



РЕЗЕРВНЫЙ ЦЕНТР ОБРАБОТКИ ДАННЫХ

ОЦЕНКА НАДЕЖНОСТИ

Не секрет, что сегодня объемы информации растут лавинообразно и обычно она обрабатывается централизованно, а затем передается в территориально распределенные филиалы и клиентам. Это означает, что каждой крупной организации необходимо единое информационное пространство. Сегодня практически все крупные ведомства и корпорации имеют информационно-телекоммуникационные системы (ИТКС), обеспечивающие скоростной доступ к коллективным информационным ресурсам и различным телекоммуникационным сервисам. Для ряда ИТКС вопросы надежности принципиальны, особенно в военном, государственном и банковском секторах. Сегодня, как правило, применяют резервирование наиболее важных узлов, устанавливая рядом резервный комплект. Но этого часто недостаточно для обеспечения непрерывности бизнес-процессов и эффективного управления компаний.

КАТАСТРОФООУСТОЙЧИВЫЕ ЦЕНТРЫ ОБРАБОТКИ ДАННЫХ

Если корпоративная сеть и информационные системы корпораций и министерств функционируют ненадежно, это большой риск, связанный с потерей данных или недоступностью информационных сервисов организации. Поэтому, чтобы обеспечить стабильность управления государственными структурами или непрерывность бизнес-процессов в коммерческих организациях, помимо основных информационных систем используют резервные центры (РЦ), основу которых составляют резервные центры обработки данных (ЦОД) или дата-центры (Data Center).

Основная задача резервного центра обработки данных – обеспечивать непрерывность вычислительного процесса при выходе из строя основных систем. Чтобы добиться толерантности основных сервисов ИТКС к различным аварийным ситуациям, нужно разместить на резервной площадке дублирующий комплект серверов и телекоммуникационного оборудования, а также внедрить механизмы переключения информационных нагрузок и механизмы репликаций для поддержания актуальной копии данных.

М.Куперман, д.т.н.,
Д.Аверьянов, к.т.н.

Объемы информации постоянно растут, требования к ее сохранности и непрерывности процесса обработки ужесточаются, поэтому компаниям приходится сочетать методы консолидации информационных ресурсов в составе ЦОД и структурного резервирования их аппаратно-программных комплексов. Наиболее эффективное решение – резервирование совместно с территориальным разнесением узлов обработки и хранения данных (последнее обеспечивает катастрофоустойчивость ядра информационно-телекоммуникационной инфраструктуры организации). Такое решение делает систему максимально жизнеспособной как при воздействии на нее локальных деструктивных факторов, в том числе отказов аппаратуры, сбоев программного обеспечения, ошибочных или злонамеренных действий пользователей, так и при воздействии различных внешних угроз – терроризма, пожара, затопления, аварий в сетях энергоснабжения или сетях передачи данных провайдера, а также других техногенных катастроф и стихийных бедствий. Резервный центр, удаленный от основной площадки, обеспечивает не только отказоустойчивость, но и катастрофоустойчивость, определяющую максимальную надежность ИТКС и, следовательно, минимальный риск остановки бизнес-процессов организации.

Обеспечение катастрофоустойчивой конфигурации ИТКС организации (ведомства) – это непростая и затратная задача, поэтому решается она только для ответственных применений. Резервные центры широко распространены в банковском секторе, где бизнес и репутация компании напрямую зависят от доступности информационной системы. Здесь стоимость простоя платежных систем настолько велика, что строительство резервного центра всегда окупается.

