CAC bulletin

The Newsletter of the Cloud and Autonomic Computing Center Vol. 4, Spring 2012

UNIVERSITY of FLORIDA
THE UNIVERSITY OF ARIZONA
RUTGERS
MISSISSIPPI STATE UNIVERSITY
NATIONAL SCIENCE FOUNDATION
intel
ERDC US Army Engineer Research and Development Center
xerox
Raytheon
Microsoft
AVIRTEK Autonomic Management Solutions

In this issue:

Save the date! The Spring 2012 CAC workshop will be held June 14 & 15 at the University of Florida in Gainesville. Stay tuned for details about the upcoming meeting and Center happenings at http://www.nsfcac.org.

Dear CAC collaborators and friends,

I am happy to share with you important news about the CAC center. Early this year, the National Science Foundation and CAC's Industry Advisory Board approved the Change of the name of the center to the Cloud and Autonomic Computing (CAC) Center, a change reflected in our new logo shown below. The new name projects more accurately both the nature of work already underway within CAC and the strategic directions for future work, as described in previous newsletters. Currently, the CAC center is the only NSF Industry-University Cooperative Research Center whose name explicitly identifies cloud computing as one of the key areas of emphasis of center activities.

Reflecting the strategic focus on cloud systems, on June 14 & 15 the CAC center will host a workshop entitled "Cloud Computing: from Cybersecurity to Intercloud" at the University of Florida. As the name suggests, the workshop will have sessions covering a range of topics—namely, cybersecurity; data clouds; cloud applications and services; cloud management; cloud standards and testbeds; and intercloud computing and cloud federation. Talks and panels by CAC members and several internationally known invited speakers, including Jeff Kephart, Manager of the Agents and Emergent Phenomena Program at IBM Research; John Howie, Director of CloudCERT and Chair of the the Cloud Security Alliance's Working Group; and Reagan Moore, Chief Scientist for Data Intensive Cyber Environments (DICE) at UNC's Renaissance Computing Institute (RENCI). More details about the workshop can be found at http://nsfcac.org/upcoming-meetings. If you are interested in attending the workshop, please contact CAC coordinator Julie Walters at jooly@acis.ufl.edu or myself at fortes@ufl.edu.

CAC
Cloud and Autonomic Computing Center

In preparation for the renewal of the NSF award that partially supports the center, we are in the process of recruiting new industry members. In particular, we are looking for industrial members who would like to both participate and be recognized as center anchors. As the name suggests, anchor members will play a special role in enabling clusters of activity in certain technical areas by holding several center memberships that support projects in those technical areas. If you are interested in becoming an anchor member or a regular member at one or more of the CAC sites, or would just like to find out more about CAC memberships, please feel free to contact me at fortes@ufl.edu or any of the CAC co-directors or site directors.

Some of you may be aware that the CAC center is looking into adding new university sites and thus expanding the number of research areas it covers within cloud and autonomic computing. At the time of this writing, we cannot yet reveal any new additions, but we are confident that we will be able to share good news about ongoing proposals for such sites in the next newsletter. Until then, please enjoy reading the contents of this newsletter, which includes feature articlesby researchers from each of the CAC sites, several reports on ongoing projects, and our recurring features on CAC personnel, news about the latest and the greatest within the CAC center, meetings between CAC personnel and the world at large, and recent publications.

As always, I welcome your feedback and inquiries. Until we meet again, enjoy your summer.

Sincerely,

José Fortes, CAC Director, on behalf of CAC co-directors and site directors:

University of Florida: Renato Figueiredo [renato@acis.ufl.edu]

University of Arizona: Salim Hariri and Ali Akoglu [{hariri, akoglu}@ece.arizona.edu]

Rutgers University: Manish Parashar and Dario Pompili [{parashar, pompili}@cac.rutgers.edu]

Mississippi State University: Ioana Banicescu [ioana@cse.msstate.edu]

Investigating the Use of Virtualization Technologies for Disaster Mitigation and Recovery, by Maurìcio Tsugawa, UF[1]

On the 11th of March 2011, a massive earthquake hit Japan, triggering a powerful tsunami that caused several nuclear accidents. Due to a large-scale electrical power failure, many IT services were unavailable after the disaster. As many critical infrastructures and systems depend on IT services, the Tohoku earthquake highlighted the need for IT services that are resilient to catastrophic events.

Virtualized IT infrastructure entails running services on virtual machines (VMs) deployed on physical systems (e.g., datacenters and clouds) and potentially distributed across multiple geographical locations. As such, it is possible to take advantage of VM mechanisms to move from one server to another without interrupting applications in execution, or to save the entire VM's system state to recover at a later time. A virtualized environment offers many advantages over conventional systems, such as improved hardware utilization and resource management, isolation among multiple consumers, and migration capability. In this project, the scalability of concurrent VM migration—transferring the execution environment of a VM from one physical server to another in order to move IT services from a site affected by a disaster to a safe location—is investigated. This international collaboration between CAC researchers at the University of Florida and the Japanese National Institute of Advanced Industrial Science and Technology (AIST) takes into account the information collected by AIST to assess the availability of IT services and infrastructure after the 2011 Tohoku Earthquake.

Typical disaster recovery (DR) practices and services for IT systems rely on proactive data backup and system state checkpointing. Although reliable, the cost of such services can be very high; with a best-effort model, VM migration can be used as a low-cost alternative. In a local area network (LAN) environment, the storage and network components of a VM migration are straightforward to address. The main challenge with storage is that both source and destination physical hosts need to have access to the VM's disks, where the guest OS and data are stored. Shared storage or file systems are used in LAN environments, but the approach is not appealing when wide-area networks (WAN) are involved, due to performance and connectivity limitations. Moreover, shared file systems are not reliable in disaster scenarios, especially when storage servers are located at or near disaster sites.

Ideally, the network configuration of a VM should remain unchanged after migration—something that is naturally achieved when migration is performed within a single broadcast domain but is difficult when subnet boundaries are crossed. This work assumes that the necessary network technology is in place to support the migration of VMs across WANs.

The goal of this work is to evaluate the scale of virtualized infrastructures that can be moved under restricting conditions, such as limited time with electrical power, varying network connectivity and performance, and long distances. Five Japanese institutions were interviewed to survey the IT infrastructure damages suffered due to the earthquake. The results are summarized in the table below. An important observation is that damage to IT equipment due to the earthquake was minimal. This can be attributed to the fact that constructions and installations in the affected area were engineered to be earthquake-resistant. In terms of electrical power, uninterruptible power supply (UPS) systems were able to keep datacenters operational for ten to sixty minutes. At Iwate Prefectural University, power generators kept IT equipment running until the regular electrical power distribution was restored.

	Distance from epicenter (km)	Electrical Power	Network connectivity
Iwate Prefectural University	220	Power generators used; power uninterrupted	Redundant links kept-connectivity alive
Tohuku University	150	UPS supplied tens of minutes	Lost after 28 minutes due to SINET shutdown
High Energy Accelerator Research Organization	310	UPS supplied tens of minutes	Not available
University of Tsukuba	310	UPS supplied tens of minutes	Lost immediately
AIST	310	UPS supplied 15 to 60 minutes	Available for 60 minutes

When an intense earthquake strikes, it is expected that a large area is affected and that migrating IT services to a distant location may help to maintain performance levels. However, the larger the distance, the less network throughput is expected, which in turn affects the migration performance. The experimental setup consists of servers running at AIST, the University of Florida (UF), and the San Diego Supercomputing Center (SDSC) facilities. Servers at each site have been exclusively allocated for this work. The servers use Intel-based multi-core x86_64 CPUs with at least 16GB of memory. Linux kernel 2.6.32 with Kernel-based Virtual Machine (KVM) enables system machine virtualization and live migration.

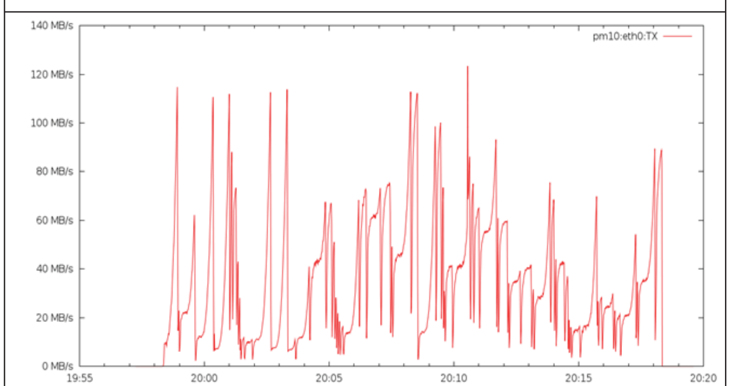

Figure: TCP throughput between AIST and UF

The round-trip latency between AIST and UF is around 170ms, and between UF and SDSC is around 128ms. The figure above illustrates the TCP network throughput measured around 8PM JST (8AM EST). Although different network param-

[1] Feature article adapted from a paper entitled " "On the Use of Virtualization Technologies to Support Uninterrupted IT Services," published with co-authors José Fortes, Renato Figueiredo, Takahiro Hirofuchi, Hidemoto Nakada, Ryousei Takano, to be presented at the IEEE ICC2012 Workshop on Re-think ICT Infrastructure Designs and Operations

eters and queue size adjustments were tried, it was not possible to obtain consistent network performance. The throughput can be as high as 117MB/s and can drop to below 5MB/s at times. This network behavior directly affects VM migration performance, so multiple runs of each experiment were executed. It is important to note that oscillating network performance is common in disaster scenarios due to a high probability of congested links. There are opportunities for autonomic computing research to adjust VM migration parameters to best use available network resources. A very stable network environment was observed on a UF-SDSC link with TCP throughput of approximately 118MB/s, which is the throughput of a gigabit Ethernet LAN. These resources are part of the FutureGrid testbed connected via the National LambdaRail and Florida LambdaRail backbone networks.

The VM memory update rate affects migration time because the migration finishes only when both source and destination servers have exact copies of the VM's memory pages. The plot at below left illustrates the measured migration time of a VM transferred from AIST to UF. Each line represents a different run of the experiment. As expected, due to differences in network throughput over time, large variance is observed. Migration time increases substantially when VMs update memory during the migration process. In disaster recovery, assuming that downtime is not important for applications, it may be better to suspend the VMs from running (consequently avoiding the memory updates), and get migration time equivalent to that of idle VMs.

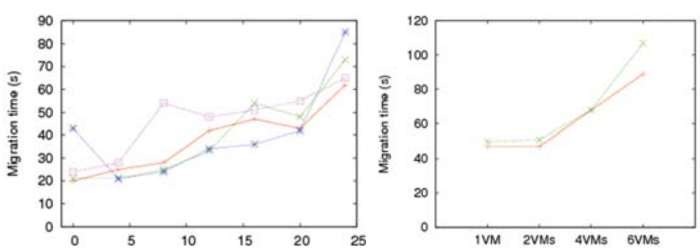

Figure: (right) 512MB VM migration time from AIST to UF with varying memory update speed; (left) concurrent migration of VMs from AIST to UF. VM memory size was fixed at 512MB, and the update size was fixed at 4MBytes/s.

