THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY

Investigating Novel Hardware Architectures

Energy Efficient Computational Chemistry: x86 and ARM

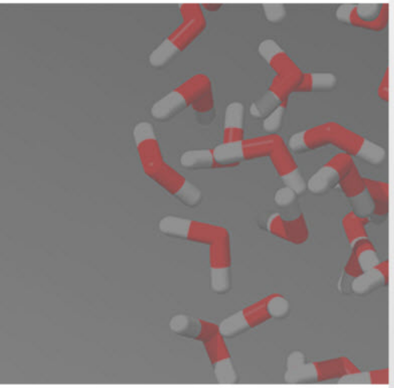
Automating Intrusion Detection and Response

A Spin-free [2]R12 basis set incompleteness correction to the Local Multi-Reference Configuration Interaction and the Local Multi-Reference Average Coupled Pair Functional methods

BACK TO MAP

Joint Frequency Scaling of Processor and DRAM

THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY



BACK TO MAP

Energy Efficient Computational Chemistry: x86 and ARM

Kristopher Keipert[a], Gaurav Mitra[b], Vaibhav Sunriyal[a], Sarom S. Leang[a],
Masha Sosonkina[c], Alistair Rendell[b], Mark S. Gordon[a]
[a]Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50010, United States
[b]Research School of Computer Science, Australian National University, Acton ACT 0200, Australia
[c]Department of Modeling and Simulation, Old Dominion University, Norfolk, VA 23529 United States

The performance and energy-to-solution of GAMESS workloads on two ARM-based systems, a 32-bit NVIDIA Jetson TK1 and a 64-bit APM Xgene1 X-C1 is compared to a 64-bit Haswell x86 Intel Xeon-E5 cpu. A set of commonly used computational chemistry techniques are evaluated, namely HF SCF, DFT (PBE0) and MP2 energy and gradient calculations.

Computational Details

The GAMESS performance evaluations used a benchmark set of molecules that contains 99-509 basis functions when using the 6-31G(d) basis set. Power measurements were obtained for DFT, HF SCF, MP2 energy, and MP2 gradient calculations. The PBE0 functional was used in all DFT calculations. In all benchmarks, two-electron integrals were calculated at each (direct) SCF step.

Hardware Details

The 32-bit ARM machine is an NVIDIA Jetson TK1, configured with a quad-core 2.35 GHz ARM Cortex-A15 CPU (ARMv7-A architecture) paired with 2 GB of LP-DDR3 RAM. The 64-bit ARM machine is an AppliedMicro (APM) X-Gene X-C1 with an 8-core 2.4 Ghz APM883208-X1 CPU (ARMv8-A architecture) and 16 GB of DDR3 memory. The Haswell x86 machine utilizes an 18-core Intel Xeon E5-2699 v3 CPU clocked to the maximum turbo frequency of 3.6 GhZ with hyperthreading disabled and 32 GB of DDR4 memory.

Computational Efficiency

The CPU wall clock times for the various methods on the different platforms are shown in Figure 1 (top right) normalized according to the number of basis functions in each molecule.

Energy Consumption

The energy consumption per basis function for the various methods on the different platforms are shown in Figure 2 (bottom right).

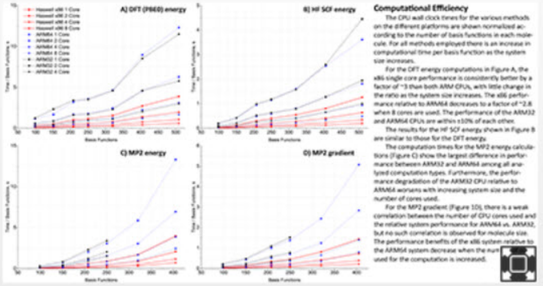


Figure 1. Computational Efficiency

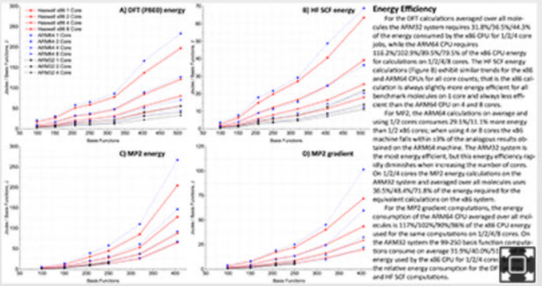
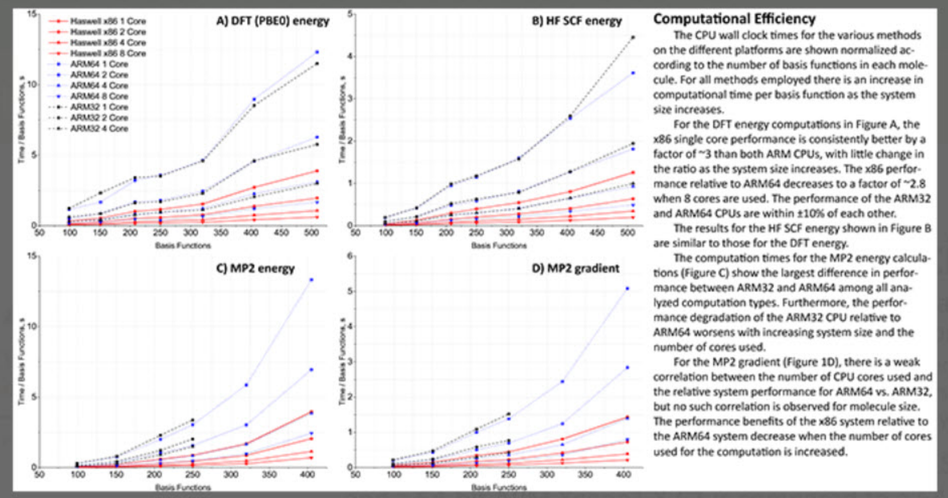


Figure 2. Energy Consumption

FUNDING & CREDITS



Computational Efficiency

The CPU wall clock times for the various methods on the different platforms are shown normalized according to the number of basis functions in each molecule. For all methods employed there is an increase in computational time per basis function as the system size increases.

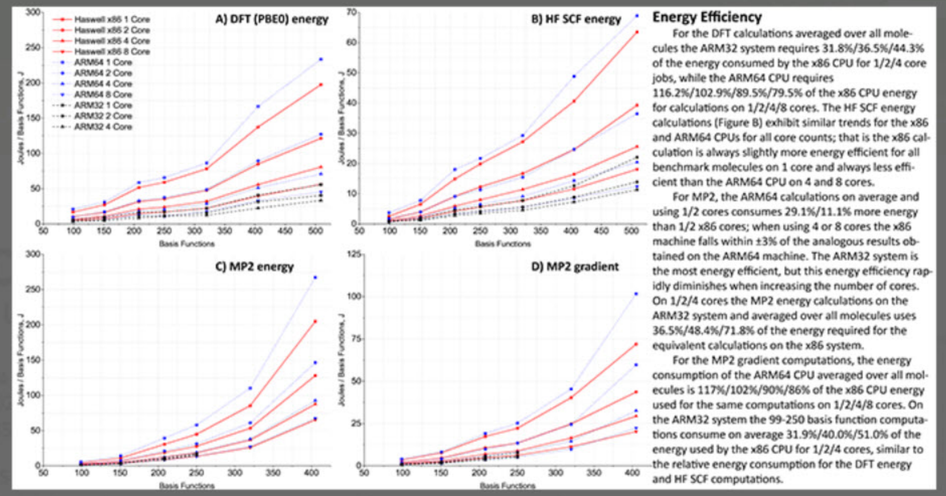
For the DFT energy computations in Figure A, the x86 single core performance is consistently better by a factor of ~3 than both ARM CPUs, with little change in the ratio as the system size increases. The x86 performance relative to ARM64 decreases to a factor of ~2.8 when 8 cores are used. The performance of the ARM32 and ARM64 CPUs are within ±10% of each other.

The results for the HF SCF energy shown in Figure B are similar to those for the DFT energy.