Обычно серьезные проблемы в банковских ИТКС тщательно скрываются, так как могут отразиться на лояльности клиентов и престиже банка в целом. Поэтому, чтобы проиллюстрировать то, как эффективная работа финансовых учреждений зависит от качества функционирования информационно-телекоммуникационных систем, мы приведем примеры остановки информационных систем крупнейших зарубежных бирж. В апреле 2000 года на восемь часов (почти на всю



торговую сессию) были парализованы торги на Лондонской фондовой бирже (LSE) из-за ошибок в работе программного обеспечения. Убытки оценили в несколько миллионов фунтов. В сентябре 2008 года там же был зафиксирован сбой, и работа биржи была остановлена на четыре часа. Это не только привело к прямым потерям – около 40% дохода биржи приходится на продажу информации о курсах акций в режиме реального времени, – но и пошло на пользу конкурентам LSE. Известны также случаи крупных аварий, ставших причиной остановки критически важных сервисов на Токийской фондовой бирже и NASDAQ*. Очевидно, что обеспечение непрерывности бизнес-процессов (business continuity) является первоочередной задачей финансовых учреждений.

Резервные центры имеют Банк России, "Альфа-Банк", Банк ВЕФК, Московский кредитный банк, подразделения Сбербанка и другие. Резервные центры широко распространены в Министерстве обороны и других силовых ведомствах, например, крупный резервный ЦОД внедряется в интересах пограничной службы для обработки данных биометрических загранпаспортов граждан России.

СТРУКТУРА ОСНОВНОГО И РЕЗЕРВНОГО ЦЕНТРА

При создании толерантных к серьезным авариям и катастрофам систем используется территориальное разнесение вычислительных платформ и информационных хранилищ: кампусных, метро-кластеров, геокластеров (geocluster). В катастрофоустойчивых геокластерах, как правило, узлы разнесены на расстояние не менее 50 км. В типовой структуре основного и резервного центра (рис.1), основанного на геокластерах, можно выделить три основных элемента обеспечения катастрофоустойчивости:

- протяженная скоростная магистраль, соединяющая обе площадки, в данном случае реализованная в виде DWDM-кольца;
- вычислительный геокластер (узлы кластера находятся на основной и резервной площадках);
- геокластер систем хранения (SAN, Storage Area Network).

В отличие от задачи обеспечения отказоустойчивости, которая не требует территориального разнесения сетевых узлов (следовательно, могут применяться технологии локальных сетей), для выполнения основной задачи резервного центра – обеспечения катастрофоустойчивости – необходимы магистральные линии связи и организация геокластера.

Часть приведенных выше подсистем, на первый взгляд, может показаться излишней даже для крупного банка и актуальной только для военных систем. Однако энергетический коллапс в мае 2005 года показал уровень толерантности эксплуатируемых систем. Торги на биржах были остановлены, а некоторые банки вынуждены были прекратить работу части своих офисов и филиалов. Сбербанк выступил с официаль-

ным обращением к гражданам о "возможных временных затруднениях в обслуживании клиентов банка".

Учитывая, что сегодня банковские системы, как правило, имеют централизованную обработку платежной и другой информации, даже наличие автономных источников электроэнергии на основной площадке не гарантирует связь с внешними объектами (филиалами, клиентами, другими платежными сетями). Нет гарантий работоспособности и всей цепочки операторов связи в проблемном регионе. Кроме того, работа руководства и структурных подразделений организации в таких условиях будет затруднена на основной площадке из-за дефицита энергии, генерируемой резервными источниками. От способности "безболезненно" переносить подобные коллапсы и любые другие катаклизмы зависит финансовая устойчивость и престиж банка. Поэтому все более широкое применение находят резервные центры, позволяющие предотвращать не только технические аварии и ошибки обслуживающего персонала, но и крупные пожары, обрушения зданий, заражение опасными вирусами, террористические акты (ложную информацию о них) и другие чрезвычайные ситуации. Но надо сказать, что целевое назначение резервного центра – повышение надежности – должно быть количественно обосновано.

