02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS

Submit only ONE copy of this form **for each PI/PD and co-PI/PD** identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. ***DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.***

PI/PD Name: Elizabeth Bradley

Gender: ☐ Male ☒ Female

Ethnicity: (Choose one response) ☐ Hispanic or Latino ☒ Not Hispanic or Latino

Race:
(Select one or more)
- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Black or African American
- ☐ Native Hawaiian or Other Pacific Islander
- ☒ White

Disability Status:
(Select one or more)
- ☐ Hearing Impairment
- ☐ Visual Impairment
- ☐ Mobility/Orthopedic Impairment
- ☐ Other
- ☒ None

Citizenship: (Choose one) ☒ U.S. Citizen ☐ Permanent Resident ☐ Other non-U.S. Citizen

Check here if you do not wish to provide any or all of the above information (excluding PI/PD name): ☒

REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project ☒

Ethnicity Definition:
Hispanic or Latino. A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.
Race Definitions:
American Indian or Alaska Native. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
Black or African American. A person having origins in any of the black racial groups of Africa.
Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

WHY THIS INFORMATION IS BEING REQUESTED:

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information recieved from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational oppurtunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

List of Suggested Reviewers or Reviewers Not To Include (optional)

SUGGESTED REVIEWERS:
Not Listed

REVIEWERS NOT TO INCLUDE:
Not Listed

COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 11-1 PD 11-7478 10/01/11	FOR NSF USE ONLY
FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.) CMMI - DYNAMICAL SYSTEMS	NSF PROPOSAL NUMBER **1162440**

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
09/30/2011	1	07030000 CMMI	7478	007431505	10/10/2012 6:44pm S

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)	SHOW PREVIOUS AWARD NO. IF THIS IS	IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY?
846000555	☐ A RENEWAL ☐ AN ACCOMPLISHMENT-BASED RENEWAL	YES ☐ NO ☒ IF YES, LIST ACRONYM(S)

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE University of Colorado at Boulder	ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE 3100 Marine Street, Room 481 572 UCB Boulder, CO 80309-0572
AWARDEE ORGANIZATION CODE (IF KNOWN) 0013706000	
NAME OF PRIMARY PLACE OF PERF University of Colorado at Boulder	ADDRESS OF PRIMARY PLACE OF PERF, INCLUDING 9 DIGIT ZIP CODE University of Colorado at Boulder 3100 Marine Street, Room 479 Boulder ,CO ,803090572 ,US.

IS AWARDEE ORGANIZATION (Check All That Apply) (See GPG II.C For Definitions)
☐ SMALL BUSINESS ☐ MINORITY BUSINESS ☐ IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE
☐ FOR-PROFIT ORGANIZATION ☐ WOMAN-OWNED BUSINESS

TITLE OF PROPOSED PROJECT **DynSyst_Special_Topics: Reduced-Order Dynamical Models for Effective Power Management in Computer Systems**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE
$ 365,522	36 months	06/01/12	

CHECK APPROPRIATE BOX(ES) IF THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW

- ☐ BEGINNING INVESTIGATOR (GPG I.G.2)
- ☐ DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.C.1.e)
- ☐ PROPRIETARY & PRIVILEGED INFORMATION (GPG I.D, II.C.1.d)
- ☐ HISTORIC PLACES (GPG II.C.2.j)
- ☐ EAGER* (GPG II.D.2) ☐ RAPID** (GPG II.D.1)
- ☐ VERTEBRATE ANIMALS (GPG II.D.6) IACUC App. Date ______
  PHS Animal Welfare Assurance Number ______
- ☐ HUMAN SUBJECTS (GPG II.D.7) Human Subjects Assurance Number ______
  Exemption Subsection ______ or IRB App. Date ______
- ☐ INTERNATIONAL COOPERATIVE ACTIVITIES: COUNTRY/COUNTRIES INVOLVED (GPG II.C.2.j) ______
- ☐ HIGH RESOLUTION GRAPHICS/OTHER GRAPHICS WHERE EXACT COLOR REPRESENTATION IS REQUIRED FOR PROPER INTERPRETATION (GPG I.G.1)

PI/PD DEPARTMENT Department of Computer Science	PI/PD POSTAL ADDRESS 430 UCB University of Colorado at Boulder Boulder, CO 803090430 United States
PI/PD FAX NUMBER 303-492-2844	

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Electronic Mail Address
PI/PD NAME Elizabeth Bradley	PhD	1992	303-492-5355	lizb@cs.colorado.edu
CO-PI/PD				
CO-PI/PD				
CO-PI/PD				
CO-PI/PD				

Electronic Signature

CERTIFICATION PAGE

Certification for Authorized Organizational Representative or Individual Applicant:

By signing and submitting this proposal, the Authorized Organizational Representative or Individual Applicant is: (1) certifying that statements made herein are true and complete to the best of his/her knowledge; and (2) agreeing to accept the obligation to comply with NSF award terms and conditions if an award is made as a result of this application. Further, the applicant is hereby providing certifications regarding debarment and suspension, drug-free workplace, lobbying activities (see below), responsible conduct of research, nondiscrimination, and flood hazard insurance (when applicable) as set forth in the NSF Proposal & Award Policies & Procedures Guide, Part I: the Grant Proposal Guide (GPG) (NSF 11-1). Willful provision of false information in this application and its supporting documents or in reports required under an ensuing award is a criminal offense (U. S. Code, Title 18, Section 1001).

Conflict of Interest Certification

In addition, if the applicant institution employs more than fifty persons, by electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative of the applicant institution is certifying that the institution has implemented a written and enforced conflict of interest policy that is consistent with the provisions of the NSF Proposal & Award Policies & Procedures Guide, Part II, Award & Administration Guide (AAG) Chapter IV.A; that to the best of his/her knowledge, all financial disclosures required by that conflict of interest policy have been made; and that all identified conflicts of interest will have been satisfactorily managed, reduced or eliminated prior to the institution's expenditure of any funds under the award, in accordance with the institution's conflict of interest policy. Conflicts which cannot be satisfactorily managed, reduced or eliminated must be disclosed to NSF.

Drug Free Work Place Certification

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant is providing the Drug Free Work Place Certification contained in Exhibit II-3 of the Grant Proposal Guide.

Debarment and Suspension Certification

(If answer "yes", please provide explanation.)

Is the organization or its principals presently debarred, suspended, proposed for debarment, declared ineligible, or voluntarily excluded from covered transactions by any Federal department or agency? Yes ☐ No ☒

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant is providing the Debarment and Suspension Certification contained in Exhibit II-4 of the Grant Proposal Guide.

Certification Regarding Lobbying

The following certification is required for an award of a Federal contract, grant, or cooperative agreement exceeding $100,000 and for an award of a Federal loan or a commitment providing for the United States to insure or guarantee a loan exceeding $150,000.

Certification for Contracts, Grants, Loans and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:

(1) No federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.

(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, ''Disclosure of Lobbying Activities,'' in accordance with its instructions.

(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements and that all subrecipients shall certify and disclose accordingly.

This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, Title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

Certification Regarding Nondiscrimination

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative is providing the Certification Regarding Nondiscrimination contained in Exhibit II-6 of the Grant Proposal Guide.

Certification Regarding Flood Hazard Insurance

Two sections of the National Flood Insurance Act of 1968 (42 USC §4012a and §4106) bar Federal agencies from giving financial assistance for acquisition or construction purposes in any area identified by the Federal Emergency Management Agency (FEMA) as having special flood hazards unless the:

(1) community in which that area is located participates in the national flood insurance program; and
(2) building (and any related equipment) is covered by adequate flood insurance.

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant located in FEMA-designated special flood hazard areas is certifying that adequate flood insurance has been or will be obtained in the following situations:

(1) for NSF grants for the construction of a building or facility, regardless of the dollar amount of the grant; and
(2) for other NSF Grants when more than $25,000 has been budgeted in the proposal for repair, alteration or improvement (construction) of a building or facility.

Certification Regarding Responsible Conduct of Research (RCR)
(This certification is not applicable to proposals for conferences, symposia, and workshops.)

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative of the applicant institution is certifying that, in accordance with the NSF Proposal & Award Policies & Procedures Guide, Part II, Award & Administration Guide (AAG) Chapter IV.B., the institution has a plan in place to provide appropriate training and oversight in the responsible and ethical conduct of research to undergraduates, graduate students and postdoctoral researchers who will be supported by NSF to conduct research.
The undersigned shall require that the language of this certification be included in any award documents for all subawards at all tiers.

AUTHORIZED ORGANIZATIONAL REPRESENTATIVE	SIGNATURE	DATE
NAME **Parker M Dougherty**	**Electronic Signature**	**Sep 30 2011 4:22PM**

TELEPHONE NUMBER	ELECTRONIC MAIL ADDRESS	FAX NUMBER
303-735-6525	**parker.dougherty@colorado.edu**	**303-492-6421**

* EAGER - EArly-concept Grants for Exploratory Research
** RAPID - Grants for Rapid Response Research

This proposal is about reduced-order modeling and forecasting techniques for reducing power use in microprocessor chips, one of the most critical classes of engineered systems in use today. Modern power management solutions for these systems employ extremely simple control strategies—e.g., lowering the clock frequency by a fixed, pre-determined amount if a processor's load crosses some threshold. The methods developed by the control theory and nonlinear dynamics communities have long since moved beyond this level of sophistication. Model-based prediction, in particular, could enable vastly improved power management, but only if the models involved are accurate. If one could predict that a particular thread of computation would be bogged down for the next 0.6 seconds waiting for data from the computer's memory, for instance, one could put that thread on a low-power hold for that time period. Prediction of the future behavior of a complex nonlinear dynamical system like a modern computer is a serious challenge, however—and it is all but impossible if one uses mathematics that assumes linearity and/or time invariance, as has been the rule until recently in the computer systems community. The mathematical machinery developed by the nonlinear dynamics community, on which the solutions proposed here are based, offers a much better starting point from which to attack this problem.

Intellectual merit:

The approach proposed here uses a novel reduced-order modeling strategy that first transforms the time-series data into a 2D representation called a τ-return map. The power of this representation is that it brings out temporal relationships explicitly, 'unfolding' the temporal patterns in the time series into a spatial dimension. Forecast models working on this transformed data can be used to create accurate predictions of processor and memory loads in multicore processors: information that can be used to dynamically adapt the computation to the resources, and vice versa.

