REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
03-01-2016	Final Report	21-Jul-2014 - 20-Jul-2015

4. TITLE AND SUBTITLE
Final Report: Hybrid Computational Architecture for Multi-Scale Modeling of Materials and Devices

5a. CONTRACT NUMBER
W911NF-14-1-0432

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611103

6. AUTHORS
Enrico Bellotti

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Boston University
Office of Sponsored Program
881 Commonwealth Avenue
Boston, MA 02215 -1300

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
65125-MA-RIP.8

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The activity carried out within this program has focused on acquiring and evaluating a hybrid computational system that has enabled the development of a new generation of multi-scale simulation tools for the design of electronic and photonic materials. This computing hardware has made it possible to test the performance of different hybrid computing architectures in solving a number of problems that are relevant to the simulation of electronic materials and devices. The system can be configured by changing the number and kind of conventional multi-core processors assigned to a certain of problem. The proposed activity has also significantly augmented the quality and quantity of

15. SUBJECT TERMS
Parallel Computing, Multiscale Modelling, Hybrid Architectures

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Enrico Bellotti
UU	UU	UU	UU		19b. TELEPHONE NUMBER 617-358-1576

Report Title

Final Report: Hybrid Computational Architecture for Multi-Scale Modeling of Materials and Devices

ABSTRACT

The activity carried out within this program has focused on acquiring and evaluating a hybrid computational system that has enabled the development of a new generation of multi-scale simulation tools for the design of electronic and photonic materials. This computing hardware has made it possible to test the performance of different hybrid computing architectures in solving a number of problems that are relevant to the simulation of electronic materials and devices. The system can be configured by changing the number and kind of conventional multi-core processors assigned to a certain of problem. The proposed activity has also significantly augmented the quality and quantity of work that the PI is doing within the collaborative research alliance (CRA) for Multi-Scale Simulation of Electronic Materials (MSME). The goal of this Army Research Laboratory's initiative is to develop the next generation of material simulation tools. The system acquired using DURIP funds has provided the Computational Electronics Group at Boston University with an unprecedented capability to design electronics and photonics materials that are needed for the next generation of defense systems. The system is currently used for production run and it is expected to continue generating results for the next few years.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received Paper

TOTAL:

Number of Manuscripts:

Books

Received Book

TOTAL:

Received Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Names of Post Doctorates

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Names of Faculty Supported

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

Technology Transfer

DURIP: Hybrid Computational Architecture for Multi-Scale Modeling of Materials and Devices

2014 DoD DURIP Program

PI: Prof. Enrico Bellotti
Department of Electrical and Computer Engineering
8 Saint Mary's Street
Boston, MA 02215
bellotti@bu.edu

ARO Technical POC: Dr. Joseph Myers
U.S. Army Research Office RDRL-ROI-M
Box 12211 Research Triangle Park, NC 27709-2211
josephd.myers@us.army.mil
Tel. (919) 549-4245

Abstract

The activity carried out within this program has focused on acquiring and evaluating a hybrid computational system that has enabled the development of a new generation of multi-scale simulation tools for the design of electronic and photonics materials. This computing hardware has made it possible to test the performance of different hybrid computing architectures in solving a number of problems that are relevant to the simulation of electronic materials and devices. The system can be configured by changing the number and kind of conventional multi-core processors assigned to a certain of problem.

The proposed activity has also significantly augmented the quality and quantity of work that the PI is doing within the collaborative research alliance (CRA) for Multi-Scale Simulation of Electronic Materials (MSME). The goal of this Army Research Laboratory's initiative is to develop the next generation of material simulation tools. The system acquired using DURIP funds has provided the Computational Electronics Group at Boston University with an unprecedented capability to design electronics and photonics materials that are needed for the next generation of defense systems. The system is currently used for production run and it is expected to continue generating results for the next few years.

Table of Contents

List of Appendices, Illustrations, and Tables:

Summary of the most important results

1 – Introduction

Recent electronic and photonic devices based on novel electronic materials are highly complex and their development has required the buildup of increasingly sophisticated applied mathematics, numerical analysis, and simulation tools. These simulation programs have provided insight into new physical phenomena and led to devices with enhanced performance, additional functionalities, and novel architectures. While the flexibility and power of modern computational resources have enabled complex numerical simulation capabilities, true "material by design" (synthesis rather than analysis) is still a significant challenge. A key issue is that one must have efficient simulation methodologies which describe physical phenomena at different spatial and temporal scales. The development of multi-scale simulation platforms is an active and on-going area of research. The Boston University Computational Electronics Group is involved with such an initiative through the Army Research Laboratory (ARL) Multi-scale Electronic Material Simulation (MSME) Collaborative Research Alliance (CRA) and was awarded $150K as part of the 2014 Defense University Research Instrumentation Program (DURIP) for the acquisition of a hybrid computational cluster. Over the past 18 months, a hybrid computational architecture has been acquired, integrated with existing computational resources, and applied to the investigation of contemporary materials science and device physics problems. In this document, the acquired computational systems will be outlined and a description of how the new equipment has been integrated with existing resources will be given. The software and simulation tools implemented on these new systems will then be shown along with their scaling capabilities.

2 - Hardware Acquisition

Given the complexity of modern numerical simulation techniques and algorithms, different types of modeling problems can show a wide range of performances depending on the system in which it is implemented. That is, the efficiency of computational simulation programs is dependent on the hardware on which they are run. In the context of developing a multi-scale simulation platform which is composed of a multitude of different techniques, it is then necessary to employ a hybrid computational architecture to test the performance of different combinations of computing units. The Boston University Computational Electronics Group was awarded $150K in the 2014 DURIP to obtain an ad hoc configurable cluster consisting of a number of conventional multi-core servers, GPU units, networking hardware, and storage solutions.

Based on performance evaluations prior to the award of the 2014 DURIP, three specific computational architectures were specified for acquisition, the (generalized) merits of which are listed below:

- High-core count conventional CPU AMD machines: The AMD processor architecture allows for higher physical core counts than Intel architectures at the cost of lower clock speeds and smaller on-chip memory. In our experience, these machines are thus well-suited for

applications in which a large portion of the simulation is parallelizable with a low amount of communication among different parallel threads.

- Fast clock conventional CPU Intel machines: The Intel architecture typically allows higher clock speeds with lower physical core counts but the option of doubling the total threads using hyper-threading technology. Our experience has led us to use these machines for applications in which significant serial scalar bottlenecks exist in the simulation algorithms.
- GPU units: GPU-accelerated architectures provide a fundamentally different platform over which to distribute computational loads. By offloading a small but compute-intensive portion of the application code to the graphics processing unit, significant speed up can be achieved in properly tuned codes. Our experience shows that GPU-accelerated processing is most efficient for codes in which a large number of independent scalar operations must be performed.

In addition to the compute nodes, peripheral hardware such as storage nodes and networking switches were required to achieve the cluster's full performance capability.

