Thermal, nonequilibrium phase space for networked computers

Mark Burgess
Oslo College, Cort Adelers Gate 30, 0254 Oslo, Norway
(Received 22 February 2000)

It is shown that networks of computers can be described by concepts of statistical physics. Computers in a network behave like systems coupled to a thermal reservoir. The role of thermal fluctuations is played by computing transactions. A thermal Kubo-Martin-Schwinger condition arises due to the coupling of a computer to a strong periodic source, namely, the daily and weekly usage patterns of the system.

PACS number(s): 05.70.Ln, 05.45.Tp, 07.05.Tp

Computer networks are cooperative systems, composed of many interacting elements. The entirety of a computer system depends upon the successful interaction of many separate information systems, via client-server transactions. In actual use, servers play the role of reservoirs of information and clients tap into these reservoirs with fluctuationlike transactions. These analogies reveal a thermodynamical quality, rooted in information theory, that is not merely of philosophical interest; it has practical uses in both computer science and physics.

The intriguing possibility of understanding the behavior of computers as dynamical systems is now starting to be appreciated [1,4]. It is motivated mainly by the possible practical rewards, which could be exploited in anomaly detection and security intrusion detection. In Refs. [2,3], it was argued that computer systems ought to behave like physical models of statistical mechanics, for the reasons above. In the present paper, drawing on recent meaurements in [4], this view is confirmed and amplified with both empirical and theoretical considerations.

There are two primary reasons for wanting to make a comparison between computers and physical models: first, practical knowledge, essential for building computer anomaly detectors, demands an understanding of the dynamical parametrization of the system; second, empirical knowledge about system dynamics arms us with methods for modeling networks of computer systems in a way that is readily accessible to researchers in both computer science and physics.

Anomaly detection in computer systems means identifying patterns of unusual activity. Unusual patterns of resource consumption or unusual trends in system variables point toward activities that could signify a fault in the system, or a potential attempt to abuse the system. The automatic detection of "anomalies" would allow automated responses, like immune systems to respond with countermeasures, where necessary. A considerable amount of effort is currently being invested in this type of technology. It has perhaps more in common with physics than with traditional computational models.

In order to detect anomalous and hence potentially threatening behavior, one first needs to characterize what is normal. Software for anomaly based intrusion detection has proved to be a difficult problem, mainly because the authors of such software do not have a sufficient concept of what characterizes normal behavior in a statistical sense. Computer programmers deal mainly with exact, microscopic logical states and transitions with Boolean thresholds. On the basis of their complexity of multiple tasks and interactions, one would expect computer systems to be fluctuating, statistical systems. Computers operate by performing microscopic transactions, which play the same formal role as fluctuations in statistical physics. The exchange of controlling information between computer processes is directly analogous to the exhange of particles between heat baths.

There is also the attractive notion of having an experimentally accessible system available to theoreticians in the field of finite temperature field theory and nonequilibrium physics. Although computer systems are relatively small in the sense of many-body systems, their behavior over time averages out to forms that can be approximated by infinite heat baths and the usual machinery of thermodynamics, in the space of weeks and months. This is a reasonable period of time over which to gather data, and it is a viable experimental arena for examining the influence of slowly varying change from a purely equilibrium situation.

Measurements show that the periodic topology of time evolution in computer networks places them in the same class of statistical systems as open thermodynamical ideal gases [4], i.e., systems that display Planckian statistics. Although thermal systems are far from unique in having these signatures, they are the most well known. The theoretical reason linking computer networks with thermal physics is that any randomly *fluctuating* system whose average behavior is constrained periodically (or, in this case, approximately periodically) will exhibit a Planck spectrum of fluctuations. This depends only on the periodic constraint. It is therefore possible to summarize the averages in terms of a temperature, resulting in a considerable compression of information.