An important element of the intellectual merit of this proposal is its cross-fertilization of nonlinear dynamics and computer systems, two fields that have to date been largely disjoint. It is important to note, however, that while the power of the approach proposed here derives from its use of nonlinear dynamics methods, this project moves well beyond straightforward application of those methods to the problems of another field. Traditional nonlinear time-series analysis models are extremely powerful, but they are generally ineffective for 'on the fly' prediction because estimation of their free parameters requires expert human interpretation. The τ-return map effectively sidesteps this issue: it has a single free parameter, its geometry is not hugely sensitive to small changes in the value of that parameter, and it appears to work quite well for predicting the complex dynamics of computer performance data. This speaks directly to the solicitation to which this proposal responds: *"When the phenomena are complicated, a compromise must be found between completeness and simplicity..."*

Broader impact:

Because computers are so common and so critical, this work has the potential to contribute to science, engineering, and well beyond. Even a short-term prediction of the load patterns in a modern microprocessor, for example, could enable power savings in any computer application by allowing the operating system software to effectively allocate loads to resources. The potential benefits of model-based power management are particularly compelling in view of the recent design evolution of multicore processors and the rapid proliferation of mobile devices. And the reduced-order modeling and forecasting strategies proposed here could be used to improve the designs of other systems as well: anywhere that a good prediction could make a difference.

This project also has potentially broad impacts in education, outreach, and human resource development. The PI is a senior woman with a diverse, vertically integrated research group and longstanding commitment to undergraduate research, outreach, and women/STEM issues. She teaches courses in computer architecture and nonlinear dynamics and has a significant prior track record of educational innovations. She will incorporate class projects in computer architecture courses where students undertake small validation projects; these will give students an appreciation for the complexity of modern hardware and software. She will also incorporate class projects in nonlinear dynamics where students will use textbook techniques for comparing, modeling, and predicting computer-performance data; these will give students a deeper understanding of the relative strengths and weaknesses of different time-series analysis algorithms.

TABLE OF CONTENTS

For font size and page formatting specifications, see GPG section II.B.2.

Appendix Items:

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

1 Overview

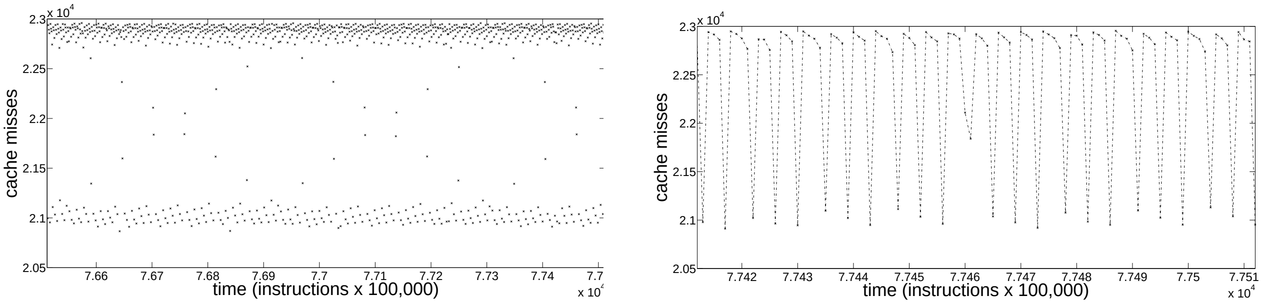
Though it is not necessarily the view taken by those who design them, modern computers are deterministic nonlinear dynamical systems. The complexity of these ubiquitous engineered artifacts has increased very rapidly over the past few decades, as processor designers sought competitive advantage, with corresponding effects on the complexity of their dynamics. Modern computers have multiple processing units and multi-layer memories, for instance, and they use complicated hardware/software strategies to move data and threads of computation across those resources. These features, along with all the others that go into the design of modern processor chips, makes the patterns of their processor loads and memory accesses very complex indeed.

This dynamical complexity has extremely important implications for computer design—particularly regarding **power management**, which is the focus of this proposal. If one could predict that a particular computational thread would be bogged down for the next 0.6 seconds waiting for data from the computer's memory, for instance, one could put that thread on hold for that time period (e.g., by migrating it to a processing unit whose clock speed is scaled back). The resulting power savings could be substantial. However, prediction of the future behavior of a complex nonlinear dynamical system is a serious challenge, and the traditional mathematics used in the computer performance community is not up to this task. In this literature, computers are typically modeled as high-dimensional stochastic processes (e.g., [25]) and the temporal details of their dynamics are lumped into aggregate metrics (e.g., [3, 33]). These strategies cannot capture the deterministic, time-varying nature of computer dynamics, which dictates the power used by these machines. And few of this community's power-management strategies use models at all, let alone prediction. Intel's Speedstep, for instance, uses a simple, non-adaptive strategy—bang-bang control [31]—and dynamic frequency/voltage scaling to manage processor power use by scaling back clock speed. This strategy involves two clock frequency settings: if the processor is idle, Speedstep sets the clock rate to the lower setting; if the processor is busy, it invokes the faster one. Enhanced Speedstep (EIST) employs finer-grained clock frequency control, but still involves static, hard-coded parameters and thresholds that must be tuned.

The mathematical machinery developed by the nonlinear dynamics community offers a much better starting point from which to attack the problem of power management in computer systems. Until recently, however, there has been almost no cross-fertilization between these two fields. This is a missed opportunity. As we were able to establish with the support of an NSF grant awarded in 2007[1], it is highly effective to treat computers as nonlinear dynamical systems [5, 23, 24]. In this view, register and memory contents and physical variables like the temperature of different regions of the processor chip define the state of the system. The logic hardwired into the computer, combined with the software executing on that hardware, defines the system's dynamics. Under the influence of these dynamics, the processor's state moves on a trajectory through its high-dimensional state space as the clock cycles progress and the program executes.

The framework outlined in the previous paragraph gives us a tremendous amount of new

[1] "CSR—SMA: Validating Architectural Simulators Using Non-Linear Dynamics Techniques;" co-PIs: E. Bradley and A. Diwan. Please see page 13 for more details.



(a) (b)

Figure 1: L2 cache misses, aggregated and averaged over 100,000-instruction periods, in an Intel Core Duo® running a simple column-major matrix initialization: (a) time-series trace (b) closeup. The almost-periodic patterns in this trace, which tracks the efficiency of the computer's memory accesses, are a result of deterministic chaos in the system.

traction on the problems of the computer systems field. Among other things, it lets us use nonlinear time-series analysis techniques to study the dynamics of these critical engineered systems. Harvesting time-series data for this purpose is comparatively straightforward, thanks to the hardware performance monitoring facilities that are built into most modern microprocessor chips. Figure 1 shows a trace of the rate of L2 cache misses—events that occur when the data needed by a processor is not immediately available—gathered from an Intel Core Duo®-based computer as it executes a four-line C program that repeatedly initializes a 256×256 matrix in column-major order. A close look at this time series reveals interesting structure: patterns that are almost periodic, but not quite. In [23], we showed that these dynamics exhibited deterministic chaos. And we were able to corroborate those results using multiple methods and replicate them using other performance metrics (e.g., instructions per cycle), so we have significant confidence in the results.

The basic premise behind this proposal follows directly from that observation: if the performance load in a microprocessor can be effectively *modeled* as a nonlinear dynamical system, then there exists a deterministic forecast rule that can be used to *predict* the future value of that load—up to a point dictated by the measurement uncertainty and the nature of the dynamics. The research goals of this proposal are to construct those forecast models and use them to dynamically allocate loads to resources in a running computer, thereby reducing power usage without impacting time to completion.

To accomplish this, we will use nonlinear time-series analysis techniques to build forecast models from data like the trace in Figure 1. The traditional first step in that process is delay-coordinate embedding [26, 28, 32], which allows one to reconstruct a system's full state-space dynamics from scalar time-series data—provided that some conditions hold regarding the data. Figure 2 shows a delay-coordinate embedding of the data from Figure 1. The coordinates of each point on such a plot are differently-delayed elements of the time-series $y(t)$: that is, $y(t)$ on the first axis, $y(t+\tau)$ on the second, $y(t+2\tau)$ on the third, and so on. Geometric structure in these kinds of plots—clearly visible in Figure 2—is an indication of determinism. A deeper analysis of Figure 2, as alluded to above, supports that diagnosis, indicating the presence of a chaotic attractor in these cache-miss dynamics, with largest

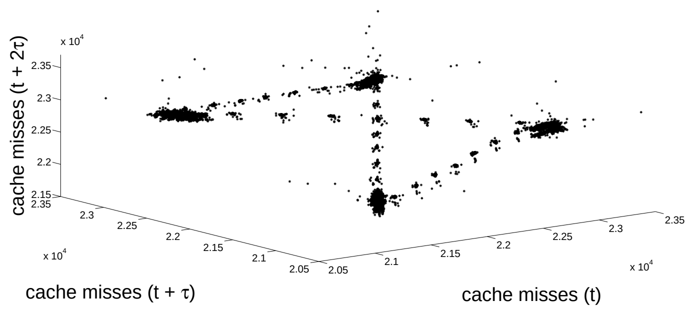


Figure 2: A 3D projection of a delay-coordinate embedding of the trace from Figure 1 with a delay (τ) of 100,000 instructions. This technique ‘unfolds’ the dynamics of a time series, finding the lowest-order model—twelve dimensions, here—that is faithful to the topology of the underlying dynamics.

Lyapunov exponent $\lambda_1 = 8000 \pm 200$ instructions, embedded in a 12-dimensional state space [23]. We repeated this experiment using different performance metrics, different computers, and different code; the dynamics differed in interesting ways as the hardware and software were varied, but topological dimension of the embedding space was similar ($m = 10$ to 15) in all of these experiments. This comparatively low dimension is somewhat surprising in systems as complex as modern computers, which contain billions of transistors. Computers are organized by design into a small number of tightly coupled subsystems, however, which greatly reduces the effective number of degrees of freedom in the system. This brings out an important feature of the delay-coordinate embedding methodology: *it finds the lowest-order model that is faithful to the topology of the underlying dynamics*. In the case of computer performance, this effects a reduction in order from $> 10^9$ to ~ 10.

The obvious next step is to use that embedding to build a forecast model. There are many possible approaches to this (e.g., [8, 34]); perhaps the simplest is the “Lorenz method of analogues” (LMA), which is essentially nearest-neighbor prediction [20]. This method, which was developed by Edward Lorenz for use in the true state space of a dynamical system, was extended to delay-coordinate embedding space by Kantz *et al.* [17]. Even this simple algorithm—which builds predictions by looking for the nearest neighbor of a given point, then taking that neighbor’s path as the forecast—works quite well on computer performance data, as shown in Figure 3. In this experiment, we held back the last 10% of the signal in Figure 1, embedded the first 90% of the signal as in Figure 2, modeled the spatiotemporal geometry of the results using LMA, used that model to predict forwards from the last point in that series, and compared the true and predicted values. The root mean squared prediction error (RMSPE) between forecast and truth is 15.2—a very good result considering that the average value of the signal is 2.2×10^4. In other words, we were able to predict cache-miss rate very accurately more than 800 million instructions ($\sim$ 7 billion clock cycles) into the future.