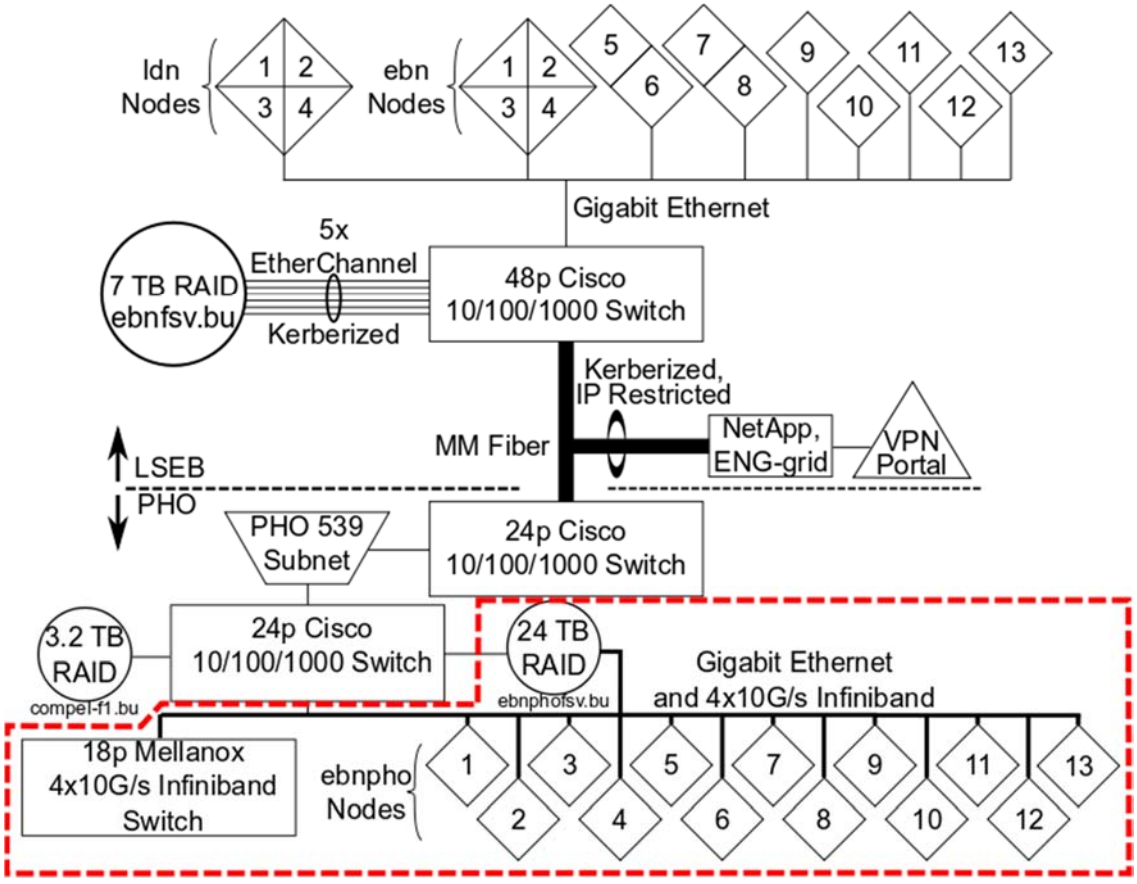


Figure 1 - Topology of the CompEl cluster. Equipment purchased through the 2014 DURIP awarded to Boston University is enclosed in the dashed red line. Refer to Table 1 for machine designations.

To target the above computational architectures, an order was placed in August 2014 for 1 GPU server, 2 AMD 32-core servers, 3 AMD 64-core servers, 3 Intel 20-core servers, a 24TB storage

node, and an 18-port Mellanox Infiniband switch. Thinkmate, Inc. (Waltham, MA) was selected as the vendor following a competitive bidding process resulting in a $120.6K purchase. Following performance testing, the remainder of the funds was used to purchase two additional Intel compute nodes in May and October of 2015. In total, the 2014 DURIP award enabled the acquisition of 440 conventional CPU cores, 5000 CUDA cores, 3.8 TB RAM, and 24TB of file storage. A detailed description of the acquired hardware can be seen in Table 1.

In August of 2014, the equipment purchased through the 2014 DURIP was integrated into the existing Computational Electronics cluster, resulting in the cluster topology shown in Figure 1. Funded by the Department of Electrical and Computer Engineering, a new server closet was constructed in the Boston University Photonics Center (PHO539G) to house the new equipment. The CompEl DURIP cluster occupies 20 units in a server rack and is redundantly interconnected with gigabit Ethernet and 4x10 gigabit Infiniband. The storage server hosts a 24TB hardware RAID5 single XFS partition which is mounted via NFS to each of the compute nodes. All machines are assigned static addresses on the PHO539 subnet and access to them is restricted via IP and Kerberos username. The remainder of the CompEl cluster is housed remotely and connection is made via multimode fiber. Existing file storage is mounted to each of the DURIP purchased machines identically to the new storage server. Boston University and College of Engineering shared computing resources are hosted with Active Directory allowing access to University software and storage solutions.

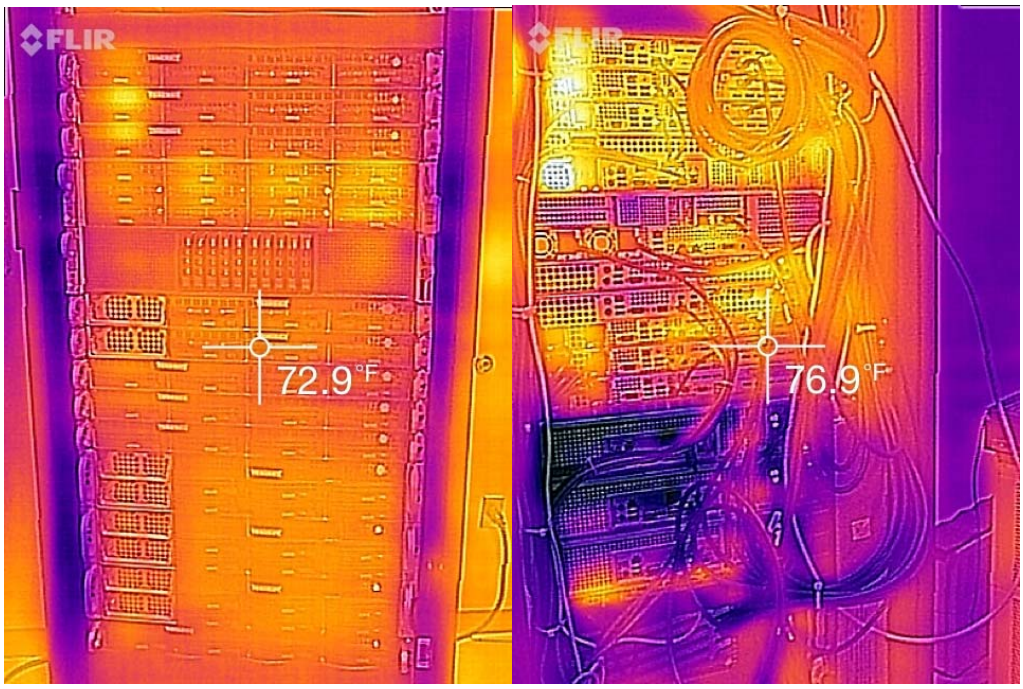


Figure 2 - Thermal maps (LWIR Lepton FLIR Camera) of the front and back side of the cluster. Efficient front-to-back cooling lead to a minimum temperature gradient in the system.