Two influences dominate the average behavior of computer systems over human time scales; these are the external reservoirs of users and network clients, which undergo transactions with the system, but these are accompanied by a sea of fluctuations, which arise from the many interacting background processes that comprise modern computer systems. Many variables might be considered to characterize the behavior of a computer system over long times. Of all the variables one might record, some prove to be relevant and some to be irrelevant to computer behavior at the time scales on which humans interact with them. For instance, the number of independent processes running, the number of network transactions to particular services, and the amount of free

1063-651X/2000/62(2)/1738(5)/$15.00

disk space are three types of variable that are found to have a direct bearing on the state of the system and its long-term behavior. On the other hand, variables like CPU load, paging rates, and free random access memory (RAM) turn out to affect only the short-term behavior [4]. This is found empirically.

Computers, like other complex systems, are characterized by qualitatively different behaviors at multiple scales. It is appropriate to refer to these as microscopic, mesoscopic, and macroscopic. Microscopic behavior refers to exact mechanisms or atomic operations; mesoscopic behavior looks at small conglomerations of microscopic processes and examines them in isolation; macroscopic processes concern the long-term average behavior of the whole system. At the microscopic level we have individual system calls (on the order of milliseconds). At the mesoscopic level we have clusters and patterns of system calls including algorithms, procedures, and even viral activity (on the order of seconds). Finally, there is the macroscopic level at which one views all the activities of all the users over scales at which they typically work and consume resources (minutes, hours, days, weeks). Since it is users who cause the most significant changes and problems in computer systems, the macroscopic scale is of special interest for the detection of anomalies.

Consider now the relationship between computer systems and statistical systems. Statistical systems may either be steady state (in equilibrium) or change appreciably over times longer than the rate of statistical fluctuations (nonequilibrium). They are characterized by the existence variables that exhibit fluctuations, i.e., they vary randomly about an average value, on a time scale that is much shorter than the time over which one observes the system. The fluctuation scale has no effect on measured values, since no microscopic dynamics are visible at the scale of observation. When a system is close to equilibrium it can be thought of as being in a quasiequilibrium, with a superposed pattern of adiabatic changes. This situation has been analyzed using many different techniques [5–8] in statistical field theory. Such systems have some specific model-dependent properties, but also many universally applicable properties. It is the latter that are interesting here. Such a slowly varying statistical system is the model with which we hope to explain the behavior of networked computers.

The justification for assuming this model of computer dynamics is rather interesting: there are strong periodic rhythms in the dynamics which are disposed around three scales. The presence of fluctuations with this causal structure leads directly to a thermal interpretation. The three macroscopic time scales over which the system changes may be represented as separate, periodic driving forces.

(1) *Daily*. Users' daily work patterns exhibit the strongest influences on system periodicity. This is typically a nine to five rhythm which follows the general level of human activity.

(2) *Weekly*. The pattern of activity over a week tends to peak around midweek and fall to a minimum at weekends.

(3) *Seasonal*. Seasonal changes in work patterns depend on the type of site. For instance, universities have periods of low activity during summer vacation.

Of the above, the most important period for the definition of statistical variation is the daily rhythm, as seen in example