While these results are encouraging, the methods that produced them are completely impractical for the problem tackled here. Delay-coordinate embedding involves two free parameters, the delay τ and the dimension m, which must be estimated from the data.

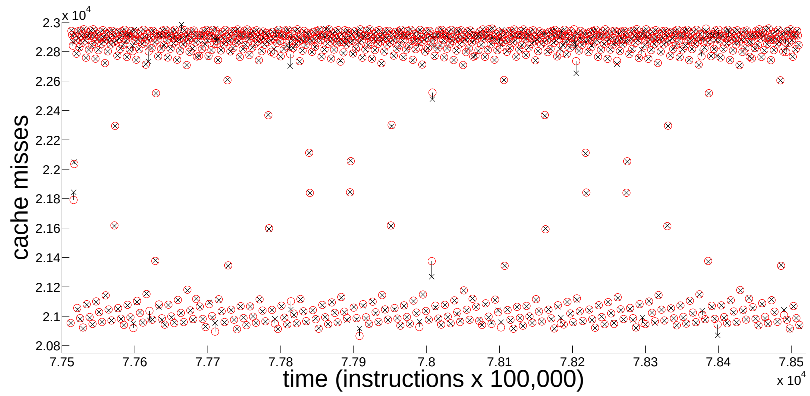


Figure 3: Forecast of last 1000 points of Figure 1 made using nearest-neighbor prediction on Figure 2. Red circles and black $\times$s are the true and predicted values, respectively; vertical bars show where these values differ. The root mean squared prediction error of this forecast is 15.2, or 0.07% of the average value of the signal.

This estimation must be redone if the dynamics changes and it requires expert interpretation (especially m), so that process cannot be automated. Unfortunately, that makes delay-coordinate embedding ineffective for on-the-fly prediction of nonstationary systems[2]. We conjecture, however, that a full formal embedding—which is mandatory for detailed dynamical analysis—*is not necessary for the purposes of prediction of computer performance dynamics*. That conjecture catalyzed this proposal, leading us to explore prediction models built in a reduced-order version of the full embedded space.

In particular, we propose to construct a τ-return map—$y(t)$ versus $y(t+\tau)$—from the time series $y(t)$ and then build the forecast model in that space. Figure 4 shows a τ-return map of the trace from Figure 1 with $\tau = 100,000$ instructions. Even though a 2D reconstruction like this is not completely faithful to the underlying dynamics, our preliminary results suggest that it is good enough to support accurate forecast models of computer performance dynamics—i.e., that carrying out nearest-neighbor prediction in this space produces accurate results. As a first piece of evidence in support of this claim, we constructed a forecast of the trace in Figure 1 using nearest-neighbor prediction on the τ-return map of Figure 4. The resulting RMPSE was 24.4. As one would expect, this is larger than the result we obtained using the same prediction algorithm on the full 12D embedding (15.2), but not *much* larger, particularly in view of the fact that the average value of the signal is 2.2×10^4. In other words, we were again able to predict cache-miss rate very accurately more than 800 million instructions into the future—but this time using a reduced-order version of the full embedding space.

This pair of experiments effectively validates the conjecture on which this proposal is based: the full complexity (and effort) of the delay-coordinate 'unfolding' process may not be strictly necessary to the success of forecast models of these dynamics. That complexity

[2]Predictions based on full delay-coordinate embeddings have seen some success in the nonlinear dynamics community (e.g., [34]), but not in real time and on systems whose dynamics varies on time scales as fast as modern computers.

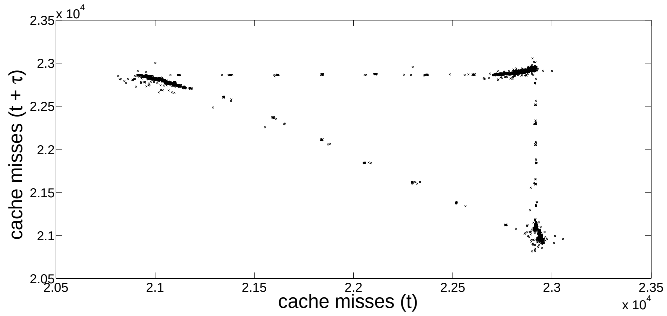


Figure 4: A τ-return map of the trace from Figure 1 with $\tau = 100,000$ instructions. The premise of this proposal is that forecast models built in this reduced-order version of the full embedding space, which only involves the single free parameter τ that can be estimated 'on the fly' from the time series, can be used for power management in computer systems.

and effort are exactly what make this order reduction important: it is all but impossible to do delay-coordinate embedding 'on the fly' because values for its dimension parameter are so hard to estimate. A τ-return map, however, has a single free parameter and its geometry is not hugely sensitive to small changes in its value. One of our research challenges in this project will be to estimate a good value for τ from the time series and adapt that estimate effectively to the rapidly changing dynamics that we have observed in computer performance traces. Our second challenge will be to develop forecast algorithms that work well in τ-return-map space. Our third challenge will to put the forecasts produced by those algorithms into physical practice to improve power management in real computers.

The following section describes our novel reduced-order modeling strategy in more depth (including the underlying details and calculations for the results mentioned in this first section), presents a few more results that validate the utility of this approach, compares it to traditional approaches, and then describes how the results can be used to support effective model-based control and power management in real-world computing applications. Since this work straddles two substantially disjoint research areas—nonlinear dynamics and computer systems—the discussion in the following section includes some basic background material that can be skimmed or skipped by experts in the corresponding area.

2 Research Plan

The specific elements of the **research plan** for the project proposed here are:

1. to develop effective reduced-order nonlinear time-series prediction models
2. to use those models to make predictions of important features of computer performance dynamics (e.g., processor loads, memory access patterns)
3. to exploit those predictions to reduce power use in a running computer

The rest of this section outlines the specifics of each element of this plan: items 1 and 2 in Section 2.1 and item 3 in Section 2.2.

2.1 Modeling and Forecasting Computer Performance

For the purposes of this proposal, we will consider a "stored-program computer," *i.e.,* a standard von Neumann architecture, as a deterministic nonlinear dynamical system. In a stored-program computer, the current state—both instructions and data—are stored in some form of addressable memory. The contents of this memory are, as established in [23], part of the state X of the computer. Other components of the computer, such as external memory and video cards, also play roles in its state. Those roles depend on the decisions made by the computer designers—how things are implemented and connected—almost all of which are proprietary. In order to distinguish known and unknown effects, we define the state space X as a composition of the addressable memory elements $\vec{x}_m$ and the unknown implementation variables $\vec{x}_u$:

$$X = \{\vec{x} \mid \vec{x} = [\vec{x}_m,\ \vec{x}_u]\} \tag{1}$$

The distinction between $\vec{x}_m$ and $\vec{x}_u$ is important because the dynamics of a running computer have two distinct sources: a map $\vec{F}_{code}$ that acts on the addressable memory $\vec{x}_m$ directly, based on the program instructions, and a map $\vec{F}_{impl}$ that captures how the implementation affects the evolution of the computer state. The overall dynamics of the computer—that is, the mapping from its state at the n^{th} clock cycle to its state at the $n+1^{st}$ clock cycle—is a composition of these two maps:

$$\vec{x}(t_{n+1}) = \vec{F}_P(\vec{x}(t_n)) = \vec{F}_{impl} \circ \vec{F}_{code}(\vec{x}(t_n)). \tag{2}$$

where $\vec{F}_P$ is the performance dynamics of the computer. An improved design for the processor, for instance—that is, a "better" $\vec{F}_P$—is a change in $\vec{F}_{impl}$. The form of the map $\vec{F}_{code}$ is dictated by the combination of the computer's formal specification (`x86`, for the Intel Core Duo®) and the software that it is running. Both $\vec{F}_{impl}$ and $\vec{F}_{code}$ are nonlinear and deterministic, and their composed dynamics must be modeled and predicted in order to produce an accurate forecast of the state X.

The framework outlined in the previous paragraph lets us use the methods of nonlinear dynamics to model $\vec{F}_P$—as long as we can gather time-series data from the computer that samples those dynamics in a manner that satisfies the requirements of the associated theorems. The hardware performance monitor registers (HPMs) that are built into modern processor chips can be programmed to count events on the chip: the total number of instructions executed per cycle (IPC), for instance, or the total number of references to the data cache. These are some of the most widely used and salient metrics in the computer performance analysis literature [3, 19, 25, 30]. IPC is a good proxy for speed because most modern microprocessors can execute more than one instruction per clock cycle. Cache-access data is an effective way to study the dynamical role of a program's memory use, another key bottleneck in computer performance. While these two metrics are not elements of the state vector $\vec{x}$, the fundamental theorem of delay-coordinate embedding only requires that one measure a quantity that is a smooth, generic function of at least one state variable. We showed in [5] that the transformation performed by the HPMs in sampling the state $\vec{x}$ is smooth and generic *unless those registers overflow*—an unlikely event, given that they are

64 bits long and that we read them every $10^5 - 10^6$ instructions[3]. The choice of that sample interval is important for another reason as well. The HPMs are part of the system under study, so accessing them can disturb the very dynamics that we use them to sample. We address this potential *observer problem* by varying the sample interval and testing to make sure that the sampling is not affecting the dynamics [24]. To further reduce perturbation, our measurement tool only monitors events when the target program is running, and not when the operating system (or the monitoring tool itself) have control of the microprocessor. Finally, we follow best practices from the computer performance analysis community [12] when measuring the system: we only use local disks and limit the number of other processes that are running on the machine (i.e., Linux `init` level 1).

Provided that the underlying dynamics and the measurement function (the mapping from the state vector $\vec{x}$ to the scalar value y measured by the HPMs) are both smooth and generic, Takens [32] formally proves that the delay-coordinate map

$$F(\tau, m)(x) = ([x(t)\ x(t+\tau)\ \ldots\ x(t+m\tau)]) \tag{3}$$

from a d-dimensional smooth compact manifold M to Re^{2d+1} is a diffeomorphism on M [28]—in other words, that the reconstructed dynamics F and the true (hidden) dynamics $\vec{F}_P$ have the same topology. Figure 2 was produced by applying such a map to the time-series data in Figure 1 with $m = 12$ and $\tau =$100,000. The Takens theorem[4] is an extremely powerful result because it guarantees that F is a good model of the system. Estimating the two free parameters m and τ, however, involves a variety of approximate heuristics have a host of parameters and are extremely sensitive to the values of those parameters—as well as to noise, data length, and numerical effects. This estimation process, particularly that of the dimension m, demands expert human judgment and interpretation, which makes the delay map F impossible to calculate in the real-time environment of a forecast model for a computer whose clock is running at several GHz.