The computation times for the MP2 energy calculations (Figure C) show the largest difference in performance between ARM32 and ARM64 among all analyzed computation types. Furthermore, the performance degradation of the ARM32 CPU relative to ARM64 worsens with increasing system size and the number of cores used.

For the MP2 gradient (Figure 1D), there is a weak correlation between the number of CPU cores used and the relative system performance for ARM64 vs. ARM32, but no such correlation is observed for molecule size. The performance benefits of the x86 system relative to the ARM64 system decrease when the number of cores used for the computation is increased.

Computational Efficiency
The CPU wall clock times for the various methods on the different platforms, normalized according to the number of basis functions in each molecule.



Energy Efficiency

For the DFT calculations averaged over all molecules the ARM32 system requires 31.8%/36.5%/44.3% of the energy consumed by the x86 CPU for 1/2/4 core jobs, while the ARM64 CPU requires 116.2%/102.9%/89.5%/79.5% of the x86 CPU energy for calculations on 1/2/4/8 cores. The HF SCF energy calculations (Figure B) exhibit similar trends for the x86 and ARM64 CPUs for all core counts; that is the x86 calculation is always slightly more energy efficient for all benchmark molecules on 1 core and always less efficient than the ARM64 CPU on 4 and 8 cores.

For MP2, the ARM64 calculations on average and using 1/2 cores consumes 29.1%/11.1% more energy than 1/2 x86 cores; when using 4 or 8 cores the x86 machine falls within ±3% of the analogous results obtained on the ARM64 machine. The ARM32 system is the most energy efficient, but this energy efficiency rapidly diminishes when increasing the number of cores. On 1/2/4 cores the MP2 energy calculations on the ARM32 system and averaged over all molecules uses 36.5%/48.4%/71.8% of the energy required for the equivalent calculations on the x86 system.

For the MP2 gradient computations, the energy consumption of the ARM64 CPU averaged over all molecules is 117%/102%/90%/86% of the x86 CPU energy used for the same computations on 1/2/4/8 cores. On the ARM32 system the 99-250 basis function computations consume on average 31.9%/40.0%/51.0% of the energy used by the x86 CPU for 1/2/4 cores, similar to the relative energy consumption for the DFT energy and HF SCF computations.

Energy Consumption
The energy consumption per basis function for the various methods on the different platforms.

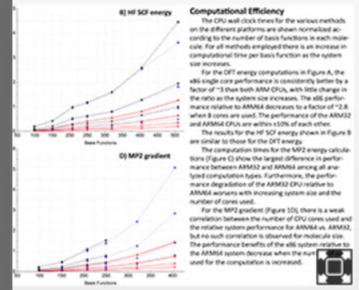


Figure 1. Computational Efficiency

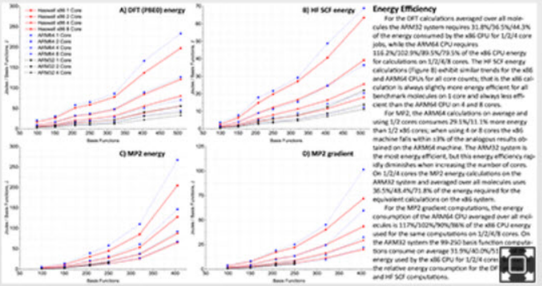


Figure 2. Energy Consumption

...onal ...
...a, Sarom S...
...ity, Ames, Iov...
...y, Acton ACT...
...Norfolk, VA 2...
...S workload...
...bit Haswel...
...namely H...

...of molecules that contains 99-509 basis functions when using the 6-31G(d) basis set. Power measurements were obtained for DFT, HF SCF, MP2 energy, and MP2 gradient calculations. The PBE0 functional was used in all DFT calculations. In all benchmarks, two-electron integrals were calculated at each (direct) SCF step.

Hardware Details

The 32-bit ARM machine is an NVIDIA Jetson TK1, configured with a quad-core 2.35 GHz ARM Cortex-A15 CPU (ARMv7-A architecture) paired with 2 GB of LP-DDR3 ... PM883208-X1 CPU (ARMv8-A architecture) ... CPU clocked to the maximum turbo freque...

Computational Efficiency

The CPU wall clock times for t... ...ording to the number of basis functions in ...

Energy Consumption

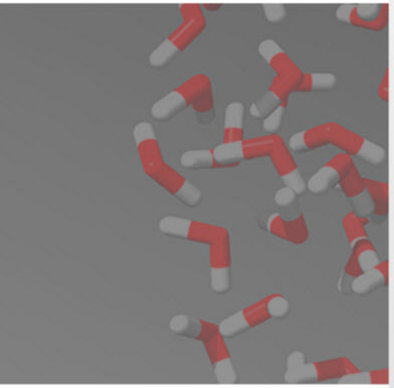
The energy consumption per basis function for the various methods on the different platforms are shown in Figure 2 (bottom right).

The authors thank Intel Corp. and NVIDIA Corp. for their support of this work.
K.K., V.S., S.L., M.S., and M.S.G. thank the Air Force Office of Scientific Research for their support of this work under AFOSR Award No. FA9550-12-1-0476.

BACK TO MAP

FUNDING & CREDITS

THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY



BACK TO MAP

Joint Frequency Scaling of Processor and DRAM

Vaibhav Sundriyal, Masha Sosonkina
Department of Modeling and Simulation, Old Dominion University, Norfolk, VA 23529 United States

To achieve sustainable exascale performance, capping power consumption of the hardware components is paramount. While many past studies have focused on reducing processor power consumption, DRAM has largely been ignored. However, depending on the workload characteristics, power saving can come from DRAM when processor gets saturated. Keeping this in mind, a runtime system was designed that would apply frequency scaling to both processor and DRAM based on the workload.

Runtime System Design

Based on the Intel RAPL server capabilities, a power model was proposed for processor and DRAM that would provide system power consumption at different processor and DRAM frequency combinations. Considering the out-of-order architecture of modern processor, a performance model was also designed to map the change in number of instructions executed per unit time with the variation in processor and DRAM frequency.

Runtime Algorithm

A timeslice based algorithm was designed that would profile the application into fixed duration timeslices, obtain the workload characteristics through the performance counters and then chooses the processor-memory frequency pair that would result in the minimum energy consumption, satisfying the user defined performance constraint. Three different frequency scaling strategies were proposed: Conservative, Moderate, and Aggressive, named so to reflect the degree of their DVFS application. They are distinguished mainly by varying the way of determining how the workload parameters are determined.

Experimental Platform

The experiments were performed on a computing platform that has an Intel i5-4590 “Haswell" quad-core processor and 8 GB of main memory. For measuring the node power and energy consumption, a Wattsup power meter is used with a sampling rate of 1 Hz. To validate the proposed energy-saving model, SPEC CPU 2006 benchmarks were used.

Performance

Fig. 1 depicts the performance loss resulting from the proposed runtime system, compared to the full-power execution in that the benchmarks are run at the highest available processor-memory frequency pair. (top right).