During disaster recovery, it is desirable to concurrently migrate as many VMs as possible. However, network bandwidth limits the number of VMs that can be efficiently migrated. The plot at above right shows the performance of migrating concurrent VMs from AIST to UF. For this set of experiments, VMs with 512MB of memory and update rates of 4MBytes/s were used. UF to SDSC migrations required substantially less time.

Available network bandwidth plays a key role in VM migration performance. In general, TCP is used to transfer the execution states of VMs since reliable transport is required. Although TCP is known not to offer good throughput in high-latency links, concurrent TCP connections better utilize available bandwidth. This is one of the reasons that concurrent migrations of VMs outperform a single VM migration. Too many transfers can saturate and congest the network and cause migration performance degradation. The approximate migration throughputs for different numbers of concurrent VMs were computed and are listed in the table below.

No. of VMs	AIST-UF	UF-SDSC
1	74 VMs/hour	194 VMs/hour
2	147 VMs/hour	378 VMs/hour
4	212 VMs/hour	612 VMs/hour
6	220 VMs/hour	654 VMs/hour

In conclusion, in order to sustain IT services in the event of a disaster by means of live VM migration, a large number of VMs must be moved. For this approach to work, several conditions need to be met in a potentially chaotic disaster site: (a) IT equipment must be operational during and after a disaster; (b) there is a window of time in which electrical power is available; (c) there is a window of time in which network connectivity is maintained; and (d) IT services can be encapsulated into virtualized systems and present a small memory and storage footprint. Experiments conducted in this work show that when the fourth condition is met, tens to hundreds of VMs can be migrated from a disaster site to a safe location. This is an encouraging result as it indicates that many useful services can be kept alive or quickly recovered in the event of a disaster. The results are also significant due to the fact that experiments were conducted across very long distances (from Japan to the east coast of the US). In practice, VMs should be migrated over shorter distances, where better network and VM migration performance are expected. Future work includes experiments with different sites (including shorter distances), high-performance WAN connections (e.g., the use of the National LambdaRail), larger scale both in number of servers and migrating VMs and in concurrent migration from one to many sites. These experiments are expected to provide the basis for the design of a system that can efficiently migrate virtualized IT infrastructures by automatically adjusting relevant VM migration parameters.

Cloud Paradigms and Practices for CDS&E—Explorations Using CometCloud, by Manish Parashar, RU

Cloud computing is revolutionizing the enterprise world, much as the Internet did in the recent past. Clouds are fundamentally changing how enterprises think about IT infrastructure, both internally and externally, by providing on-demand access to always-on computing utilities; an abstraction of unlimited resources; a potential for scale-up, scale-down and scale-out as needed; and for IT outsourcing and automation. Finally, clouds provide a usage-based payment model where users essentially "rent" virtual resources and pay for what they use. Typically underlying these cloud services are consolidated and virtualized data centers that exploit economies of scale to provide attractive cost-benefit ratios. In spite of being in its early stages, cloud computing is already reshaping the IT world—a recent report by Gartner estimates that cloud services will be a $150 billion industry by 2015. At the same time that cloud computing is redefining IT, extreme data and compute scales are transforming science and engineering research by enabling new paradigms and practices—those that are fundamentally information- or data-driven and collaborative. Complex computational and data-enabled science and engineering (CDS&E) applications are providing unprecedented opportunities for understanding and managing natural and engineered systems and providing unique insights into complex problems.

It is expected that cloud services will join more traditional cyberinfrastructure components—such as high-performance computing systems, clusters and grids—as part of NSF's Cyberinfrastructure Framework For 21st Century Science and Engineering (CIF21) Program supporting scientific exploration and discovery. Analogous to their role in enterprise IT, clouds can enable the outsourcing of many of the mundane and tedious aspects of research and education, including the deployment, configuration and management of infrastructure, allowing scientists to focus on science. Cloud services and associated standardization can also improve productivity, facilitate sharing of research results, and enable the reproducibility of associated computations. Furthermore, clouds can democratize access to computational and data resources (by providing access to researchers who don't have adequate local infrastructure), a benefit that has been shown to significantly impact research productivity.

However, it is also critical to look beyond the benefits of outsourcing and understand application formulations and usage modes that are meaningful in a hybrid HPC/grid + cloud cyberinfrastructure—for example, how emerging data- and compute-intensive application workflows can effectively utilize these resources, and how this hybrid cyberinfrastructure can enable new practices in science and engineering. We have used the CometCloud framework developed at the Rutgers CAC site to explore such usage modes.

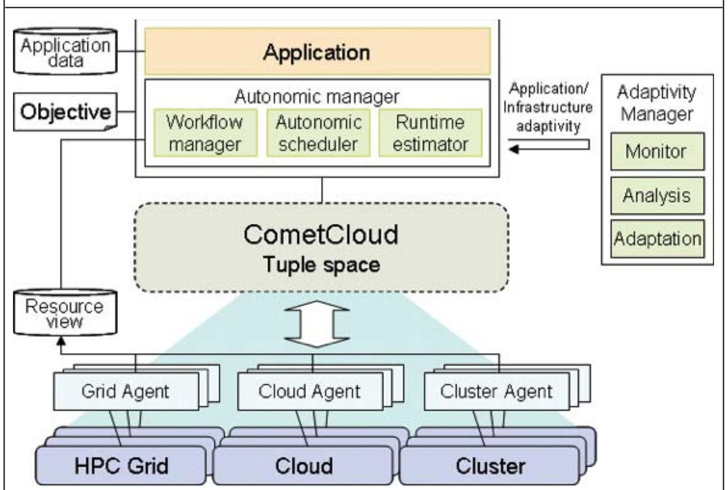

Figure: Architectural overview of the autonomic application management framework and supported platforms

CometCloud is an autonomic computing engine that enables dynamic and on-demand federation of clouds and grids as well as the deployment and robust execution of applications on these federated environments. It supports highly heterogeneous and dynamic cloud/grid infrastructures, enabling the integration of public and private clouds and autonomic cloudbursts, i.e., dynamic scale-out to clouds to address changing requirements resulting from heterogeneous and dynamic workloads, spikes in demand, and other extreme circumstances. We have used CometCloud to demonstrate how a hybrid HPC/grid + cloud cyberinfrastructure can be effectively used to support real-world science and engineering applications on existing cyberinfrastructure and on public and private clouds. Our experiments have explored a broad set of federated resources spanning existing public commodity clouds, grids, supercomputing resources, and local clusters. We have also explored a broad set of applications classes including business intelligence, financial analytics, oil reservoir simulations, medical informatics, and document management. Specifically, we explore three key usage modes listed below:

HPC in the Cloud using CometCloud: HPC in the cloud explores outsourcing entire applications to current public and/or private cloud platforms. We have studied the performance and performance fluctuation of the ensemble Kalman filter (EnKF)-based history-matching workflow on Amazon EC2 and FutureGrid. We have also explored the Replica Exchange Molecular Dynamics application, which traditionally has not been well suited to cloud environments. Specifically, we reformulated this application to be asynchronous and resilient so that it can be effectively run on public and private clouds using the messaging and coordination abstractions provided by CometCloud.

HPC plus Cloud using CometCloud: HPC + Cloud explores scenarios where clouds can complement HPC/grid resources to support science and engineering application workflows—for example, to support heterogeneous requirements or unexpected spikes in demand. We have used CometCloud to explore meaningful usage modes for a hybrid HPC plus cloud infrastructure. In particular, we used a reservoir-characterization application workflow, which uses the EnKF for history matching as the driving application, and we complemented TeraGrid resources with EC2 public cloud instances.

HPC as a Cloud using CometCloud: HPC as a cloud focuses on exposing HPC/grid resources using elastic on-demand cloud abstractions, aiming to combine the flexibility of cloud models with the performance of HPC systems. In an early experiment, we explored how a cloud abstraction can be effectively used to provide a simple interface for current HPC resources and support real-world applications. In particular, we experimentally validated the benefits of the cloud paradigm, such as ease of use and dynamic allocation, and its application to supercomputers, specifically on an IBM Blue Gene/P system. This experiment was demonstrated at the 4th IEEE SCALE Challenge, and received the first place award. During the experiment, Blue Gene/P resources varied from 640 to 22,016 processors, spanning two Blue Gene systems on two different continents.

Aggressive cloud computing technology development has resulted in many different classes of cloud service that provide attractive solutions for many different types of business applications. It is clear from previous research and our own experiences that there are real benefits in using cloud computing abstractions as part of a hybrid cyberinfrastructure to support CDS&E—for example, to simplify the deployment of applications and the management of their execution, to provide more attractive cost/performance ratios, and to improve their efficiency, effectiveness and/or productivity. However, before CDS&E can fully realize the potential benefits of a hybrid cyberinfrastructure that integrates cloud services, several research challenges remain spanning multiple levels, including algorithms, programming abstractions and systems, middleware, application formulation, security, and autonomic management.

Autonomic Cyber-Physical System for Biosphere 2 Landscape Evolution Observatory Testbed, by Salim Hariri, UA

Our understanding of how coupled water, energy and carbon cycles will respond to future climate scenarios is currently limited by uncertainties in observational studies and in process-based numerical models. An increased understanding of behaviors and interactions between different components of Earth systems requires the development of a new observational, experimental, and computational cyber-physical system testbed. The next generation of studies of highly complex Earth systems requires several components: first, an experimental infrastructure that encompasses realistically complex Earth systems is needed to allow for precise manipulation of these systems. Second, these studies require an information integration and processing system to ingest the massive datasets that capture large-scale dynamic Earth system complexity, and to instantly process and update it in order to develop dynamic, data-driven models that will drive the next set of experiments. Finally, the data, models, and knowledge generated need to be easily accessible to the entire scientific and educational community.