As described in Section 1, we propose to sidestep this problem using a novel modification to the traditional approach: rather than building a forecast model in the full embedding space, we will build it in τ-return map space. The power of this representation—essentially a reduced-order version of the Takens machinery—is that it brings out temporal relationships explicitly, via the second axis of the plot—without all of the complexity and subtlety that is involved in the full embedding process. If the system dynamics are deterministic, characteristic patterns appear on a τ-return map like Figure 4, and those patterns can be used to predict future values of the time series. That is, even though this 2D unfolding of the dynamics may not be accurate up to the diffeomorphism promised by the Takens theorem, we conjecture that it is nonetheless useful for the purposes of forecasting computer performance dynamics—and it has the huge advantage of having only a single free parameter that can be estimated 'on the fly' from the time series.

The example in Section 1 demonstrated the potential of this reduced-order modeling approach using a specific forecast method—the Lorenz method of analogues (LMA)—but any method that can capture the spatiotemporal patterns in a τ-return map could produce

[3]This process entails subtracting successive HPM readings, checking for overflow and adjusting accordingly.

[4]generally known only by this name, though it also rests heavily on work by Whitney and Mane

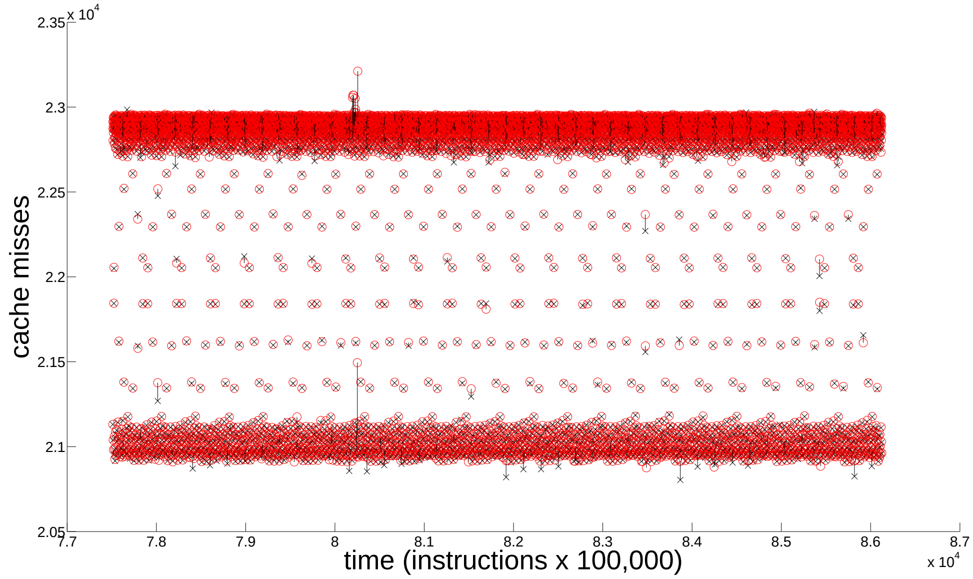
good forecasts. One of the challenges in meeting the research goals of this proposal will be to develop methods that work well in this space, and on this kind of data. Like most real-world data, computer performance traces are noisy. The characteristics of the noise somewhat unusual; it is not gaussian, but rather occasional and large—e.g., when some essential operating-system function intervenes. We found that LMA is (not surprisingly) very sensitive to noise, so we developed a novel modification of that algorithm that is more robust and more accurate. To predict the future time course of a given point, this algorithm finds the k nearest neighbors to that point, extracts their paths, and then averages them together to produce the forecast. For the time-series data in Figure 1, the forecast generated by this *k-ball LMA* algorithm has a root mean squared prediction error of 16.1—a substantial improvement on basic LMA (24.4 for the same trace and same period). See Figure 5(a) for a time-series plot of this forecast. LMA-based algorithms are only one possibility for modeling the geometry of a τ-return map of computer performance data; we will also evaluate the utility of algorithms from the machine-learning literature, such as neural nets or support-vector machines. These methods do not always work well on deterministic nonlinear dynamical systems [13], but the transformation into τ-return space might improve that situation. Finally, the PI's group is also working on topology-based techniques for identifying and removing the kind of 'shot' noise that appears in computer-performance data; we will evaluate those techniques for possible incorporation in our prediction algorithms as well [4].

Any discussion of new prediction technology is incomplete, of course, without a solid comparison to traditional techniques like autoregressive/moving-average models. One builds ARMA models by fitting the parameters of two series approximations—one for the AR part and one for the MA part—to the data, generally using a least-squares regression. The order of the model is defined by the number of terms in each series. Traditionally, one chooses it by plotting the autocorrelation versus the length of the time window spanned by the model's terms and seeing where that plot tails off. The autocorrelation of the data in Figure 1 remained high out to the full length of the time series, so we explored a wide range of orders for the AR and MA parts of the model by hand, applying each one to the first 90% of the trace in Figure 1 and calculating the RMSPE between true and predicted values of the rest of that signal. The best results were obtained with ten terms in each part of the model—i.e., a (10,10) ARMA model. This forecast is shown in Figure 5(b). Above that order, the error did not improve:

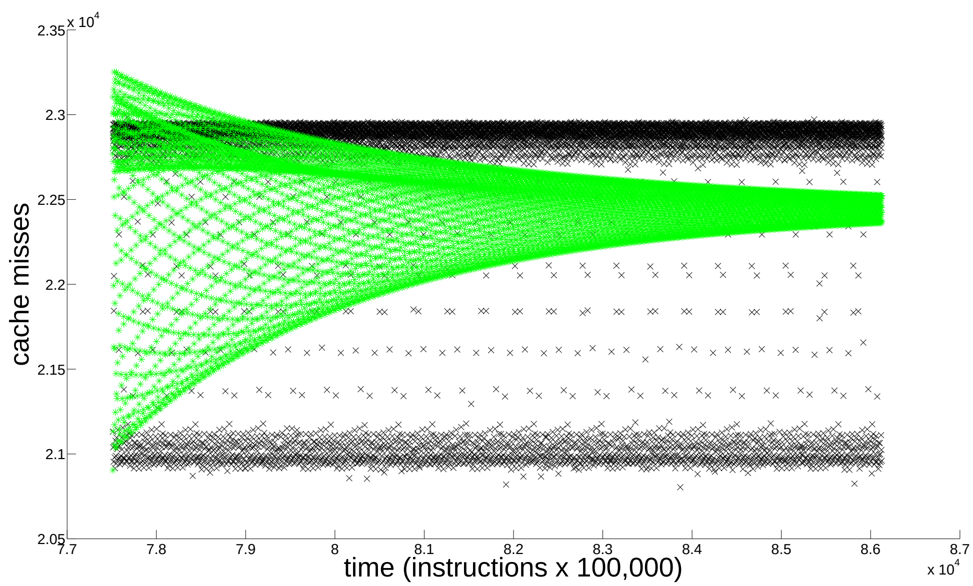
Order(p,q)	1,1	1,5	1,10	5,1	5,5	5,10	10,1	10,5	10,10	20,20
RMSPE	775.17	775.1	774.8	774.5	629.0	742.3	773.2	770.7	597.7	597.7

For comparison, the LMA and k-ball LMA methods, operating on a τ-return map of the same data, produced forecasts with RMSPE values of 24.4 and 16.1, respectively. That is, the error values for the ARMA forecasts—a respected traditional approach to modeling scalar time-series data like this—are at least two orders of magnitude worse than the results of the LMA-based algorithms acting on τ-return maps of the same data, the approach proposed in this document. ARMA models also share the main shortcoming of delay-coordinate embedding-based models: they are time consuming to build and thus problematic in the context of real-time prediction of complex, nonstationary dynamics.

There is one very important loose end remaining here: how to estimate τ. Computer



(a)



(b)

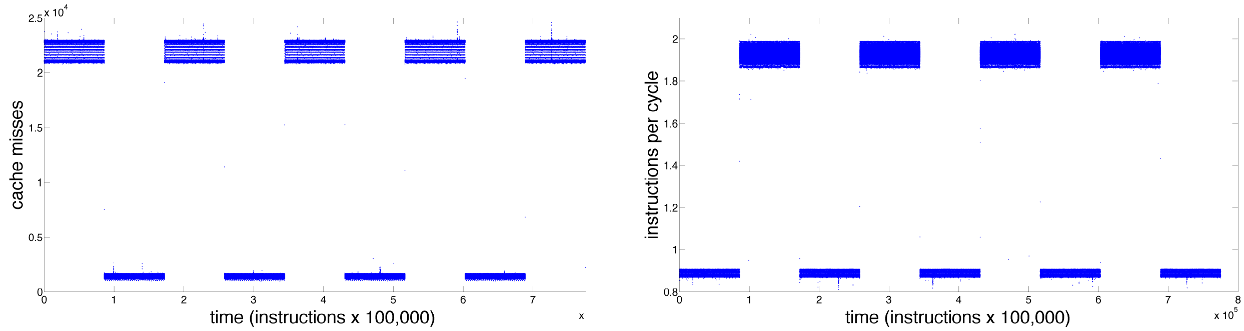
Figure 5: Forecasts of the last 10% of trace in Figure 1 with (a) a k-ball version of the nearest-neighbor prediction algorithm and (b) an order (10,10) ARMA model, both built from the first 90% of the signal. RMSPE values are 16.1 and 597.7, respectively. Black $\times$ points are the true values.

dynamics are highly nonstationary, so we will have to "learn" (and adapt) τ on the fly in order to build good τ-return maps with which to model them. The Takens theorem only requires $\tau > 0$ for a successful embedding; in practice, however, noise and numerical precision place more constraints on this value. One way to estimate τ from experimental data is to calculate some sort of correlation statistic on pairs of τ-separated points in the time series—e.g., average mutual information—and vary τ. For very small τ, these statistics will be close to 1.0, since successive points are very close to one another. For larger τ, successive points become increasingly uncorrelated. The first minimum in the distribution is a sensible choice for τ: qualitatively, it corresponds to the smallest τ for which the dynamics has caused nearby trajectory points to become *somewhat* uncorrelated (i.e., new information has been introduced between samples) [11]. That is how we chose τ for all of the models in this proposal, but that kind of *post facto* analysis is impractical in the context of real-time prediction.

