УДК 519.872.7
Ю.Е. Шишкин, студент,
С.А. Чорномыз, аспирант
Севастопольский национальный технический университет
ул. Университетская 33, г. Севастополь, Украина, 99053
E-mail: root@sevgtu.sebastopol.ua

ИМИТАЦИОННАЯ МОДЕЛЬ КЛАСТЕРА ДЛЯ ОПТИМИЗАЦИИ ПРОЦЕССА МОНИТОРИНГА С ЦЕЛЬЮ ЛИКВИДАЦИИ ПОСЛЕДСТВИЙ ВИРУСНЫХ АТАК

Проведены исследования параметров компьютерного кластера в условиях действия вирусных атак, с применением дискретно-событийного имитационного моделирования. Разработанная модель позволила минимизировать затраты на восстановление системы путём выработки оптимальной стратегии принятия решений.

Ключевые слова: *компьютерный кластер, вирусная атака, имитационное моделирование, дискретно-событийное моделирование.*

Введение. На сегодняшний день для обеспечения обработки больших объёмов информации в таких направлениях как экономка, физика, биология, военное дело используются компьютерные кластеры (КК). В общем случае, вычислительный кластер – это набор компьютеров (вычислительных узлов), объединенных некоторой коммуникационной сетью. С точки зрения пользователя они представляют собой – единый аппаратный ресурс [1, 2]. Отказ или нарушение работы узлов КК может привести к дезорганизации работы всей системы и, следовательно, критическим ситуациям, которые могут иметь катастрофические последствия.

Очевидно, что высокий уровень гарантоспособности системы передачи данных не может быть достигнут без обеспечения требуемого уровня безопасности. Несмотря на интенсивное развитие компьютерных систем и информационных технологий защиты, КК продолжают быть уязвимыми для внешних и внутренних вирусных атак (ВА). Основной угрозой безопасности КК являются акты несанкционированного перехвата каналов обмена информацией и управления. В настоящее время разработаны и эффективно применяются многие методы и технологии противодействия вирусным атакам. К этим методам следует отнести широкий спектр антивирусных программных средств. Однако остаётся комплексно нерешённая проблема минимизации времени обнаружения факта снижения производительности, вследствие воздействия вирусных атак. Следует заметить, что минимизация времени обнаружения ВА позволит минимизировать расходы от простоя и затраты на восстановление КК. Сопутствующей проблемой является задача минимизации потерь, связанных с допущением ошибок при принятии решений о наличии ВА в сети.

Все вышесказанное подтверждает, что разработка информационных технологий детекторов вирусных атак с целью повышения функционального быстродействия и гарантоспособности КК является актуальной.

В свою очередь, преодоление последствий вирусных атак позволяет упростить и оптимизировать процесс распределения поступающих данных между членами кластера – компьютерами и разработать оптимальную стратегию детектирования вирусного заражения с целью своевременной активации механизма восстановления системы.

В качестве наиболее значимого критерия эффективности СМО в [3] предложено использовать экономическую прибыль системы за единицу времени.

Все вышесказанное подтверждает, что разработка оптимальной стратегии принятия решений для детектирования вирусных атак с целью повышения функционального быстродействия, экономических показателей, гарантоспособности КК, обладает определённой научно-практической новизной и актуальностью.

Цель статьи — найти пороговые значения снижения производительности СМО, достижение которых определяет факт наличия вирусного заражения, а также интенсивность детектирования производительности СМО, при которых средняя прибыль, характеризуемая критерием k, будет максимальной, с применением дискретно-событийного имитационного моделирования.

Элементы, характеризующиеся свойствами отказа и восстановления, будем называть ресурсами, они составляют множество $r \in R$. Элементы обрабатываемые системой, т.е. заявки, составляют подсистему заявок $w \in W$. Организация работы системы координируется подсистемой диспетчеризации на основе эвристики «Монте-Карло», учитывающей размеры очередей к ресурсам [4].

Система обладает свойствами частичного отказа и восстановления.