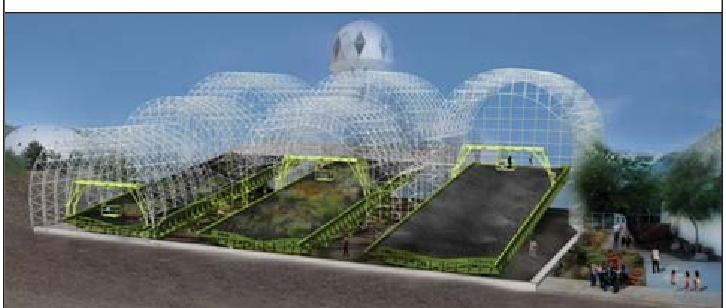

Figure: Artist's rendering of the Biosphere 2 Landscape Evolution Observatory

The newly developed Biosphere 2 Landscape Evolution Observatory (LEO) consists of three full-scale physical models of complex landscapes and a parallel effort to develop the process-based Terrestrial Integrated Modeling System (TIMS). Our research approach will tightly integrate cyberinfrastructure resources (high-performance computing and storage resources, high-speed networks, autonomic computing tools, and sensor and effector technologies) with the LEO facility to realize a fully controllable and observable cyber-physical system testbed to bridge the gap between data collection and development of Earth science models. This will enable us to dynamically couple and autonomically configure, in real-time, complex and customized interactions between ecohydrological models, environmental sensors and samplers, and experimental controls and protocols. The autonomic cyber-physical system (ACPS) architecture will be designed such that it will provide value for future complex experimental programs and for the rapidly developing community of interdisciplinary, field-based, environmental observatories such as the NSF Critical Zone Observatory projects.

The Biosphere 2 Landscape Evolution Observatory: The cyber-physical system developed in this project will significantly enhance and accelerate research and discovery in the LEO, the first large-scale, controllable Earth science experiment ever built. LEO has no analogue in the world in terms of physical scale, level of environmental control, and density and diversity of measurements in an Earth science infrastructure. LEO consists of three watersheds that will be completed by December 2012. Each of the three LEO watersheds has a surface area of 360 square meters, has a convergent form like "zero-order" basins (the portion of the landscape above the highest river channel-head) found in nature, and stands over 8 meters tall, filling a 0.2-hectare greenhouse facility at Biosphere 2, depicted in the rendering at left. This facility stands poised to contribute fundamental, transformative knowledge of how the terrestrial water cycle operates, how biological systems modify physical systems, and how these processes interact at realistic spatial scales. Our primary goals in this endeavor are, first, to observe the development of hydrologic flow paths and physical structure driven by physical, chemical and microbial transformation of an initially homogenous and relatively isotropic substrate, and, second, after 2–3 years, to observe the co-evolution of physical and biological systems by introducing plants and measuring ecosystem development. The range of ancillary goals and research activities on these landscapes includes studies of hydrology, biogeochemistry, geomorphology, microbiology, ecology, atmospheric science, and soil physics. The LEO facility will quantify hydrological partitioning in real time and allow dense measurement of the states and fluxes of water, carbon, and energy in an environmentally controlled facility. Each landscape is a realistic physical system with a remarkable embedded environmental sensor network (see photos below). Additional measurement capabilites within the LEO include quantification of precipitation, atmospheric temperature, humidity, atmospheric chemistry, soil water outflow, soil overland flow, detailed topographic change, light, radiation, soil gas fluxes, leaf-level gas exchange, and subsurface imaging.

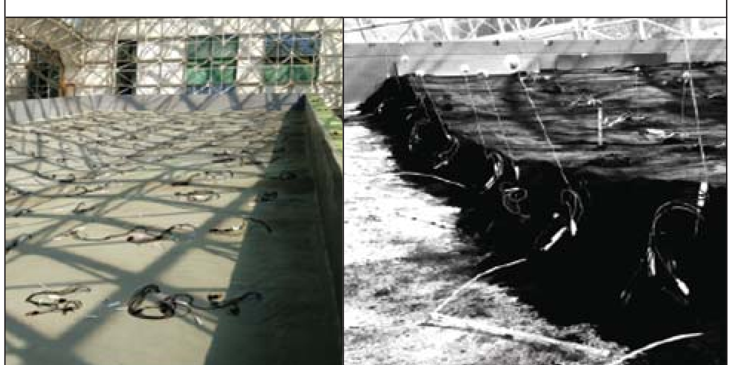

Figure: Sensor networks installed in the Landscape Evolution Observatory

Our approach to integrating and coupling the numerical model (TIMS) with the unique experimental infrastructure of LEO is based on the development of an adaptive and self-managing cyber-physical system, depicted in the figure below. The ACPS architecture is based on an innovative autonomic computing design paradigm inspired by nature and biological systems. The proposed ACPS research will focus on robust data acquisition, automated control and management of sensors, adaptive parameter estimation for model selection, model-based data analysis, and online analysis and visualization.

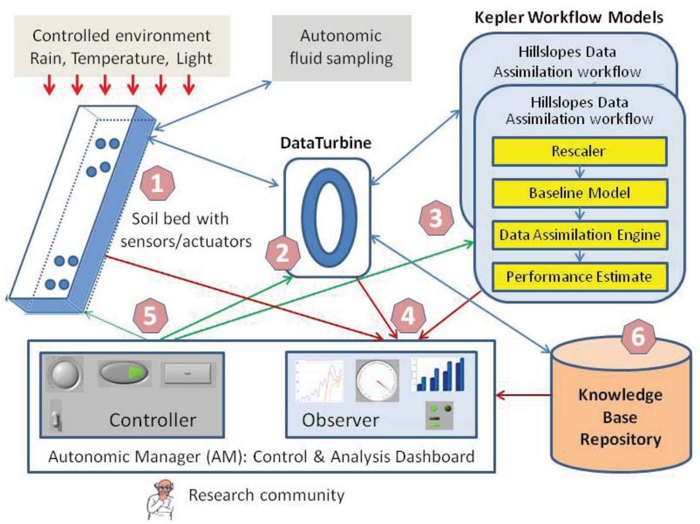

Figure: The ACPS architecture for LEO research

The Autonomic Manager (AM) dashboard provides multidisciplinary researchers with the tools required to dynamically set up and configure sensor networks, data to be collected in support of a given experiment (Step 5 above), and the coupled models (Step 3) that will be used to guide the experiments. To maximize the scientific benefit of measurements made in the LEO watersheds, we are developing a Kepler-based autonomic workflow for iteratively integrating measurements with numerical model simulations. This workflow will integrate component models for subsurface water flow, subsurface biogeochemistry, surface water flow, sediment transport, and land-surface vegetation and energy exchange processes using a modular, parallel structure. The numerical models will be tested and improved via iterative estimation and data analysis in real time during the course of the physical experiment. The coupled model and software infrastructure will be used by LEO researchers and will be shared with the broader community for use in similar projects that integrate geospatial measurements and models (e.g., NSF Critical Zone Observatories) (Step 6).

Requirements of a Generic Framework for Service Deployment in a Cloud Infrastructure, by Rajat Mehrotra, MSU

Clouds are increasingly used in hosting financial, health, e-commerce, and other mission-critical IT services. According to a survey[1] published in 2011, many organizations are migrating from a typical data center environment towards cloud infrastructure. In the same survey, however, only 26% of surveyed organizations reported that performance of their applications improved in a cloud infrastructure, down from 33% in 2009 (see figure below). The primary reason for decreased application performance in the cloud is the relative complexity associated with configuring resources required to host applications. Excessive use of virtualization technologies by cloud providers on a single physical machine can also degrade the quality of service (QoS) of deployed applications.

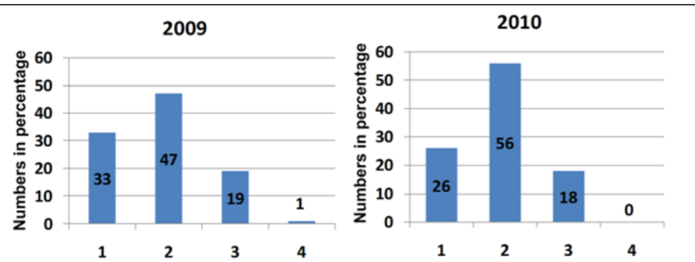

Figure: 2009 survey results. Columns represent percentage of respondents reporting application performance (1) improved in the cloud vs. in traditional datacenters, (2) unchanged in cloud vs. datacenter environments, (3) degraded in cloud environments, and (4) unacceptable in the cloud environment.

Currently, service-level agreements (SLAs) between application providers and cloud providers cover only pay-per-use, and are subject to the availability of cloud resources—there is no agreement regarding application-level performance or QoS provided to applications and underlying services. The designers of service deployment schemes consider multidimensional and mutually conflicting objectives with regard to operating costs, SLAs, configuration constraints, resource utilization, and resource availability. The effective operation of deployed services is extremely dependent upon the deployment scheme: underestimating resource requirements in the deployment scheme may result in excessive scarcity of resources, leading to a SLA violation. On the other hand, overestimating resource requirements may use resources less efficiently, increasing cloud operation costs without increasing revenue. Moreover, the objectives of cloud and service providers are at odds as each tries to maximize their respective profits. The service provider aims to satisfy the SLA using the minimum amount of hardware resources from the cloud provider, whereas the cloud providers benefit from increased use of computing resources. Also, since the cloud provider may be hosting multiple services, several service providers may have to compete for cloud resources.

Therefore, there is a tremendous need for a generic framework that can derive an optimal deployment scheme for hosting a large class of services in the cloud. Such a framework would be useful in addressing trade-offs between efficiency and reliability in the execution environment by considering an exhaustive set of computation and communication attributes that affect the performance of both the deployed service and the cloud infrastructure. The system QoS can be specified using multiple metrics defining operation efficiency, system reliability, strict SLA requirements, and operational constraints. The proposed generic framework will be a step towards a model-based autonomous development of a robust, adaptive, and reliable cloud infrastructure.

[1] http://reports.informationweek.com/abstract/5/5116/Cloud-Computing/research-2011-state-of-cloud.html [Mar2012]

Hamid Alipour
University of Arizona

Hamid is a Ph.D. student and graduate research assistant in the Department of Electrical and Computer Engineering at University of Arizona. He joined the CAC site at the University of Arizona when he started his Ph.D. and has been involved in several cybersecurity projects. His main research focus is on developing autonomic network protection systems based on protocol behavior analysis. He designed the DNS Autonomic Protection System (DNS-APS) and the Wireless Autonomic Protection System (WAPS) and is currently working on secure cloud computing based on a behavior-analysis approach. Hamid spent last summer at an internship with Microsoft designing a monitoring system for the Office 365 cloud environment. As an active member of CAC, Hamid has published in two peer-reviewed conferences, and has a transaction journal under revision.

Dario Pompili
Rutgers University

Dario is the Site Director for the CAC site at Rutgers University, a faculty member at Rutgers' Department of Electrical and Computer Engineering, and Director of the Cyber Physical Systems (CPS) Laboratory. His research interests include wireless sensor networks, underwater acoustic communication and coordination of underwater vehicles, green computing, and network optimization and control.

Dario received the prestigious CAREER award from the National Science Foundation in 2011 for his work on underwater multimedia acoustic communication. In March, he was awarded an Office of Naval Research Young Investigator Program (YIP) grant for a three-year project entitled "Investigating Fundamental Problems for Real-time In-situ Data Processing in Heterogeneous Mobile Computing Grids". Only 26 grants nationwide were funded from this program in 2012.