Energy Saving

Fig. 2 depicts the energy savings obtained from the proposed runtime system, with respect to the full-power execution for the seven SPEC benchmarks. (bottom right)

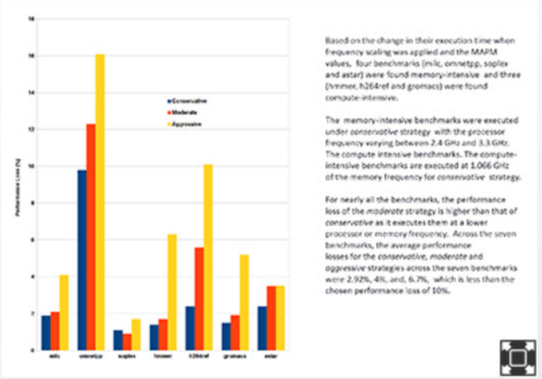


Figure 1. Performance

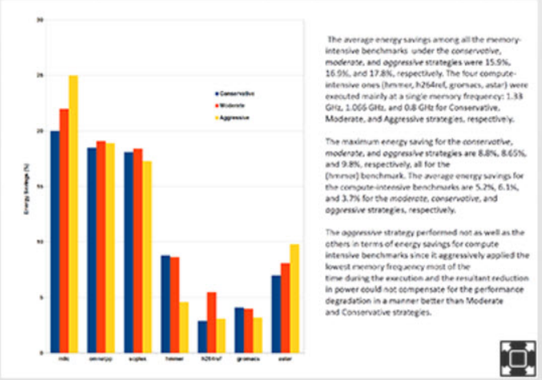


Figure 2. Energy Savings

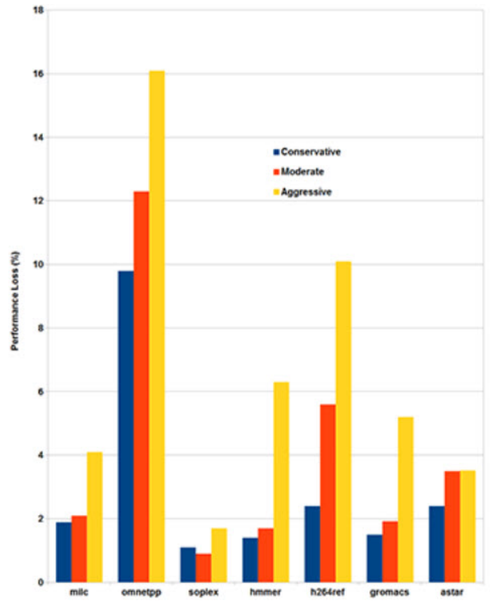
FUNDING & CREDITS

g of Proc

iversity, Norfolk, V

, capping powe
ducing processo
teristics, powe
stem was desig

model was propo
ency combinatio
hange in number

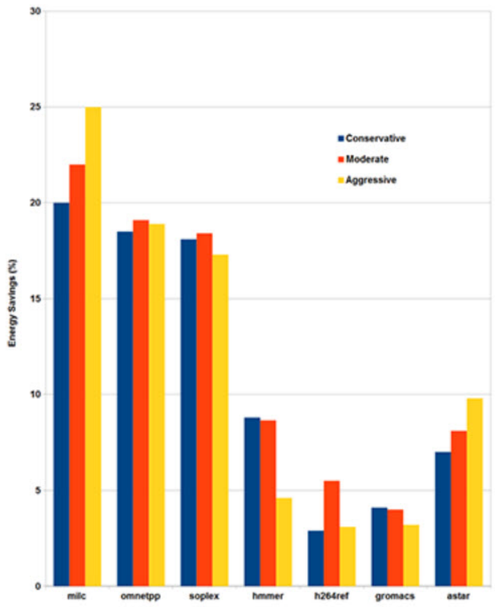


Based on the change in their execution time when frequency scaling was applied and the MAPM values, four benchmarks (milc, omnetpp, soplex and astar) were found memory-intensive and three (hmmer, h264ref and gromacs) were found compute-intensive.

The memory-intensive benchmarks were executed under *conservative* strategy with the processor frequency varying between 2.4 GHz and 3.3 GHz. The compute intensive benchmarks. The compute-intensive benchmarks are executed at 1.066 GHz of the memory frequency for *conservative* strategy.

For nearly all the benchmarks, the performance loss of the *moderate* strategy is higher than that of *conservative* as it executes them at a lower processor or memory frequency. Across the seven benchmarks, the average performance losses for the *conservative, moderate* and *aggressive* strategies across the seven benchmarks were 2.92%, 4%, and, 6.7%, which is less than the chosen performance loss of 10%.

Performance
Fig. 1 depicts the performance loss resulting from the proposed runtime system, compared to the full-power execution in that the benchmarks are run at the highest available processor-memory frequency pair.



The average energy savings among all the memory-intensive benchmarks under the *conservative*, *moderate*, and *aggressive* strategies were 15.9%, 16.9%, and 17.8%, respectively. The four compute-intensive ones (hmmer, h264ref, gromacs, astar) were executed mainly at a single memory frequency: 1.33 GHz, 1.066 GHz, and 0.8 GHz for Conservative, Moderate, and Aggressive strategies, respectively.

The maximum energy saving for the *conservative*, *moderate*, and *aggressive* strategies are 8.8%, 8.65%, and 9.8%, respectively, all for the (hmmer) benchmark. The average energy savings for the compute-intensive benchmarks are 5.2%, 6.1%, and 3.7% for the *moderate*, *conservative*, and *aggressive* strategies, respectively.

The *aggressive* strategy performed not as well as the others in terms of energy savings for compute intensive benchmarks since it aggressively applied the lowest memory frequency most of the time during the execution and the resultant reduction in power could not compensate for the performance degradation in a manner better than Moderate and Conservative strategies.

Energy Saving
Fig. 2 depicts the energy savings obtained from the proposed runtime system, with respect to the full-power execution for the seven SPEC benchmarks.

Figure 1. Performance

Figure 2. Energy Savings

d profile the appl
nd then chooses
efined performance constraint. Three different frequency scaling strategies were proposed: Conservative, Moderate, and Aggressive, named so to reflect the degree of their DVFS application. They are distinguished mainly by varying the way of determining how the workload parameters are determined.

Experimental Platform

The experiments were performed on a computin
memory. For measuring the node power and en
the proposed energy-saving model, SPEC CPU 20

Performance

Fig. 1 depicts the performance loss resulting fro
benchmarks are run at the highest available pro

Energy Saving

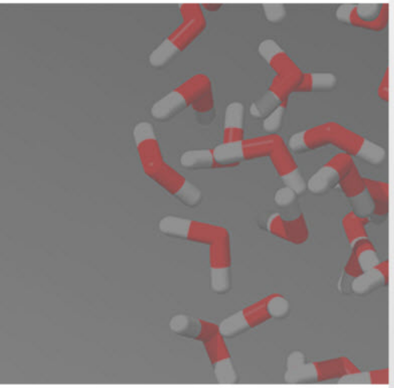
Fig. 2 depicts the energy savings obtained from the proposed runtime system, with respect to the full-power execution for the seven SPEC benchmarks. (bottom right)

This work was supported in part by Ames Laboratory and Iowa State University under the contract DE-AC02-07CH11358 with the U.S. Department of Energy, by the Air Force Office of Scientific Research under the AFOSR award FA9550-12-1-0476, and by the National Science Foundation grants NSF/OCI-0941434, 0904782, 1047772.

BACK TO MAP

FUNDING & CREDITS

THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY



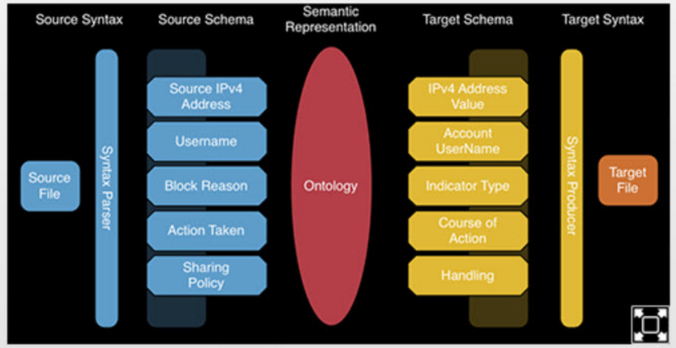
BACK TO MAP

Automating Intrusion Detection and Response

CFM – FlexT

https://github.com/anl-cyberscience/FlexTransform

Problem
Responding to cyber threats effectively as a community (e.g. the Energy Sector, DOE Labs) requires a (near-) real-time information sharing framework. However many organizations have cyber threat information pipelines which rely on home-grown or proprietary data formats, complicating the communication of information between them. This complication can slow, or even entirely prevent the participation of smaller organizations, those who can benefit most from such a framework.



Solution
We consider a cyber threat document to be comprised of three components: syntax, schema, and semantics. Using semantic web language (RDF), we map the syntax and schema of existing documents once to a common semantic representation. This mapping can then be used to both consume and produce cyber threat data in any supported format.