3 - Software Applications

The Computational Electronics high performance computing cluster has been connected to shared Boston University resources to enable access to a wide variety of software. Here, we will show scaling results across different computational architectures from three software packages used routinely in our research: VASP density functional theory for evaluating the electronic structure of semiconducting materials, Synopsys EMW finite difference time domain for electromagnetic scattering and absorption, and Synopsys SDEVICE for finite element solutions of the drift-diffusion semiconductor device equations. These applications have been chosen as they cover wide spatial scales, from the quantum to classical, and demonstrate the needs of multi-scale simulation hierarchies.

3.1 - Density Functional Theory Modelling

First-principles density functional theory (DFT) is routinely used to investigate the electronic structure of materials which lack long-range order. We have extensively investigated dislocations in GaN, a model material for opto- and high-power electronics, which requires an accurate atomistic description of the dislocation core and a large domain to capture long-range elastic field. System size is however generally limited to several hundred atoms even on modern supercomputers since the computational complexity increases exponentially with the number of electrons, N_{el}. For VASP (our choice for implementation), the complexity of the standard DFT calculation is $O(N_{el}^2 \log N_{el})$. To reduce run-time, the problem is distributed across a number of multi-core, multi-node systems via parallelization. In order to develop an efficient way to distribute problems ad hoc across our computational cluster, we characterized the speed-up achieved as a function of the number of cores used during simulation. Results are shown in Figure 3, compared to results obtained on Army HPC resources. Generally, the majority of speed-up is achieved over the first 20 cores. Beyond this, the benefit from further parallelization is marginal. Interestingly, in several cases (for example, "Xeon" and "ebn9") there is a discrete drop in the speed-up when an additional core is added. We attribute this effect to hyper-threading into virtual cores; the problem is efficiently parallelized over the physical cores, but when it is passed to virtualized cores the resources available to each core are decreased resulting in slower performance. These results lead us to the conclusion that DFT simulations are best performed on low core-count machines with fast processors to best capture the scaling before the onset of diminishing returns. Using machines with higher core counts will add only marginal benefits and make them unavailable for applications that could better utilize them.

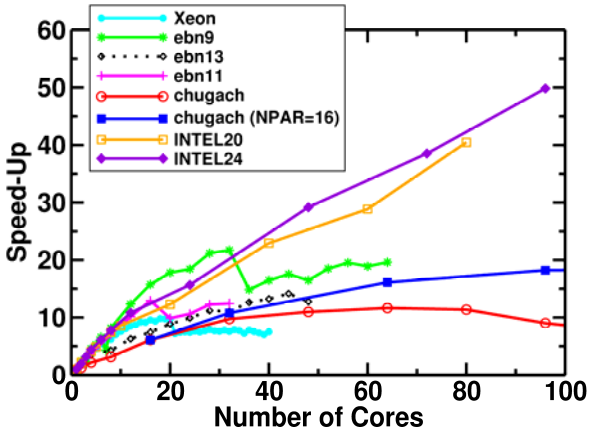


Figure 3 - Scaling of density functional theory code (Vasp) on CompEL cluster and DoD HPC systems.

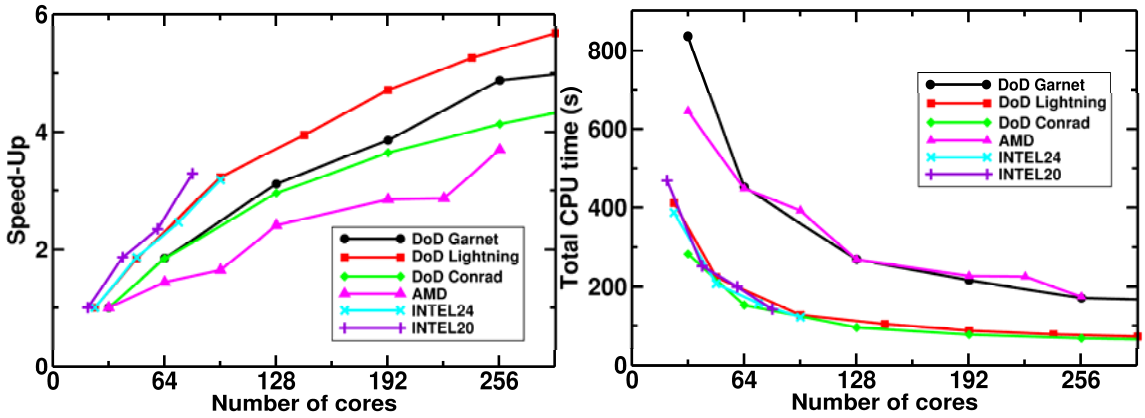


Figure 4 – Scaling of density functional theory code (Vasp) on CompEL cluster and DoD HPC systems Garnet, Lightning and Conrad.

Additional tests have been performed to investigate the scaling properties of the newly acquired systems with the latest HPC systems installed by the DoD. These are Garnet (AMD based machine) and Lightning and Conrad (Intel based systems). Figure 4 presents the scaling properties of both CompEL AMD and Intel nodes compared to Garnet, Lightning and Conrad. It can be seen that similar scaling properties are obtained for systems with the same processor family.

3.2 - Finite Difference Time Domain

The Computational Electronics group frequently uses the finite-difference time domain (FDTD) method (implemented in Synopsys EMW) to determine the electromagnetic response of optoelectronic devices. The FDTD method uses a direct-time approach to solve Maxwell's curl equations by splitting them into three scalar partial differential equations and replacing the partial derivatives with first order central differences. The result is a set of six algebraic update equations at each spatial point on a structured grid. The update equations are used with a time-stepping algorithm to propagate a solution through a simulation domain. The efficiency of the FDTD algorithm is directly related to the computational mesh; an update equation must be solved at each grid point the FDTD method is therefore O(N) where N is the number of points in the grid. Furthermore, a physical steady-state solution must causally link one side of the domain to the other causing additional scaling with n_{tot}, the number of time steps.

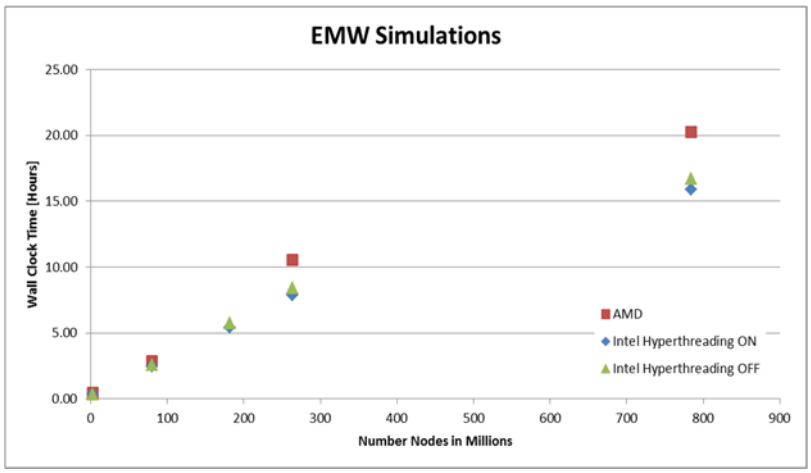


Figure 5 - Scaling of solution time with size of finite difference time domain structured mesh.