Prapaporn Rattanatamrong
University of Florida

Prapaporn recently completed her Ph.D. studies under the guidance of CAC Director José Fortes in the Department of Electrical and Computer Engineering at the University of Florida. She has been actively involved in the design and implementation of a cyber-workstation portal and middleware management system for brain-machine interface architecture with real-time performance guarantees. Her research focuses primarily on real-time scheduling and optimized resource management for distributed systems.

As a graduate student, Prapaporn co-authored 11 articles published in refereed journals and conference proceedings. She received the Engineering Outstanding International Student Award by UF's International Center in Fall 2010. She has taken a position as a faculty member in the Department of Computer Science at Thammasat University in Thailand.

Srishti Srivastava
Mississippi State University

Srishti Srivastava is a Ph.D. student in the Department of Computer Science and Engineering at Mississippi State University. Her research interests include robustness analysis and dynamic scheduling of scientific applications on high-performance computing systems and model-based autonomic management of cloud computing systems. Her research work has been accepted in the form of technical papers at a number of renowned IEEE/ACM international conferences and has been published in conference proceedings through the IEEE computer society press. In 2011, she received the best paper award at the IEEE International Symposium of Parallel and Distributed Computing (ISPDC). She has also received the Bagley College of Engineering (BCoE) graduate fellowship award (2011-2012) and NSF and IEEE/TCPP travel grants for presenting her selected papers at various international conferences.

Autonomic Cloud Management System (ACMS)

PI: Salim Hariri [hariri@email.arizona.edu]
Lead students: Zhitao Li [zhitaoli@email.arizona.edu]
Farah Alfay [falfay@email.arizona.edu
Cihan Tunc [cihantunc@email. arizona.edu]

Cloud computing enhances cooperation, scaling, performance, and accessibility by reducing cost, improving performance and providing on-demand computing, storage and network resources that can be accessed using heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs). While current high-performance computing systems are designed to handle peak workloads, the workload varies greatly during runtime. Many studies have shown that data servers typically operate at a low utilization of 10% to 15%, while their consumption of power is typically close to those at peak loads.

Our goal is to minimize power consumption while maintaining required performance levels for platform as a service (PaaS) by scaling up/down the hardware resources at runtime. In this project, we are developing an autonomic power and performance management system based on "AppFlow-based reasoning" targeting cloud systems and data centers. AppFlow is an n-dimensional data structure that is being developed at the NSF Cloud and Autonomic Computing (CAC) Center to characterize the current operational points of hardware and software resources as well as projecting their next operational points. In our approach, we classify the workloads into a set of workload types; for each workload type, we model the behavior of this workload into one AppFlow type. Similar to case-based reasoning, the online monitoring and analysis of the workload will then aim at determining the appropriate AppFlow that can be used to accurately describe the current workload. Once that is determined, we can then configure the datacenter or cloud system resources according to the determined AppFlow type such that the workload performance is maximized and power consumption is minimized, as shown in the figure below. Our experimental results show that our approach can reduce power consumption by 20% with very little overhead cost and minimal performance degradation.

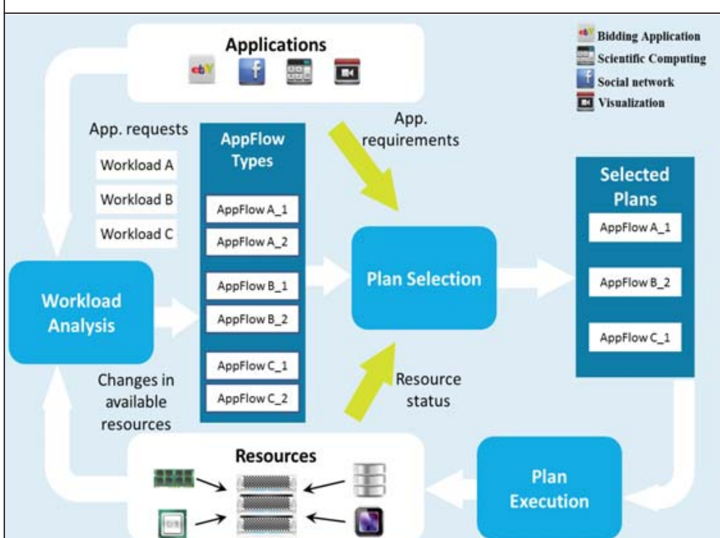

Figure: The Autonomc Cloud Management System (ACMS) architecture

Personal Device Social Networking with End-to-End Trusted Communication

PI: Renato Figueiredo [renato@acis.ufl.edu]
Lead student: Pierre St. Juste [ptony82@acis.ufl.edu]

Mobile devices such as smartphones and tablets are increasingly the platform of choice for many everyday tasks. The emerging application area of mobile augmented reality (MAR) capitalizes on the ubiquity of these devices to create rich user experiences. This project will research and develop systems that address the communication and information processing needs of future mobile augmented reality devices through three primary tasks:

Social routing through adaptive connections: This task involves adapting SocialVPN—a protocol developed at CAC to create P2P networks using data from social networks such as Facebook—to work with mobile devices and use both proximity-based and social networking-based peer discovery.

Power-efficient connections for mobile devices: Because mobile devices employ multiple types of connections, this task will involve analyzing traffic patterns of various workloads and understanding the costs associated with each type of connection in mobile environments.

Integration with mobile processing framework: This task involves enabling mobile applications to dispatch workloads to trusted computing nodes without the need to rewrite source code.

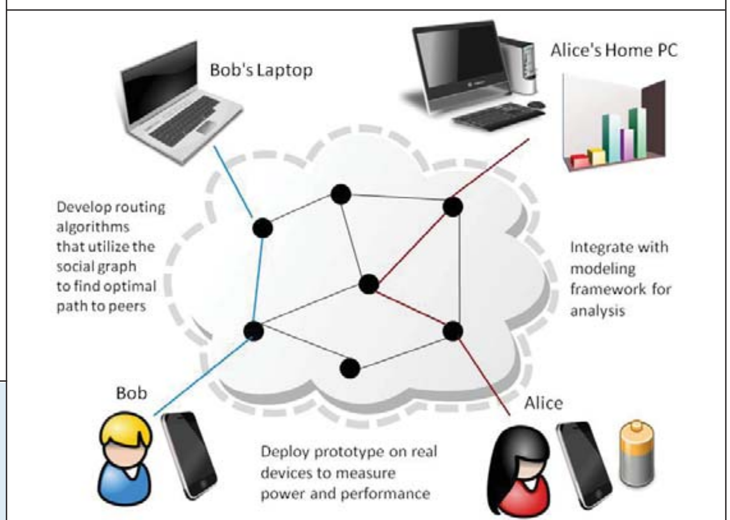

Figure: Networking personal devices using social networks to establish trust

Progress has been made towards all three tasks. First, towards the goal of social routing through adaptive connections, we have successfully tested a base implementation of our SocialVPN software stack on the Android operating system and implemented seamless transitions from IP tunneling over LAN/ WiFi and Bluetooth. The next steps in this task will be to port SocialVPN to mobile platforms and add support for unstructured overlays and probabilistic social routing capabilities.

Second, in an effort to characterize the power associated with maintaining SocialVPN connections on mobile devices, we have characterized the network traffic of a SocialVPN node in terms of the distribution of packets sent/received and their size as a function of the number of overlay connections associated with a node; created a synthetic packet generator that reproduced the distribution observed from measurements; and used it in experiments on an Android HTC smartphone running the PowerTutor tool. These measurements have recorded the energy savings that can be achieved by limiting the number of overlay connections; based on these experiments, work will focus on understanding performance/power trade-offs that arise as a function of the number of connections.

Finally, we have collaborated on a conceptual framework for a remote offload architecture that allows for seamless discovery of services in cloud providers or within a user's social network (e.g., within a LAN or personal area network). Services are then used to offload computation and information-processing tasks in cases where a mobile device does not have a functional capability, or when device power efficiency can be improved if tasks are offloaded. Focusing on the network layer of the framework, we have conducted experiments on a wide-area testbed (FutureGrid) that demonstrate the feasibility of using widely-used IP-layer resource discovery protocols (such as UPnP) to work seamlessly across SocialVPNs that may span multiple geographical domains.

Investigating Fundamental Problems for Real-time In-situ Data Processing in Heterogeneous Mobile Computing Grids

PI: Dario Pompili [pompili@cac.rutgers.edu]

RU Site Director Dario Pompili has won a Young Investigator Program (YIP) grant from the Office of Naval Research (ONR), one of only 26 awarded nationwide in 2012, for his proposal titled "Investigating Fundamental Problems for Real-time In-situ Data Processing in Heterogeneous Mobile Computing Grids". The YIP program invests in academic scientists and engineers who show exceptional promise for creative study.

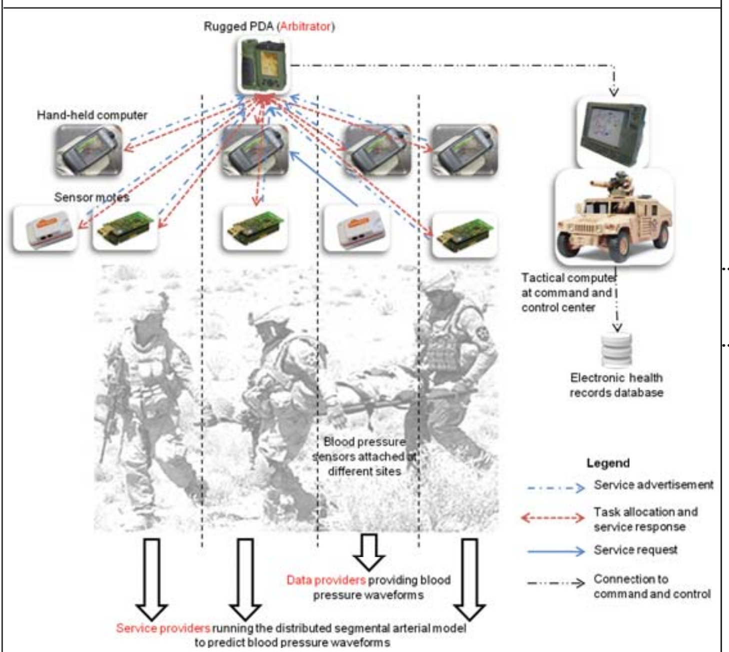

Figure: The in-situ data processing project aims to provide actionable medical information using both local devices and remote resources.

The objective of this three-year project is to enable real-time in-situ vital sign data processing to extract non-measurable physiological parameters, to interpret this data under context, and to acquire actionable knowledge about a deployed soldier's health. To realize this objective, which requires computing capabilities that go beyond those of an individual sensor mote and/or portable device, the collective computational capabilities of local resources—such as hand-held computers, rugged PDAs, and tactical computers carried by soldiers and/or armored vehicles—as well as remote computing clusters need to be exploited.