Contacts
Chris Strasburg (cstras@ameslab.gov)
Maheedhar Gunasekharan (mahee@iastate.edu)
Mark Gordon (mgordon@iastate.edu)

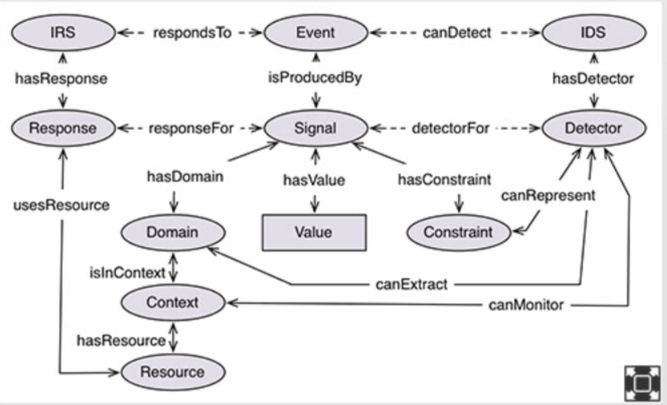
FUNDING & CREDITS

S-MAIDS

C. Strasburg, S. Basu, J.S. Wong, "S-MAIDS: A Semantic Model for Automated Tuning, Correlation, and Response Selection in Intrusion Detection Systems", COMPSAC, 2013.

Problem
As cyber environments and the threats against them increase in complexity, systems of Intrusion Detection need to be able to dynamically correlate and self-adjust to maintain optimal performance. Existing models of IDS are too course grained, to support this, however, and existing measures cannot be compared across data domains (e.g. between network packets and application logs) due to the Unit of Analysis problem. A new model and measure of detection, attacks, and response is needed.

Solution
We define a new model formally which unified detection, attacks, and response in terms of observable components called 'signals'. Formalized in the Web Ontology Language (OWL), automated reasoning can be applied to determine which signals are supported by detectors, and which resources provided the needed information for automated response. Combined with a modified information theoretic measure using the multi-instance learning technique of 'bagging', we provide a framework to automated IDS which is both granular enough to support automated tuning and correlation, and which overcomes the Unit of Analysis problem.



CyDIME

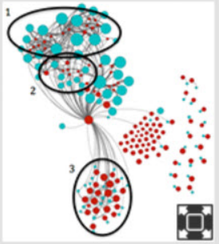
https://github.com/Ames-Laboratory-Cyber-Group/Cydime

Problem
Intelligent response to detected attacks and other malicious activity requires both knowledge of the characteristics of the attack as well as how resources involved in the attack related to the mission of the organization. Existing IDRS systems do not take into account the relationship between the target organization and the resources involved in the attack.

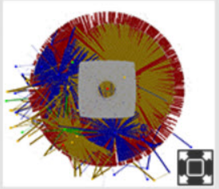
Solution
We propose an approach which allows an organization to assess and rank the degree to which an external resource impacts the mission of the organization. Specifically we: (i) formulate the resource impact ranking problem; (ii) adapt several representative machine learning based algorithms that predict impact ranks of the external resources; and (iii) introduce features that constitute the semantic context of a resource, and develop a parallel and scalable extraction method to extract such features.

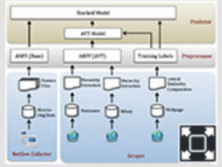
Achievements and work in progress
[Fig.Graph Analysis I] Identifying clusters of external resources (blue) having common communication with internal groups of resources (red) using network flow frequency as well as data volume. In the figure, some interesting characteristics immediately stand out, circled in black:
1. Externally visible services intended for the general public (SMTP, Web, DNS)
2. Externally visible services intended for employees (VPN, SSH, Webmail, POP/IMAP)
3. Scanning activity

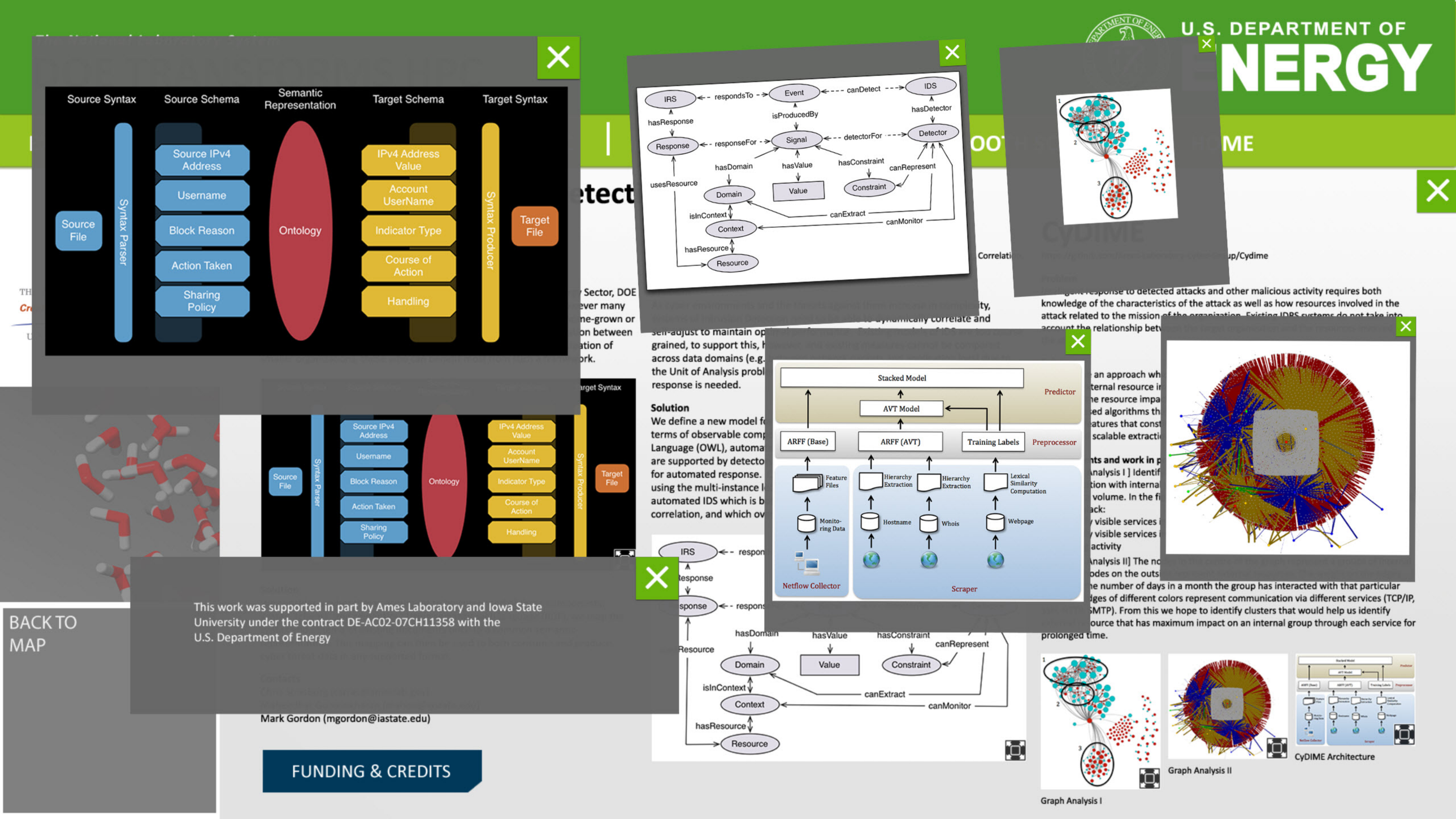
[Fig.Graph Analysis II] The nodes in the centre of the graph represent a groups of internal hosts and nodes on the outside represent external resources. The weight on the edges represent the number of days in a month the group has interacted with that particular resource. Edges of different colors represent communication via different services (TCP/IP, SSH, HTTP, SMTP). From this we hope to identify clusters that would help us identify external resource that has maximum impact on an internal group through each service for prolonged time.