In three-dimensional simulations, it is assumed that n_{tot} is proportional to the third root of the mesh size causing overall $O(N^{4/3})$. Figure 3 shows the wall clock time required for a steady-state solution as a function of the structured mesh size for a number of different computational architectures. Results have been normalized by the number of cores used during the calculation. We have found that on a per-core basis, our Intel machines outperform the AMD nodes. However, since our AMD machines generally have a higher core count, it is more efficient to use the AMD machines for the FDTD simulations. If Intel machines are used, it is efficient to enable hyper-threading.

3.3 - Finite Element Drift Diffusion Code

We have performed a similar analysis of the scaling of the finite element method (FEM) used for solving the drift-diffusion formulation of the semiconductor device equations, the results of which are shown in Figure 4. Unlike the previous FDTD simulations, the FEM method uses an unstructured mesh which must be carefully designed with consideration of the physics of the device. In general, the computational requirements (both CPU hours and memory) scale linearly with the size of the domain, but there is some problem-to-problem variation depending on the specific physics of the device under consideration. Comparing the results of Intel and AMD machines, it is seen again that on a per-core basis, the Intel CPUs are more efficient than AMD. This is likely a direct consequence of the different clock speeds of the two processors. Again, since the AMD machines have a higher per-node core count, it is more efficient to distribute our FEM problems to the AMD servers. Unlike in the case of FDTD simulations, it is disadvantageous to enable hyper-threading if Intel machines are to be used.

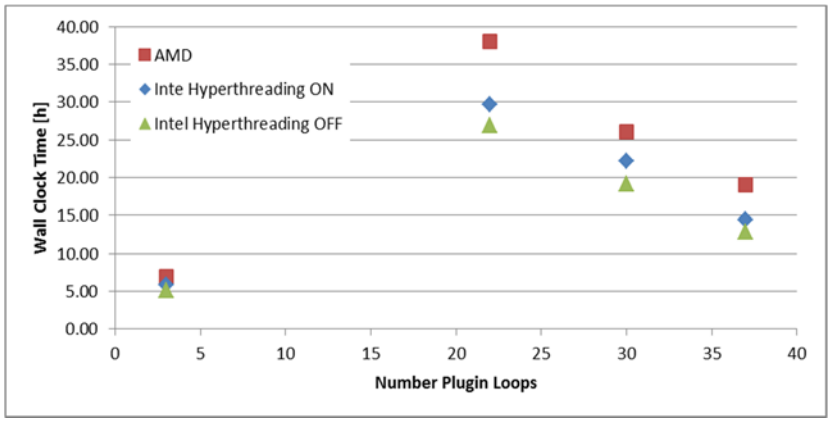


Figure 6 - Weak scaling of semiconductor device simulations using the finite element method (implementation in Sentaurus TCAD SDEVICE).

3.4 Applications of Hybrid GPU/Multi-Core Systems

During the first phase of the hardware acquisition process we have obtained a computing node that included two GPU processors (TESLA K40). We did initially test the GPUs for a number of software applications that we normally use for production runs.

We have performed an initial code porting of our standard transport Monte Carlo code. Based on the resulted obtained we have decided that the GPU architecture is unsuitable to run this kind of application due to the single-instruction multiple-data (SIMD) paradigm that cannot be matched to the software structure of the Monte Carlo applications.

We have subsequently used the GPUs for a specific FDTD packages that was provided to us for evaluation by Synopsys. As expected the speed-up for this application is significant, but unfortunately Synopsys no longer provides this application with this licensing option.

As a result we have not acquired any additional GPU nodes. We are currently testing a version of VASP that runs partially on GPUs and based on the outcome we will decide how to proceed with the use of this computer architecture

4 - Educational Activity

The equipment purchased through the 2014 DURIP award to the Boston University Computational Electronics group has supported the ongoing research activities of two post-doctoral associates and five PhD students actively involved in DoD funded programs. Besides supporting their ongoing research, the acquisition of the new computational resources provided impetus to investigate the development of computationally efficient software. For example, the increasing availability of machine time across the cluster led to the use of MPI and OpenMP to distribute programs across a number of physically separate compute nodes. These techniques have been integrated into existing software. Additionally, having removed the machine availability bottleneck, a significant effort was devoted towards developing codes and methods for automating designs to increase overall throughput. The students have also been able to present their work at several high-profile conferences (SPIE Photonics West, Defense, Security + Sensing, Optics and Photonics) where it is possible to engage researchers from various DoD organizations. As a result of on-going collaborative efforts, two PhD students have graduated and taken positions at DoD laboratories.

5 - Bibliography of Work Supported

The following DoD funded publications benefited from the equipment purchased through the 2014 DURIP award to the Boston University Computational Electronics Group. Although some of the publication do not specifically acknowledge the DURIP award, the work described has been performed using the hardware procured using DURIP funding.

5.1 Manuscripts Published

1) A.R. Wichman, B. Pinkie, E. Bellotti, “Negative differential resistance in dense short wave infrared HgCdTe planar photodiode arrays” IEEE Trans. Electron. Dev. **62**, pp 1208 (2015).
2) A. R. Wichman, B. Pinkie, E. Bellotti, “Dense array effects in SWIR HgCdTe photodetecting arrays” J. Electron. Mater. **44**, pp 3134 (2015).
3) H. Wen, B. Pinkie, E. Bellotti, “Direct and phonon-assisted indirect Auger and radiative recombination lifetime in HgCdTe, InAsSb, and InGaAs computed using Green’s function formalism” J. Appl. Phys. **118**, pp 15702 (2015).
4) B. Pinkie, A. R. Wichman, E. Bellotti, “Modulation transfer function consequences of planar dense array geometries in infrared focal plane arrays” J. Electron. Mater. **44**, pp 2981 (2015).
5) Hanqing Wen and Enrico Bellotti, “Optical absorption and intrinsic recombination in relaxed and strained InAs1-xSbx alloys for mid-wavelength infrared application”, Appl. Phys. Lett., 107, 222103 (2015)
6) Alexandros Kyrtsos, Masahiko Matsubara and Enrico Bellotti, “First-principles study of migration mechanisms and diffusion of carbon in GaN”, Journal of Physics: Conference Series **633** (2015) 012143.
7) B. Pinkie, E. Bellotti, “Numerical simulation of the modulation transfer function in HgCdTe detector arrays” J. Electron. Mater. **43,** pp 2864 (2014).

5.2 - Manuscripts in press

1) B. Pinkie, E. Bellotti, “A failure mode in dense infrared focal plane arrays” J. Electron. Mater. In press. (2015).

5.3 - Manuscripts under review

1) A. Kyrtsos, M. Matsubara and E. Bellotti, “Migration mechanisms and diffusion barriers of carbon and native point defects in GaN”, Submitted to Phys. Rev. B. Under review.
2) M. Matsubara and E. Bellotti, “A first-principles study of carbon-related energy levels in GaN: Complexes formed by substitutional/interstitial carbons and gallium/nitrogen vacancies”. Submitted to Phys. Rev. B. Under review.