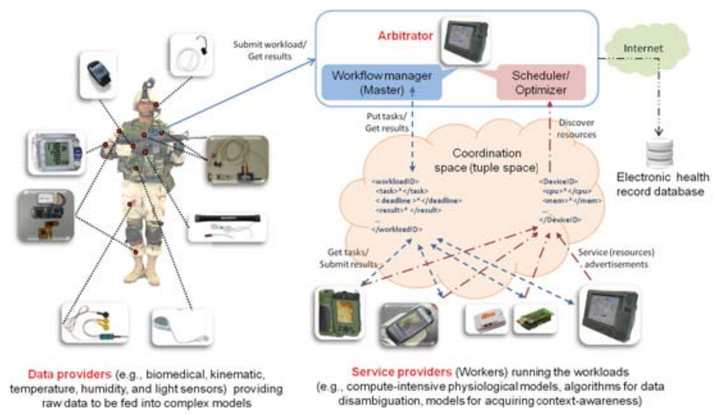

Figure: In this project, sensors provide data from the field in real time. Data processing and modeling is dynamically allocated to available resources.

This research project focuses on the fundamental research challenges associated with organizing these resources into an elastic resource pool (a hybrid computing grid). The most significant challenge is presented by the inherent uncertainty in the environment that can be attributed to unpredictable node mobility, varying rate of battery drain, and high susceptibility to hardware failures. The significant contributions of this research are i) a role-based architectural framework for reliable grid coordination under uncertainty, i.e., for handling resource and service discovery, service request arrivals, and workload distribution and management, and ii) a novel uncertainty- and energy-aware resource allocation engine, which will distribute workload tasks optimally among networked computing devices so to ensure quality of service (QoS) in terms of application response time and energy consumption.

A Model-Based Framework for Autonomic Performance Management of Cloud Computing Systems

PI: Ioana Banicescu [ioana@cse.msstate.edu]
Lead students: Rajat Mehrotra [rm651@msstate.edu]
Srishti Srivastava [ss878@msstate.edu]

Traditional applications address service level agreements (SLAs) regarding the performance, reliability, and availability between application providers and end users. In a cloud environment, there is an additional entity: the cloud provider. SLAs between the application provider and the cloud provider cover only pay-per-use and are subject to the availability of cloud resources. However, there is no agreement regarding the application level performance or quality of service (QoS) that the given resources will provide to the application. As a result, cloud applications do not currently provide the same availability or QoS guarantees as their counterparts hosted in traditional datacenter environments. Therefore, there is a need to include these application performance objectives into SLAs between application providers and cloud providers. Cloud services are also subject to outages or even data losses that could result from reasons varying from hardware and/or software failures to acts of nature. Thus, cloud providers must address these challenges in order to provide reliable delivery of application performance to the end user.

The main goal of this project is to employ a model-based management approach to enabling self-management attributes in a cloud management system such that the overall performance of the system (in terms of QoS, availability, and reliability) meets SLA specifications. The key aspects for addressing this goal include (i) establishing an SLA between the cloud provider and the application provider to guarantee a certain level of application performance, (ii) developing a model to address the behavior of the hosted applications with respect to available resources and control decisions, (iii) investigating and extending current cloud computing tools for monitoring the infrastructure and identifying and projecting the possible faults in the system, and (iv) designing a model-based controller to take decisions regarding placement of the applications and number of instances to be allocated based on the application requirements and the monitored data provided by the cloud monitoring tool.

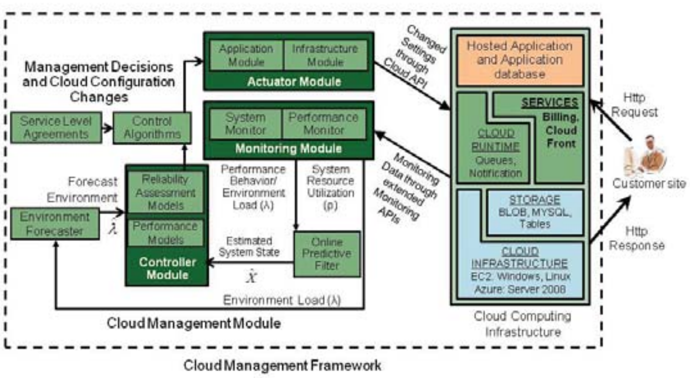

Figure: a schematic of the cloud management framework

The focus of this project is on the following tasks, which are identified to introduce new SLA aspects for performance guarantees of application execution in cloud environments. These tasks are based upon the assumption that QoS management in cloud computing environments can be performed through various small but interdependent control decisions. These control decisions include the placement of applications, number of instances of individual applications, and allocation of computational, memory, and storage resources to the application. These tasks are:

1. Define the structure of advanced SLAs between application provider and cloud provider which will contain pay-per-use pricing strategy and QoS constraints (application performance, throughput, and penalty in case of QoS violation). Additionally, identify the capabilities of current enterprise and open-source cloud management products to deliver the advanced SLAs, monitor the infrastructure, identify and analyze faults in the system, and apply control actions in case of SLA violation.
2. Develop a comprehensive monitoring framework for cloud computing systems that can extensively monitor system variables, system health, and application performance variables accurately, with minimum latency at a specified rate and with minimal resource utilization. Furthermore, an effective and efficient fault identification, diagnosis, and projection module will be developed to analyze the cloud computing infrastructure for possible performance issues.
3. Develop reliability metrics and models with respect to potential configuration and control changes in the system (to measure their short- and long-term impact on application performance), resource allocation, and the hardware profile of the system. Existing hardware reliability models will be utilized in this step. Delay in actuating various configuration or control changes will also be studied in this phase.
4. Solve the dynamic management problem in cloud computing systems via a feedback-based control framework, which will be developed using the monitoring and fault diagnosis modules developed in previous steps. Additionally, use application performance models to predict the changes in application performance with respect to changes in control parameters. Furthermore, use the reliability models to derive the optimal control changes for maintaining the QoS objectives. This framework will be then added to an open-source cloud management product as a library.

At the end of the project, we will deliver the following:

1. A distributed monitoring framework with fault diagnosis modules to monitor SLAs in cloud computing environments.
2. Performance models, reliability metrics, and models for measuring short- and long-term impact of configuration and control changes on application performance.
3. An autonomic cloud management framework, as an open source library, for managing SLA agreements between application provider and cloud provider.
4. Publications describing our framework.

Wireless Autonomic Protection System

PIs: Salim Hariri [hariri@ece.arizona.edu]
Youssif Al Nashif [alnashif@ece.arizona.edu]
Lead student: Hamid Alipour [hra@email.arizona.edu]

The rapid and vast deployment of wireless networks and their applications have changed the structure of ubiquitous network services and the defense strategies for network security. The range of wireless networks can vary from 10 meters, as in the case of wireless personal area networks (WPAN), to 100km, as in the case of wireless regional area networks (WRAN). Advances in wireless technology have made portable devices an essential part of our daily life, as we use them to access our critical information, such as banking accounts, credit cards or email addresses. While wireless networks are taking a progressively crucial role in modern society, the lack of effective and affordable security measures make them easy targets for intruders. Current wireless network defense tools rely primarily on predefined signatures for attack detection. Although these methods can successfully detect commonly known attacks, they cannot recognize abnormal behaviors caused by new and sophisticated attacks until it is likely too late to take any useful action. Recent research studies show that current wireless intrusion detection systems can detect a maximum of around 40% of current wireless attacks and fail to detect complex, dynamic, and knowledgeable attacks. Our goal is to develop an innovative technology to secure wireless networks and their services.

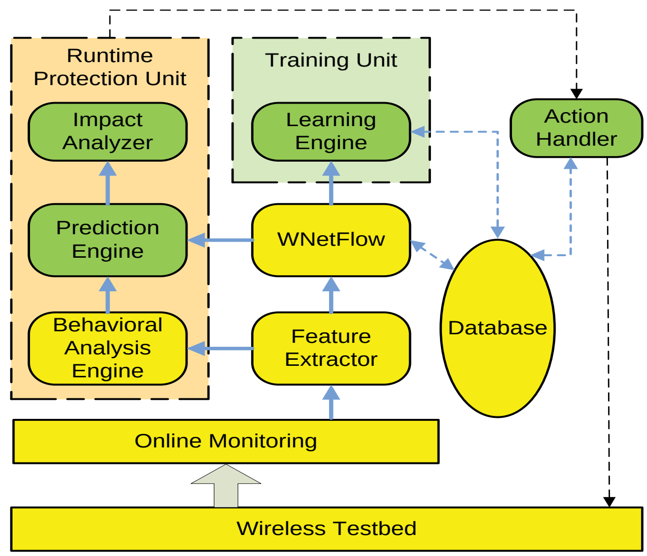

Figure: The Wireless Autonomic Protection System (WAPS) architecture

Our approach is based on implementing an anomaly-based wireless autonomic protection system (WAPS); the main modules of the WAPS architecture are shown in the figure above. WAPS protection will be achieved by 1) monitoring multiple channels concurrently in real time, 2) detecting anomalies resulting from exploitation of the wireless communication protocol by the intruder, and 3) taking proactive actions to isolate and quarantine the intruder.

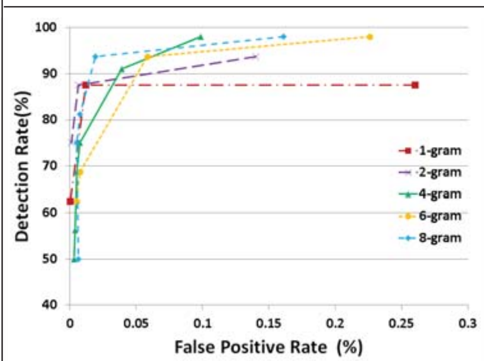

Figure: Observed detection rate vs. false positive rate using the WAPS system on a wireless testbed.

We have built a wireless testbed consisting of two access points (APs) operating at channels three and eight, respectively. These APs are configured to work with both WEP and WPA security enabled. The testbed uses high-gain antennas for channel monitoring. The normal data is collected using some normal stations, and the attacker can launch attacks on APs. Our behavioral analysis engine relies on sequential pattern analysis to characterize the behavior of wireless traffic. We apply statistical techniques to quantify any n consecutive transitions in the protocol, referred to as n-gram patterns, and determine the probabilities of these transitions being normal. We have successfully demonstrated the accuracy of our detection technique and its low rate of false alarms. Our results indicate that our system can detect different wireless attacks (e.g., deauthentication, association and authentication flooding and injection test) with a false positive rate of less than 0.01%.

Self-caring IT

PI: José Fortes [fortes@ufl.edu]
Lead student: Selvi Kadirvel [selvi@acis.ufl.edu]

The MapReduce programming paradigm provides fault-tolerance as an attractive feature in addition to its simple programming model for large-scale job parallelism and data distribution. However, the occurrence of faults can lead to high performance penalties as determined by the characteristics of the job. In this project, we leverage the elastic properties of a virtualized environment to overcome these performance penalties. The end goal of effective fault and performance management is to proactively mitigate SLA-based job completion time violations.