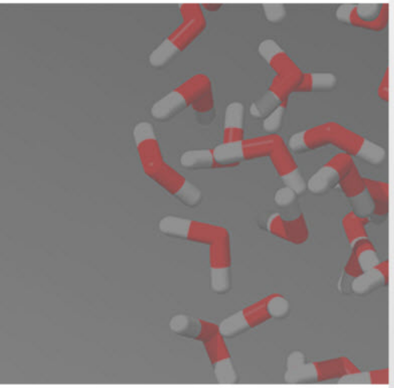
Graph Analysis I


Graph Analysis II


CyDIME Architecture


U.S. DEPARTMENT OF
NERGY
Source Syntax
Source Schema
Semantic Representation
Target Schema
Target Syntax
Source File
Syntax Parser
Source IPv4 Address
Username
Block Reason
Action Taken
Sharing Policy
Ontology
IPv4 Address Value
Account UserName
Indicator Type
Course of Action
Handling
Syntax Producer
Target File
IRS
respondsTo
Event
canDetect
IDS
hasResponse
isProducedBy
hasDetector
Response
responseFor
Signal
detectorFor
Detector
hasDomain
hasValue
hasConstraint
canRepresent
usesResource
Domain
Value
Constraint
isInContext
canExtract
Context
canMonitor
hasResource
Resource
Stacked Model
Predictor
AVT Model
ARFF (Base)
ARFF (AVT)
Training Labels
Preprocessor
Feature Files
Hierarchy Extraction
Hierarchy Extraction
Lexical Similarity Computation
Monito-ring Data
Hostname
Whois
Webpage
Netflow Collector
Scraper
Graph Analysis I
Graph Analysis II
CyDIME Architecture
BACK TO MAP
This work was supported in part by Ames Laboratory and Iowa State University under the contract DE-AC02-07CH11358 with the U.S. Department of Energy
Mark Gordon (mgordon@iastate.edu)
FUNDING & CREDITS

THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY



BACK TO MAP

Investigating Novel Hardware Architectures

The hardware used in high performance computing systems is constantly evolving. At Ames Laboratory, new hardware architectures are analyzed in terms of performance and energy efficiency. These studies help to ensure that scientific codes used in Department of Energy efforts are compatible with the most cutting edge computer hardware.

Energy Efficiency of Modern CPUs

The power consumption of an exascale system built from today's hardware would have unrealistic energy costs. The importance of the energy-to-solution metric is becoming increasingly significant for scientific codes. The widely-used GAMESS computational chemistry software package was used to measure performance and energy efficiency of modern ARM32, ARM64 and x86 processors[1]. The best performance is achieved with the x86 CPU, while the ARM32 CPU is shown to be the most energy efficient.

Using Accelerators in Quantum Chemistry

Hardware trends indicate that heterogeneous computer architectures (CPU + accelerator) are here to stay. Graphics processing units (GPUs) and the Xeon Phi are the two most dominant accelerator platforms in high performance computing today. Some routines in GAMESS can already be used with GPUs via the LIBCCHEM library. A preliminary investigation of utilizing the Xeon Phi in quantum chemistry[2] inspired work to optimize performance-critical portions of GAMESS for the Xeon Phi. In a collaborative effort with Argonne National Laboratory, the libERD integral library package is being interfaced with GAMESS. This interface will be extended to the OptERD library: a version of the libERD library which is optimized for the Xeon Phi architecture.

Energy-Savings in Applications Offloaded to Intel Xeon Phi

We used time on-chip to time off-chip ratio to determine whether applications are compute- or memory-intensive of applications and found that the offload mode remains compute-intensive and yields energy gains due to possibility of the frequency scaling (DVFS) on the host CPU. DVFS during the offload is beneficial: Energy savings of 9% were achieved with only 1% performance increase.

[1]Keipert, K.; Mitra, G; Sundriyal, V.; Leang, S. S.; Sosokina, M.; Rendell, A. P.; Gordon, M.S. J. Chem. Theory Comput. 2015, Article ASAP
[2]Leang, S. S.; Rendell, A. P.; Gordon, M. S. J. Chem. Theory Comput. 2014, 10, 908–912.

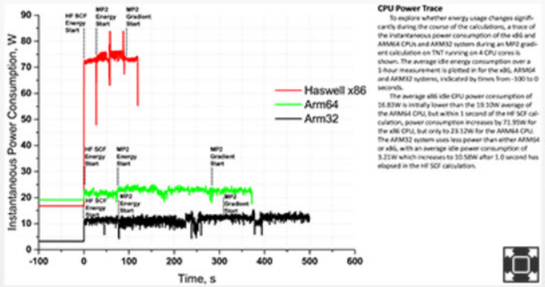


Figure 1. CPU Power Trace

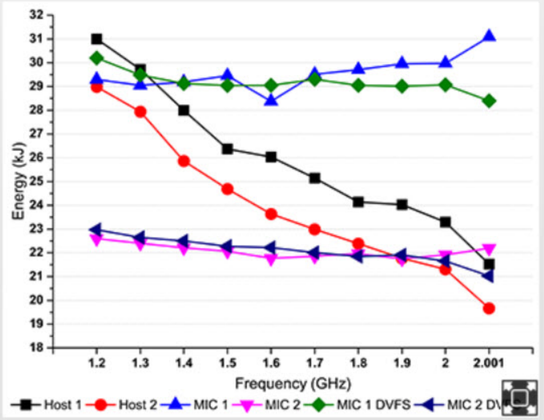
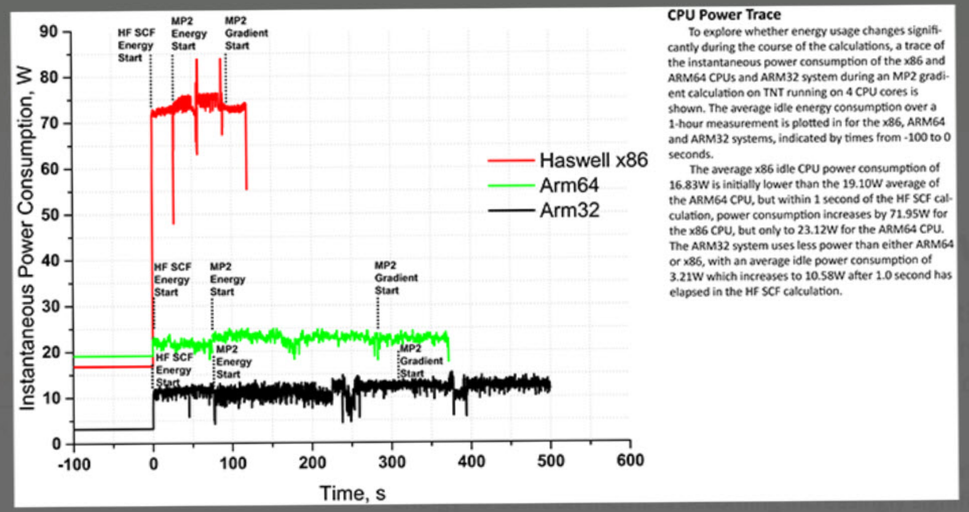


Figure 2. MIC Energy vs. Core Frequency

FUNDING & CREDITS



CPU Power Trace

To explore whether energy usage changes significantly during the course of the calculations, a trace of the instantaneous power consumption of the x86 and ARM64 CPUs and ARM32 system during an MP2 gradient calculation on TNT running on 4 CPU cores is shown. The average idle energy consumption over a 1-hour measurement is plotted in for the x86, ARM64 and ARM32 systems, indicated by times from -100 to 0 seconds.

The average x86 idle CPU power consumption of 16.83W is initially lower than the 19.10W average of the ARM64 CPU, but within 1 second of the HF SCF calculation, power consumption increases by 71.95W for the x86 CPU, but only to 23.12W for the ARM64 CPU. The ARM32 system uses less power than either ARM64 or x86, with an average idle power consumption of 3.21W which increases to 10.58W after 1.0 second has elapsed in the HF SCF calculation.