Для параметрической оптимизации работы система располагает ресурсом $M \in R$ и набором оптимизируемых параметров $\gamma \geq 0$ и $\varphi \in [0\ldots1]$.

Пакеты данных поступают на распределительный узел с интенсивностью *I*. Если отсутствуют свободные ресурсы, пакет становится в очередь (по принципу FIFO/FCFS). Производительность обрабатывающей системы под воздействием вирусной атаки изменяется согласно равенства

$$\rho_t = \rho_{t-1} - \Phi(\rho_{t-1} - R)\cdot R, \qquad (1)$$

где Ф – функция Хевисайда, *R* – величина уменьшения производительности системы за одну единицу времени.

Предложено процесс изменения системных показателей КК характеризовать наличием трех непересекающихся зон (A, B, C) (рисунок 1).

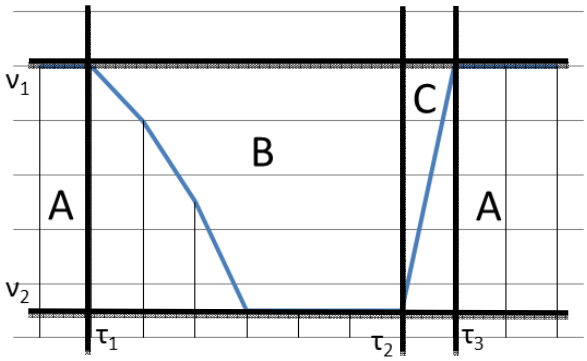

Рисунок 1 – Изменение относительного быстродействия КК в условиях наличия вирусной атаки и восстановления

Здесь $(\tau_1;\tau_2)$ – фаза действия ВА в КК до момента ее обнаружения, при этом функциональное быстродействие понижается с ν_1 до ν_2; τ_2 – момент обнаружения ВА; $(\tau_2;\tau_3)$ – фаза восстановления системы, т.е. устранение последствий ВА. Размер площади зоны В пропорционален объему потерь до обнаружения ВА. Размер площади С определяет эффективность восстановления, связанного с повышением быстродействия с ν_2 до ν_1. Зона А – нормальное функционирование КК.

Обработка одной заявки приносит *p* условных единиц (у.е.) прибыли. С интенсивностью γ происходит оценка производительности системы и если она ниже φ , то система считается заражённой, что инициализирует процесс восстановления производительности системы до уровня ρ_0. Стоимость проведения каждой процедуры детектирования и восстановления системы составляют *s* и *r* у.е. соответственно.

Требуется определить временной интервал между проверками системы и уровень производительности, после которого требуется считать систему заражённой, а также найти область эффективных решений в пространстве параметров (γ,φ).

Для стационарного режима работы критерий эффективности определяется аддитивной функцией

$$k = \mathrm{M}\left[\min(I,\mu)\cdot p - \gamma\cdot s - \Phi(\varphi-\mu)\cdot r\right] \to extr, \qquad (2)$$

где M – математическое ожидание; μ – производительность КК; *p* – доход от обработки одного пакета данных; γ – интенсивность детектирования КК; *s* – стоимость детектирования КК; φ – допустимый уровень снижения производительности; *r* – стоимость восстановления КК.

Определим основную задачу как максимизацию удельного дохода за единицу времени, т.е.

$$Q_t = \frac{k_t}{T} \to \max, \qquad t \in [t_0 \ldots T],$$

где t_0 – время, необходимое для стабилизации системы, $t_0 \approx \left[1-\frac{I}{\rho}\right]^{-1}$. При ограничениях: $0 \le s \le r \le p$, $0 < \mu < \mu_{гр}$, $0 \le \gamma \le 1$, $0 < I$, $0 < \rho < 1$.

Цель оптимизации – определить интенсивность детектирования системы и порог φ, при достижении которого следует запускать механизм восстановления КК, обеспечивающие максимальное значения заданного критерия эффективности *k*.

При неоправданно высокой интенсивности детектирования КК удельный вес затрат превышает прибыль, получаемую от его работы; при неоправданно низкой производительность КК значительно снижается вследствие воздействия ВА, в результате чего происходит уменьшение критерия эффективности.