Table 1 – Summary of equipment obtained through 2014 DURIP awarded to Boston University

Designation	Chassis	Power	Chipset	Processor	RAM	Disk	Network	Misc
32c AMD Compute Nodes								
ebnpho1	1U Server	1400W	AMD SR5690	4x8 core Opteron 6328 3.2GHz	32x16 GB 1600MHz ECC	500GB SATA	Infiniband	
ebnpho2	1U Server	1400W	AMD SR5690	4x8 core Opteron 6328 3.2GHz	32x16 GB 1600MHz ECC	500GB SATA	Infiniband	
20c Intel Compute Nodes								
ebnpho3	1U Server	700W	Intel C602	2x10 core Xeon E5-2690v2 3.0GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
ebnpho4	1U Server	700W	Intel C602	2x10 core Xeon E5-2690v2 3.0GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
ebnpho5	1U Server	700W	Intel C602	2x10 core Xeon E5-2690v2 3.0GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
ebnpho12	1U Server	700W	Intel C602	2x10 core Xeon E5-2690v2 3.0GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
ebnpho13	1U Server	700W	Intel C602	2x10 core Xeon E5-2690v2 3.0GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
64c AMD Compute Nodes								
ebnpho6	2U Server	1400W	AMD SR5690	4x16 core Opteron 6386 2.8GHz	16x16 GB 1600MHz ECC	500GB SATA	Infiniband	
ebnpho7	2U Server	1400W	AMD SR5690	4x16 core Opteron 6386 2.8GHz	16x16 GB 1600MHz ECC	500GB SATA	Infiniband	
ebnpho8	2U Server	1400W	AMD SR5690	4x16 core Opteron 6386 2.8GHz	16x16 GB 1600MHz ECC	500GB SATA	Infiniband	
Intel Chipset NVIDIA GPU Node								
ebnpho9	2U Server	1600W	Intel C602	2x10 core Xeon E5-2690v2 3.0GHz	8x16 GB 1866MHz ECC	250GB SATA	Infiniband	2x NVIDIA Tesla K40M GPU
24c Intel Compute Nodes								
ebnpho10	1U Server	700W	Intel C602	2x12 core Xeon E5-2690v3 2.6GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
ebnpho11	1U Server	700W	Intel C602	2x12 core Xeon E5-2690v3 2.6GHz	16x16 GB 1866MHz ECC	500GB SATA	Infiniband	
24TB Storage Node								
ebnphofsv	2U Server	920W	AMD SR5690	2x8 core Opteron 6328 3.2GHz	8x8 GB 1600MHz ECC	12x2 TB SATA	Infiniband	21TB RAID5 mountable XFS
18 port Infiniband Switch								
IBSwitch	1U Half length	120W					18xInfiniband	1.44Tb/s switching capacity
TOTAL:	20U	13.14kW		440 cores	3.8 TB RAM	29.25 TB Storage		

Appendix - 2

This section provides the results of the benchmark tests performed on various nodes, and their combinations, of the system to understand the scaling properties for DFT calculation using the code VASP. The system used for this test is a GaN supercell composed of 73 atoms with a Carbon interstitial. This is a prototype structure that we currently used to investigate various types of defects. The numerical model relies on an integration scheme based on four special k-points, and at least 391 bands (depends on band parallelization) and a total of 124416 plane waves. In the test we consider both conventional exchange correlation DFT-PBE and hybrid DFT-HSE functionals.

The following nodes have been used, both in combinations of similar (same processor type) and of different nodes:

AMD32: 2 nodes (2×32=64 cores)

AMD64: 3 nodes (3×64=192 cores)

INTEL20: 4 nodes (4×20=80 cores)

INTEL24: 4 nodes (4×24=96 cores)

TEST – 1 Standard DFT-PBE

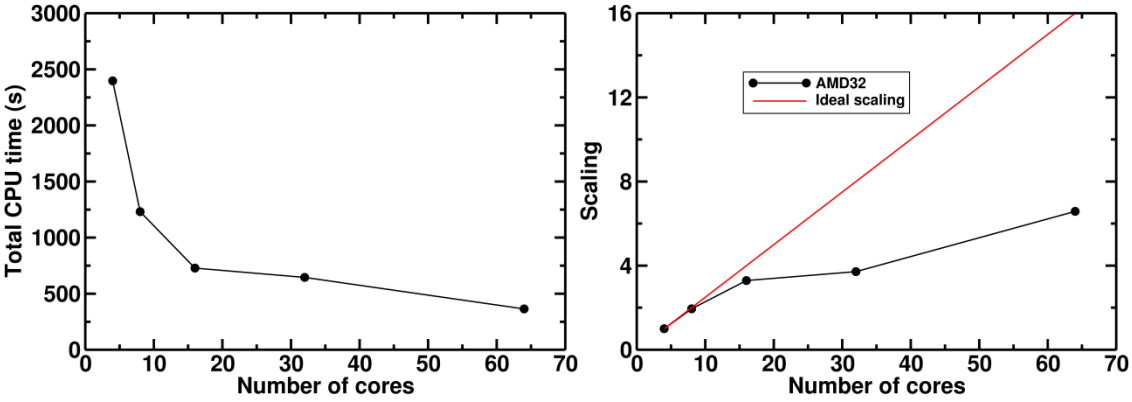
AMD32: AMD Opteron processor 6328, 3.2 GHz, 32 cores

Single node performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
AMD32	32	645.019	640.360	4.659	649.980
	16	728.065	723.676	4.389	732.042
	8	1230.987	1224.482	6.505	1235.485
	4	2397.054	2385.142	11.911	2402.595

2 nodes performance

2×AMD32	64 (2×32)	364.259	363.009	1.250	369.319
	32 (2×16)	400.947	400.045	0.902	405.854
	16 (2×8)	657.958	656.902	1.056	662.554
	8 (2×4)	1291.781	1290.758	1.023	1297.382



AMD64: AMD Opteron processor 6386 SE, GHz, 64 cores

Single node performance

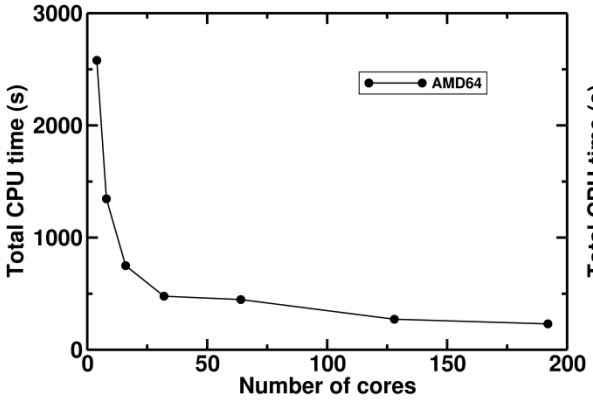
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
AMD64	64	447.762	441.818	5.944	456.069
	32	477.494	473.422	4.072	488.166
	16	749.804	744.640	5.164	753.843
	8	1345.821	1339.088	6.733	1350.025
	4	2580.176	2574.319	5.857	2586.285
	2	5167.265	5149.462	17.802	5177.814
	1	11679.138	11675.759	3.378	11706.126