Figure 1. CPU Power Trace

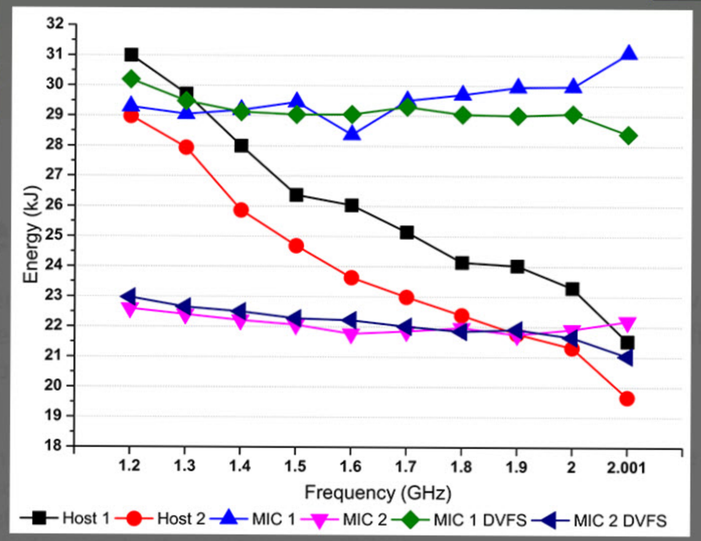


Figure 2. MIC Energy vs. Core Frequency

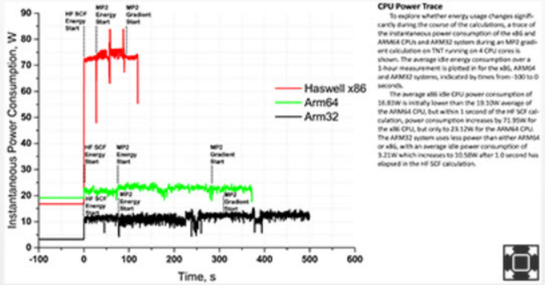


Figure 1. CPU Power Trace

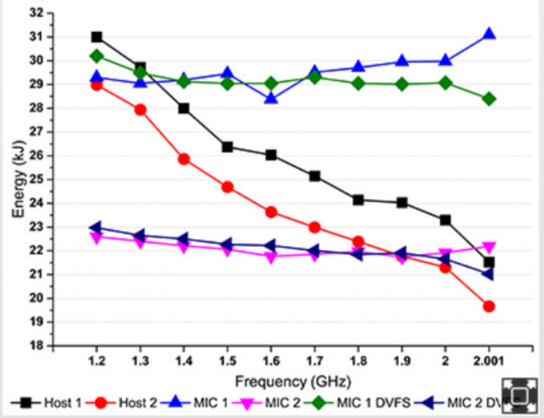
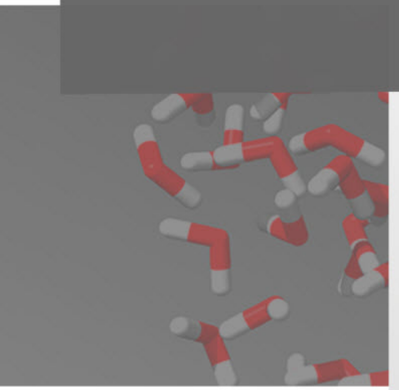


Figure 2. MIC Energy vs. Core Frequency



BACK TO MAP

Using Accelerators in Quantum Chemistry

Hardware trends indicate that heterogeneous computer architectures (CPU + accelerator) are here to stay. Graphics processing units (GPUs) and the Xeon Phi are the two most dominant accelerator platforms in high performance computing today. Some routines in GAMESS can already be used with GPUs via the LIBCCHEM library. A preliminary investigation of utilizing the Xeon Phi in quantum chemistry[2] inspired work to optimize performance-critical portions of GAMESS for the Xeon Phi. In a collaborative effort with Argonne National Laboratory, the libERD integral library package is being

libERD library which is optimized for th

Energy-Savings in Applications

We used time on-chip to time off-chip

found that the offload mode remains

host CPU. DVFS during the offload is b

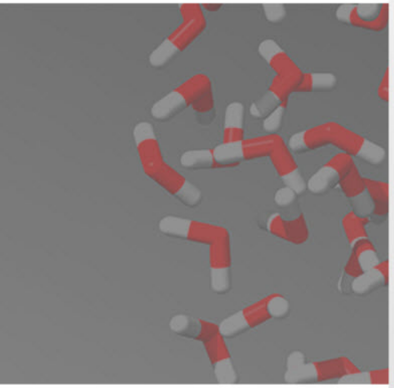
The authors thank Intel Corp. and NVIDIA Corp. for their support of this work.
K.K., V.S., S.L., M.S., and M.S.G. thank the Air Force Office of Scientific Research for their support of this work under AFOSR Award No. FA9550-12-1-0476.

[1]Keipert, K.; Mitra, G; Sundriyal, V.; Leang, S

[2]Leang, S. S.; Rendell, A. P.; Gordon, M. S. J. Chem. Theory Comput. 2014, 10, 908–912.

FUNDING & CREDITS

THE Ames Laboratory
Creating Materials & Energy Solutions
U.S. DEPARTMENT OF ENERGY



BACK TO MAP

A Spin-free $[2]_{R12}$ basis set incompleteness correction to the Local Multi-Reference Configuration Interaction and the Local Multi-Reference Average Coupled Pair Functional methods

Luke B. Roskop[a], Edward F. Valeev[b], Emily A. Carter[c], Mark S. Gordon[a], and Theresa L. Windus[a]

[a]Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50010, United States
[b]Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States
[c]Department of Mechanical and Aerospace Engineering, Program in Applied and Computational Mathematics, and Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ, 08544, United States

One goal of electronic structure theory is to improve the affordability of existing highly accurate methods to allow them to be applied to larger chemical systems. Local multi-reference configuration interaction (LMRCI) and local multi-reference averaged coupled pair functional (LMRACPF) are two such methods. In an effort to reduce the error associated with truncated basis sets, the LMRCI and LMRACPF methods have been extended to include explicit correlation via the spin-free $[2]_{R12}$ basis set incompleteness correction.

Background

The local electron correlation approximation circumvents the computational overhead of conventional (i.e. non-local) methods through neglect of the interactions between electrons that occupy distant pairs of localized molecular orbitals (LMOs) and neglect of excitations to localized unoccupied orbitals that are distant from a given LMO. Here, the LMRCI and the LMRACPF methods have been augmented with the explicitly correlated $[2]_{R12}$ basis set incompleteness correction. The explicitly correlated $[2]_{R12}$ method accelerates the convergence of the correlation energy (that is computed with a finite basis set) to the complete basis set (CBS) limit by including terms that depend explicitly on the inter-electronic distance $_{r12}$. This advance is important for local methods since the use of large basis sets can lead to numerical issues (e.g. linearly dependent AOs), which complicates orbital localization and impedes application of the local method.

Basis Set Convergence

The convergence of conventional and explicitly correlated bond dissociation energies (BDE) for local and non-local treatments are shown in Chart 1 (top right).

DDI Parallel Performance

Parallel scaling for the OpenMP and the distributed data interface (DDI) TigerCI codes is shown in Chart 2 (bottom right).