ПОДХОДЫ К ОЦЕНКЕ НАДЕЖНОСТИ

Различают два подхода к оценке надежности резервного центра: оценка катастрофоустойчивой конфигурации и оценка отказоустойчивой конфигурации. При оценке катастрофоустойчивости, обеспечиваемой резервным центром, обычно рассматривается схема, приведенная на рис.2. Часто при

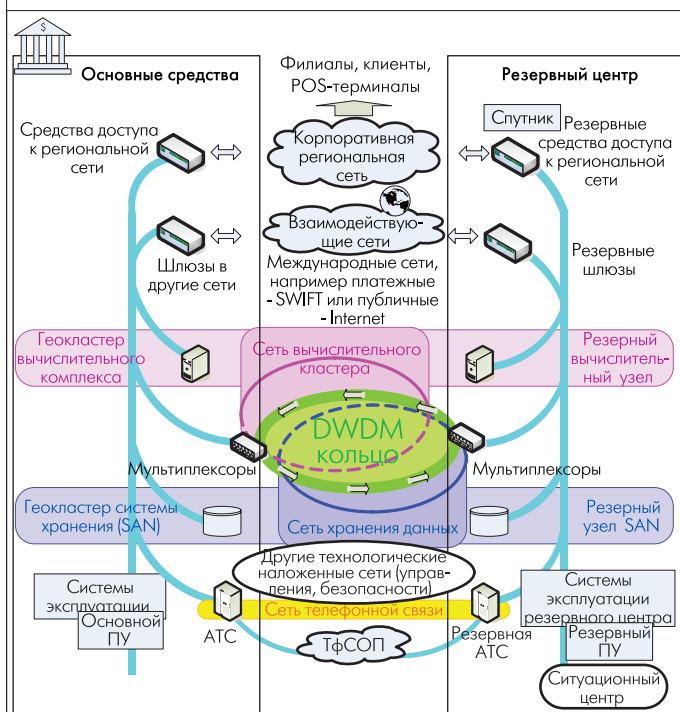


Рис.1. Типовая структура основного и резервного центра

* Газета "Коммерсантъ" № 68/П (3644) от 23.04.2007. <http://www.kommersant.ru/doc.aspx?DocsID=761230&print=true>

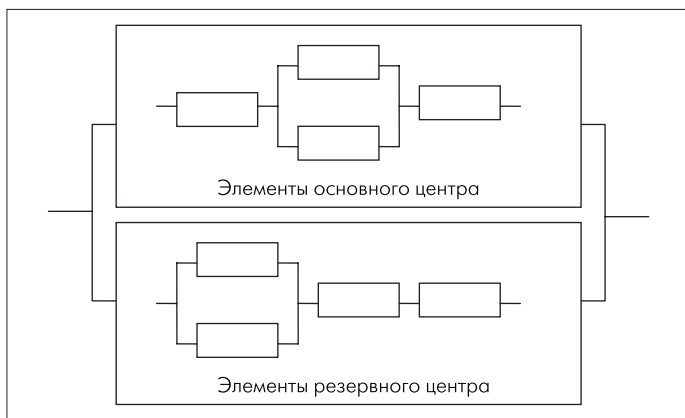


Рис.2. Схема для оценки катастрофоустойчивости

этом оценивается не надежность резервируемой структуры при выходе из строя отдельного элемента основного центра, а вероятность крупной аварии или возникновения чрезвычайной ситуации (например, катастрофы), при которых неработоспособными считаются все элементы основной площадки.

При оценке надежности отказоустойчивой конфигурации может рассматриваться поэлементная декомпозиция площадок с агрегированием элементов обеих площадок в составе одной функциональной подсистемы (рис.3).

Заметим, что с позиции надежности эта схема обладает лучшими показателями надежности по сравнению со схемой на рис.2. В этом случае имеет место гибкое перераспределение резервов между основной и резервной площадками. Так, например, допустимы одновременный выход из строя основного вычислительного центра и резервного узла связи.

При оценке надежности очень важна формулировка критерия отказа системы, в зависимости от которого разрабатывается надежностная схема. При задании критерия надежности указывается допустимое число отказавших устройств от их общего числа, предельное время неработоспособности (недоступности сервиса, например, при переключении на резерв) и другие параметры и условия.

