01

Оценка вычислительных возможностей классических компьютеров

© *Ю.Н. Зайко*

Поволжская Академия государственной службы им. П.А. Столыпина, Саратов
E-mail: zyrnick@lycos.com

Поступило в Редакцию 1 марта 2005 г.

Рассматривается модель для описания вычислений, проводимых классическим (необратимым) компьютером в виде ансамбля частиц в потенциальной яме, обладающей двумя потенциальными минимумами, разделенными барьером, предложенная Ландауэром. Показано, что оценка частоты переключения такого устройства в виде экспоненты от величины диссипируемой энергии приводит к заключению относительно вычислительных возможностей классических компьютеров, противоречащему реальности. Для устранения противоречия исходная система уравнений для заселенностей ям дополняется уравнением для энтропии. Решение этой системы уравнений приводит к степенной зависимости частоты переключения от диссипируемой энергии, что подтверждается на практике.

Классические компьютеры — это компьютеры, работающие в соответствии с законами классической физики. Главной их особенностью является необратимость вычислений, т. е. невозможность восстановления входных данных по результатам вычислений. Поэтому их поведение, как и поведение всех макроскопических физических систем, подчиняется второму началу термодинамики, гласящему, что энтропия любой замкнутой системы при всех процессах не убывает со временем. Это связано с рассеянием энергии при вычислениях. Дж. фон Нейман предположил, что минимальное количество рассеянной энергии за один шаг вычислений равно $kT \cdot \ln 2$, где k — постоянная Больцмана, T — абсолютная температура вычислительной среды [1]. Тем самым он связал термодинамическую и логическую необратимость.

Р. Ландауэр в работе [2] подробно рассмотрел различные модели вычислительных ячеек в виде бистабильных переключающих устройств и получил выражения для времени их переключения, которое естественно считать временем одного такта вычислений (такта процессора). Одна из наиболее полно рассмотренных им моделей представляет собой симметричную бистабильную двойную потенциальную яму по одной из координат (информационной степени свободы).

Уравнение, описывающее поведение ансамбля частиц в такой яме, имеет вид [2]:

$$\frac{dn_A}{dt} = -\nu n_A \cdot \exp\left(-\frac{U - U_A}{kT}\right) + \nu n_B \cdot \exp\left(-\frac{U - U_B}{kT}\right). \quad (1)$$

Здесь $n_A, n_B = N - n_A$ — числа частиц соответственно в ямах A и B, N — полное сохраняющееся число частиц на ячейку, ν — частота перехода частиц между ямами, U, U_A и U_B — энергии максимума межъямного барьера и минимумов каждой из ям. Разность $\Delta = 1/2(U_A - U_B)$ представляет собой половину энергии, рассеиваемой в процессе переключения и доставляемой некоторой внешней силой, управляющей переключением. В результате переключения система выходит из равновесия, характеризуемого для одинаковых ям $(U_A = U_B)$ значениями $n_A = n_B$ и релаксирует к новому равновесному распределению

$$n_A = n_B \cdot \exp\left(\frac{U_B - U_A}{kT}\right) \quad (2)$$

по закону $\sim \exp(-\lambda t)$ за время τ, где $\tau^{-1} = |\lambda|$, λ — характеристическое число уравнения (1):

$$\lambda = -\nu \cdot \exp\left(-\frac{U - U_A}{kT}\right) - \nu \cdot \exp\left(-\frac{U - U_B}{kT}\right). \quad (3)$$

Путем несложных преобразований можно показать [2], что

$$\frac{1}{\tau} = \frac{1}{\tau_0} \cdot \mathrm{ch}\left(\frac{\Delta}{kT}\right); \quad \frac{1}{\tau_0} = 2\nu \cdot \exp\left(-\frac{U - U_0}{kT}\right); \quad U_0 = \frac{1}{2} \cdot (U_A + U_B), \quad (4)$$

где τ_0 имеет смысл времени жизни информации, τ — время переключения, а $\mathrm{ch}(\Delta/kT)$ определяет число переключений, т. е. длину программы, выраженную в тактах процессора. При условии $\Delta \gg kT$, выполнение которого обычно предполагается, число это велико, что и является залогом полезности вычислительных устройств описываемого типа.