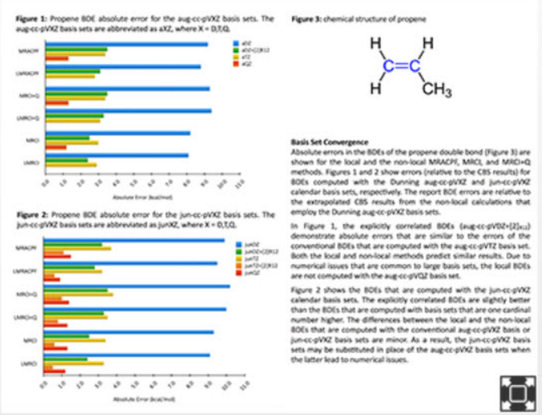


Chart 1: Basis Set Convergence

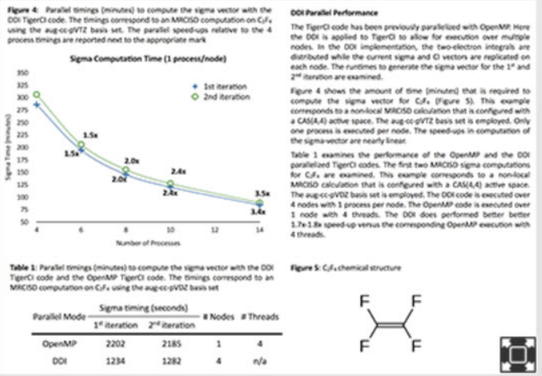
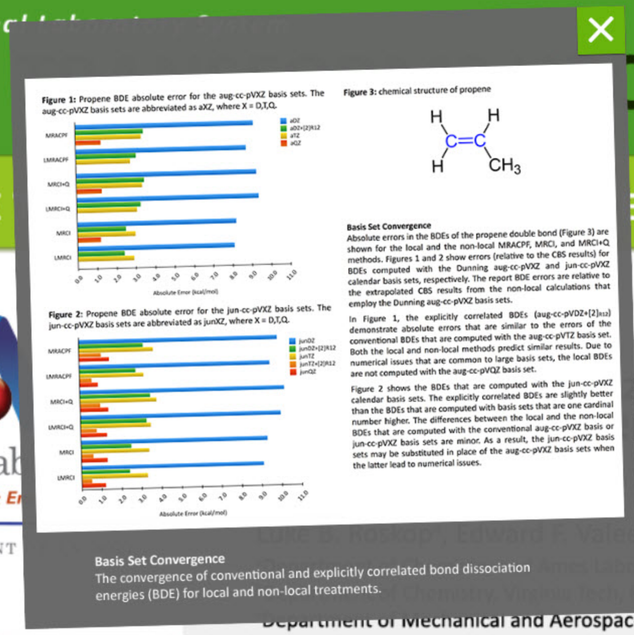
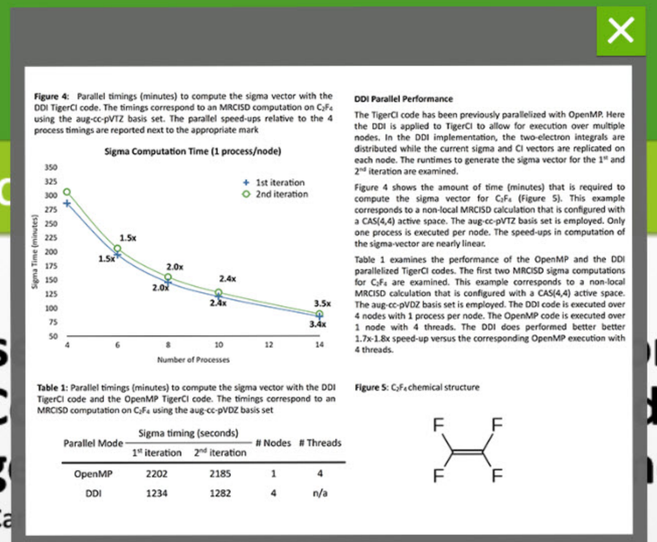


Chart 2: DDI Parallel Performance

FUNDING & CREDITS



Basis Set Convergence
The convergence of conventional and explicitly correlated bond dissociation energies (BDE) for local and non-local treatments.



Parallel Mode	Sigma timing (seconds) 1st iteration	2nd iteration	# Nodes	# Threads
OpenMP	2202	2185	1	4
DDI	1234	1282	4	n/a

DDI Parallel Performance
Parallel scaling for the OpenMP and the distributed data interface (DDI) TigerCI codes.

... basis s... n to the ... rence C... d the Local ... Averag... ethods

... v[b], Emily A. Ca...
... ratory, Iowa State ...
... lacksburg, Virginia ...
... Department of Mechanical and Aerospace Engineering, Prog...
and Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ, 08544, United States

One goal of electronic structure theory is to improve the affordability of existing highly accurate methods to allow them to be applied to larger chemical systems. Local multi-reference configuration interaction (LMRCI) and local multi-reference averaged coupled pair functional (LMRACPF) are two such methods. In an effort to reduce the error associated with truncated basis sets, the LMRCI and LMRACPF methods have been extended to include explicit correlation via the spin-free ... basis set incompleteness correction.

Funding for this research was provided by a National Science Foundation Software Infrastructure (SI2) award, OCI-1047772.

Background

The local electron correlation ap... through neglect of the interactions betwe... itations to localized unoccupied orbitals tha... nted with the explicitly correlated $[2]_{R12}$ bas... gence of the correlation energy (that is computed with a finite basis set) to the complete basis set (CBS) limit by including terms that depend explicitly on the inter-electronic distance $_{r12}$. This advance is important for local methods since the use of large basis sets can lead to numerical issues (e.g. linearly dependent AOs), which complicates orbital localization and impedes application of the local method.

Basis Set Convergence

The convergence of conventional and explicitly correlated bond dissociation energies (BDE) for local and non-local treatments are shown in Chart 1 (top right).

DDI Parallel Performance

Parallel scaling for the OpenMP and the distributed data interface (DDI) TigerCI codes is shown in Chart 2 (bottom right).

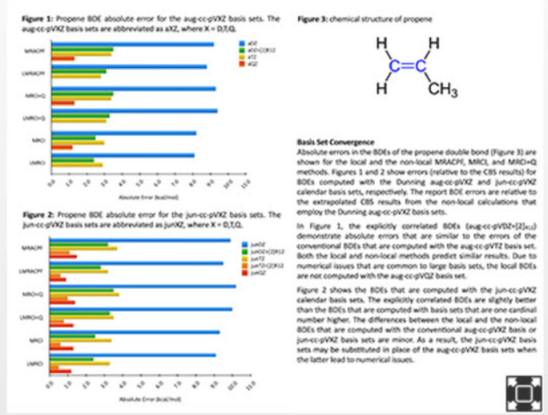


Chart 1: Basis Set Convergence

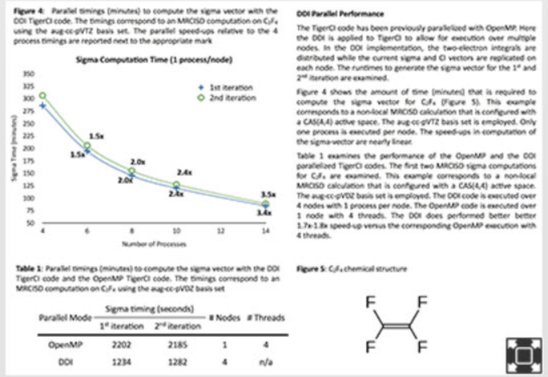
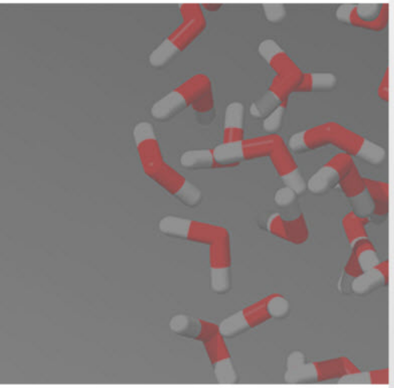


Chart 2: DDI Parallel Performance



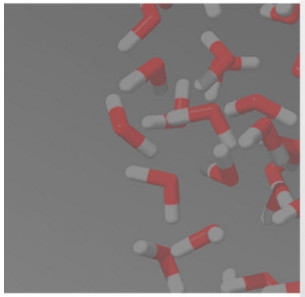
BACK TO MAP

FUNDING & CREDITS

THE Ames Laboratory

Creating Materials & Energy Solutions

U.S. DEPARTMENT OF ENERGY



BACK TO MAP

Scalable *ab initio* Dynamics with the Fragment Molecular Orbital Method

Alexander D. Findlater[a],, Spencer R. Pruitt[b], Justin A. Conrad[a], Olaseni Sode[c], Yuri Alexeev[b], Graham D. Fletcher[b], Gregory A. Voth[c], Mark S. Gordon[a]

a Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa, 50011, 201 Spedding hall, Iowa state University, Ames, IA 50011

b Argonne Leadership Computing Facility, Argonne National Laboratory, 9700 S. Cass Avenue, Lemont, IL 60439

c Department of Chemistry, James Franck Institute, Institute for Biophysical Dynamics, and Computation Institute, University of Chicago, Chicago, Illinois 60637