Для оценки надежности отказоустойчивой конфигурации используется схема (рис.3), элементы которой представляют собой последовательно соединенные укрупнен-

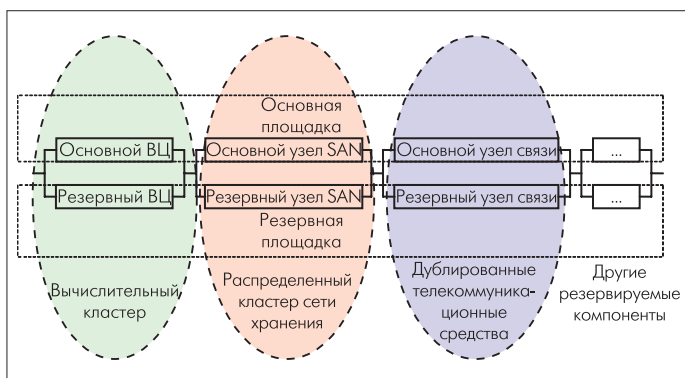


Рис.3. Надежностная схема для оценки отказоустойчивости

ные кластерные структуры: вычислительный кластер, распределенный кластер системы хранения, телекоммуникационный кластер. В практической реализации каждый из них может представлять собой набор нескольких кластеров: кластер платежного приложения, кластер платежного приложения (кластеры серверов приложений), кластеры СУБД, кластеры сетевой инфраструктуры (контроллеры доменов, DNS) и др. В простейшем случае оценка надежности сводится к расчету схемы, составленной по заданному критерию отказа последовательно соединенных кластерных структур, представленных дублированной группой узлов. Не следует забывать о необходимости учета надежности инфраструктуры ЦОД, в первую очередь системы электропитания и кондиционирования, без которых основные системы неработоспособны. Как правило, инфраструктурные элементы также задублированы, и рассчитывать их можно аналогично.

МАРКОВСКИЕ ПРОЦЕССЫ

Статические модели, которые построены на методах, использующих основные формулы теории вероятности, комбинаторики, других логико-вероятностных методов, применяемых, главным образом, для описания последовательно-параллельных структур, не позволяют учитывать изменения в характеристиках и процессах в зависимости от уже происходящих событий, отказов. Поэтому выбор модели надежности для описания резервированной структуры резервного центра, его составных частей, ЦОД, кластерных структур обусловлен классом динамического моделирования.

Моделирование кластерных структур марковскими процессами позволяет отразить в модели и учесть изменения процессов и отказов элементов во времени, временные условия реализации других событий. Марковский процесс обладает характерными свойствами, которые определяются, в первую очередь, экспоненциальными распределениями времени пребывания в каждом состоянии.

Отказ дублированной группы наступает тогда, когда во время восстановления одного из узлов отказывает второй узел. Возможные состояния:

- "0" – оба узла исправны;
- "1" – отказ в одном узле;
- "2" – отказ в обоих узлах.

Таким образом, состояния исправности системы – "0", "1", отказа – "2".

В случае отказа одного из элементов группы отказавший узел ремонтируется (заменяется) без остановки системы и после восстановления через случайный промежуток времени, распределенный по экспоненциальному закону с параметром μ , включается в состав дублированной группы: $\mu = 1/T_B$, где T_B – среднее время восстановления. Допустим, что одновременно может восстанавливаться один узел.



При расчете модели кластера с двумя узлами и идеальной системой контроля (рис.4) получаются явно завышенные значения показателей надежности, не отражающие реальную надежность системы. При исходной интенсивности отказов $\eta = 0,00005$ 1/ч (наработка на отказ составляет 20 000 ч) и интенсивности восстановления $\eta = 0,25$ 1/ч (4 ч восстановления) получим из расчета графа (см. рис.4) значение $K_r = 0,999\ 999\ 92$ (семь девяток). Подчеркнем, что взятая наработка 20 000 ч является нижней планкой MTBF (MeanTime Before Failure, средняя наработка на отказ) серверных платформ. Обычно для серверов приводятся значения 50–100 тыс. часов и, следовательно, получаются чрезмерно "хорошие" результаты.