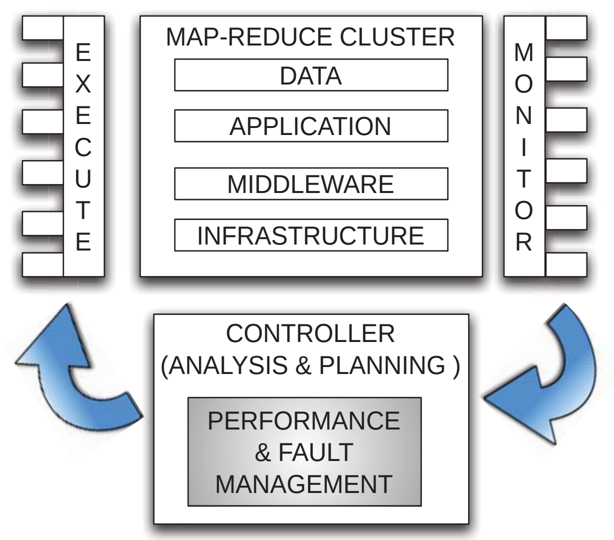

Figure: The fault testbed deployed for the self-caring IT project

Although the need for effective fault management has been increasing, the availability of real-world failure data sets is rare in the public domain. Datasets available in repositories are usually outdated and static, and hence cannot be used to deploy and validate management techniques. So, as a first step toward deploying an autonomic approach to handle faults, we have designed and deployed FaultPlay. FaultPlay is a testbed and live fault benchmark that is used to characterize the effect of faults and determine factors that influence penalties. Performance characterization results have been generated using this testbed for jobs that differ in their resources, datasets, configurations and fault occurrences. Penalties ranging up to 155% have been recorded for single-node faults. Using this fault testbed, the next critical component that was designed is a performance prediction module. We propose a grey-box approach to predict job performance using supervised machine-learning regression techniques both in the presence and absence of faults. Using the factors identified from the characterization studies, we have been able to show that these regression techniques can predict job performance with a mean prediction error of ~12%. Finally, we have deployed a prototype of the dynamic resource scaling solution to mitigate penalties both on the empirical testbed as well as on the MapReduce MRPerf simulator. Using our approach for the benchmarks and conditions studied, we were able to constrain penalties due to single-node faults to within 5% of the fault-free execution time.

Moving Target Defense Middleware (MTDM) for Intrusion Resilient Cloud Services

PI: Salim Hariri, hariri@ece.arizona.edu]
Lead student: Glynis D'souza [glynisdsouza@email.arizona.edu]

Cloud computing is emerging as a new paradigm that aims at delivering computing as a utility. For cloud computing to be fully adopted and effectively used, it is important that security mechanisms are robust and resilient to faults and attacks. Securing cloud applications and services is a challenging research problem because it involves many interdependent tasks including vulnerability scanning, application-layer firewalls, configuration management, alert monitoring and analysis, source code analysis, and user identity management. Most of these challenges are due to the monoculture of cloud software, dynamic environments where resources and services are constantly changing, and social networking technologies.

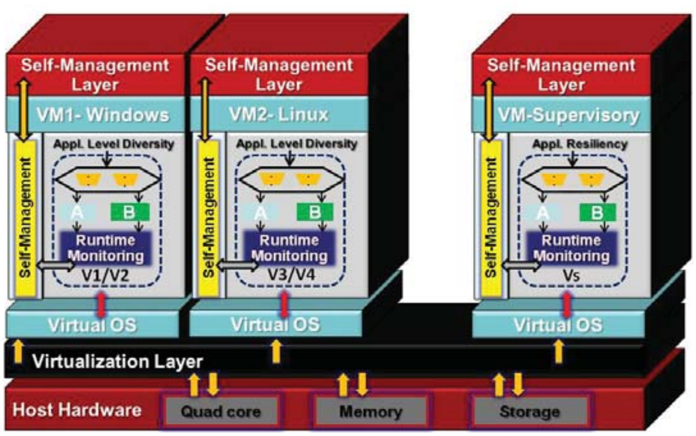

Figure: A schematic of the moving target defense middleware system.

In this project, we are developing moving target defense middleware that can provide cloud services that are resilient against anomalous events that might be triggered by malicious attacks and/or faults. The main MTDM capabilities are software behavior encryption (SBE) and self-management (SM). Software behavior encryption employs spatiotemporal behavior encryption and a moving target defense to make active software components change their implementation variants and resources continuously to evade attackers. This approach will make it extremely difficult for an attack to disrupt the normal operation of an application. Also, the dynamic change in the execution environment will hide software flaws that would otherwise be exploited by a cyber-attacker. Self management is critical in order to deliver automatic detection and recovery capabilities to enhance the resiliency of software systems and services. To validate our approach, we use a cloud application based on Hadoop MapReduce as a running example to experiment with and evaluate the resiliency of the MTDM services against attacks. We employ n-version programming by having three physical machines independently run different versions of each task, thus employing spatial diversity. The selection of the task version and its execution environment (type of operating system, programming language, etc.) are randomized at run-time using the SBE algorithm. Our experimental results show that the cloud application can continue to operate normally in spite of cyber-attacks including denial of service (DoS) and insider attacks.

Get involved!

CAC projects are developed in collaboration with industry participants to produce focused, industry-relevant research. CAC faculty and students gather with industry representatives twice a year to review progress on existing projects and to explore new research directions that best meet member companies' goals in the areas of cloud and autonomic computing. Our next semiannual meeting will be held June 14 & 15 at the University of Florida in Gainesville. We invite industry representatives to attend the meeting and tell us how CAC research can address your company's needs.

Face to Face

CAC researchers are committed to promoting the Center, addressing industry-relevant research problems, and furthering collaboration. We achieve this in part by attending conferences to increase the visibility of CAC projects and attract new members, and by visiting CAC member companies to boost industry participation. This section of the Bulletin highlights the meetings between CAC personnel and the world at large.

August 2011

CAC @ MSU student Rajat Mehotra presented a talk entitled "Automated Model-based Security Management of Web Services" at the 20th USENIX Security Symposium, held August 8-12 in San Francisco, CA.

September 2011

Rutgers graduate student Baozhi Chen delivered a talk on "Inter-glider Underwater Communication and Coordination for Ocean Monitoring and Coastal Tactical Surveillance" at the a meeting of the Rutgers Engineering Society on September 7.

CAC Director José Fortes, UF Site Director Renato Figueiredo, and RU CAC Co-director Manish Parashar attended an invitation-only workshop on US-China collaboration hosted by the National Science Foundation, held September 23-30 in San Diego, CA.

October 2011

MSU Site Director Ioana Banicescu presented a keynote/plenary talk entitled "Advances towards a Technology for Robust and Cost-Effective Autonomic Execution of Scientific Applications" on October 14 at the 15th IEEE International Conference on System Theory, Control and Computing (ICSTCC 2011), held in Sinaia, Romania.

CAC's Fall 2011 Semiannual Meeting was held at Biosphere 2 at the University of Arizona October 19-21. Representatives from member companies, the National Science Foundation and the four CAC sites were in attendance to discuss ongoing and proposed projects at the Center and discuss future research directions.

November 2011

Rutgers Site Director Dario Pompili gave a talk entitled "A Reliable Geocasting Solution for Underwater Acoustic Sensor Networks," at the IEEE Military Communications Conference (MILCOM), held in Baltimore, MD, November 7-10.

December 2011

CAC Director José Fortes gave an invited keynote talk entitled "Towards the Ubiquitous Cloud: Multi-Cloud Systems" at the 6th International Conference on Ubiquitous Information Technologies & Applications (CUTE 2011), held in Seoul, Korea, December 15-17.

Baozhi Chen presented "Extended Abstract: Uncertainty-Based Localization Solution for Under-Ice Autonomous Underwater Vehicles," at the ACM International Workshop on UnderWater Networks (WUWNet) held December 1-2 in Seattle, WA.

February 2012

CAC @ UF faculty members José Fortes and Maurício Tsugawa attended a workshop for a collaboration with AIST on an NSF Rapid Response Research (RAPID) grant, held February 9 & 10.

CAC @ MSU student Srishti Srivastava presented "A Model-based Framework for Autonomic Performance Management of Cloud Computing Systems" at the 15th SIAM Conference on Parallel Processing for Scientific Computing, held February 15 -17 in Savannah, GA.

A group from Raytheon Missile Systems visited the University of Arizona CAC site on February 28.

The CAC site at the University of Florida hosted Dr. Takahiro Hirofuchi, a researcher from the National Institute of Advanced Industrial Science and Technology (AIST) in Japan. Takahiro visited CAC to collaborate on a project investigating the use of virtualization technologies for preserve IT systems in the event of a natural disaster. Read more about this project in the "Feature Articles" section on page 3.

March 2012

Dario Pompili delivered a talk on "Underwater Communication and Coordination Among Gliders for Ocean Monitoring and Coastal Surveillance" in the Physical Ocean Science and Engineering (POSE) Seminar at the College of Earth, Ocean, and Environment, University of Delaware, Newark, DE on March 9.

CAC @ MSU student Qian Chen presented a talk entitled "On the State of the Art in Virtual Machine Security," IEEE South East Conference (SouthEastCon-2012), Orlando FL, March 15-18.

MSU student Rajat Mehrotra presented "RFDMon: A Real-Time and Fault-Tolerant Distributed System Monitoring Approach" at the Eighth International Conference on Autonomic and Autonomous Systems (ICAS-2012) held in St. Maarten, Netherlands Antilles, March 25-30.

April 2012

Eight representatives from IBM attended a meeting at the University of Arizona CAC site on April 5. Attendees included technical and sales managers visited the site for an overview of the Center and ongoing projects.

May 2012

Ioana Banicescu presented a paper titled "A Utility Based Power-Aware Autonomic Approach for Running Scientific Applications" at the 13th IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS-PDSEC) 2012, held May 21-25 in Shanghai, China.

June 2012

Dario Pompili will present a talk entitled "Uncertainty-Aware Localization Solution for Networked Under-Ice Autonomous Underwater Vehicles," at the IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), to be held in Seoul, Korea, June 2012.

Service

Rutgers Site Director Dario Pompili served as the Technical Program Committee (TPC) co-chair at the ACM Workshop on UnderWater Networks (WUWNET), held December 1-2.

Dario Pompili served as a member of the TPC at the Institute of Electrical and Electronics Engineers' (IEEE) Conference on Computer Communications (INFOCOM), held March 25-30 in Orlando, Florida.

Grants

Dario Pompili was awarded an Office of Naval Research Young Investigator Program (YIP) grant for a three-year project entitled "Investigating Fundamental Problems for Real-time In-situ Data Processing in Heterogeneous Mobile Computing Grids". Read more about this project on page 10 of this newsletter.