Полученные результаты справедливы при определенных предположениях, часть из которых перечислена выше, о других — см. [2]. Одним из предположений является то, что входящие в выражение (4) величины можно менять в некоторых пределах, иначе бы его ценность резко упала. Например, величину Δ можно менять, по крайней мере, в несколько раз. Если мы примем это во внимание, то немедленно придем к противоречию с твердо устоявшейся точкой зрения на классические компьютеры, как на устройства, плохо справляющиеся с так называемыми неполиномиальными задачами, т. е. задачами, сложность которых, измеряемая временем работы программы по нахождению их решения, не выражается никаким полиномом от размерности входных данных. Примерами таких задач являются задача коммивояжера или задача раскрытия достаточно длинного пароля методом тотального перебора [3]. Сложность таких задач, как правило, оценивается экспоненциальной функцией от размерности входных данных — числа городов, которые должен посетить коммивояжер, или длины пароля. Действительно, как следует из (4), экспоненциальный рост сложности задачи может быть компенсирован экспоненциальным же сокращением времени машинного такта и, хотя число инструкций (команд) экспоненциально возрастет, общее время на их выполнение останется неизменным.

Причина этого обстоятельства связана, очевидно, с тем, что в [2] при решении (1) температура T считается постоянной, тогда как из-за упоминавшейся выше диссипации энергии она должна расти ввиду конечной теплоемкости переключающей ячейки c. Чтобы учесть это, уравнение (1) надо дополнить уравнением для энтропии S и решать получившуюся систему совместно. Собственно говоря, это и является смыслом работы [2]. Это уравнение имеет вид:

$$\frac{dS}{dt} = \frac{2\Delta}{T \cdot \tau(T)}, \tag{5}$$

где $\tau(T)$ — время переключения, зависящее от температуры, а $2\Delta/T$ — энтропия, производимая за один такт вычисления.

Трудность поставленной задачи заключается в том, что точный вид зависимости $\tau(T)$ для произвольных T неизвестен. Поэтому мы ограничимся исследованием задачи при малых временах t, когда можно еще использовать для $\tau(T)$ выражение (4). Подставляя его в (5), выполним интегрирование с учетом соотношения $dS/dt = c/T \cdot dT/dt$. Это даст зависимость $T(t)$, определяемую уравнением

$$\exp\left(\frac{E}{kT}\right) = \exp\left(\frac{E}{kT_0}\right) - \alpha \cdot t,$$

$$E = U - U_A, \qquad \alpha = \frac{2\nu E\Delta}{ckT_0^2}, \qquad T_0 = T(0). \tag{6}$$

Подставляя (6) в (1), найдем значение $n_A(t)$:

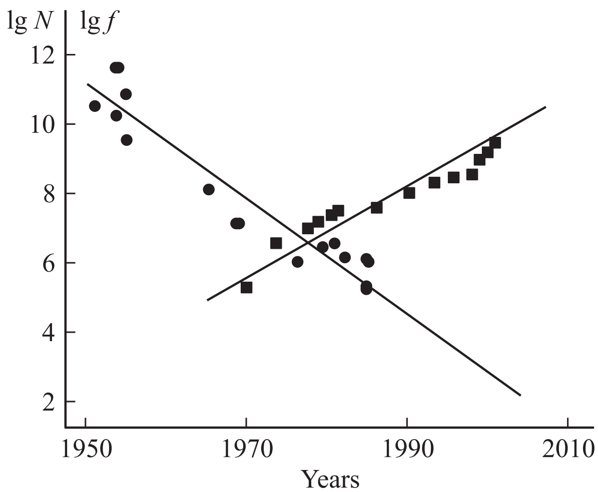
$$n_A \sim \left|\frac{t'}{\delta_0}\right|^{\frac{\nu}{\alpha}} \cdot \exp\left[-\frac{\sigma}{(t')^{\frac{2\Delta}{E}}}\right] + \ldots; \qquad t' = t - \frac{1}{\alpha} \cdot \exp\left(\frac{E}{kT}\right),$$

$$\delta_0 = \frac{1}{\alpha}, \qquad \sigma = \frac{\nu E}{2\Delta\alpha^{1+\frac{2\Delta}{E}}}. \tag{7}$$

В (7) выписана только релаксационная часть $n_A(t)$.