The Fragment Molecular Orbital (FMO) Method

The fragment molecular orbital (FMO) method [1-4] divides a large molecular system into fragments and performs ***ab initio*** calculations for individual fragments (monomers) and fragment pairs (dimers) in the presence of an electrostatic potential (ESP) due to the remaining fragments. Individual fragment calculations are calculated in parallel.

$$E = \sum_{I}^{N} E_I + \sum_{I>J}^{N} (E_{IJ} - E_I - E_J) + \sum_{I>J>K}^{N} \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J) - (E_{JK} - E_J - E_K) - (E_{KI} - E_K - E_I)\} + \ldots$$

FMO Multi-State Reactive MD (FMO-MS-RMD)

The current implementation of FMO-MD does not allow for "refragmentation" of a chemical system during a given computation. While this does not pose a problem for many chemical systems, within the realm of ms or chemical dynamics the ability to redefine fragments on the fly is crucial. In an effort to address this problem, a new refragmentation scheme called Multi-State Reactive Molecular Dynamics [5](MS-RMD) has been implemented into the GAMESS program package. New levels of parallelization were added to enhance the scalability of the method using the recently developed "multi-level" parallelization scheme in GAMESS (GDDI/3).

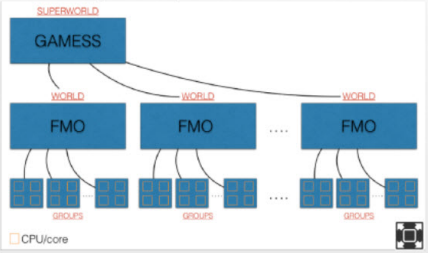
Argonne Leadership Computing Facility INCITE Project: State-of-the Art Simulations of Liquid Phenomena

The ongoing ALCF INCITE project is focused on modeling a variety of liquids with the FMO-MD and FMO-MS-RMD fragmentation methods in the GAMESS software package on the Production-Capability (~50,000 nodes) queue on the ALCF Blue Gene/Q MIRA machine. The transport properties of a variety of room-temperature ionic liquids, with applications in dye-sensitized solar cells (DSSC), are studied with FMO-MD while the study of proton transport in water requires the FMO-MS-RMD "on-the-fly refragmentation" method. The importance of proton transport, along with proton solvation, has been recognized in many different disciplines including acid base chemistry, energy storage, and biochemistry.

Interfacing the LibERD integral library with GAMESS

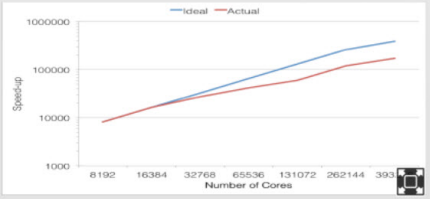
In collaboration with the Argonne Leadership Computing Facility, the libERD integral library was interfaced with GAMESS. Using libERD to calculate only the high angular momentum integrals in GAMESS, speedups of ~7x are achieved relative to the default GAMESS integral scheme for the heavy elements (Lanthanides). Even for molecules made entirely of light atoms, speedups of ~1.3x are achieved when using high-quality, high angular momentum basis-sets. The interface will be extended to OptERD, a version of libERD library optimized for Intel Xeon Phi architecture.

Figure 1. GDDI/3 Multi-Level Parallelism



A third level of MPI communicators (SUPERWORLD) allow multiple concurrent instances of the FMO method to run, each with a different fragmentation scheme determined by the FMO-MS-RMD state search algorithm.

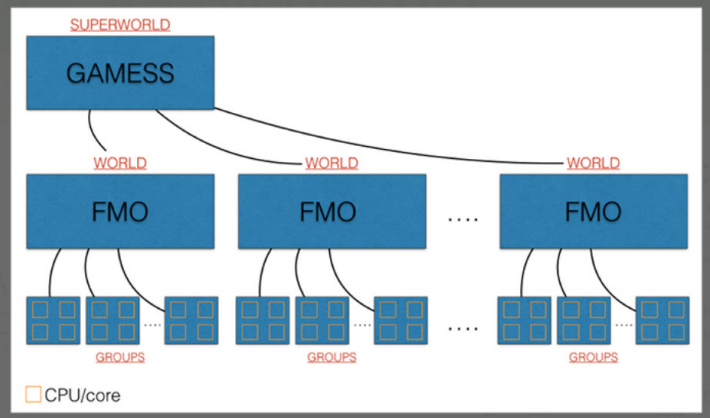
Figure 2. FMO-MS-RMD Scaling



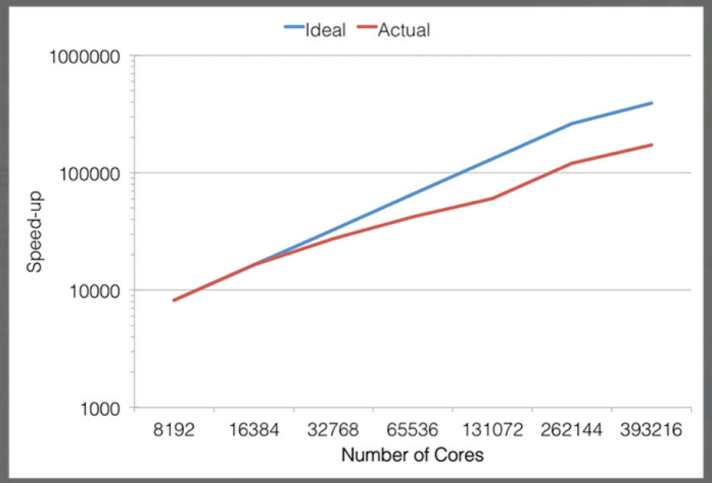
Scalability of the FMO-MS-RMD method using GDDI/3 can reach 393,216 cores, with additional test calculations performed on the entire ALCF BG/Q system *Mira* (49,152 nodes/786,432 cores).

[1] Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M., Chemical Physics Letters **1999,** 313, 701-706.
[2] Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K., Chem. Phys. Lett. **2002,** 351, 475-480.
[3] Fedorov, D. G.; Kitaura, K., J. Phys. Chem. A **2007,** 111, 6904-6914.
[4] The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems. CRC Press: Boca Raton, FL, 2009.
[5] Lange, A. W.; Voth, G. A., J. Chem. Theor. Comp. **2013,** 9 (9), 4018-4025.

FUNDING & CREDITS



A third level of MPI communicators (SUPERWORLD) allow multiple concurrent instances of the FMO method to run, each with a different fragmentation scheme determined by the FMO-MS-RMD state search algorithm.



Scalability of the FMO-MS-RMD method using GDDI/3 can reach 393,216 cores, with additional test calculations performed on the entire ALCF BG/Q system Mira (49,152 nodes/786,432 cores).

BACK TO MAP

of a variety of room-temperature ionic liquids, with applications in dye-sensitized solar cells (DSSC), are studied with FMO-MD while the study of proton transport in water requires the FMO-MS-RMD "on-the-fly refragmentation" method. The importance of proton transport, along with proton solvation, has been recognized in many different disciplines including acid base chemistry, energy storage, and biochemistry.

Interfacing the LibERD integral library with GAM

In collaboration with the Argonne Leadership Co

only the high angular momentum integrals in GA

elements (Lanthanides). Even for molecules ma

momentum basis-sets. The interface will be exte

An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. M.S.G. was supported by a grant from the Air Force Office of Scientific Research under AFOSR Award No. FA9550-11-1-0099.

[1] Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M., Chemical Physics Letters **1999,** 313, 701-706.
[2] Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K., Chem. Phys. Lett. **2002,** 351, 475-480.
[3] Fedorov, D. G.; Kitaura, K., J. Phys. Chem. A **2007,** 111, 6904-6914.
[4] The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems. CRC Press: Boca Raton, FL, 2009.
[5] Lange, A. W.; Voth, G. A., J. Chem. Theor. Comp. **2013,** 9 (9), 4018-4025.

FUNDING & CREDITS