Модель (см. рис.4) применима, например, для систем класса Stratus Continuum, в которых каждые два физических процессора объединяются парами и одновременно выполняют одну и ту же команду. При этом схема сравнения в каждом такте проверяет и устанавливает факт, что оба процессора пары дают одинаковый результат. Если результаты в паре различаются, то принимается решение о сбое, а пользователь использует результаты другой пары. Даже при такой организации вычислительного процесса FT-системы Stratus Continuum, обеспечивающие непрерывную готовность (Continuous Availability), декларируют коэффициент готовности, равный 99,999% (пять девяток – время недоступности системы 5 мин/год). При этом четыре процессора выполняют единственную команду с потактовой синхронизацией и сравнением результата.

Несмотря на такие фантастические значения, рассмотренная выше модель для расчета кластерных структур является самой распространенной и очень удобной для подтверждения высокой надежности создаваемых "отказоустойчивых" систем. Однако она слишком упрощена и не соответствует реальным процессам. Модель не учитывает ряд факторов, существенно влияющих на надежность кластера. Речь не идет о надежности внешних по отношению к узлам кластера элементов (в принципе, они также должны быть учтены), в частности о коммуникационной среде, – например об элементах СКС и ЛВС. Будем считать, что такие факторы либо не оказывают влияния на состояние кластера, заданное критерием отказа, либо надежностью этих элементов можно пренебречь.

Важно, что модель (см. рис.4) не учитывает надежность программного обеспечения (ПО), так как в качестве параметра потока отказов задается наработка на отказ аппаратных средств. Вообще оценка надежности программного обеспечения – отдельная малоизученная тема, до сегодняшнего дня не имеющая эффективных методов оценки. Однако некоторые подходы к оценке отказоустойчивых структур позволяют существенно повысить адекватность уточненных моделей по сравнению с идеализированными.

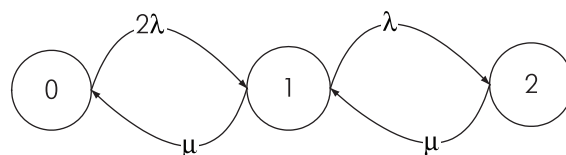


Рис.4. Граф состояний кластера с двумя узлами и идеальной системой контроля

СОСТОЯНИЕ НЕОБНАРУЖЕННОГО ОТКАЗА

В каждом элементе могут быть скрытые отказы. Модель с двумя узлами и идеальной системой контроля не учитывает вероятность обнаружения отказа, надежность "переключателя" резерва, задержку при переключении резервов и др. Учесть эти факторы можно только при введении дополнительных параметров модели. В реальности не существует идеальных программ, и следовательно, не существует идеальных систем контроля и управления. Рассмотрим возможность возникновения отказа, не обнаруженного системой управления кластера. Логическая структура дублированной системы с выделенным элементом управления и функциональная схема кластера с распределенной системой управления приведены на рис.5а,б. Элемент управления кластером размещен в каждом экземпляре кластерного ПО, выполняемого в каждом узле. Основная задача этого ПО – контроль работоспособности собственного узла и "соседа", информирование его о собственном состоянии, передача или принятие управления (переключение нагрузки на другой узел). Информация о собственном состоянии передается посредством служебных пакетов: "сердцебиения", пакетов пульса (heartbeat packets), сообщениями типа