Построим имитационную модель в среде AnyLogic™.

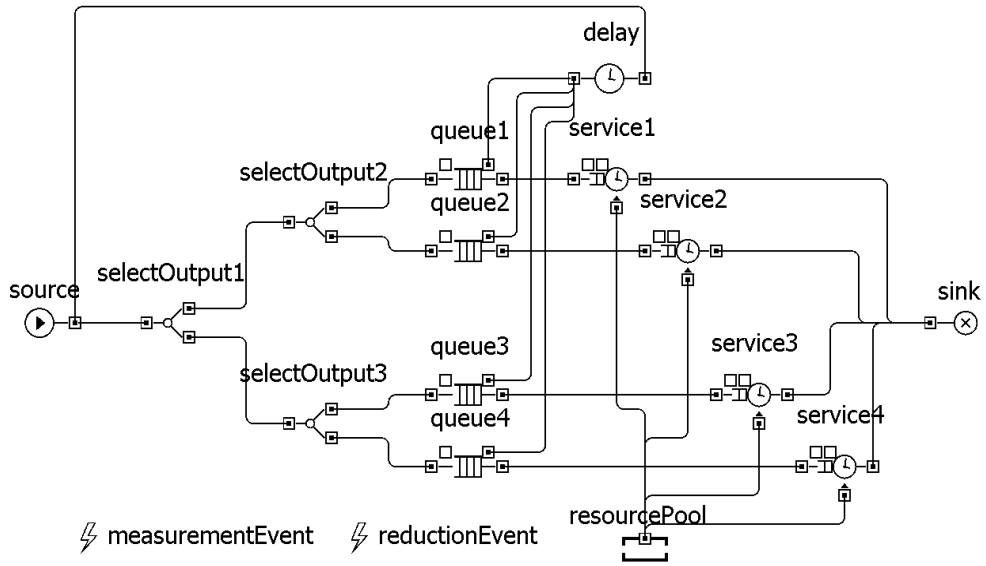

Рисунок 2 – Схема имитационной модели компьютерного кластера

Задачи, которые должны быть решены, состоят в следующем: построить имитационную модель системы, найти область эффективных решений в пространстве параметров (γ, φ), провести оптимизацию целевой функции $k(p, s, r, \gamma, \varphi, I) \to extr.$

Анализ результатов оптимизационных экспериментов. Произведём серию оптимизационных экспериментов и найдём область значений параметров системы, обеспечивающих максимальную эффективность работы компьютерного кластера, изменяя интенсивность детектирования системы и уровень производительности, при достижении которого будем считать, что эффективность работы системы достигло критически низкого уровня и необходимо активировать процесс её восстановления.

Произведём серию прогонов построенной модели (рисунок 2), и отобразим на декартовой плоскости такие параметры как интенсивность проверки состояния системы и уровень, при котором активируется процесс восстановления системы.

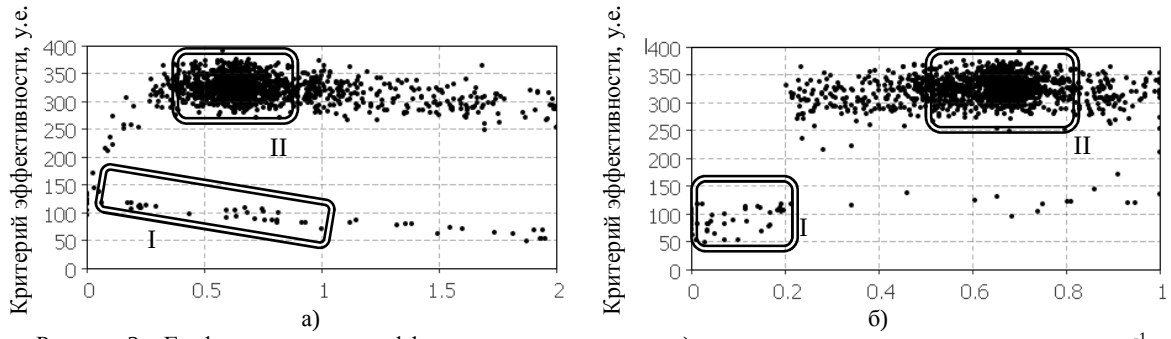

Рисунок 3 – График зависимости эффективности системы от: а) интенсивности детектирования системы, $с^{-1}$; б) минимального уровня производительности системы