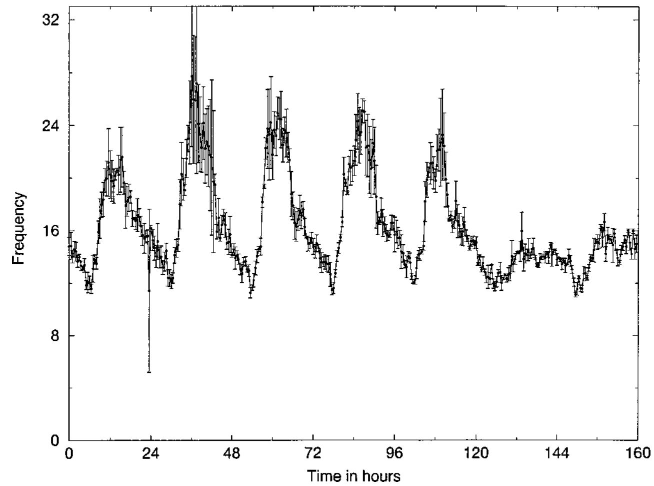

FIG. 1. The weekly average of nonprivileged (user) processes shows a constant daily pulse, quiet at the weekends, strong on Monday, rising to a peak on Tuesday, and falling off again toward the weekend. This graph also shows a conspicuous anomalous point, well outside the statistical tolerances of the error bars. The x axis measures time in hours over a weekly period, while the y axis is a frequency count. The solid line is the average value for a given time of week, while the error bars show the standard deviation of fluctuations about the mean.

Fig. 1, taken from Ref. [4]. This shows how, over a 24 hour period, one aspect of computer activity peaks during working hours. One sees also a strong weekly rhythm, characterized by a peak in activity around midweek, and a quiescent time during weekends. These are the unambiguous signatures of human work patterns. Data collected over several months are displayed in the figures. There are insufficient data, at present, to infer seasonal rhythms, but there are compelling reasons to suppose that these would also follow a periodic pattern in a stable environment. It is the existence of the daily pattern that is probably of greatest interest, since it is a short period whose repeating topology allows us to average over periodically identified points.

The magnitude of these periodic influences affects different measurements in different ways. Some activities, particularly network transactions, originate from many different physical locations around the world and are thus a superposition of work practices from several parts of the globe, i.e., daily rhythms which are time shifted with respect to one another. This tends to smear out the observed periodicity, depending on the relative numbers of transactions from different sources. Some variables are not directly coupled to user habits and do not exhibit periodicity at all. Such variables are not thermal in nature. It is the periodic variables that can be modeled using the notions of thermal physics.

The periodic nature of the influences on the system has a profound effect on the average behavior of the variables to which it applies. If we consider the average over many weeks (Fig. 1) and the average over daily periods (Fig. 1), we find stable patterns, within the tolerances expressed by standard deviation error bars. These data were collected by a monitoring process which measured values every two minutes over several months; the measurement process itself was shown to have no biasing effect on the measured values. The *average behavior* follows a predictable periodic pattern,

TABLE I. The correspondence between thermal physics and computer metrics $\phi_i(t)$. An arbitrary scale (the temperature) must be introduced in each case, on dimensional grounds. This characterizes the size and magnitude of the fluctuations from the mean values $\overline{\phi}_i$. An energy can be defined in each case. In the case of radiation, energy is inversely proportional to the wavelength of the radiation. In the computer case, an energy can be defined as being inversely proportional to the deviation from the mean. The effect of increasing the temperature in each case is to increase the width and height of the Planck distribution about the mean, i.e., to increase the fluctuations.

Planck law	Networked computer
KMS periodic imaginary time	Daily real time period
Temperature T	Temperature T
$E \propto \lambda^{-1}$	$E \propto \Delta \phi_i^{-1}$

within the tolerances. The distribution of these data will be discussed below. Any relevant system variable $\phi(t)$ that is strongly affected by periodicity can be expressed as an average (slowly varying part $\overline{\phi}$) plus a fluctuation (rapidly varying part $\delta\phi$): [9]

$$\overline{\phi}(t+\beta)=\overline{\phi}(t)+\delta\phi(t). \tag{1}$$

This expresses the fact that, after an elapsed period, the system returns to the same state, up to a statistical fluctuation $\delta\phi$. Both the averages and their fluctuations are dynamical (time-dependent) quantities. This is not typical of a strict equilibrium system, but of a system that varies on approximately the same time scale as the period β itself (note the time-evolution error bars in the figure). The variable $\overline{\phi}(t)$ is driven by a pseudoperiodic source $J(t)$,

$$J(t+\beta)\simeq J(t). \tag{2}$$

In dynamical language, this coupling would be expressed by a kernel or Green function $G(t,t')$ satisfying a Kubo-Martin-Schwinger (KMS) condition, in a linear response approximation:

$$\overline{\phi}(t)=\int dt' G(t,t')J(t') \tag{3}$$

(See Table I). The fact that the error bars $\delta\phi(t)$ are not constant in time means that the fluctuations themselves have slow variations. Thus the system is not in a perfect steady state: it is only an approximately thermal, nonequilibrium system.