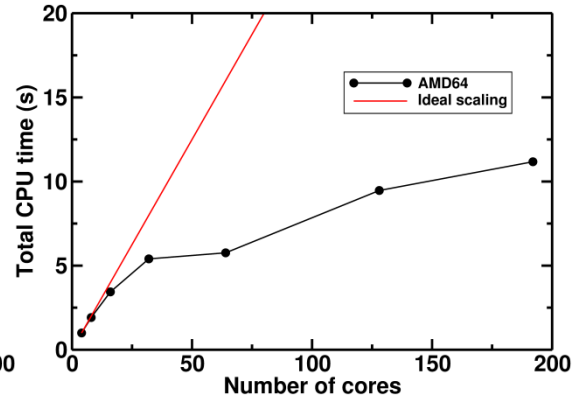
2 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
2×AMD64	128 (2×64)	272.594	270.801	1.793	276.982
	64 (2×32)	274.969	273.646	1.323	282.087
	32 (2×16)	408.190	407.184	1.006	412.016
	16 (2×8)	711.881	710.894	0.987	715.647

3 nodes performance

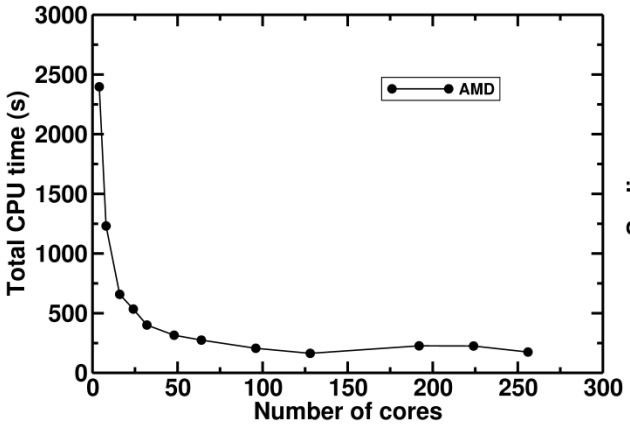
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
3×AMD64	192 (3×64)	230.929	228.668	2.261	237.480
	96 (3×32)	206.053	204.681	1.372	210.342
	48 (3×16)	315.557	314.346	1.211	320.147
	24 (3×8)	534.787	533.124	1.663	539.317
	12 (3×4)	983.020	980.218	2.802	987.582

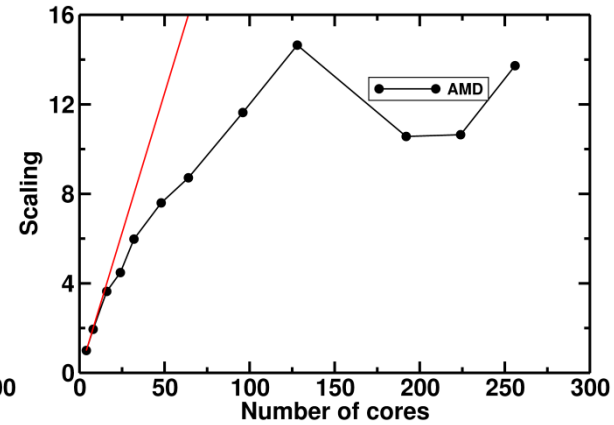




Combination of AMD32 and AMD64

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
3×AMD64+2×AMD32	256 (3×64+2×32)	174.612	171.934	2.679	178.441
3×AMD64+1×AMD32	224 (3×64+1×32)	225.228	222.148	3.080	231.679
2×AMD64+2×AMD32	192 (2×64+2×32)	226.872	224.766	2.106	231.925
2×AMD64+1×AMD32	160 (2×64+1×32)	279.675	210.356	69.319	356.723
1×AMD64+2×AMD32	128 (1×64+2×32)	267.714	212.580	55.135	357.147
3×AMD64+2×AMD32	128 (3×32+2×16)	163.652	161.941	1.711	243.939
1×AMD64+1×AMD32	96 (64+32)	392.472	389.250	3.223	399.713





INTEL20: Intel Xeon E5-2690 v2, 3.0 GHz, 20 cores (40 with hyper-threading (HT))

Single node performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	40 (with HT)	534.785	529.984	4.800	541.179
	20	468.873	466.119	2.754	476.878
	10	671.798	669.653	2.145	680.510
	8	772.269	770.256	2.013	780.571
	5	1180.662	1172.341	8.322	1187.173
	4	1415.014	1408.080	6.934	1484.055

2 nodes performance

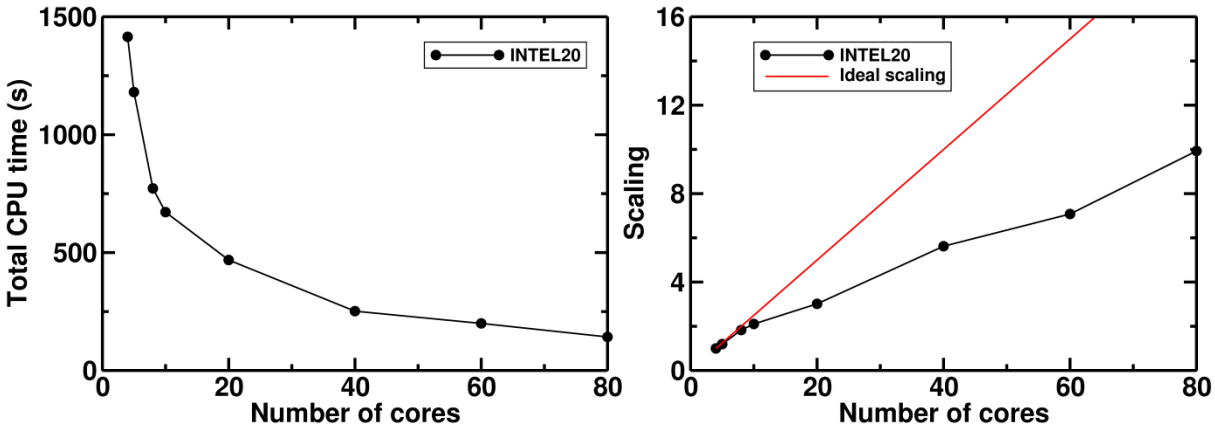
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	80 (2×40, HT)	287.626	286.523	1.103	294.537
	40 (2×20)	251.713	250.944	0.769	261.120
	20 (2×10)	354.498	353.249	1.249	364.755

3 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	60 (3×20)	199.821	199.052	0.769	326.152

4 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	80 (4×20)	142.494	141.865	0.629	263.431



INTEL24: Intel Xeon E5-2690 v3, 2.6 GHz, 24 cores

Single node performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	24	387.360	384.776	2.585	391.474
	12	560.915	557.547	3.367	565.338
	8	782.074	780.268	1.806	785.230
	6	992.365	988.981	3.384	1006.985
	4	1385.792	1383.808	1.985	1389.080
	3	1896.226	1887.266	8.960	1899.875

