

Континуальный анализ стоимости большемасштабных распределённых вычислительных систем^{*)}

Д.С. Никитин

Современный инструментарий обработки информации – это большемасштабные распределённые вычислительные системы (ВС) высокой производительности (с количеством процессоров до 1 000 000). В работе рассмотрен континуальный подход к расчёту технико-экономических показателей функционирования большемасштабных распределённых ВС (функций дохода и расходов) в стационарном и переходном режимах. Расчёт осуществлён с учётом коэффициента стоимости эксплуатации сети межмашинных связей и времени реконфигурации структуры ВС после отказов.

Ключевые слова: вычислительная система, технико-экономический анализ.

1. Введение

Одной из основных тенденций современной индустрии высокопроизводительной обработки информации является построение большемасштабных пространственно-распределённых вычислительных и GRID-систем. В архитектурном плане распределённая вычислительная система (ВС) представляется как множество элементарных машин (ЭМ), взаимодействие между которыми осуществляется через коммуникационную среду [1, 2]. Существует большое количество распределённых ВС (см. Top 500), следовательно, нужны нетрудоёмкие методы расчёта доходов и расходов, связанных с их эксплуатацией.

По своей природе распределённая ВС – стохастический объект (например, вследствие отказов её компонентов). Нужно заметить, что процесс восстановления обнаруженных отказавших ЭМ предусматривает не их ремонт, а обязательно реконфигурацию систем. При этом проверка работоспособности ВС и поиск отказавших машин выполняются соответственно средствами самоконтроля и самодиагностики. Последние средства для краткости называют *контролёром* и *диагностом* [1].

Реконфигурация системы заключается в программной настройке новой конфигурации с заданным числом исправных ЭМ; она осуществляется *реконфигуратором*. Для создания новой конфигурации могут быть использованы, в общем случае, машины из избыточности и/или резерва. Контролёр, диагност и реконфигуратор являются компонентами распределённой операционной системы (ОС). Эта композиция, по сути, является *виртуальным восстанавливающим устройством* (ВУ) для распределённой ВС; её будем называть просто ВУ. В рамках сказанного можно говорить о виртуальном переключении исправных ЭМ и об интенсивности их переключения. Задачам расчёта и оценки стоимости эксплуатации ВС посвящено множество работ как отечественных, так и зарубежных научных

^{*)} Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты № 10-07-00157, 09-07-00185, 09-07-00126) и Совета по грантам Президента РФ (ведущая научная школа НШ-5176.2010.9).

школ [5 – 8]. В области анализа технико-экономической эффективности функционирования большемасштабных вычислительных систем работают различные подразделения крупных мировых компаний и институтов, например компании IBM, Intel, Sun Microsystems. В качестве базового показателя для анализа по эксплуатации распределённых вычислительных систем используется TCO (Total Cost of Ownership). Западный термин TCO означает суммарную (или полную) стоимость владения, т. е. явные и скрытые издержки на владение тем или иным оборудованием. Впервые методика расчёта TCO была предложена аналитической компанией Gartner Group в конце 1980-х годов как средство оценки расходной части информационных активов компании, включая прямые и косвенные расходы. Список категорий затрат при расчёте TCO может варьироваться, но можно выделить основные: первоначальные затраты на систему, затраты на обновления аппаратного обеспечения системы, затраты на поддержку системы, затраты на обработку данных (зарплата), затраты на обработку данных (иные), затраты на учёт труда (зарплата). Расчёт TCO – это сложная и дорогостоящая процедура.

2. Расчёт показателей стоимости большемасштабных ВС

В данной работе рассматривается подход, не требующий значительных затрат, чтобы произвести технико-экономический экспресс-анализ работы ВС. Данный подход базируется на стохастических моделях, описывающих процесс функционирования ВС [2, 4]. Стохастические модели приводят к простым расчётным формулам для координат вектор-функции $\Gamma(t)$ и $D(t)$; соответственно, расходов и доходов ВС. Традиционный подход, базирующийся на марковских процессах с доходами, при тех же результатах является более трудоёмким [6, 9]. В отличие от работы [1], здесь будет осуществлён расчёт технико-экономических показателей с учётом стоимости эксплуатации сети связей и времени реконфигурации структуры системы после отказов. Итак, требуется рассчитать затраты и доход, которые последуют в процессе эксплуатации распределённой ВС.