From studies of general nonequilibrium field theories, it is possible to relate the nonequilibrium system to the equilibrium system they resemble. This is achieved using a locally varying scale, or *conformal transformation*. This transformation is a mapping that relates the actual time-varying distribution of fluctuating values to an idealized static or thermal distribution. The scaling is single-valued reparametrization:

$$\psi(t)=\frac{\phi(t)}{\sigma_\phi(t)}. \tag{4}$$

$\sigma(t)$ is the time-dependent standard deviation, or error bar in the figure. This rescaling is the identical procedure to the field theoretical method described in Ref. [8], applied at a more pedestrian level. It shows how small deviations from equilibrium can be viewed as a dynamical transformation (not unlike a gauge transformation [8]) that is, as perturbations around a simpler equilibrium model. The gradients that guide the evolution of the distribution therefore follow the form of a conformal or pseudogauge connection to the time derivative, as described in Ref. [8], owing to the local rescaling by $\sigma(t)$. In field theoretical language, the effect of a derivative on ψ is to make the effective change

$$D_t\rightarrow\partial_t-\frac{\partial_t\sigma}{\sigma(t)}. \tag{5}$$

This reproduces the scaling connections described in Ref. [8].

The fact that fluctuations or transactions take place with respect to a periodic background topology in time gives the phase space of this dynamical system the special character of an idealized black body. This might strike one as being peculiar at first; in fact, it is only a consequence of the quasiperiodic boundary conditions together with the presence of fluctuations. The periodicity itself associates a special wavelength in phase space with each level of excitation. Assuming that all excitations are equally likely, a general superposition generates a Bose-Einstein fluctuation spectrum. Periodic systems and their relationships to statistical physics are well known in physics [10–12].

The relationship between periodic systems and finite temperature statistical mechanics is neatly summarized by the KMS relation in statistical mechanics [10]. The KMS condition is really a group theoretical restatement of the equilibrium condition for a two-level system, as used originally by Einstein to derive his A and B coefficients for a two-level system. The KMS relation states that the ratio of absorption to emission (R) should be given by a Boltzmann probability:

$$R(\text{emission/absorption})=e^{\Delta E/kT} \tag{6}$$

for some energy spacing E and temperature T. When expressed in terms of linear response theory, using Green functions, this statistical assertion can be recast as

$$G(t+i\beta,t')=e^{-\beta E}G(t,t'), \tag{7}$$

which expresses the periodicity of the system up to a thermal fluctuation. This, in turn, leads to the observed finite temperature behavior [13].

For computer networks, the appearance of a thermal spectrum also arises from a periodicity arising from the computer's being strongly coupled to its pseudoperiodic source of users and network activity. The average transactions of many days make system activity approach the smooth envelope of a temperature distribution. For small times one sees jagged features due to the small numbers involved.

In natural units, the thermal distribution of an n-dimensional blackbody system is

$$E(\lambda)\propto\frac{\lambda^{-(n+2)}}{e^{1/\lambda T}-1}, \tag{8}$$

where $\lambda = 1/\omega$ is the wavelength of light emitted or absorbed. This can be derived from the fluctuation energy for any noninteracting gas or free field on a periodic Euclidean space. Why choose a free field theory to model a computer system? Because it is the simplest physical model exhibiting Gaussian fluctuations about a mean, and it is well known in the literature. Any universal features that owe their existence to the causal structure of the system should already be apparent here. One need not dream up detailed interactions in order to identify such a basic concept as a temperature for the system. It remains then only to compare the model to the measured data. The effective action or free energy for a system is the sum over fluctuations (transaction bubbles) and may be used to determine the distribution signature. The energy due to such random fluctuations on a periodic space is a Gaussian functional integral:

$$\Gamma = -\ln \int d\phi(t) e^{-\int d^n x \phi(-\Box \phi)}. \tag{9}$$

The exponent is the Lagrangian density for a free Euclidean scalar field. This is the simplest physical system that exhibits a dynamical evolution in time and whose fluctuations can be summed. It makes no assumptions about the nature of the field, other than the fact that fluctuations are free to commute. This reflects the fact that the occupation of a particular value is not limited to a maxmimum number, as it would be in Fermi-Dirac statistics. Since the labels that characterize each special event are irrelevant to the total count over time, the fluctuations are bosonic, i.e., the order in which equivalent events actually occurred is irrelevant to the count. The free energy functional integral is a known functional integral [12]:

$$\Gamma \sim \mathrm{Tr} \ln(-\Box) \propto \int d\omega \frac{\omega^n}{\exp(\beta\omega)-1} = \int d\omega E(\omega), \tag{10}$$

and yields the integral of the Planck formula, on changing variables $\omega = 1/\lambda$. The overall constant of proportionality is uninteresting and must be fitted to the data. This fluctuation distribution may be seen fitted to the measured example data in Fig. 2, using the formula with $n=1$:

$$D(\lambda) = A e^{-(\lambda-\bar{\lambda})^2/2\sigma^2} + \frac{B}{(\lambda-\lambda_0)^3 (e^{1/(\lambda-\lambda_0)T}-1)}. \tag{11}$$

The values of the constants A, B and λ_0 are chosen to fit the data. Their absolute values have no significance, since there is no ''standard candle'' computer system to compare to, but changes relative to the local norm could be interpreted as anomalies. Nonzero A allows for the presence of additional Gaussian noise in some measurements. Although there is room for only one graph here, the result has been reproduced across many machines and for many variables at two widely different network sites.

The question remains, however, as to precise identification of a suitable detailed model. Would a more detailed statistical field theory be an adequate model? What then would be the correct configuration (position) space for such a theory? If we consider every variable as a generic field

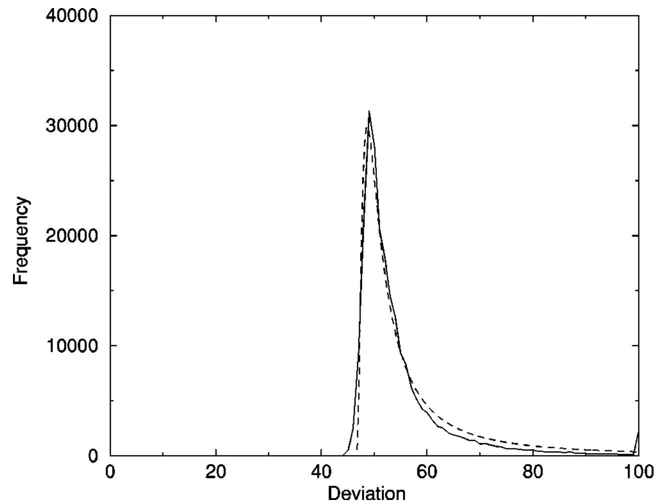

FIG. 2. The distribution of fluctuations in network sockets from world wide web connections averaged over the daily period. The x axis shows the deviation about the scaled mean value of 50 $\delta\phi/\sigma$ and the y axis shows the number of points measured in class intervals of a half σ. The solid line shows that the distribution of values about the mean is an almost pure Planckian black-body distribution. The dotted line shows the theoretical Planckian fit with $A=0$.