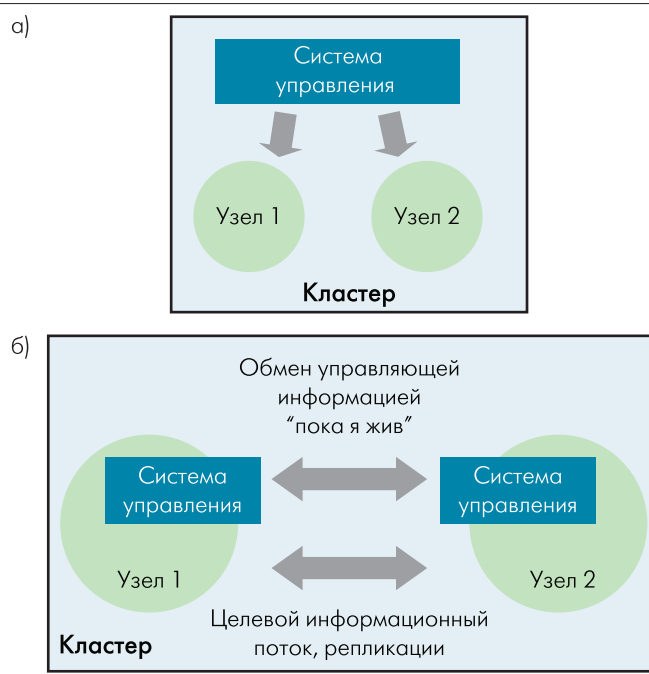


Рис.5. Граф состояний кластера: логическая структура (а) и функциональная схема (б)

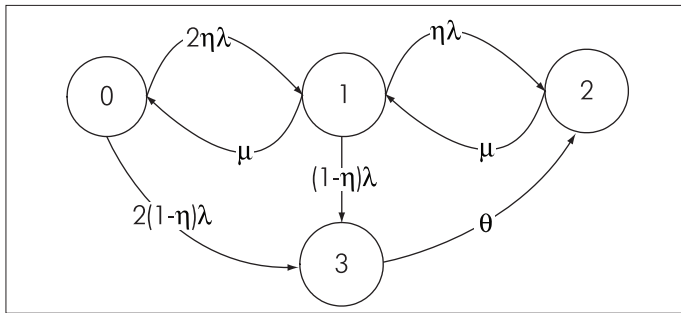


Рис.6. Граф состояний кластера с двумя узлами и идеальной системой контроля

"пока я жив" (keep alive). Передача управления реализуется пакетами, свидетельствующими о переключении узлов (fail-over packet).

Чтобы модель могла учитывать отказ, не обнаруженный системой контроля, введем соответствующее дополнительное состояние. Полученная модель, учитывающая идеальный контроль, приведена на рис.6.

Множество состояний модели с идеальной системой контроля (см. рис.4) дополняется состоянием "3", т.е. состоянием, когда отказ не обнаружен средствами внутреннего (внутрикластерного) контроля.

Таким образом, с точки зрения контролируемости, кластер представляет собой дублированную структуру с непрерывным неполным контролем (внутренними средствами кластера), заданным η , и периодическим – внешним полным контролем работоспособности узлов, заданным θ , причем отказ узла с вероятностью η обнаруживается мгновенно, а отказ с вероятностью $1-\eta$ – с задержкой на время $1/\theta$. Таким образом, время задержки обнаружения скрытых отказов имеет экспоненциальное распределение с параметром θ .

Путем расчета полученной модели (графа на рис.4) для той же наработки на отказ и времени восстановления определим значение коэффициента готовности. Предположим, что для того, чтобы обнаружить "необнаруженный" внутренними средствами кластера отказ, потребуется 15 мин (за это время клиенты, убедившись в отсутствии сервисов, начинают звонить в техническую поддержку, вынуждая администраторов вручную проверять работоспособность кластера). Здесь сознательно используются заниженные значе-

