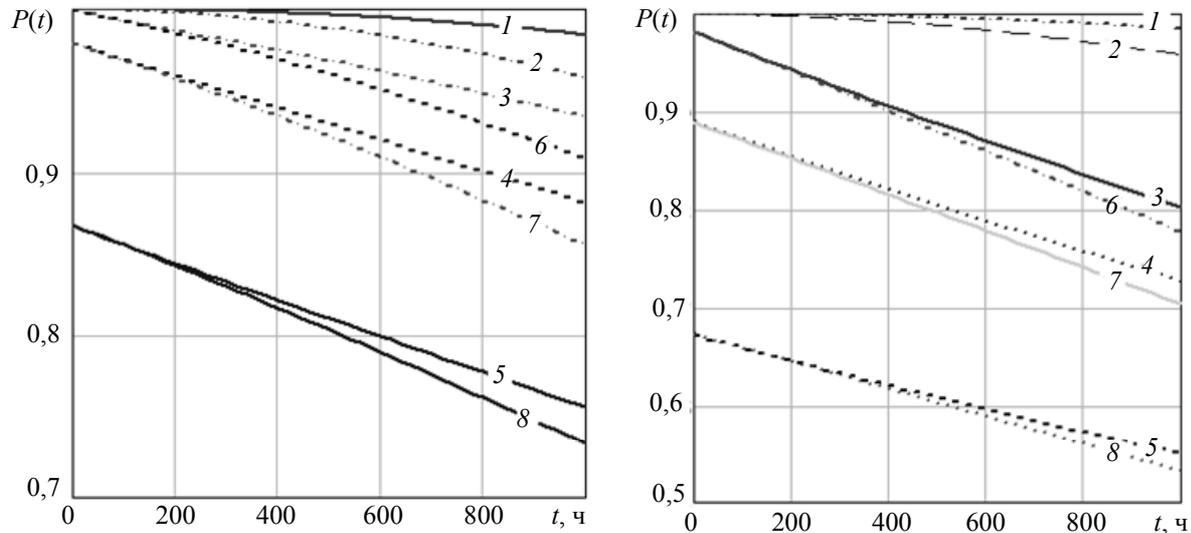


Рис. 3

Проведенные расчеты подтвердили существенную зависимость надежности кластеров от порядка прямого подключения устройств хранения к серверам приложений при явном преимуществе конфигурации S3 по отказоустойчивости, ВБР и надежности в случае выполнения запросов в установленные сроки.

**Заключение.** Предложен подход к оценке надежности кластерных вычислительных систем с прямым подключением двухвходовых устройств хранения и серверов, позволяющий учитывать вероятность выполнения запросов при непревышении предельно допустимого времени их пребывания в системе. Продемонстрировано влияние вариантов подключения серверов и устройств хранения на надежность системы при одинаковых издержках на ее реализацию. Даны рекомендации по выбору вариантов конфигурации резервированных кластеров.

#### СПИСОК ЛИТЕРАТУРЫ

1. Juud J. Principles of SAN Design. San Jose: Brocade Bookshelf, 2008. 589 p.
2. Clark T. The New Data Center. New technologies are radically reshaping the data center. Brocade Bookshelf. San Jose, 2010. 156 p.
3. Богатырев В. А. Надежность функционально-распределенных резервированных структур с иерархической конфигурацией узлов // Изв. вузов. Приборостроение. 2000. Т. 43, № 4. С. 67—70.
4. Богатырев В. А. О влиянии размещения функциональных ресурсов на отказоустойчивость информационно-вычислительных систем с функциональной реконфигурацией // Информационные технологии. 2002. Т. 45, № 2. С. 10—15.
5. Богатырев В. А. Безотказность систем из функционально неоднородных модулей // Приборы и системы. Управление, контроль, диагностика. 2002. № 3. С. 6—8.
6. Богатырев В. А. К размещению резервированных функциональных ресурсов в системах с функциональной реконфигурацией // Управляющие системы и машины. 2003 № 3. С. 42—45.
7. Богатырев В. А. Надежность вариантов размещения функциональных ресурсов в однородных вычислительных сетях // Электронное моделирование. 1997. № 3. С. 21—25.
8. Богатырев В. А. К оценке надежности систем из многофункциональных модулей // Автоматизация и современные технологии. 2001. № 6. С. 12—15.

9. Богатырев В. А., Богатырев С. В. Объединение резервированных серверов в кластеры высоконадежной компьютерной системы // Информационные технологии. 2009. № 6. С. 41—47.
10. Bogatyrev V. A. Fault Tolerance of Clusters Configurations with Direct Connection of Storage Devices // Automatic Control and Computer Sciences. 2011. Vol. 45, N 6. P. 330—337.
11. Богатырев В. А., Богатырев А. В., Богатырев С. В. Надежность кластерных вычислительных систем с дублированными связями серверов и устройств хранения // Информационные технологии. 2013. № 2. С. 27—32.
12. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979. 432 с.

**Сведения об авторах**

- Владимир Анатольевич Богатырев** — д-р техн. наук; Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра безопасности информационных систем; профессор;  
E-mail: Vladimir.bogatyrev@gmail.com
- Станислав Владимирович Богатырев** — Санкт-Петербургский государственный университет аэрокосмического приборостроения, кафедра безопасности информационных систем; младший научный сотрудник;  
E-mail: realloc@gmail.com
- Анатолий Владимирович Богатырев** — аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники;  
E-mail: ganglion@gmail.com

Рекомендована кафедрой  
№ 51 безопасности информационных систем

Поступила в редакцию  
01.02.13 г.