The Center received funding for a proposal to the Fundamental Research Program in 2011. The grant, entitled "Unified Cloud Computing and Management," supports a project that spans all four CAC university sites and addresses security and resource management aspects of cloud computing that are important to a variety of industry-relevant applications.

In August 2011, Dario Pompili was awarded a National Science Foundation grant by the NSF for carrying out fundamental research in the area of "Sensor-driven Thermal-aware Autonomic management of Green Datacenters".

Awards

CAC @ MSU student Sristi Srivastava (with co-authors Ioana Banicescu, Florina Ciorba and Wolfgang Nagel) won the Best Paper Award at the IEEE International Symposium on Parallel and Distributed Computing (ISPDC2011) for their paper entitled "Enhancing the Functionality of a GridSim-based Scheduler for Effective Use with Large-Scale Scientific Applications." ISPDC2011 was held in Cluj-Napoca, Romania from July 6 to July 8.

Ioana Banicescu and CAC @ MSU faculty member Sherif Abdelwahed were each awarded a State Pride Award from MSU's Bagley College of Engineering for the 2011-2012 academic year.

CAC Director José Fortes was presented with an American Association for the Advancement of Science (AAAS) Fellowship Award at the organization's 178th Annual Meeting in February 2012. According to the AAAS website, the title of AAAS Fellow is bestowed on select members "for meritorious efforts to advance science or its applications."

Rutgers student Baozhi Chen received the ACM Student Travel Grant to attend the Workshop on UnderWater Networks (WUWNET) in Seattle, WA, December 2011.

Dario Pompili received the Rutgers/ECE Outstanding Young Researcher Award in Electrical Engineering for "demonstrating outstanding research performance as well as deep insight of the Electrical and Computer Engineering Research area," in 2011.

High school student Ben Englard won a first-place Superior Award in the Miami-Dade County Science Fair and will continue on to compete in the State Science Fair, with a chance to continue on to the Intel Science and Engineering Fair. Ben was hosted at the UF CAC site in the summer of 2011 to work on a project titled "Litter: A Lightweight Peer-to-Peer Microblogging Service" as part of UF's Student Science Training Program (SSTP). He was one of only four students chosen by Miami-Dade County to present their research at the International Science and Engineering Fair, sponsored by Intel.

Graduations

UF student Prapaporn Rattanatamrong completed her Ph.D. degree. Read about Prapaporn's research in the "Spotlight on CAC People" section on page 8 of this newsletter.

CAC Arizona student Shafiul "Jacky" Islam will complete his master's degree this summer. Jacky has been involved with the Autonomic Cyber Physical System project; more information on this project can be found on page 6 of this newsletter.

In the news

Research on resource provisioning in mobile grids at Rutgers' Cyber Physical Systems (CPS) Lab was featured on International Science Grid this Week (isgtw) on October 12, 2011. The article is archived at http://www.isgtw.org/feature/research-report-fare-share-mobile-grid-computing.

Become a CAC industrial member

CAC members collaborate with and advise researchers to create a diverse, industry-relevant research program. Members are afforded access to leading-edge developments in autonomic computing and to knowledge accumulated by academic researchers and other industry partners. An annual membership fee allows industry partners to reap the full benefits of CAC's full research program, with an operating budget of over $1 million annually. To inquire about membership in CAC, please contact Center Director Jose Fortes at fortes@ufl.edu.

Related Publications by CAC Researchers

1. B. Chen and D. Pompili, "Localization in Underwater Acoustic Sensor Networks," in Computing Handbook Set—Computer Science (Volume I), Teofilo Gonzalez, Ed., CRC Press, 2013.
2. Ioana Banicescu, Florina M. Ciorba, Srishti Srivastava. "Performance Optimization of Scientific Applications using an Autonomic Computing Approach", in Scalable Computing and Communications: Theory and Practice, U. Khan, Lizhe Wang and Albert Y. Zomaya, Eds., John Wiley & Sons, 2012 (to be published October 2012).
3. H. Alipour, Y. Al. Nashhif, S. Hariri, "DNS Anomaly-based Behavior Analysis against Cyber Attacks," to appear in IEEE Transactions on Dependable and Secure Computing (accepted, under revision), 2012.
4. B. Chen and D. Pompili, "A Communication Framework for Networked Autonomous Underwater Vehicles," in The Art of Wireless Sensor Networks, Habib Ammari, Ed., Springer, 2012.
5. H. Viswanathan and D. Pompili, "Emergency Networking in Licensed Spectrum using Cognitive Radios—Challenges and Insights," in Cognitive Radio for Wireless Cellular and Vehicular Networks, H. Venkataraman and G-M. Muntean, Eds., Springer, 2012.
6. I. Rodero, H. Viswanathan, E. K. Lee, M. Gamell, D. Pompili and M. Parashar, "Energy-efficient Thermal-aware Autonomic Management of Virtualized HPC Cloud Infrastructure," revised for Journal of Grid Computing (Springer), 2012.
7. H. Viswanathan, B. Chen and D. Pompili, "Research Challenges in Computation, Communication, and Context Awareness for Ubiquitous Healthcare," to appear in IEEE Communication Magazine, 2012.
8. H. Liu, E. K. Lee, D. Pompili, and X. Kong, "Thermal Camera Networks for Large Datacenters using Real-Time Thermal Monitoring Mechanism," revised for Journal of Supercomputing (Springer), 2012.
9. S. Oh, M. Gruteser, and D. Pompili, "Coordination-free Safety Messages Dissemination Protocol for Vehicular Networks," to appear in IEEE Transactions on Vehicular Technology, 2012.
10. H. Alipour, Y. Al. Nashhif, S. Hariri, " IEEE 802.11 Anomaly-based Behavior Analysis," submitted to the 2012 IEEE Global Communication Conference (GLOBCOM 2012), Anaheim, CA, December 3-7, 2012.
11. B. Chen and D. Pompili, "Uncertainty-Aware Localization Solution for Networked Under-Ice Autonomous Underwater Vehicles," to appear in Proc. of IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), Seoul, Korea, June 2012.
12. Girish Venkatasubramanian, Ramesh Illikal, Don Newell, and Renato Figueiredo, "TMT—A TLB Tag Management Framework for Virtualized Platforms," International Journal on Parallel Programming, vol. 40, no. 3, pp. 353-380, June 2012. (doi: 10.1007/s10766-011-0189-y)
13. Mahadevan Balasubramaniam, Nitin Sukhija, Ioana Banicescu, Florina Ciorba and Srishti Srivastava, "Towards the Scalability of Dynamic Loop Scheduling Techniques via Discrete Event Simulation," in Proceedings of the 13th IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS-PDSEC), Shanghai, China, May 21-25, 2012. (On CD-ROM, IEEE Computer Society Press)
14. Rajat Mehrotra, Ioana Banicescu, and Srishti Srivastava, "A Utility Based Power-Aware Autonomic Approach for Running Scientific Applications," in Proceedings of the 13th IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS-PDSEC) 2012, Shanghai, China, May 21-25, 2012. (On CD-ROM, IEEE Computer Society Press)
15. Florina Ciorba, Timothy Hansen, Srishti Srivastava, Ioana Banicescu, Tony Maciejewski, H. J. Siegel, "A Combined Dual-stage Framework for Robust Scheduling of Scientific Applications in Heterogeneous Environments with Uncertain Availability," in Proceedings of the 13th IEEE/ACM International Parallel and Distributed Processing Symposium, Shanghai, China, May 21-25, 2012.
16. Y. Jararweh, M. Jarrah, and S. Hariri, "Exploting GUPS for Compute-Intensive Medical Applications," in Proceedings of the IEEE International Conference on Multimeida Computing and Systems, Tangiers, Morocco, May 10-12, 2012.
17. H. Alipour, Y. Al. Nashhif, S. Hariri, "Anomaly-based Behavior Analysis of DNS traffic," in Proceedings of the 1st IBM Cloud Academy Conference (ICA CON 2012), Research Triangle Park, NC, April 19-20, 2012
18. G. Dsouza, H. Alipour, S. Hariri, Y. Al-Nashif, and M. Eltoweissy, "Cloud Resilient Architecture," in Proceedings of the 1st IBM Cloud Academy Conference (ICA CON 2012), Research Triangle Park, NC, April 19-20, 2012
19. Rajat Mehrotra, Abhishek Dubey, Sherif Abdelwahed, and Krisa Rowland, "RFDMon: A Real-Time and Fault-Tolerant Distributed System Monitoring Approach," in Proceedings of the Eighth International Conference on Autonomic and Autonomous Systems (ICAS-2012), St. Maarten, Netherlands Antilles, March 25-30, 2012.
20. Qian Chen, Rajat Mehrotra, Abhishek Dubey, Sherif Abdelwahed, and Krisa Rowland, "On State of The Art in Virtual Machine Security," in Proceedings of the IEEE South East Conference (SouthEastCon-2012), Orlando, FL, March 15-18, 2012.
21. S. Srivastava, R. Mehrotra, I. Banicescu, and S. Abdelwahed, "A Model-based Framework for Autonomic Performance Management of Cloud Computing Systems," in Proceedings of the 15th SIAM Conference on Parallel Processing for Scientific Computing, Savannah, GA, February 15-17, 2012.
22. Mark R. Zonfrillo, Mahendra Kumar, Jose A. Fortes, Flaura K. Winston, "Telecenter for secure, remote, collaborative child fatality review," Injury Prevention, published online first: February 10, 2012. (doi:10.1136/injuryprev-2011-040283)
23. Justin Sanchez, William W. Lytton, Jose M. Carmena, Jose C. Principe, José Fortes, Randall L. Barbour, Joseph T. Francis, "Dynamically Repairing and Replacing Neural Networks," IEEE Pulse, February 6, 2012. (doi: 10.1109/MPUL.2011.2175640)