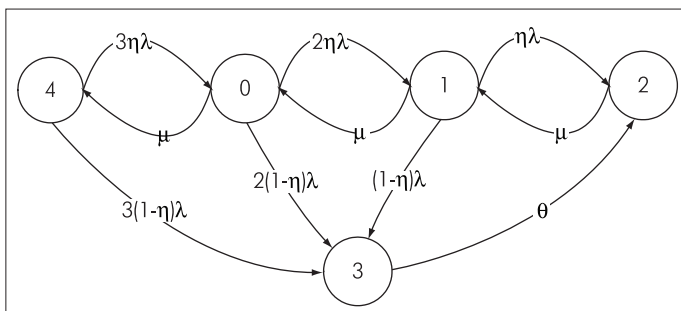


Рис.7. Граф состояний троированной системы с идеальной системой контроля

ния параметров модели, чтобы получить предельное значение K_r . В действительности рассматриваемый показатель может значительно превышать 15 мин.

При подстановке значений получим, что уже при одном необнаруженном отказе на 100 отказов ($\eta = 0,99$), обнаруженных системой управления, резко снижается значение коэффициента готовности: с семи девяток до пяти! При каждом десятом необнаруженном отказе K_r составит уже менее пяти девяток, а при каждом втором необнаруженном – менее четырех девяток.

Здесь показан только один шаг повышения адекватности модели: состояние отказа, не обнаруженного внутрикластерной системой контроля при условии, что он произошел. Следующим шагом может быть введение состояния ложного "обнаружения" отказа, т.е. когда будет переключена нагрузка, но фактического отказа узла не было.

ДУБЛИРОВАНИЕ ИЛИ ТРОИРОВАНИЕ?

Как было показано выше, при расчете дублированных структур без учета состояния необнаруженного отказа значения показателей надежности оказываются завышенными. При увеличении кратности резервирования увеличиваются и значения показателей надежности, при этом для K_r получаются еще более фантастические (точнее астрономические) цифры.

Однако при использовании модели с идеальным контролем (см. рис.6) ситуация кардинально меняется.

Введение состояния необнаруженного отказа позволяет наблюдать не только количественное, но и качественное изменение зависимости значений показателя надежности. Имеет место "феномен", когда повысить надежность путем повышения кратности резервирования (добавления нового резервного узла) уже невозможно, учитывая фактор неполноты контроля функционирования, т.е. коэффициента η . Результаты расчетов K_r , выполненные по представленным выше моделям дублирования и троирования с идеальной системой контроля (рис.6 и 7 соответственно), приведены на рис.8.

Схемы двухузловой организации кластера могут показывать лучшие показатели надежности, чем трех-, четырех- и т.д. узловые. Так, при равных системных параметрах λ , μ , θ и $\eta \leq 0,999$ двухузловой кластер обеспечивает лучший K_r по сравнению с аналогичной моделью трехузлового кластера. При одном необнаруженном отказе на тысячу обнаруженных ($\eta = 0,999$) значение K_r двухузлового кластера составит 0,999 999 49, а трехузлового – всего 0,999 999 36. При ухудшении контролируемости, т.е. при уменьшении η , преимущества двухузловых конфигураций перед трехузловыми неоспоримы.

Это объясняется тем, что определяющую роль при малых значениях интенсивности отказов играет составляющая



$m(1 - \eta)^*$, где $m = 1, 2$ для схемы дублирования (см. рис.6) и $m = 1, 2, 3$ для схемы троирования (см. рис.7). Иными словами, для резервируемых систем, претендующих на высокие показатели надежности, основной вклад в результирующее значение K_g вносит вероятность обнаружения отказа, а не добавление нового резервного узла.

СКОЛЬКО ДЕВЯТОК НУЖНО?

Приведенные примеры показывают, что популярные модели расчета надежности резервируемых структур, в том числе кластеры, являются очень упрощенными и не соответствуют реальным процессам, происходящим в исследуемых системах. Вследствие этого мы получаем не только явно завышенные значения показателей надежности, но и неверные качественные зависимости функции показателя надежности от параметров системы, например кратности резервирования.