Even though we are not using the full delay-coordinate embedding process and thus are not bound by the Takens requirements, we still need to find a τ that brings out the dynamical patterns effectively on a τ-return map. The reasoning outlined in the previous paragraph will guide our initial efforts in this regard: we will calculate average mutual information versus τ from the last l points measured by the HPMs and choose the first minimum. Choice of the window size l for this calculation is an obvious research issue here: longer windows will make the calculation less responsive to dynamical changes, but smaller windows (viz., less data) will make for lower-quality calculations. We will then turn to other measures—e.g., choosing the τ value that maximizes the information entropy or Shannon information [29] between τ-separated points in a similar window. It may also be possible to use Kalman filters [16] to choose τ, much as the data assimilation community does to estimate parameters for weather models. The topology-based signal separation methods mentioned above [4] may also prove to be useful in adapting τ on the fly, since they can detect changes between different dynamical regimes, an effect discussed in the following section. Finally, note that while methods like singular value decomposition or principal component analysis might appear useful here—viz., finding the two most meaningful axes of a delay vector—they are not practical in a real-time application. Worse yet, their truncation (removing the other axes) is not invertible, so transforming the prediction back to the time domain is problematic.

2.2 Model-Based Control of Computer Performance

There are legions of creative tactics for improving the performance of stored-program computers. Computer architects effectively speed up memory access, for instance, using a multi-level design: a large main memory, plus one or more levels of *cache*, small, fast memories that store copies of any main-memory data that is likely to be needed during a program's execution. When a piece of data is needed by a processor, it first checks the cache. If the data is available there, it is loaded and used; this is referred to as a *cache hit*. Otherwise, the data must be fetched from main memory, a slow operation that is referred to as a *cache miss*—the performance metric plotted in Figure 1. The key to multi-level memory effectiveness is to correctly anticipate what will be needed and "pre-fetch" it into the cache. Most of



(a) (b)

Figure 6: Cache misses (a) and instructions per cycle (b) in an Intel Core Duo® running a program that alternates between two loops that access memory very differently.

the design tactics used to accomplish this rely on *locality of reference*: the observation that a program that touches memory location k is likely to touch location $k+1$ soon thereafter. This works fairly well if programs step through memory in a linear fashion; if they do not, the code hangs up waiting for the data that it needs to proceed. This can not only slow things down, but is actually wasteful, since processors that are idle still burn power.

Another way that computer designers improve performance is by making more processor cycles available per unit time. Many modern computers have multiple processing units (or "cores"), for instance, among which the computation can be apportioned. The challenge in making this work is to determine a partial ordering on the "threads" of computation—which ones depends on which others, and which ones can be executed in parallel—and dynamically allocate threads to cores according to that partial ordering. Some of this can be done at compile time; this is an important aspect of modern compiler research. But some aspects of the partial ordering can only be determined at run time, since the interdependencies of the threads can be influenced by the program's input and other processes on the computer. Processor and memory performance are coupled, of course; when a piece of code is running efficiently, the cache-miss rate in the processor(s) is low and the number of instructions executed per cycle (IPC) is high. Figure 6 shows a pair of performance traces that demonstrate this effect. The code involved in this experiment alternates between two loops: one that initializes a matrix in column-major order and another that initializes the same matrix in row-major order. Row-major memory-access patterns are a natural match to cache design, so there are few cache misses during those phases of the code and the IPC is high. Traversing a matrix in column-major order works very badly with cache design, so those phases of this code exhibit high cache-miss rates and low IPC.

These traces bring out another extremely important issue: computer performance dynamics are not stationary. The trace in Figure 1 was produced by a simple microkernel that has a single dynamical attractor, but real programs are far more complicated—far more so, even, than Figure 6. The framework proposed in [23] gives us a useful way to think about this from a dynamical systems standpoint: the dynamics of the computer—F_P in equation (2)—undergoes a series of bifurcations as the execution moves through different parts of the code (F_{code}). (The computer performance community sometimes refers to these dynamical shifts as "phases."[14, 30].) A model of one attractor will obviously not be useful in predicting the evolution of a trajectory on a different attractor; this is exactly why we need the adaptive τ-estimation algorithms described in the previous section. The topology-based signal separation methods that our group is developing [4]—which model changes in continuity in order to detect these dynamical shifts—could be a useful addition to these methods: i.e.,, to quickly flag when it is necessary to re-estimate.

One of the standard design tradeoffs in engineering is that more speed generally requires more power. *Our goal is to use reduced-order nonlinear forecast models to finesse that tradeoff: that is, to reduce power use without impacting time to completion.* The key in doing so, as in many engineering applications, is to be creative about tailoring the resources to the loads, and vice versa. To first approximation, the power used by an active core is proportional to its clock frequency[5]. That is, the cores in a processor chip use power even when they are not executing instructions unless one puts them into a 'halt' state or turns down their clock rate. Halting a core, which is accomplished by sending it a 'halt' instruction, can reduce its power use by an order of magnitude or more. This strategy is used widely in modern operating system software. Clock frequency provides even more leverage. On many modern processors, the clock speed of each core can be controlled individually by the "kernel" of the operating system software. This opportunity, which has not been exploited beyond the simple strategies described in the second paragraph of this proposal (Speedstep *et al.*), is a real opportunity for model-based control. If one could predict that the overall IPC level would be 0.2-0.8 for the next few million clock cycles, for instance, one could sharply scale back the clock rates of all but one of the cores in the machine. And if one could predict that a particular computational thread would be bogged down for the next 0.6 seconds waiting for data from the computer's main memory, one could migrate it for that period to a core whose clock speed is scaled back. Cores, halts, and clocks are not the only opportunity here; if one could predict that none of the computations on the processor would need data from the hard drive for the next several seconds, one could save power by spinning down that drive. Putting these kinds of ideas into physical practice in a running computer is the end goal of this project.

Our first round of experiments will employ a Hewlett-Packard Pavilion Elite computer whose Intel Core i7-2600 processor (code named Nehalem) has four physical cores. This processor is an ideal testbed for this research because of its "Dynamic scalability, managed cores, threads, cache, interfaces, and power for energy-efficient performance on-demand [1]." We will instrument the kernel of its operating system software[6] to use the chip's performance monitoring facilities to capture cache-miss rates from each thread, run the application code, and gather the time-series data. We will set aside the second half of each data set, build a $\tau-$return map from the first half, apply the prediction algorithm (LMA, k-ball LMA, etc.) to capture the geometry of that map, use the results to create the forecast, and then calculate the RMSPE between the forecast and the second half of the signal. We will begin with simple microkernels like the one that produced Figure 1, then move to real-world computing applications with complex, nonstationary dynamics, such as the SPEC benchmark suite [2].

[5]Power is also proportional to the supply voltage, but clock frequency is easier to manipulate, so we will focus on it for the purposes of this project.

[6]Linux Ubuntu version 11.04, with the 2.6.38-8 kernel.

Milestones for this part of the project include hardware/software tools for gathering data from this machine, as well as methods for estimating τ, algorithms for making forecasts on τ-return maps, and a collection of forecast results (with RMSPEs) for a substantive range of computing applications. This effort will occupy years one and two of this project and will be documented by a paper in the dynamics literature.

After these modeling strategies have been worked out and tested, we will move towards implementation by estimating the power savings that would accrue if the kernel were to use the associated predictions to migrate memory-bound threads to lower-speed, lower-power cores. This will require research into the thread migration costs and timing in the Nehalem processor, as well as the power usage of cores at different clock frequencies. These estimates will be the primary milestones for this second phase of the project. This effort will occur in year two of the timeline; at the end of this phase, we hope to publish a paper in the computer systems literature about the overall strategy and its potential power savings.

The third phase of this project will involve understanding the Nehalem processor's dynamic clock scaling facilities and then using those facilities to put our model-based control strategies into physical practice. This will entail rebuilding the kernel to incorporate code that instantiates our modeling, forecast, and control algorithms. Working with the monitoring facilities that we will have built into the kernel in the first year of the project, these algorithms will produce predictions of critical processor metrics and take appropriate action based on those predictions—e.g., scaling back clock frequencies of different cores based on projected cache-miss rates and IPC values, migrating stalled threads to cores with lower clock speeds, and so on. We will compute the actual power savings that result from these actions, then repeat the whole exercise on at least one and hopefully two or three other computers, with the intent of broadening the exploration across the range of modern computer architecture. These efforts will take place during years two and three of the project; the milestones involved are the power savings estimates, another paper in the computer systems literature, and successful defense of Mr. Joshua Garland's Ph.D. dissertation.

Results of Prior NSF Support

The PI and her group have been supported since 2007 by an NSF grant entitled"CSR—SMA: Validating Architectural Simulators Using Non-Linear Dynamics Techniques (30 July 2007 through 31 August 2010 with a no-cost extension into 2012; $577730 plus $15K REU supplement; co-PIs: E. Bradley and A. Diwan; NSF project #1543868). Under the aegis of that grant, we developed the nonlinear dynamics-based framework for modeling and analyzing computer systems that is mentioned at several points in this proposal. Working with that framework, we were able to develop and deploy a custom measurement infrastructure and delay-coordinate embedding for studying the dynamics of these complex nonlinear systems. We found strong indications, from multiple corroborating methods, of low-dimensional dynamics in the performance of a simple program running on a popular Intel microprocessor—including the first experimental evidence of chaotic dynamics in real computer hardware. We also found that the dynamics changed completely when we ran the same program on a different Intel microprocessor, or when that program was changed slightly. This not only validated our framework; it also raised important issues about computer analysis and design—some of which led to the current proposal.

The core of the vertically integrated team that carried out this research included four undergraduates, three Ph.D students, two M.S. students, and three faculty members from two departments[7]. This group represents substantial current and future contributions to the development of human resources in science and engineering. Ph.D. student Todd Mytkowicz defended his thesis on this work in the summer of 2010 and took a position at Microsoft Research in Redmond. M.S. students Stephen Heck and David Cheeseman graduated and joined Sandia National Labs and the Navy, respectively. REU student Amber Roche is now in graduate school; the other three undergrads are still engaged with this project, as are the two Ph.D. students besides Dr. Mytkowicz.