24. B. Chen and D. Pompili, "Team Formation and Steering Algorithms for Underwater Gliders using Acoustic Communications," Computer Communication (Elsevier), January 2012. (doi:10.1016/j.comcom.2012.01.001)
25. C. Jeffery, R. Figueiredo, “A Flexible Approach to Improving System Reliability with Virtual Lockstep”, IEEE Transactions on Dependable and Secure Computing, vol. 9, no. 1, pp. 2-15, January 2012.
26. R. Mehrotra, A. Dubey S. Abdelwahed, and A. Tantawi "A Power-aware Modeling and Autonomic Management Framework for Distributed Computing Systems," in Handbook of Energy-Aware and Green Computing, S. Ranka and I. Ahmad Eds. Chapman and Hall/CRC Press, January 2012.
27. D. Wolinsky, K. Lee, P. Chuchaisri and R. Figueiredo, “Experiences with Self-Organizing, Decentralized Grids Using the Grid Appliance”, Cluster Computing, pp. 1-19, January 2012. (doi:10.1007/s10586-011-0195-2).
28. D. J. Hill and D. Pompili, "Pervasive Sensing and Computation for Sustainable Urban Systems: Trends and Opportunities," submitted to Pervasive and Mobile Computing (Elsevier), 2011.
29. B. Chen and D. Pompili, "A QoS-aware Underwater Optimization Framework for Inter-vehicle Communication using Acoustic Directional Transducers," submitted to IEEE Journal of Selected Areas in Communications, 2011.
30. B. Chen and D. Pompili, "Extended Abstract: Uncertainty-Based Localization Solution for Under-Ice Autonomous Underwater Vehicles," in Proc. of ACM International Workshop on UnderWater Networks (Doctoral Symposium), Seattle, WA, Dec. 2011.
31. Ram Prasad Viswanathan, Youssif Al-Nashif, Salim Hariri, "Application Attack Detection System (AADS): An Anomaly Based Behavior Analysis Approach," in Proceedings of the 9th ACS/IEEE International Conference.on Computer Systems and Applications (AICCSA’11), December 2011.
32. Y. Yuan, E. K. Lee, D. Pompili, and J. Liao, "Thermal Anomaly Detection in Green Datacenters," in Proc. of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, first available online December 2011.
33. B. Chen and D. Pompili, "A Reliable Geocasting Solution for Underwater Acoustic Sensor Networks," in Proc. of IEEE MILCOM, Baltimore, MD, November 2011.
34. B. Chen and D. Pompili, "Prioritization of Patient Vital Signs Transmission Using Wireless Body Area Networks," E-Healthcare Systems and Wireless Communications: Current and Future Challenges, Mohamed K. Watfa, Ed., October 2011.
35. S. Hariri, M. Eltoweissy and Y. Alnashif, “BioARC:Bilogically-inspired Autonomic Resilient Cloud, “ in Proceedings of the 7th Cyber Security and Information Intelligence Workshop (CSIIRW’2011), October 2011.
36. Rajat Mehrotra, Abhishek Dubey and Sherif Abdelwahed, "RFDMon: A Real-Time and Fault-Tolerant Distributed System Monitoring Approach," Technical Report for the Institute for Software Integrated Systems (ISIS), Vanderbilt University, October 2011.
37. A. Matsunaga, P. Riteau, M. Tsugawa, J. A. B. Fortes, "Crosscloud Computing," Advances in Parallel Computing, High Performance Computing: From Grids and Clouds to Exascale, pp. 109-123, September 2011. (ISBN 978-1-60750-802-1)
38. J. P. Varkey, D. Pompili, T. Walls, "Human Motion Recognition Using a Wireless Sensor-based Wearable System," Personal and Ubiquitous Computing (Springer), pp. 1-14, September 2011.
39. P. Rattanatamrong and J. A. B. Fortes, "Mode Transition for Online Scheduling of Adaptive Real-Time Systems on Multiprocessors," The 17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), Toyama, Japan, August 28-31, 2011.
40. Rajat Mehrotra, Qian Chen, Abhishek Dubey, Sherif Abdelwahed, and Krisa Rowland, "Automated Model-based Security Management of Web Services," in Proceedings of the 20th USENIX Security Symposium, San Francisco, CA, August 8-12, 2011.
41. S. Kadirvel and J.A.B. Fortes, “Towards Self-Caring MapReduce: Proactively Reducing Fault-Induced Execution-Time Penalties,” The International Conference on High Performance Computing & Simulation (HPCS), Istanbul, Turkey, July 4-8, 2011.
42. P. Rattanatamrong and J. A. B. Fortes, "Improved Real-Time Scheduling for Periodic Tasks on Multiprocessors,” The International Conference on High Performance Computing & Simulation (HPCS), Istanbul, Turkey, July 4-8, 2011.
43. S. Kadirvel and J. Fortes, "Towards IT Systems Capable of ManagingTheir Health," in Foundations of Computer Software: Modeling, Development, and Verification of Adaptive Systems, R. Calinescu and E. Jackson, Eds., pp. 77-102, Springer-Verlag, 2011.
44. H. Viswanathan, E. K. Lee, and D. Pompili, "Self-organizing Sensing Infrastructure for Autonomic Management of Green Datacenters," IEEE Network Magazine, July 2011.
45. Girish Venkatasubramanian, Renato Figueiredo and Ramesh Illikkal, "On the Performance of Tagged Translation Lookaside Buffers: A Simulation-Driven Analysis," in Proceedings of the 19th International IEEE Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), July 2011.
46. Sristi Srivastava, Ioana Banicescu, Florina Ciorba and Wolfgang Nagel, "Enhancing the Functionality of a GridSim-based Scheduler for Effective Use with Large-Scale Scientific Applications", in Proceedings of the IEEE International Symposium on Parallel and Distributed Computing (ISPDC 2011), July 2011.
47. B. Chen and D. Pompili, "QUO VADIS: QoS-aware Underwater Optimization Framework for Inter-vehicle Communication using Acoustic Directional Transducers," Proc. of IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), Salt Lake City, UT, June 2011.
48. David Wolinsky and Renato Figueiredo, "Experiences with Self-Organizing, Decentralized Grids Using the Grid Appliance," in Proc. of the 20th International ACM Symposium on High-Performance Parallel and Distributed Computing, June 2011.

49. L. Wang, M. Zhao, J. Xu and J. A. B. Fortes, "Adaptive Virtual Resource Management with Fuzzy Model Predictive Control," 6th International Workshop on Feedback Control Implementation and Design in Computing Systems and Networks (FeBID 2011) held in conjunction with The 8th International Conference on Autonomic Computing, Karlsruhe, Germany, June 14, 2011.
50. E. K. Lee, H. Viswanathan, and D. Pompili, "SILENCE: Distributed Adaptive Sampling for Sensor-based Autonomic Systems," in Proceedings of the IEEE International Conference on Autonomic Computing (ICAC), Karlsruhe, Germany, June 2011.
51. J. Xu, J. A. B. Fortes, "A Multi-objective Approach to Virtual Machine Management in Datacenters", The 8th International Conference on Autonomic Computing, Karlsruhe, Germany, June 14-18, 2011.
52. B. Chen and D. Pompili, "Advanced MAC Protocols for ZigBee Networks," ZigBee Network Protocols and Applications, C. Wang, T. Jiang, and Q. Zhang, Eds., Auerbach Publications, CRC press, May 2011.
53. H. Viswanathan, E. K. Lee, I. Rodero, D. Pompili, M. Parashar, and M. Gamell, "Energy-Aware Application-Centric VM Allocation for HPC," in Proceedings of High-Performance Grid and Cloud Computing Workshop (HPGC), Anchorage, AK, May 2011.
54. E. K. Lee, J. P. Varkey, and D. Pompili, "On the Impact of Neighborhood Discovery on Geographical Routing in Wireless Sensor Networks," Proc. of IEEE Sarnoff Symposium, Princeton, NJ, May 2011.
55. Yonggang Liu, Renato Figueiredo, Dulcardo Clavijo, Yiqi Xu, and Ming Zhao," Towards Simulation of Parallel File System Scheduling Algorithms with PFSsim," in Proceedings of the 7th IEEE International Workshop on Storage Network Architecture and Parallel I/O (SNAPI), May 2011.
56. Kyungyong Lee, Tae Woong Choi, Arijit Ganguly, David I Wolinsky,Oscar Boykin, Renato Figueiredo, "Parallel Processing Framework on a P2P System Using Map and Reduce Primitives," in Proceedings of the Eighth International Workshop on Hot Topics in Peer-to-Peer Systems (HotP2P), May 2011
57. R. Mehrotra, A. Dubey S. Abdelwahed, and A. Tantawi "A Power-aware Modeling and Autonomic Management Framework for Distributed Computing Systems," in Handbook of Energy-Aware and Green Computing, S. Ranka and I. Ahmad Eds. Chapman and Hall/CRC Press, to be published in 2011.
58. Girish Venkatasubramanian, Renato Figueiredo, Ramesh Illikkal and Donald Newell, "TMT—A TLB Tag Management Framework for Virtualized Platforms," accepted for publication in the International Journal on Parallel Programming (IJPP).
59. Venkata Krishna Nimmagadda, Ali Akoglu, Salim Hariri, and Talal Moukabary, "Cardiac Simulation on Multi-GPU Platform," to appear in the Springer Journal of Supercomputing.
60. M. Mallouhi, Y. Al-Nashif, D.Cox, T. Chadaga, and S. Hariri, "A Testbed for Analyzing Security of SCADA Control Systems (TASSCS)," the Second IEEE PES Innovative Smart Grid Technologyies Conference (ISGT 2011), Anaheim, CA, January 2011.
61. S. Srivastava, F. M. Ciorba and I. Banicescu, "Employing a Study of the Deterministic Metrics to Assess the Robustness of Dynamic Loop Scheduling," in Proceedings of the 17th Annual IEEE High Performance Computing Symposium(HiPC), Goa, India, December 2010.
62. M. Tsugawa, P. Riteau, A. Matsunaga, J. Fortes, 'User-level Virtual Networking Mechanisms to Support Virtual Machine Migration Over Multiple Clouds', IEEE International Workshop on Management of Emerging Networks and Services,December 2010.
63. Jing Xu and José Fortes, "Multi-objective Virtual Machine Placement in Virtualized Data Center Environments," in Proceedings of the 2010 IEEE International Conference on Green Computing and Communications (GreenCom 2010), Hangzhou, China, December 2010.
64. Bithika Khargharia, Haoting Luo, Youssif Alnashif, Salim Hariri, "AppFlow: Autonomic Performance-per-Watt Management of Large-Scale Data Centers," in Proceedings of the 2010 IEEE/ACM International Conference on Green Computing and Communications (GreenCom2010), Hangzhou, China, December 2010.
65. David Wolinsky, Kyungyong Lee, P. Oscar Boykin, Renato J. Figueiredo, "On the Design of Autonomic, Decentralized VPNs," presented at CollaborateCom 2010, Chicago, IL, October 2010.
66. I. Rodero, E. K. Lee, D. Pompili, M. Parashar, M. Gamell, and R. J. Figueiredo, "Towards Energy-Efficient Reactive Thermal Management in Instrumented Datacenters," Proceedings of the IEEE/ACM International Conference on Energy Efficient Grids, Clouds and Clusters Workshop (E2GC2), Brussels, Belgium, October 2010.
67. Pierre St. Juste, David Wolinsky, Kyungyong Lee, P. Oscar Boykin, Renato J. Figueiredo, "SocialDNS: A Decentralized Naming Service for Collaborative P2P VPNs," presented at CollaborateCom 2010, Chicago, IL, October 2010.

To request additional copies of this newsletter, please send an email to Julie at jooly@acis.ufl.edu