$\phi(t,x)$, of the type well known in field theories, how shall we assign meaning to the position x? How many dimensions are appropriate? The effective dimensionality of a configuration space would determine the nature of any phase transitions that could occur in the statistical system. Far from being of purely whimsical interest, phase transitions in system variables could be an important phenomenon as network interactions increase, forcing us to think of multiple host conglomerates as large virtual machines with potentially unstable behavior.

One should not be tempted to think of a position variable in terms of the physical location of different computer systems, or the connectivity of the network, since we have already partitioned into local system and external reservoir (sources): there is no relevant position variable in this model. (Intercomputer positions refer to an equally valid but different problem.) Rather, the configuration space refers to the number of ways the system can change, i.e., its number of internal degrees of freedom, as viewed by each host. The dimensionality will have to be understood in each case separately, from the set of relevant parameters that controls the dynamics (this is different in different computer systems). It is unclear whether a continuous variable or a lattice (cellular automaton) model would be more appropriate at present. A careful group theoretical analysis of the variables would resolve this.

There are many problems remaining for a long-term practical anomaly detection scheme and physicists can contribute significantly to this discussion. One is how to compute the relevant averages without storing large quantities of data. Another is how to analyze the time series with respect to the averages in real time. With the present work, one has a partial parametrization of the expected behavior in terms of a temperature T and a scale $\sigma(t)$ that summarize the full statistical data set, i.e., a framework in which to make predictions. As far as physics is concerned, it is gratifying to see well-established theory validated in such an experimentally accessible system. The analogy goes deeper than it has been possible to present here, and there is room for additional work which clarifies the issue of entropy production. This will be presented in a more detailed paper presently.

In summary, computer systems are open thermodynamical systems, exchanging information with reservoirs of users and network clients. The methods and notions of statistical physics can be applied to understand their average behavior (macrostate). The system rescales, approximately conformally, as one would expect of an envelope described by a nonequilibrium field theory. The observations thus confirm and link two separate pieces of work [8,3]. This should be of interest both to computer scientists and to physicists alike.

The empirical analyses supporting this work were collected in collaboration with Sigmund Straumsnes and Hårek Haugerud in Ref. [4], and Tim Bower.

[1] N. Glance, T. Hogg, and B.A. Huberman, Int. J. Mod. Phys. C **2**, 735 (1991).

[2] M. Burgess, Software Prac. Exper. **28**, 1519 (1998).

[3] M. Burgess, *Proceedings of the 12th Systems Administration Conference (LISA)* (USENIX, San Francisco, 1998), p. 283.

[4] M. Burgess, H. Haugerud, and S. Straumsnes (unpublished).

[5] J. Schwinger, J. Math. Phys. **2**, 407 (1961).

[6] R. Feynman and F.L. Vernon, Ann. Phys. (N.Y.) **24**, 118 (1963).

[7] E. Calzetta and B.L. Hu, Phys. Rev. D **37**, 2872 (1988).

[8] M. Burgess, Phys. Rev. D **52**, 7103 (1995).

[9] F. Reif, *Fundamentals of Statistical Mechanics* (McGraw-Hill, Singapore, 1965).

[10] P.C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).

[11] L. Ford, Phys. Rev. D **21**, 933 (1980); A. McLachlan, Nucl. Phys. B **338**, 188 (1990).

[12] M. Burgess and D.J. Toms, Ann. Phys. (N.Y.) **210**, 438 (1991); M. Burgess, A. McLachlan, and D.J. Toms, Phys. Rev. D **43**, 1956 (1991).

[13] Briefly, the KMS condition notes that, in a dynamical theory whose time evolution progresses by a Hamiltonian h, with an Abelian group transformation $\exp(-iht)$, a periodicity in the imaginary part of the time t generates a static Boltzmann distribution $\exp(-h\beta)$. Strictly, these two viewpoints are mutually orthogonal, since an equilibrium system has no time evolution, but in quasistatic systems, approximate meaning can be given to adiabatic time evolution, alongside thermal equilibrium.