В результате оказывается, что процесс релаксации распределения частиц в ямах к равновесному значению выражается в интересующей нас области времен $t \geqslant 0$ не экспоненциальной, а степенной функцией, причем характерный временно́й масштаб аргумента $\delta_0 = \delta(T_0) = ckT_0^2/(2\nu E\Delta)$, который естественно считать временем переключения, теперь также степенным образом зависит от величины диссипируемой энергии 2Δ, правильно отражая эту зависимость: убывание δ с ростом Δ.

Временно́й масштаб τ изменения степенной функции определяется ее аргументом и оценивается по значению производной $|n'_A| \sim n_A/\tau$ по времени. Величина $\tau \approx (\alpha/\nu \cdot |t'|)$ и в интересующей нас области малых t, где $t' \sim -1/\alpha \cdot \exp(E/kT)$, приводит к прежнему значению $\tau = 1/\nu \cdot \exp(E/kT)$, с помощью которого было получено (7). Для других значений t, τ будет другим, что вызывает ряд вопросов при использовании $\tau(t)$ в качестве меры времени релаксации в случаях, когда релаксация имеет неэкспоненциальный характер, как, например в (7). Исследование этих вопросов, как и различных зависимостей $\tau(T)$, представляет самостоятельный интерес и не является целью настоящей работы.



Зависимость числа примесей N в основаниях биполярных транзисторов (точки, график заимствован из [5]) и тактовой частоты CISC процессоров f (Hz) [4] (квадраты) по годам. В случаях, когда в [4] приводятся несколько различных значений f, значение $\lg f$ усреднялось. Прямые проведены для демонстрации группировки точек.

Использование в качестве меры времени переключения δ_0 можно обосновать следующими рассуждениями. Для частоты переключения $f = \alpha$ (тактовой частоты процессора) получаем оценку:

$$f = \nu \cdot \frac{1}{N} \cdot \frac{E}{kT} \cdot \frac{2\Delta}{kT}, \qquad N = \frac{c}{k}, \qquad T \equiv T_0, \tag{8}$$

где N имеет смысл числа частиц (электронов) на переключающую ячейку (транзистор). Такая зависимость f от N действительно имеет место, что подверждается графиками на рисунке [4,5]. Заметим, что число электронов на транзистор в [5] оценивается как число примесей в основаниях биполярных транзисторов, требуемых для логических операций. Экстраполяция данных по примесям до 2010 г. сделана автором [5]. Результаты сравнения зависимостей N и f по годам позволяют утверждать, что произведение Nf очень слабо зависит от времени. Это может быть обусловлено, во-первых, постоянством характеристик материала (ν) и незначительными изменениями характеристик транзисторов (E, Δ).

В заключение, возвращаясь к работе [2], следует сказать, что результаты настоящей работы ни в коем случае не отменяют ни результатов [2], не ее значения как первой работы, поставившей вопрос об обратимых вычислениях.

Автор выражает признательность В.В. Попову за обсуждение результатов.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 03–02–17219).

Список литературы

[1] *Фон Нейман Дж.* Теория самовоспроизводящихся автоматов. М.: Мир, 1971.
[2] *Landauer R.* // IBM J. Res. Dev. 1961. V. 5. P. 183. (Пер.: Р. Ландауэр. Необратимость и выделение тепла в процессе вычисления. Квантовый компьютер и квантовые вычисления. Ижевск, 1999).
[3] *Бондарев В.М., Рублинецкий В.И., Качко Е.Г.* Основы программирования. Харьков: «Фолио», 1997. 368 с.
[4] *Бройдо В.Л.* Вычислительные системы, сети и телекоммуникации. СПб.: Питер, 2002. 688 с.
[5] *Keyes R.W.* // IBM J. Res. Develop. 1988. V. 32. P. 24.