2 nodes performance

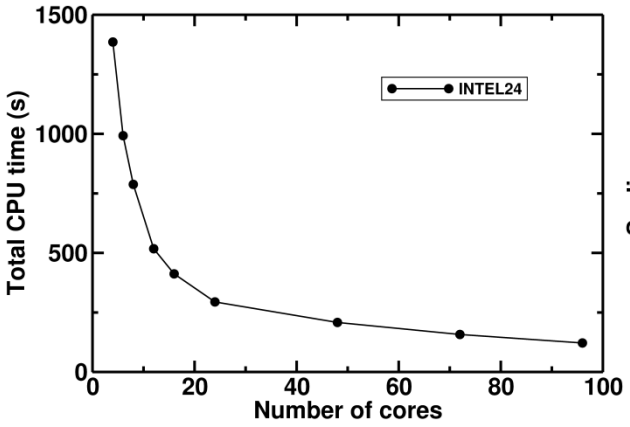
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	48 (2×24)	208.015	207.454	0.561	211.831
	24 (2×12)	294.281	293.700	0.581	297.542
	16 (2×8)	412.098	410.901	1.198	415.646
	12 (2×6)	517.695	516.080	1.616	522.117

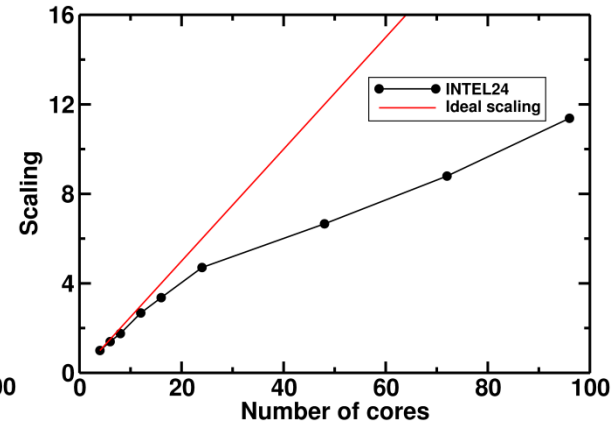
3 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	72 (3×24)	156.960	155.623	1.337	253.919

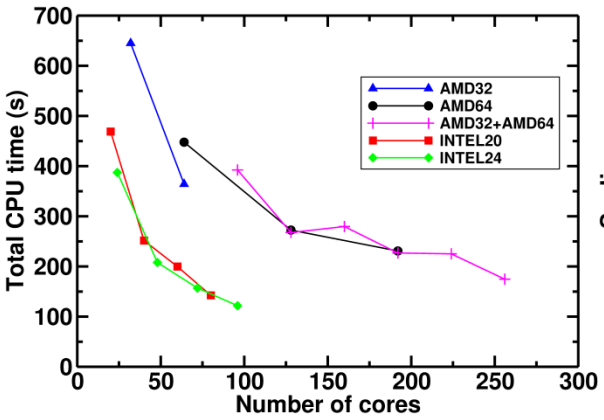
4 nodes performance

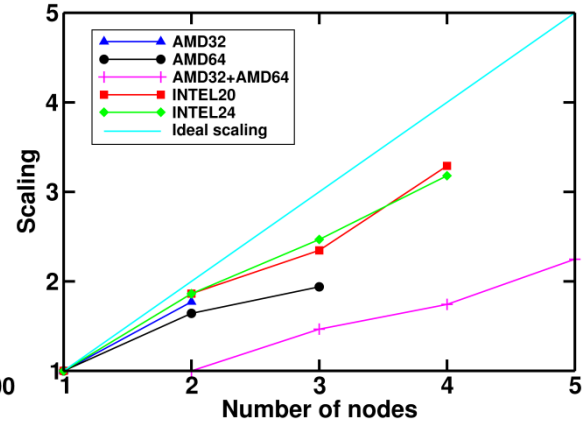
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	96 (4×24)	121.798	120.949	0.850	235.431





Benchmark Summary for DFT PBE





Nodes with Intel Xeon E5-2690 v3, 2.6 GHz and 24 cores counts are the fastest systems

Nodes with INTEL processors scale better than AMD processors.

TEST – 2 HYBRID DFT-HSE

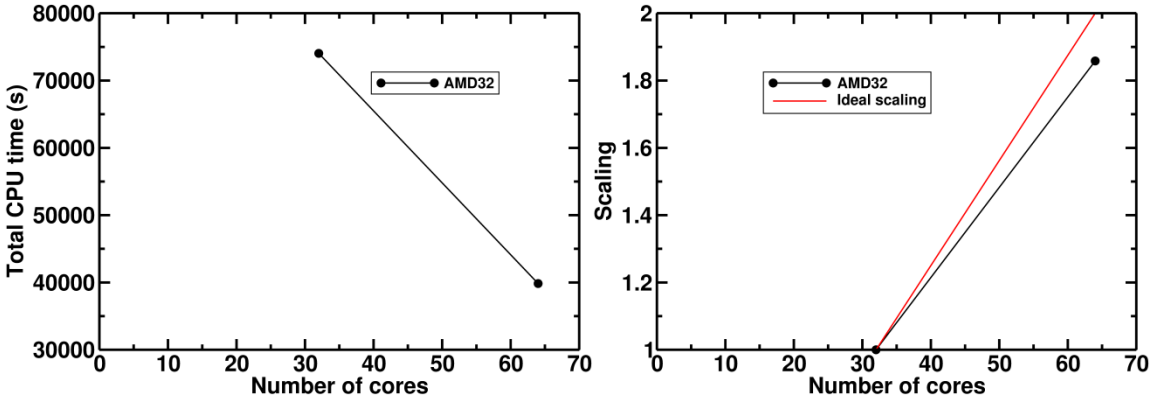
AMD32: AMD Opteron processor 6328, 3.2 GHz, 32 cores

Single node performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
AMD32	32	74040.359	73864.176	176.181	74200.343

2 nodes performance

2×AMD32	64 (2×32)	39844.418	39813.236	31.182	39979.152



AMD64: AMD Opteron processor 6386 SE, GHz, 64 cores

Single node performance

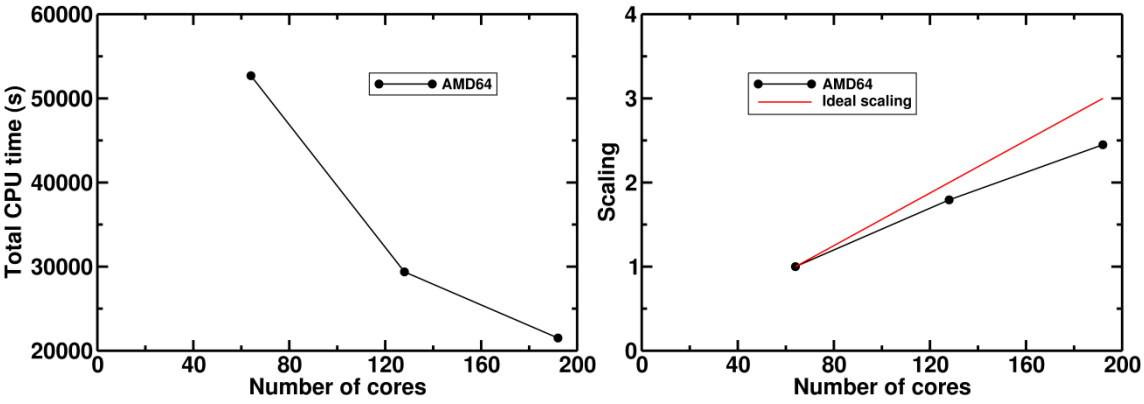
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
AMD64	64	52693.699	52560.150	133.548	52851.703
	32	57163.688	57112.549	51.138	57247.225