Пусть имеются вычислительная и восстанавливающая системы, состоящие соответственно из N элементарных машин (ЭМ) и m устройств, $N \geq m \geq 0$. Пусть λ и μ – интенсивности соответственно отказов ЭМ и восстановления отказавших ЭМ одним ВУ; c_1 и c_2 – стоимости соответственно эксплуатации одной ЭМ и содержания одного ВУ в единицу времени. Требуется оценить средние эксплуатационные расходы при работе ВС в переходном и стационарном режимах. $K_i(t)$ – математическое ожидание исправных ЭМ в ВС, $M_i(t)$ – математическое ожидание занятых ВУ в восстанавливающей системе, в момент времени $t \geq 0$ при условии, что $K_i(0) = i$, $i \in E = \{0, 1, 2, \dots, N\}$. Считается, что если ВС находится в состоянии $j \in E$, то имеется j исправных ЭМ. Время реконфигурации системы обозначим через ν^{-1} . Тогда для координат вектор-функции расходов $\Gamma_i(t)$ справедливы следующие уравнения [2, 3].

$$\left. \begin{aligned} \frac{d}{dt} \Gamma_i(t) &= c_1 [N - K_i(t)] + c_2 [m - M_i(t)], \\ \Gamma_i(0) &= 0, \quad i \in E. \end{aligned} \right\} \quad (1)$$

Решение системы (1) представляется функциями

$$\Gamma_i(t) = -\beta_i + \gamma + \beta_i \delta(t),$$

Пусть восстанавливающая система имеет высокую производительность, т. е. выполняется неравенство

$$N\lambda \leq m\mu.$$

Тогда решение системы дифференциальных уравнений (1) запишется в виде

$$\beta_i = \frac{i\lambda - (N-i)\mu}{(\lambda + \mu)^2} (c_1 - c_2), \quad \gamma = \frac{N\lambda}{\lambda + \mu} (c_1 - c_2) + mc_2, \quad \delta(t) = e^{-(\lambda + \mu)t}, \quad i \in E^1$$

а при $N\lambda > m(\lambda + \mu)$

$$\beta_i = \frac{i\lambda - m\mu}{\lambda^2} c_1, \quad \gamma = \frac{N\lambda - m\mu}{\lambda} c_1, \quad \delta(t) = e^{-\lambda t}, \quad i \in E^2,$$

где $E^1 = (N-m, N-m+1, \dots, N)$, $E^2 = (0, 1, \dots, N-m-1)$.

Интерес также представляет вектор-функция бесполезных эксплуатационных расходов с учётом времени реконфигурации системы. Тогда для вектор-функции $\Gamma(t)$ решение системы уравнений (1) в случае высокопроизводительной системы в переходном режиме функционирования ВС запишется в следующем виде:

$$\Gamma(i, t) = -[\varepsilon_i' - \varepsilon_i''] + \gamma t + \varepsilon_i' \exp[\alpha_1 t] - \varepsilon_i'' \exp[\alpha_2 t],$$

где γ имеет вид

$$\gamma = \frac{N\lambda\nu}{\lambda\nu + \nu\mu + \mu\lambda} (c_1 - c_2) + mc_2,$$

а коэффициенты имеют следующий вид:

$$\varepsilon_i' = \frac{\alpha_2 \exp[\alpha_1 t]}{\alpha_1 \alpha_2 (\alpha_1 - \alpha_2)} A (c_1 - c_2), \quad \varepsilon_i'' = \frac{\alpha_1 \exp[\alpha_2 t]}{\alpha_1 \alpha_2 (\alpha_1 - \alpha_2)} B (c_1 - c_2),$$

$$A = \frac{N(\alpha_1 + \lambda)(\alpha_1 + \nu)}{\alpha_1} - i(\alpha_1 + \lambda + \nu) + j\lambda, \quad B = \frac{N(\alpha_2 + \lambda)(\alpha_2 + \nu)}{\alpha_2} - i(\alpha_2 + \lambda + \nu) + j\lambda,$$

$$\alpha_{1,2} = -0.5 \left[\lambda + \mu + \nu \mp \sqrt{\lambda^2 + \mu^2 + \nu^2 - 2(\lambda\nu + \nu\mu + \mu\lambda)} \right].$$

Для стационарного режима справедливо следующее выражение

$$\Gamma(t) = \gamma t.$$

Это объясняется тем, что для больших t справедливо $\beta \ll \gamma$ и $\delta(t) \rightarrow 0$. Величины γ определяются следующим образом.