Ситуацию усугубляет положение в IT-сфере, когда "правдивость пяти девяток" остается "на совести" вендора, т.е. заявленные значения показателей надежности системы, как правило, ничем не обосновываются. С одной стороны, отказоустойчивые системы и системы высокой готовности с высокими декларируемыми значениями не могут годами испытываться для получения необходимой статистики (пять девяток – время недоступности системы 5 мин/год) и после этого допускаться на рынок с подтвержденными испытаниями результатами, так как в таком случае они устареют до появления на рынке. С другой стороны, компании-производители не предоставляют математических моделей, по которым была проведена оценка надежности предлагаемых изделий.

Можно предположить, что такие модели являются упрощенными, идеальными и не позволяют получать близкие к реальным значениям показателей надежности. При просьбе о предоставлении моделей компании, как правило, ссылаются на то, что это закрытая информация или значения взяты по аналогии с подобными комплексами. Сегодня отсутствуют международные или фирменные общедоступные методики и модели расчета кластерных структур, позволяющие стандартизовать хотя бы подход к подобным вычислениям. Поэтому при сравнении изделий двух производителей даже более низкое значение коэффициента готовности может характеризовать более надежную систему – например, когда один из них предоставил адекватную методику, а другой идеализированную модель или вообще отказал в ее предоставлении. Целесообразно отдать предпочтение изделию с "честными" четырьмя девятками ($K_g = 0,9999$), полученными из понятной модели, учитывающей максимальное число влияющих на систему факторов, нежели изделию с "сомнительными" шестью девятками, подтвержденными только рекламными заверениями

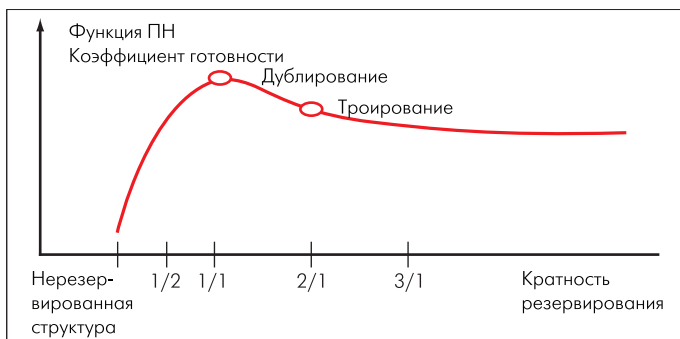


Рис.8. Зависимость K_g от кратности резервирования при $\eta \leq 0,999$

вендора без убедительных обоснований декларируемой цифры. Следовательно, стоит хорошо подумать, выбирая на богатом предложениями рынке высоконадежную вычислительную и телекоммуникационную систему, либо систему хранения данных. Но следует помнить, что этот рынок также "богат" фантастическими оценками надежности таких систем, в том числе "виртуальными" девятками коэффициента готовности.



Новая флеш-память

компании Microchip Technology

Компания Microchip Technology расширила SPI-семейство флеш-памяти серии 25, выполненное по технологии суперфлеш (SuperFlash), выпустив последовательную флеш-память типа SST25WF080 емкостью 8 Мбит. В результате теперь компания предлагает набор микросхем последовательной флеш-памяти на напряжение 1,8 В емкостью от 512 Кбит до 8 Мбит. Новая флеш-память отличается чрезвычайно малым током, потребляемым в режиме ожидания, – 5 мкА. Ток при считывании данных на частоте 33 МГц составляет 2 мА. Работает микросхема на тактовой частоте 75 МГц, имеет четырехпроводной SPI-совместимый интерфейс. Диапазон рабочих температур флеш-памяти составляет $-40...85^{\circ}\text{C}$. Выпускается в восьмивыводном корпусе SOIC-типа или восьмиконтактном корпусе типа XFBGA.



Новая микросхема предназначена для планшетных компьютеров, головной гарнитуры стандарта Bluetooth, беспроводных устройств управления и Wi-Fi сети, а также видеорекамера. Поставляются как опытные, так и промышленные образцы.

www.eetasia.com