This team has published two peer-reviewed journal and conference papers on this work [5, 23], plus a tech report [14] and Mytkowicz's Ph.D. thesis [24]. One more journal paper in review at the time of this submission [4]. Three grant proposals have been nucleated by this project, along with several new collaborations with researchers at the University of Colorado, University of Massachusetts, Google, and Microsoft Research. We are still publishing our original results, so we have not yet made the raw or processed data available on the web.

The broader impacts of this work are both technical and educational. In showing that computers are deterministic nonlinear dynamical systems, we alerted the computer systems community to the need to use better mathematical tools. By demonstrating the success of nonlinear time-series analysis on computer performance data, we offered some prospects for those better tools. Both PIs incorporated this research into their classrooms in meaningful ways; indeed, that was what attracted Mr. Garland, the Ph.D. student who will work on these prediction algorithms, to the project. This extends beyond University of Colorado; PI Bradley developed homework and projects based on this work for the students in the Santa Fe Institute Complex Systems Summer School. Again, that has played a formative role in the research trajectories of several graduate students.

Please note that this proposal is *not for renewed support of the line of research covered by project #1543868.* This is a wholly new project that was catalyzed by (and builds upon) that previous work.

3 Conclusion

The reduced-order forecast models and model-based control strategies proposed in this document, if successful, could improve power management in modern multi-core microprocessors. Current approaches to this important problem use simple, non-adaptive control strategies and do not involve any kind of model that captures the nonlinear, nonstationary dynamics of modern computers. One exception to this is [10], which uses a regression—but a linear one—to predict time on/off board for a particular process. While computer architects do realize the effects of nonlinearity and nonstationarity, they lack a truly principled framework for dealing with their effects. As a partial solution, the field has employed knowledge-based approaches (e.g. [7, 15, 22, 27]) and techniques grounded in search and operations research (e.g. [9, 18, 21]). Even so, the outcome of a new design feature routinely surprises its creators—sometimes in very unpleasant ways [6]. The line of research proposed here advances the

[7] Two other graduate students and one other undergradute worked on the project for shorter periods of time without making meaningful contributions.

state of the art in this area by using the *time history* of an important variable to *adaptively* build *nonlinear* forecast models that accurately capture that variable's *dynamics*—in ways that static and/or linear models simply cannot.

The power of the approach proposed here derives from its use of nonlinear dynamics methods, but this research project moves well beyond straightforward application of those methods to the problems of another field. Rather, we have proposed a novel reduced-order modeling strategy that builds upon traditional nonlinear dynamics methods: geometric forecast methods working in a 2D version of the delay-coordinate embedding space. Our preliminary results suggest that this approach captures the dynamics well enough to enable effective prediction, even though the τ-return map trajectory may not be completely faithful to the underlying topology of the dynamics. The advantage of this is that this 2D map has only a single free parameter that, we believe, can be estimated 'on the fly' from the data. The regular delay-coordinate embedding process, in contrast, has a second free parameter (the dimension) that requires expert human judgement to estimate, making that method all but useless for adaptive modeling and control in the real-time environment of a modern computer. All of this speaks directly to the words in the solicitation to which this proposal responds: *"When the phenomena are complicated, a compromise must be found between completeness and simplicity..."* Of course, taking action based on a prediction changes the dynamics, and changing clock frequencies alters time scales. This is another reason why continually re-estimating τ and rebuilding the τ-return map will be so critical to the success of our proposed scheme.

This project's leadership, interdisciplinary nature, and compelling focus should make it appealing to women and members of other under-represented groups. It is led by a senior woman with a diverse, vertically integrated research group and longstanding commitment to outreach and women/STEM issues. It applies cutting-edge mathematics to an interesting problem regarding the one of the most ubiquitous, critical engineered systems of our time. These features will also make for compelling classroom learning, in both computer science and mathematics courses, at the graduate and undergraduate levels. As in our prior NSF-supported work, we will incorporate this research into the classroom in meaningful ways: not only topical presentations, but use of the data and models in assignments and availability of the hardware and software for student projects. The research training provided by this project will be highly interdisciplinary, spanning the levels from undergraduates to the end-stage Ph.D. level. As mentioned above, the research team on the previous project included four undergraduates, three Ph.D. students, two M.S. students, and three faculty members from Computer Science and Applied Mathematics. We have the same model in mind for this project, though the intellectual focus is tighter and only one Ph.D. student is involved.

This research project not only represents compelling opportunities for education and diversity, and a new interdisciplinary research approach to an important engineering problem. Because computers are so common and critical, this work has the potential to contribute broadly to many applications in science, engineering, and beyond—wherever computers are used. And the reduced-order modeling strategies proposed here could be used to improve the designs of other systems as well—in computers and in other engineered systems.

References

[1] `www.intel.com/technology/architecture-silicon/next-gen`.

[2] `www.spec.org/benchmarks.html`.

[3] A. Alameldeen and D. Wood. IPC considered harmful for multiprocessor workloads. *IEEE Micro*, 26:8–17, 2006.

[4] Z. Alexander, E. Bradley, J. Garland, and J. Meiss. Iterated function system models in data analysis: Detection and separation. *CHAOS*. In review; preprint available as Technical Report CU-CS-1087-11 (`www.cs.colorado.edu/publications`) and on `arXiv`.

[5] Z. Alexander, T. Mytkowicz, A. Diwan, and E. Bradley. Measurement and dynamical analysis of computer performance data. In N. Adams, M. Berthold, and P. Cohen, editors, *Advances in Intelligent Data Analysis IX*, volume 6065. Springer Lecture Notes in Computer Science, 2010.

[6] Z. Ansari. Intel Corporation, Personal Communication, 2009.

[7] M. Arnold, S. Fink, D. Grove, M. Hind, and P. Sweeney. Adaptive optimization in the Jalapeno JVM. In *Proceedings of the ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)*, pages 47–65, 2000.

[8] M. Casdagli and S. Eubank, editors. *Nonlinear Modeling and Forecasting*. Addison Wesley, 1992.

[9] J. Cavazos and J. Moss. Inducing heuristics to decide whether to schedule. In *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, 2004.

[10] K. Choi, R. Soma, and M. Pedram. Fine-grained dynamic voltage and frequency scaling for precise energy and performance trade-off based on the ratio of off-chip access to on-chip computation times. In *Proceedings of the Conference on Design, Automation and Test in Europe*, page 10004, 2005.

[11] A. Fraser and H. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2):1134–1140, 1986.

[12] A. Georges, D. Buytaert, and L. Eeckhout. Statistically rigorous Java performance evaluation. In *Proceedings of the ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)*, pages 57–76, 2007.

[13] N. Gershenfeld and A. Weigend. The future of time series. In *Time Series Prediction: Forecasting the Future and Understanding the Past*. Santa Fe Institute Studies in the Sciences of Complexity, Santa Fe, NM, 1993.

[14] S. Heck. Detecting phases in cache-miss traces with wavelets. Technical Report CU-CS (Department of Computer Science) 1071-10, University of Colorado, 2010.

[15] D. Jimenez. Code placement for improving dynamic branch prediction accuracy. In *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, 2005.

[16] R. Kalman. Phase-plane analysis of automatic control systems with nonlinear gain elements. *Transactions of the AIEE*, 73:383, 1955.

[17] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, 1997.

[18] D. Knights, T. Mytkowicz, P. Sweeney, M. Mozer, and A. Diwan. Blind optimization for exploiting hardware. In *Proceedings of the International Conference on Compiler Construction*, 2009.

[19] A. Lebeck, J. Koppanalil, T. Li, J. Patwardhan, and E. Rotenburg. A large, fast instruction window for tolerating cache misses. In *Proceedings of the International Symposium on Computer Architecture (ISCA)*, pages 59–70, 2002.

[20] E. N. Lorenz. Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, 26:636–646, 1969.

[21] H. Massalin. Superoptimizer: A look at the smallest program. In *Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, pages 122–126, 1987.

[22] S. McFarling. Procedure merging with instruction caches. *SIGPLAN Notices*, 26:71–91, 1991.

[23] T. Myktowicz, A. Diwan, and E. Bradley. Computers are dynamical systems. *Chaos*, 19:033124, 2009. doi:10.1063/1.3187791.

[24] T. Mytkowicz. *Supporting experiments in computer systems research*. PhD thesis, University of Colorado, November 2010.

[25] S. Nussbaum and J. Smith. Modeling superscalar processors via statistical simulation. In *Proceedings of the 2001 International Conference on Parallel Architectures and Compilation Techniques (PACT)*, pages 15–24, 2001.

[26] N. Packard, J. Crutchfield, J. Farmer, and R. Shaw. Geometry from a time series. *Physical Review Letters*, 45:712, 1980.

[27] K. Pettis and R. C. Hansen. Profile guided code positioning. In *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, pages 16–27, 1990.

[28] T. Sauer, J. Yorke, and M. Casdagli. Embedology. *Journal of Statistical Physics*, 65:579–616, 1991.

[29] C. Shannon. *The Mathematical Theory of Communication*. University of Illinois Press, 1964.

[30] T. Sherwood, E. Perelman, G. Hamerly, and B. Calder. Automatically characterizing large scale program behavior. In *Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, pages 45–57, 2002.

[31] L. Sonneborn and F. Van Vleck. The bang-bang principle for linear control systems. *SIAM J. Control*, 2:151–159, 1965.

[32] F. Takens. Detecting strange attractors in fluid turbulence. In D. Rand and L.-S. Young, editors, *Dynamical Systems and Turbulence*, pages 366–381. Springer, Berlin, 1981.

[33] M. Vachharajani, N. Vachharajani, D. Penry, J. Blome, S. Malik, and D. August. The Liberty simulation environment: A deliberate approach to high-level system modeling. *ACM Trans. Comput. Syst.*, 24:211, 2006.

[34] A. Weigend and N. Gershenfeld, editors. *Time Series Prediction: Forecasting the Future and Understanding the Past*. Santa Fe Institute Studies in the Sciences of Complexity, Santa Fe, NM, 1993.