При $N\lambda \leq m\mu$

$$\gamma = \frac{(c_1 - c_2)}{\lambda\mu + \lambda\nu + \mu\nu} N\lambda\nu + c_2 m,$$

при $N\lambda > m\mu$

$$\gamma = c_1 \left[N - m\mu \left(\frac{1}{\lambda} + \frac{1}{\nu} \right) \right].$$

Расчёты показывают, что для вычислительных систем средней и высокой производи-

тельности и стационарного режима, при оценке наиболее оптимального числа восстанавливающих устройств, можно использовать нижеследующие рассуждения. Из формулы для расчёта u нельзя увидеть оптимальное значение числа m^* восстанавливающих устройств.

Однако при $N \rightarrow \infty$ имеет место

$$\frac{m^*}{N} \rightarrow \frac{\lambda v}{\lambda \mu + v \lambda + v \mu},$$

причём

$$\frac{m^*}{N} \leq \frac{\lambda v}{\lambda \mu + v \lambda + v \mu}.$$

Поэтому для ВС средней и высокой производительности при расчёте значения m^* можно использовать следующую формулу:

$$m^* = \left] N \frac{\lambda v}{\lambda \mu + v \lambda + v \mu} \right[,$$

где $\left] Z \right[$ означает такое ближайшее к Z целое число, что $\left] Z \right[\geq Z$.

Для координат вектор-функции дохода $D_i(t)$ справедливы следующие уравнения.

$$\left. \begin{aligned} \frac{d}{dt} D_i(t) &= c_1 K_i(t) - c_2' m - c_3 \mu M_i(t) - c_4 N, \\ D_i(0) &= 0, i \in E. \end{aligned} \right\} \quad (2)$$

Здесь c_2' – себестоимость содержания восстанавливающего устройства в единицу времени, c_3 – стоимость запасных деталей, расходуемых при однократном восстановлении отказавшей ЭМ, c_4 – постоянные издержки, возникающие при эксплуатации ЭМ в единицу времени.

Решением системы (2) будут функции

$$D_i(t) = D_i + gt - D_i \delta(t),$$

где при $N\lambda \leq m(\lambda + \mu)$

$$D_i(t) = \frac{i\lambda - (N-i)\mu}{(\lambda + \mu)^2} (c_1 + c_3 \mu), \quad i \in E^1, \quad g = \frac{N\mu}{\lambda + \mu} (c_1 - c_3 \lambda) - (c_2' m + c_4 N),$$

а при $N\lambda > m(\lambda + \mu)$

$$D_i(t) = \frac{i\lambda - (N-i)\mu}{(\lambda + \mu)^2} (c_1 - c_2), \quad i \in E^2, \quad g = \frac{m\lambda}{\mu} (c_1 - c_3 \lambda) - (c_2' m + c_4 N).$$

Определим вектор-функцию дохода $D_i(t)$ с учётом показателя v^{-1} – времени реконфигурации системы и величины c_5 , характеризующей стоимость пересылки данных по каналам связи. Вектор-функция $D_i(t)$ описывается следующим образом:

$$\left. \begin{aligned} \frac{d}{dt} D_i(t) &= c_1 K_i(t) - c_2' m - c_3 \mu M_i(t) - c_4 N - c_5, \\ D_i(0) &= 0, i \in E. \end{aligned} \right\} \quad (3)$$

Решением системы уравнений (3) в стационарном режиме для вектор-функции дохода будут функции

$$D_i(t) = gt,$$

где при $N\lambda \leq m\mu$

$$g = \frac{N\mu}{\lambda\nu + \lambda\mu + \mu\nu} [c_1(\nu + \mu) - c_3\nu\lambda] - (c_2m + c_4N + c_5),$$

при $N\lambda > m\mu$

$$g = c_1m\mu \left(\frac{\lambda + \nu}{\lambda\nu} \right) - (c_2m + c_3m\mu + c_4N + c_5).$$

Итак, сложность расчёта математического ожидания дохода, приносимого ВС, остаётся такой же, что и при расчёте средних эксплуатационных расходов.

3. Численный анализ эффективности ВС

Распространённой конфигурацией является однородная ВС в составе $N = 1024$ ЭМ (процессоров, узлов, ядер). Как показывает практика, отказы элементарных машин случаются раз в три-четыре года, поэтому интенсивность потока отказов в данном случае положим $\lambda = 0.00002$ 1/ч. Время на устранение неполадок отказавшей ЭМ занимает около 1 часа, поэтому $\mu = 1$ 1/ч. В начальный момент времени количество исправных ЭМ $i = 1000$. Время, которое требуется на реконфигурацию данной системы, составляет 0.2 – 0.3 минуты, поэтому $\nu = 240$ 1/ч. Пусть имеется одно восстанавливающее устройство, $m = 1$, показатели стоимости принимают значения $c_1 = 3$ руб/ч, $c_2 = 2$ руб/ч, $c_2' = 1$ руб/ч, $c_3 = 500$ руб, $c_4 = 0,5$ руб/ч, $c_5 = 3$ руб/ч. На рис. 1 приведены графики вектор-функций доходов и расходов ВС; видно, что с течением времени доходы существенно превышают эксплуатационные расходы системы.