2 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
2×AMD64	128 (2×64)	29378.096	29373.741	4.354	29776.784

3 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
3×AMD64	192 (3×64)	21514.910	21496.408	18.502	21690.149
	96 (3×32)	21564.871	21534.848	30.022	21597.838



Combination of AMD32 and AMD64

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
1×AMD64+1×AMD32	96 (64+32)	(40450.645	37527.325	2923.320	40621.032
1×AMD64+2×AMD32	128 (1×64+2×32)	(29535.982)	23022.852	6513.130	29633.719
2×AMD64+1×AMD32	160 (2×64+1×32)	(22596.400)	20279.796	2316.604	22662.790
2×AMD64+2×AMD32	192 (2×64+2×32)	(22757.861)	18324.464	4433.397	22901.310
3×AMD64+1×AMD32	224 (3×64+1×32)	ERROR	-	-	-
3×AMD64+2×AMD32	256 (3×64+2×32)	15981.390	12504.571	3476.818	16071.557

INTEL20: Intel Xeon E5-2690 v2, 3.0 GHz, 20 cores (40 with hyper-threading (HT))

Single node performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	40 (with HT)	63815.359	63685.676	129.684	64181.441
	20	58212.188	58148.166	64.023	58326.631

2 nodes performance

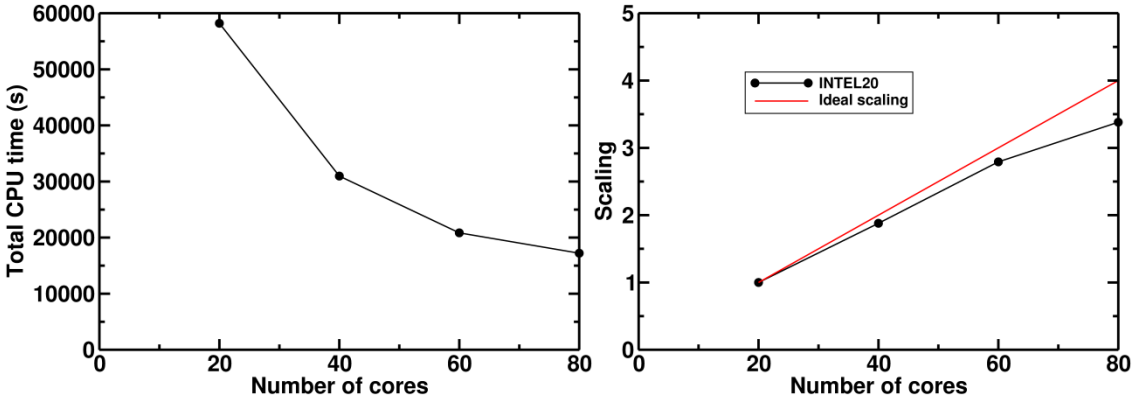
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	80 (2×40, HT)	32714.654	32683.697	30.957	32871.050
	40 (2×20)	30975.283	30946.126	29.157	31067.303
	20 (2×10)	44606.312	44576.150	30.161	44686.182

3 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	60 (3×20)	20847.502	20830.404	17.097	20899.068

4 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL20	80 (4×20)	17218.004	17213.010	4.993	17260.109



INTEL24: Intel Xeon E5-2690 v3, 2.6 GHz, 24 cores

Single node performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	24	51349.039	51297.965	51.074	51450.314

2 nodes performance

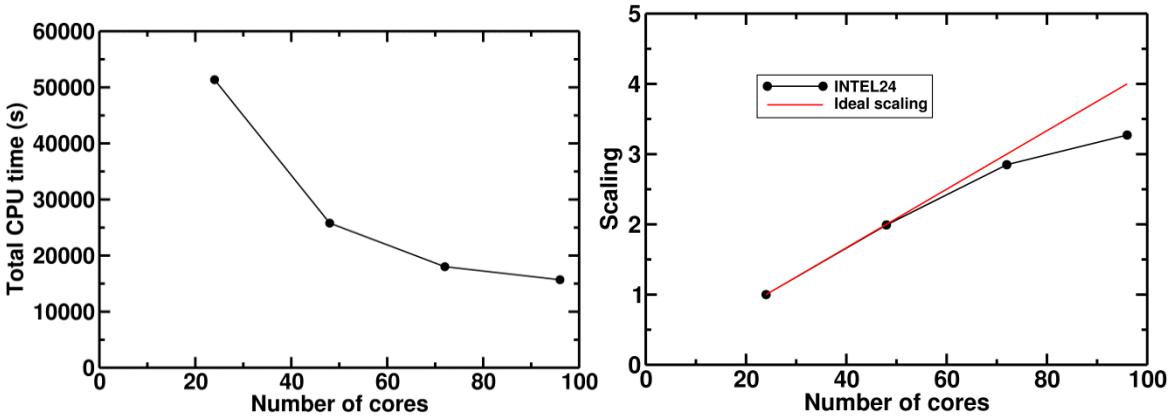
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	48 (2×24)	25787.641	25776.344	11.296	25846.938
	24 (2×12)	36766.418	36750.616	15.802	36780.497

3 nodes performance

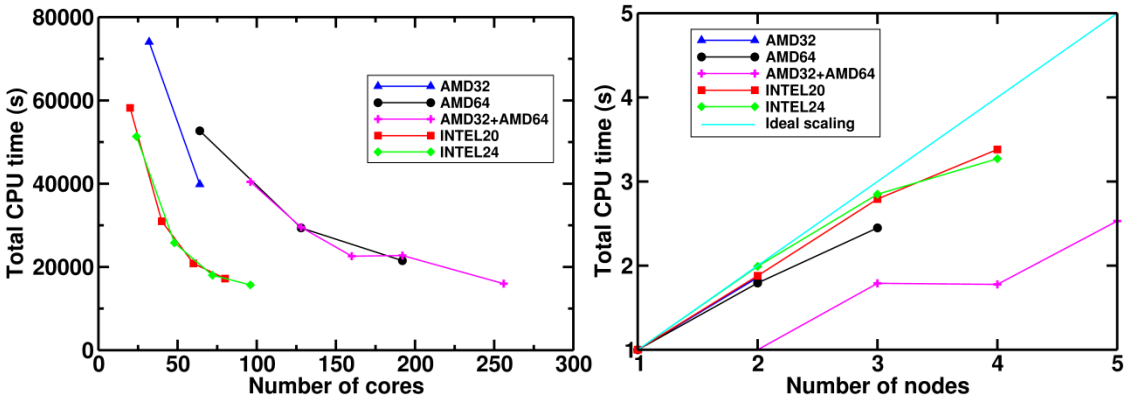
Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	72 (3×24)	18023.561	17991.738	31.823	18063.586

4 nodes performance

Node	# of cores	Total CPU time	User CPU time	System CPU time	Elapsed time
INTEL24	96 (4×24)	15696.682	15651.570	45.112	15759.997



Benchmark Summary for DFT-HSE



Nodes with Intel Xeon E5-2690 v3, 2.6 GHz and 24 cores counts are the fastest systems

Nodes with INTEL processors scale better than AMD processors.