Как видно из результатов моделирования (рисунок 3), найдено две области решений, обеспечивающих достижение максимального значения критерия эффективности. Область I, где коэффициент эффективности принимает невысокие положительные значения (рисунок 3, а и 3, б). Существование этой области обусловлено тем, что согласно ограничениям, введённым в данной работе, в системе не может произойти полный отказ, но так как восстановление системы происходит достаточно редко по сравнению с функцией, имитирующей постепенный отказ – происходит переход в режим с минимально возможной производительностью. Как следствие, прибыль, полученная от обработки небольшого объёма данных, обеспечивает некоторый доход, в то время как проверка и ремонт практически не производятся, что минимизирует расходы. Данная область решения является оптимальной в том случае, если удельные мониторинговые затраты значительно превышают доход по обработке заявок, либо при низкой интенсивности входного потока заявок. Если наилучшее решение будет найдено в этой области, то это сигнализирует о наличии избыточных ресурсов системы или о несоизмеримо больших затратах на проведение детектирования производительности КК вследствие ВА по сравнению с доходами, получаемыми от ее работы.

Вторая область (II) представляет собой параболу с экстремумом в области [0,4…0,8] (рисунок 3,а и 3,б). Именно данное решение характеризуется наибольшим значением целевой функции.

При повторных запусках модели со случайными исходным параметрами в пределах введённых ограничений было экспериментально выявлено наличие двух областей решений, одно из которых характеризовалось наибольшим значением целевой функции. В некоторых случаях области совпадали, например, как видно на рисунке 3, а обе области решения пересекаются при нулевом значении параметра интенсивности.

Возьмём выборку из 0,025 % решений с наибольшим критерием эффективности и отобразим на графике соответствующие данному критерию значения интенсивности проверки и уровня активации механизма восстановления системы.

Выполним аппроксимацию двух полученных областей по алгоритму выпуклого многоугольника в следующей последовательности:

1) создать треугольник, соединяющий крайние точки;
2) если вне многоугольника точек нет – конец;
3) найти очередную наиболее удаленную от стороны многоугольника точку;
4) если точка лежит вне многоугольника – присоединить ее к крайним вершинам;
5) перейти к п.2.

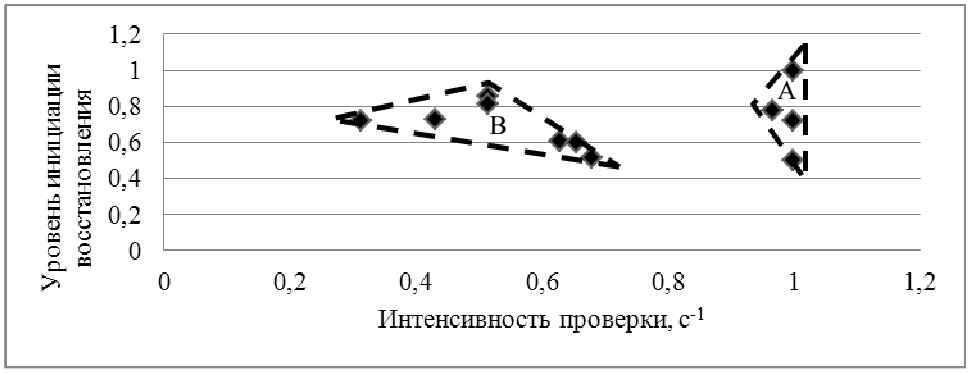

Рисунок 4 – Область эффективных решений системы: A – при высокой интенсивности входного потока; B – при низкой интенсивности входного потока

В результате моделирования был получен результат, изображённый на рисунке 4. Как видно, область решения находится в достаточно ограниченном промежутке параметров: (0,3…0,7) для интенсивности поступления заявок меньшей, чем производительность системы и (0,9…1,0) для интенсивности поступления заявок большей, чем производительность системы, в данном случае система находится в нестационарном режиме и рекомендуемая интенсивность проведения проверок системы совпадает с интенсивностью вирусных атак, что является ожидаемым.

