**ACIS/CAC RESEARCH ACTIVITIES**

The mission of the Advanced Computing and Information Systems Laboratory (ACIS) is to conduct fundamental and applied research on all aspects of systems that integrate computing and information processing. Current ACIS research falls under the broad categories of Cloud/Grid-computing Middleware, Cyberinfrastructure for e-science, Autonomic computing and Peer-to-peer computing.

**Faculty:** José A. B. Fortes, Renato Figueiredo, Andy L, Tao L, Mauricio Tsugawa, Andréa Matsunaga

**Students:** Prapraporn Rattanatamrong, Selvi Kadivrel, Tao Seung Kang, Meng Han, Gillia Lee, Joseph Makkar, Pierre St. Juste, Jianguyan Xu, Kyungyong Lee, Yonggang Liu, Heungsik Eom

---

**ACIS FACILITIES**

- State-of-the-art computing, storage and networking facilities
- Unique environment for experimental research and design of distributed systems that use virtualization software developed by commercial and open-source projects
- >200 servers, >1250 cores, >4.8TB of memory, >260TB storage
- FutureGrid cluster: IBM DataFLEX connected to Florida Lambda Rail.
- NIMDcloud: allows up to 64 cores and 512MB of memory in a single image
- Autonomic Testbed: enables research on cross-layer automations for datacenter management
- VM and cloud: rich set of VMs: (VMware, Xen, KVM, QEMU), and cloud software (Nimbus, OpenStack)
- Storage: centralized (IBM DS4800) and cloud-based (OpenStack)

---

**FEDERATED COMPUTE CLOUD (DiG)**

- **Need:** Platform to perform operations on large digitized collections housed in US institutions

  **Objective:** Design, build and support Platform-as-a-Service integration modules for domain-specific scientific workflows and data parallel applications (e.g., Camera, CloudBLAST, Condor, Kepler, Pegasaus, Pig, Specify)

  **Case Study:** Emphasis on platforms that use standard vocabularies for metadata, interoperable APIs for software and data management, and perform operations on geographically distributed data

---

**DATA CENTER RESOURCE MANAGEMENT**

- **New VM requests**
- **Power model**
- **Temperature model**

  - **Controllers predict + provision virtual resources for applications**
  - **Multiojective optimization (30% faster with 20% less power)**
  - **Use fuzzy logic, genetic algorithms and optimization methods**
  - **Use cross-layer information to manage virtualized resources to minimize power, avoid hot spots and improve resource utilization**

  - **J. Xu, J. Fortes, “Multi-Objective Virtual Machine Placement in Virtualized Data Center Environments” Int. Conf. on Green Computing and Communications, 2010.**

---

**SELF-CARING IT SYSTEMS**

- **Need:** High likelihood of failures due to scale, heterogeneity, distribution, configuration errors, complexity, etc.
- **Objective:** Self-Caring IT systems capable of proactively managing their health in addition to reactively responding to failures.
- **Case Study:** MapReduce on a cloud environment, mitigate performance (job runtime) penalties under node and task crashes through dynamic resource scaling, migration and rejuvenation.


---

**CROSS-LAYER AUTOMONICS INTERCLOUD TESTBED**

- **Distributed testbed at US, UA and RU**
- **Framework for cross-layer optimization studies**
- **Thermal modeling**
  - **Sensors in different locations**
  - **Report temperature and humidity**
- **Short-term/long-term predictions to improve controllers’ performance**
- **Intercloud protocols**
- **Virtual networking**
- **Interoperability studies**
- **Sky computing**

---

**FUTUREGRID 2009-2013 (OCI-0901212)**

http://futuregrid.org

- **FutureGrid is an international testbed that supports computer science and computational science research in cloud, grid and parallel computing (HPC)**
- **Provides a flexible development and testing platform for middleware and application users looking at interoperability, functionality, performance or evaluation, and a rich education and teaching platform for advanced cyberinfrastructure classes**

---

**UNIVERSITY OF FLORIDA ROLES IN FUTUREGRID**

- **Key roles in middleware, teaching, education and outreach with emphasis on overlay networks and virtual appliances.**
- **Lead roles in operations team, and education and outreach team**
- **UF FutureGrid team: IBM DataFLEX – 256 cores, 768GB of RAM, 20TB of storage**

---

**VIRTUAL NETWORK (ViNE) PROJECT**

- **Need:** Enable communication among cloud resources overcoming limitations imposed by firewalls.
- **Objective:** Develop an easy-to-manage intercloud communication infrastructure.
- **Case Study:** Interconnect VMs deployed on different FutureGrid clouds.


---

**SKY COMPUTING**

- **Need:** Large scale scientific computation requires resources that a single cloud provider cannot offer.
- **Objective:** Efficiently combine cloud technologies (ViNe, Nimbus, Hadoop, etc.) to form an intercloud virtual cluster.
- **Case Study:** Execution of CloudBLAST on 750 VMs (across 3 FG sites and 3 Grid’5000 sites), and 1500 cores


---

**IDGBIO 2011-2016 (EF-1151210)**

http://idgbio.org

- **The Home Uniting Biocollections (HUB) established and funded by the NSF. Advancing Digitization of Biological Collections (ADDLC)**
- **ADBC is a national effort to support digitization of all vouchered biological and paleontological collections housed in US institutions**
- **IDGBio oversees data integration, support for computational needs and assessment of new technologies and programs to facilitate digitization efforts including directing the development of “appliances” and the development of cloud storage capabilities**

---

**APPARLACE-BASED CLOUD-ORIENTED ARCHITECTURE**

- **Need:** Software appliances and reliance on cloud computing to be adaptable and handle a large set of tools, scenarios and partners.
- **Objective:** Creation, dissemination, deployment and sharing of “virtual appliances” encapsulating software environments and tools found in the IDGbio in a “sandbox” environment.
- **Case Study:** Platform for cyber-learning packages with self-contained executable educational modules tailored by educators and seamlessly deployed by students

---

**FEDERATED STORAGE CLOUD (DiG)**

- **Need:** Distributed storage infrastructure with large capacity and common namespace that is reliable, fast, accessible and affordable
- **Objective:** Design, build and deploy a customized cloud storage environment for collections based on off-the-shelf disks and offering distributed object storage services including distributed federated namespace registry/mapper and bucket/object caching and replication manager
- **Case Study:** Support efficiently different platforms with heterogeneous access patterns (e.g., email, management platform, distributed content management platform, distributed computation platform, federated appliance repository).