Biographical Sketch

Elizabeth BRADLEY

a. Professional Preparation: M.I.T.; S.B. in Electrical Engineering (1983), S.M. in Computer Science (1986), Ph.D. in Electrical Engineering and Computer Science (1992).

b. Appointments:

• Fellow, Radcliffe Institute for Advanced Study, AY 2006/2007.
• Professor, University of Colorado, Department of Computer Science, May 2004 to present. Associate Professor, August 1999 to May 2004. Assistant Professor, January 1993 to August 1999. Chair, July 2003 to March 2006. Joint appointment in the Department of Electrical and Computer Engineering.
• Visiting Scholar, Harvard University, Division of Engineering and Applied Sciences, Spring 1997 and AY 1999/2000.

c. Publications: `www.cs.colorado.edu/~lizb/publications.html`

- **most closely related to this proposal:**
  - Z. Alexander, T. Mytkowicz, A. Diwan and E. Bradley, "Measurement and Dynamical Analysis of Computer Performance Data," *IDA-10 (Proceedings of the 9th International Symposium on Intelligent Data Analysis)*, Tucson; May 2010; Springer *Lecture Notes in Computer Science* volume 6065
  - E. Bradley, "Time-Series Analysis," in M. Berthold and D. Hand, editors, *Intelligent Data Analysis: An Introduction*, Springer-Verlag, 2000; second edition, 2003
  - T. Mytkowicz, E. Bradley, and A. Diwan, "Computers Are Dynamical Systems," *Chaos* **19**:033124 (2009); doi:10.1063/1.3187791
  - V. Robins, J. Meiss, and E. Bradley, "Computing Connectedness: Disconnectedness and Discreteness," *Physica D* **139**:276-300 (2000)
  - R. Stolle, M. Easley, and E. Bradley, "Reasoning about Models of Nonlinear Systems," in *Logical and Computational Aspects of Model-Based Reasoning*, L. Magnani *et al.*, eds. Kluwer, 2002
- **other significant:**
  - E. Bradley, M. Easley, and R. Stolle, "Reasoning About Nonlinear System Identification," *Artificial Intelligence* **133**:139-188 (2001)
  - R. Hoenigman, A. Lim, and E. Bradley, "Cooperation in Bike Racing: When to Work Together and When to Go It Alone," *Complexity* DOI: 10.1002/cplx.20372 (2011)
  - T. Peacock and E. Bradley, "Going with (or Against) the Flow," *Science* **320**:1302-1303 (2008)
  - V. Robins, J. Meiss, and E. Bradley, "Computing Connectedness: An Exercise in Computational Topology," *Nonlinearity*, **11**:913-922 (1998)
  - N. Ross, J. Hertzberg, and E. Bradley, "Discretization of the Vorticity Field of a Planar Jet," *Experiments in Fluids* **49**:1161 (2010)

d. Synergistic Activities:

- I am consistently involved in research fairs, professional panels, industrial liaison activities, and the planning and execution of other events designed to encourage high-school, women, and minority students in their pursuit of science and engineering.

- I serve on the Science Board of the Santa Fe Institute, and I teach at their summer school every year to a broad audience from across the country and across all disciplines of science, engineering, and the social sciences. In 1999, I received the yearly student-voted College of Engineering "Innovation in Teaching" Award for my work in coupling research and undergraduate teaching.
- I collaborate with students and faculty in eight departments, ranging from Astrophysics to Dance. My students and I present the resulting work at conferences that cross the traditional discipline boundaries (e.g., the *International Dance and Technology* workshop, where we presented a paper on the mathematics of movement sequences).
- I am an editor of *Chaos*, the American Institute of Physics's interdisciplinary journal of nonlinear science, and I review articles and programs for a broad selection of journals, conferences, and organizations: *Artificial Intelligence*, *Info. Proc. Lett.*, *Int. J. of Par. Prog.*, *Consciousness & Cognition*, *IEEE Control Systems Mag.*, *IEEE Trans. Auto. Control*, *Physica D*, the *IEEE American Control Conference*, *IEEE Trans. Circuits and Systems*, *Computers and Electrical Engineering*, *Physical Review*, *Phys. Rev. Lett.*, *Phys. Lett. A*, *Geophys. Res. Lett.*, and the International (*IJCAI*) and National(*AAAI*) conferences on Artificial Intelligence, as well as the US Geological Survey and the National Science Foundation.
- As department chair, two of my primary goals were to revamp the undergraduate teaching laboratories and to institute a junior-faculty mentoring program. Both projects were judged successful by our external review committee in the spring of 2006.

e. Collaborators & Other Affiliations:

- **collaborators and co-editors:** Niall Adams, University College London; Ken Anderson, Computer Science, UColorado; Michael Berthold, University of Konstanz; David Campbell, Boston University; David Capps, Hunter College; Paul Cohen, University of Arizona; John Clyne, NCAR; Amer Diwan, University of Colorado; Jerome Dorignac, Boston University. Mark Dubin, University of Colorado; Jean Hertzberg, Mechanical Engineering, UColorado; Alain Karma, Northeastern University; Jeffrey Luftig, University of Colorado; Pablo Mininni, NCAR; Thomas Peacock, MIT; Rob Sharman, NCAR; Marek Zreda, Hydrology, UArizona; Chris Zweck, Hydrology, UArizona.
- **graduate and postdoctoral advisors:** H. Abelson, MIT; G. J. Sussman, MIT.
- **thesis advisor and postgraduate-scholar sponsor:**
  - *undergraduate research/thesis students:* Patrick Clary (Ph.D. student, Carnegie Mellon), Matt Culbreth (Ph.D. student, Stanford), Eric Horacek, Apollo Hogan (Ph.D. student, Berkeley), Connor Janowiak, Asim Khwaja (Professor, Kennedy School of Govt, Harvard), Sebastian Kuzminsky (UColorado BioServe Center), Nikki Look, Aaron Sheppard, Josh Stuart (Professor, UC Santa Cruz).
  - *graduate students:* Jenny Abernethy (NCAR), Zach Alexander, Stephanie Boyles (local industry), Jon Dixon (Lincoln Labs), Matt Easley (Rockwell), Joshua Garland, James Garnett (Secure64), John Giardino (Tennessee Valley Infrastructure Group), Kenny Gruchalla (NREL), Stephen Heck (Sandia), Rhonda Hoenigman, Joe Iwanski (Dwight Englewood School), Zhichun Ma (Korean University), Jon Marbach, Todd Mytkowicz (Microsoft Research), Tom Nelson, Laura Rassbach (family leave), Vanessa Robins (Australian Natl U.), Natalie Ross (IBM), Matt Schwieterman (local industry), Reinhard Stolle (BMW Research), Doug Straub (local industry), Elizabeth White (CU Medical School). Total number of theses supervised: seven masters and 15 doctoral.
  - *postdoctoral associates within last five years:* Hyoung-Bum Kim (Professor, Gyeongsang National University), Tom Peacock (Professor, MIT). Total number supervised: six.

People whose affiliations are unspecified are still at CU.

SUMMARY PROPOSAL BUDGET

YEAR 1

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months)	
	Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Colorado at Boulder

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Elizabeth Bradley

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. Elizabeth Bradley - PI	0.00	0.00	1.00	15,216	
2.					
3.					
4.					
5.					
6. (0) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	0	
7. (1) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	1.00	15,216	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (0) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	0	
2. (0) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	0.00	0.00	0.00	0	
3. (1) GRADUATE STUDENTS				34,041	
4. (1) UNDERGRADUATE STUDENTS				3,960	
5. (0) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				0	
6. (0) OTHER				0	
TOTAL SALARIES AND WAGES (A + B)				53,217	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				6,780	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				59,997	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				0	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				2,000	
2. FOREIGN				3,000	
F. PARTICIPANT SUPPORT COSTS 1. STIPENDS $ 0 2. TRAVEL 0 3. SUBSISTENCE 0 4. OTHER 0					
TOTAL NUMBER OF PARTICIPANTS (0) TOTAL PARTICIPANT COSTS				0	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				11,000	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				0	
3. CONSULTANT SERVICES				0	
4. COMPUTER SERVICES				0	
5. SUBAWARDS				0	
6. OTHER				12,871	
TOTAL OTHER DIRECT COSTS				23,871	
H. TOTAL DIRECT COSTS (A THROUGH G)				88,868	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) On-Campus IDC Rate (Rate: 52.5000, Base: 75997)					
TOTAL INDIRECT COSTS (F&A)				39,898	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				128,766	
K. RESIDUAL FUNDS				0	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				128,766	

M. COST SHARING PROPOSED LEVEL $ 0 AGREED LEVEL IF DIFFERENT $

PI/PD NAME	FOR NSF USE ONLY		
Elizabeth Bradley	INDIRECT COST RATE VERIFICATION		
ORG. REP. NAME*	Date Checked	Date Of Rate Sheet	Initials - ORG
Parker Dougherty			

 *ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET

SUMMARY PROPOSAL BUDGET

YEAR 2

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months)	
	Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Colorado at Boulder

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Elizabeth Bradley

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. Elizabeth Bradley - PI	0.00	0.00	1.00	15,672	
2.					
3.					
4.					
5.					
6. (0) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	0	
7. (1) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	1.00	15,672	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (0) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	0	
2. (0) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	0.00	0.00	0.00	0	
3. (1) GRADUATE STUDENTS				35,062	
4. (1) UNDERGRADUATE STUDENTS				3,960	
5. (0) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				0	
6. (0) OTHER				0	
TOTAL SALARIES AND WAGES (A + B)				54,694	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				6,982	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				61,676	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				0	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				2,000	
2. FOREIGN				3,000	
F. PARTICIPANT SUPPORT COSTS 1. STIPENDS $ 0 2. TRAVEL 0 3. SUBSISTENCE 0 4. OTHER 0					
TOTAL NUMBER OF PARTICIPANTS (0) TOTAL PARTICIPANT COSTS				0	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				1,000	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				0	
3. CONSULTANT SERVICES				0	
4. COMPUTER SERVICES				0	
5. SUBAWARDS				0	
6. OTHER				13,515	
TOTAL OTHER DIRECT COSTS				14,515	
H. TOTAL DIRECT COSTS (A THROUGH G)				81,191	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) On-Campus IDC Rate (Rate: 52.5000, Base: 67676)					
TOTAL INDIRECT COSTS (F&A)				35,530	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				116,721	
K. RESIDUAL FUNDS				0	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				116,721	

M. COST SHARING PROPOSED LEVEL $ 0 | AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Elizabeth Bradley

ORG. REP. NAME*
Parker Dougherty

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 *ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET

SUMMARY PROPOSAL BUDGET

YEAR 3

ORGANIZATION **University of Colorado at Boulder**				FOR NSF USE ONLY PROPOSAL NO.	DURATION (months) Proposed	Granted
PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR **Elizabeth Bradley**				AWARD NO.		
A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)	
1. **Elizabeth Bradley - PI**	0.00	0.00	1.00	**16,142**		
2.						
3.						
4.						
5.						
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**		
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	1.00	**16,142**		
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)						
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**		
2. (**0**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	0.00	0.00	0.00	**0**		
3. (**1**) GRADUATE STUDENTS				**36,114**		
4. (**1**) UNDERGRADUATE STUDENTS				**3,960**		
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**		
6. (**0**) OTHER				**0**		
TOTAL SALARIES AND WAGES (A + B)				**56,216**		
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**7,190**		
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**63,406**		
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)						
TOTAL EQUIPMENT				**0**		
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**2,000**		
2. FOREIGN				**3,000**		
F. PARTICIPANT SUPPORT COSTS 1. STIPENDS $ **0** 2. TRAVEL **0** 3. SUBSISTENCE **0** 4. OTHER **0**						
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**		
G. OTHER DIRECT COSTS						
1. MATERIALS AND SUPPLIES				**1,000**		
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**		
3. CONSULTANT SERVICES				**0**		
4. COMPUTER SERVICES				**0**		
5. SUBAWARDS				**0**		
6. OTHER				**14,191**		
TOTAL OTHER DIRECT COSTS				**15,191**		
H. TOTAL DIRECT COSTS (A THROUGH G)				**83,597**		
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **On-Campus IDC Rate (Rate: 52.5000, Base: 69406)**						
TOTAL INDIRECT COSTS (F&A)				**36,438**		
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**120,035**		
K. RESIDUAL FUNDS				**0**		
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				**120,035**		
M. COST SHARING PROPOSED LEVEL $ **0**	AGREED LEVEL IF DIFFERENT $					
PI/PD NAME **Elizabeth Bradley**				FOR NSF USE ONLY INDIRECT COST RATE VERIFICATION		
ORG. REP. NAME* **Parker Dougherty**				Date Checked	Date Of Rate Sheet	Initials - ORG

 *ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET

SUMMARY PROPOSAL BUDGET

Cumulative

ORGANIZATION University of Colorado at Boulder				FOR NSF USE ONLY PROPOSAL NO.	DURATION (months) Proposed	Granted
PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR Elizabeth Bradley				AWARD NO.		
A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)	
1. Elizabeth Bradley - PI	0.00	0.00	3.00	47,030		
2.						
3.						
4.						
5.						
6. ( ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	0		
7. (1) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	3.00	47,030		
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)						
1. (0) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	0		
2. (0) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	0.00	0.00	0.00	0		
3. (3) GRADUATE STUDENTS				105,217		
4. (3) UNDERGRADUATE STUDENTS				11,880		
5. (0) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				0		
6. (0) OTHER				0		
TOTAL SALARIES AND WAGES (A + B)				164,127		
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				20,952		
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				185,079		
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)						
TOTAL EQUIPMENT				0		
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				6,000		
2. FOREIGN				9,000		
F. PARTICIPANT SUPPORT COSTS 1. STIPENDS $ 0 2. TRAVEL 0 3. SUBSISTENCE 0 4. OTHER 0						
TOTAL NUMBER OF PARTICIPANTS (0) TOTAL PARTICIPANT COSTS				0		
G. OTHER DIRECT COSTS						
1. MATERIALS AND SUPPLIES				13,000		
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				0		
3. CONSULTANT SERVICES				0		
4. COMPUTER SERVICES				0		
5. SUBAWARDS				0		
6. OTHER				40,577		
TOTAL OTHER DIRECT COSTS				53,577		
H. TOTAL DIRECT COSTS (A THROUGH G)				253,656		
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE)						
TOTAL INDIRECT COSTS (F&A)				111,866		
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				365,522		
K. RESIDUAL FUNDS				0		
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				365,522		
M. COST SHARING PROPOSED LEVEL $ 0	AGREED LEVEL IF DIFFERENT $					
PI/PD NAME Elizabeth Bradley				FOR NSF USE ONLY INDIRECT COST RATE VERIFICATION		
ORG. REP. NAME* Parker Dougherty				Date Checked	Date Of Rate Sheet	Initials - ORG

C *ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET

Budget Justification

Personnel

The bulk of this budget covers summer salary for the PI and tuition/stipend (AY and summer) for PhD student Joshua Garland, the author of the code that produced the figures in this proposal. The PI will guide the intellectual effort of the project. This includes managing and mentoring the research team, writing research papers, and disseminating the results of the project at technical conferences.

Finally, we have also requested funding for an undergraduate research assistant to participate in this project. The PI has a long history of involving undergraduates in meaningful ways in her work and launching all of her alumni into successful careers. This project will be no exception.

Travel

This research straddles two fields—nonlinear dynamics and computer systems—and the PI and PhD student will present the results of the proposed research at conferences in both research areas. To support this, funding has been requested to support travel to one domestic and one international conference per year of the project. (Since all project personnel are in Boulder, no travel funding is requested for project meetings.)

Other Direct Costs

We request $1,000 for materials & supplies in each year of the project. This will be used to cover page charges, books, computer supplies, and data-storage media.

We also request $10,000 to purchase computers and equipment to carry out the experiments. These machines will not be used for normal computing purposes, but rather reconfigured as subjects of study; their operating systems will be modified to support the kinds of data-gathering and control that are the goals of our research. We plan to buy at least two or three machines so as to effectively span the range of modern computer architecture.

Elizabeth Bradley

- Current Support
  - *Project/Proposal Title:* CSR—SMA: Validating Architectural Simulators Using Non-Linear Dynamics Techniques
    Source of Support: NSF
    Total Award Amount: 577730
    Total Award Period Covered: 7/30/2007-8/31/2012
    Location of Project: University of Colorado at Boulder
    Person-Months/Year Committed to the Project: Cal: 0 Acad: 0 Sumr: 0.5
  - *Project/Proposal Title:* REU Supplement to CSR—SMA: Validating Architectural Simulators Using Non-Linear Dynamics Techniques
    Source of Support: NSF
    Total Award Amount: 15000
    Total Award Period Covered: 7/30/2009-8/31/2012
    Location of Project: University of Colorado at Boulder
    Person-Months/Year Committed to the Project: Cal: 0 Acad: 0 Sumr: 0
  - *Project/Proposal Title:* Applications of artificial intelligence techniques to the computation of reachability sets
    Source of Support: University of Colorado Innovative Seed Grant Program
    Total Award Amount: 44000
    Total Award Period Covered: 7/1/2011-6/30/2012
    Location of Project: University of Colorado at Boulder
    Person-Months/Year Committed to the Project: Cal: 0 Acad: 0 Sumr: 0
- Pending Support
  - *Project/Proposal Title:* Autonomous Navigation and Planning in Extreme Orbital Environments
    Source of Support: NASA
    Total Award Amount: 100000
    Total Award Period Covered: 1/01/11-12/31/11
    Location of Project: University of Colorado at Boulder
    Person-Months/Year Committed to the Project: Cal: 0 Acad: 0 Sumr: 0.75
  - *Project/Proposal Title:* DynSyst-Special-Topics: Reduced-Order Dynamical Models for Effective Power Management in Computer Systems
    Source of Support: NSF
    Total Award Amount: 365522
    Total Award Period Covered: 6/1/12-5/31/15
    Location of Project: University of Colorado at Boulder
    Person-Months/Year Committed to the Project: Cal: 0 Acad: 0 Sumr: 1.0

FACILITIES, EQUIPMENT & OTHER RESOURCES

FACILITIES: Identify the facilities to be used at each performance site listed and, as appropriate, indicate their capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed and at sites for field studies. USE additional pages as necessary.

Laboratory: N/A

Clinical: N/A

Animal: N/A

Computer: Adequate shared computing resources (software, printers, hardware and software support, etc.) are available to this project through the Department of Computer Science.

Office: Office space for all project personnel is also available through the Department of Computer Science.

Other: With previous NSF support, the PI's research group has developed a significant amount of expertise regarding gathering traces of important computer-performance metrics and nonlinear time-series analysis of those data.

MAJOR EQUIPMENT: List the most important items available for this project and, as appropriate identifying the location and pertinent capabilities of each.

Our previous NSF grant (CSR---SMA: Validating Architectural Simulators Using Non-Linear Dynamics Techniques, NSF project #1543868) covered the purchase of a HP Pavilion Elite computer with a four-core Intel i7-2600 Nehalem processor, which we will use for our initial experiments. This machine is dedicated to this project, allowing major modifications to its operating systems software. The budget for this proposal includes funds for several other computer platforms with which to test our results.

OTHER RESOURCES: Provide any information describing the other resources available for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available for the project. Include an explanation of any consortium/contractual arrangements with other organizations.

Professor Amer Diwan, the co-PI on project #1543868, is on leave at Google but is available on a cost-free basis for consultation and collaboration on this project, as is Dr. Todd Mytkowicz, the lead author on our 2009 CHAOS paper, who is now at Microsoft Research.

Data Management Plan

This work straddles two fields—dynamical systems and computer systems—and we plan to publish conference and journal papers about our results in the literatures of both of those communities, as well as in the form of PhD (and hopefully BS) theses of the students involved in the project. All students involved in the project will share authorship on all publications.

The raw data created in the course of this project will consist of time-series traces of computer performance metrics—like Figure 1 in this proposal—captured, as described in Section 2 of that document, using the PAPI performance monitoring API software, which reads the hardware performance monitors on the processor chip.

The PI's group is currently using a shared Dropbox folder to house these raw data; once the research results have been published, these data will be moved to a permanent public website hosted on the University of Colorado Department of Computer Science's web server.

These data are two-column text files; they will be stored in ascii format with a metadata header that explains the experimental conditions that produced them, including:

- Date
- Author of the run
- Computer (e.g., `bit.cs.colorado.edu`)
- Processor (e.g., Intel Core i7-2600)
- Operating system information (what OS, what boot level, etc.)
- The code running on that system while the data were gathered
- Metric (e.g., L2 cache misses, instructions per cycle)
- Sampling interval (in instructions or cycles)
- PAPI information (what version of the software, compiled how, etc.)

Aside from this header information, the only metadata required for another person to duplicate or extend our results will be the modelling algorithms. We will describe them carefully enough in our papers that anyone could implement them, and also make the associated source code available on our website under standard open source licensing procedures (`www.opensource.org`).

There will likely be hundreds or thousands of traces and an infinite number of possible models that one could build from them (since at least some of the parameters of the modelling algorithms will be continuous variables). Generating and storing every possible combination of data and model would obviously be prohibitive. We have chosen instead to provide the raw data, the algorithms, and the source code in order to provide users with the flexibility to replicate any of our results—and hopefully to extend them.

There are no legal/ethical restrictions associated with any of this information. We do not plan to restrict access to it in any way, nor to file for patents on the results.

There are no special costs associated with this activity; the PI will create and host this website and the student researchers will archive the data. The PI will be responsible for the correctness and completeness of the information on this website. If she leaves the University of Colorado, she will make arrangements for the Department's webmaster to maintain this website until at least 2020.