Выводы. При разработке модели, реализующей моделирование работы компьютерного кластера под воздействием вирусных атак, в качестве основного критерия, характеризующего эффективность его работы, была выделена экономическая целесообразность использования и эксплуатации КК с учётом затрат на его проверку и восстановление.

Разработана модель компьютерного кластера с использованием программного пакета имитационного моделирования AnyLogic™, с применением дискретно-событийного подхода к построению модели.

Экспериментально показано, что для СМО с ожиданиями без потерь, обладающей свойствами частичного отказа и восстановления при решении задачи оптимизации контроля над системой, существуют две области решения. Первая всегда позволяет получить неотрицательное значение эффективности, при установке параметров, обеспечивающих минимальную производительность и низкую интенсивность поступления входных заявок, что достигается за счёт использования внутренних ресурсов системы, когда восстановление происходит крайне редко, сводя затраты на проверку и восстановление системы к минимуму. Вторая представляет собой область, обеспечивающую высокую производительность системы, так как критерий эффективности есть разность между доходом, прямо пропорциональным производительности КК, и расходом, экспоненциально возрастающим по мере приближения в максимальной производительности.

В качестве дальнейшего направления исследований планируется внедрение комплекса поддержки принятия решений для организации системы обнаружения и ликвидации последствий воздействия вирусных атак.

Библиографический список использованной литературы

1. Лацис А.О. Как построить и использовать суперкомпьютер / А.О. Лацис. — М.: Бестселлер, 2003. — 274 с.

2. Лаборатория Параллельных информационных технологий – официальный сайт научно-исследовательского вычислительного центра Московского государственного университета имени М.В. Ломоносова [Электронный ресурс]. — Электрон. текстовые данные (184578 bytes). — Режим доступа: http://parallel.ru/cluster/beginner_guide.html Monday, 12 Dec 2012 12:10:35

3. Шишкин Ю.Е. Оптимизация функционирования супермаркета на основе процедур эвристической диспетчеризации / Ю.Е. Шишкин, Д.Ю. Воронин // Мир компьютерных технологий: материалы внутривузовской научно-технической конференции, г. Севастополь, 2–5 апреля 2012 г. / СевНТУ. — Севастополь, 2012. — 61 с.

4. Шрайбер Т.Д. Моделирование на GPSS / Т.Д. Шрайбер. — М.: Машиностроение, 1980. — 498 с.

5. Гнеденко Б.В. Введение в теорию массового облуживания / Б.В. Гнеденко, И.Н. Коваленко. — М.: Наука, 1966. — 429 с.

Поступила в редакцию 25.12.2012 г.

Шишкін Ю.Є., Черномиз С.А. Імітаційна модель комп'ютерного кластера що оптимізує процес моніторингу з метою ліквідації наслідків вірусних атак

Проведено дослідження параметрів комп'ютерного кластера, як системи масового обслуговування з очікуванням, із застосуванням дискретно-подійного імітаційного моделювання. Модель дозволила мінімізувати витрати на відновлення системи завдяки виробленню оптимальної стратегії прийняття рішень.

Ключові слова: комп'ютерний кластер, імітаційне моделювання, дискретно-подійне моделювання.

Shishkin Y.E., Chernomyz S.A. Computer cluster simulation model designed to optimize monitoring for elimination of viral attacks

A computer cluster parameters is analyzed as a queuing system by using discrete-event simulation. A model enables minimize the cost of system restoring by choosing the optimal decisions policy.

Keywords: computer cluster, a distributed system, simulation, discrete-event simulation.