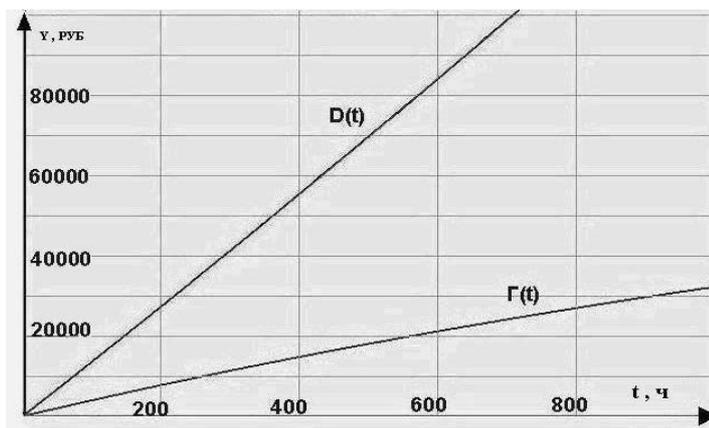


Рис. 1. Пример поведения вектор-функций $\Gamma(t)$ и $D(t)$

4. Заключение

Рассчитаны показатели технико-экономической эффективности функционирования большемасштабных вычислительных систем как ансамблей элементарных машин, учитывающие характеристики сети связей и время реконфигурации. Показатели устанавливают взаимосвязь между надёжностью и стоимостью ВС и позволяют анализировать работу сис-

тем в условиях и переходного, и стационарного режимов. Полученные формулы вполне удовлетворяют требованиям практики, они просты и позволяют производить экспресс-анализ ВС с произвольным числом машин без применения средств вычислительной техники. Результаты данной работы применимы к другим вычислительным средам и средствам обработки информации. Они будут полезны при организации работ на вычислительных центрах, эксплуатирующих группы одинаковых машин и кластерные ВС.

Литература

1. Хорошевский В.Г. Архитектура вычислительных систем. – 2-е изд., перераб. и доп. – М.: МГТУ им. Н.Э. Баумана, 2008. – 520 с.
2. Евреинов Э. В., Хорошевский В. Г. Однородные вычислительные системы. Новосибирск: Наука, 1978. – 320 с.
3. Хорошевский В. Г. Инженерный анализ функционирования вычислительных машин и систем. – М.: Радио и связь, 1987. – 256 с.
4. Вентцель Е. С. Исследование операций. –М.: Сов. Радио, 1972. – 551 с.
5. Лопато Г. П. Исследование производительности и технико-экономической эффективности универсальной ЭВМ ЕС-1020 // Алгоритмы и организация решения экономических задач. М.: Статистика, 1973. № 2. С. 23–31.
6. Хорошевский В. Г., Хорошевская Э. Г., Голоскокова Т. М. Расчёт технико-экономических показателей однородных вычислительных систем высокой производительности // Вычислительные системы. Новосибирск, 1970. № 39. С. 38–60.
7. Page H. What price PC? // Entrepreneur, 1997. № 10. С. 87–89.
8. Smith S. The smart way to invest in computers // Journal of Accountancy. 1997. № 5. С. 63–65.
9. Ховард Р. А. Динамическое программирование и марковские процессы. М.: Сов. Радио, 1964. – 188 с.

Статья поступила в редакцию 01.06.2010

Никитин Дмитрий Сергеевич

Младший научный сотрудник Лаборатории вычислительных систем Института физики полупроводников им. А.В. Ржанова СО РАН. Область научных исследований – архитектура распределённых вычислительных систем (ВС), технико-экономический анализ эффективности функционирования ВС.

Тел. & факс: (383) 333-21-71, (383) 269-82-75; e-mail: 777666@ngs.ru

The continual analysis of cost of large-scale distributed computer systems

D.S. Nikitin

Modern tools for information processing are distributed computer systems (DCS) of high performance. They are MPP-systems and contain high number of processors (up to 1 000 000). The exploitation of such systems is always connected with financial expenditures. The project contains examination of a continuous method of calculation of technical and economic indices of functioning distributed CS (income and expenditure functions) in stationary and transient condition.

Keywords: computer systems, technical and economic analysis.