

# IEEE Transaction on Cloud Computing, Special Issue on Many-Task Computing in the Cloud

<http://www.computer.org/tcc>

The Special Issue on Many-Task Computing (MTC) in the Cloud will provide the scientific community a dedicated forum, within the prestigious IEEE Transactions on Cloud Computing journal, for presenting new research, development, and deployment efforts of loosely coupled large scale applications on Cloud Computing infrastructure. MTC, the theme of this special issue, encompasses loosely coupled applications, which are generally composed of many tasks to achieve some larger application goal. This special issue will cover challenges that can hamper efficiency and utilization in running applications on large-scale systems, such as local resource manager scalability and granularity, efficient utilization of raw hardware, parallel file-system contention and scalability, data management, I/O management, reliability at scale, and application scalability. We welcome paper submissions in theoretical, simulations, and systems topics with special consideration to papers addressing the intersection of petascale/exascale challenges with large-scale cloud computing. We seek submission of papers that present new, original and innovative ideas for the "first" time in TCC (Transactions on Cloud Computing). That means, submission of "extended versions" of already published works (e.g., conference/workshop papers) is not encouraged unless they contain significant number of "new and original" ideas/contributions along with more than 49% brand "new" material. For more information on this special issue, please see <http://datasys.cs.iit.edu/events/TCC-MTC15/>.

## Special Issue Overview

The advent of computation can be compared, in terms of the breadth and depth of its impact on research and scholarship, to the invention of writing and the development of modern mathematics. Scientific Computing has already begun to change how science is done, enabling scientific breakthroughs through new kinds of experiments that would have been impossible only a decade ago. As computing becomes a pervasive part of the scientific process, there is a great opportunity to make powerful computing techniques, previously reserved for projects with only the largest investments, available to a broad scientific community.

The massive increase in concurrency provided by modern hardware presents a challenge to scientific applications with large existing investments in previously developed software and limited ability to redesign from scratch using the latest programming models. Many-task computing (MTC) studies technologies, simple and advanced, to rapidly compose highly scalable applications from existing sequential codes. MTC encompasses loosely coupled applications, which are generally composed of many tasks (both independent and dependent tasks) to achieve some larger application goal. Growing from the successes of Globus, Condor, and national-scale grid computing infrastructures, MTC techniques have been deployed on many systems from single many-core systems (leveraging GPGPUs and Intel MIC accelerators), to the largest multi-petascale high-performance computing (HPC) systems. The development and deployment of these MTC systems have expanded the utility of the underlying technologies and fed back to improve the performance and usability of the technologies themselves. Similarly, technologies developed for cloud computing (including MapReduce-based models) can provide additional connections and innovations in computing techniques. **This special issue is a unique venue to promote HPC-related concepts to the broader scientific and cloud computing communities.**

We are entering into a "big data" era, as advances in networking, instrumentation, simulation technologies, Internet computing and social networks are producing data at an unprecedented rate. The collection, storage, analysis and sharing of this data are thus one of the greatest challenges in the 21st century. Support for data intensive computing is critical to advancing modern science as storage systems have experienced an increasing gap between its capacity and its bandwidth by more than 10-fold over the last decade. There is an emerging need for advanced techniques to manipulate, visualize and interpret large datasets. While commonly associated with Hadoop and related systems, technologies from HPC and MTC are also applicable. This provides an opportunity to exchange large-scale data management technologies between scientific applications and industrial techniques, which is another emphasis of this special issue.

Scientific Computing is the key to many domains' "holy grail" of new knowledge, and comes in many shapes and forms. Exchange of ideas from HPC, MTC, and cloud communities is a critical path to the adoption of advanced techniques to best utilize emerging, highly concurrent systems. Underlying techniques for concurrency and data processing originating in the HPC space must be delivered to the broader community to promote future investment in HPC research programs and, more generally, advance scientific investigations.

## Topics of Interest

We seek submission of papers that present new, original and innovative ideas for the "first" time in TCC (Transactions on Cloud Computing). That means, submission of "extended versions" of already published works (e.g., conference/workshop papers) will only be encouraged if they contain

significant number of "new and original" ideas/contributions along with more than 49% brand "new" material. TCC expects submissions to be complete in all respects including author names, affiliation, bios etc. We aim to cover topics related to Many-Task Computing on Clouds.

- **Compute resource management**
  - Scheduling
  - Job execution frameworks
  - Local resource manager extensions
  - Performance evaluation of resource managers in use on large scale systems
  - Dynamic resource provisioning
  - Techniques to manage extreme concurrency and accelerators
  - Challenges and opportunities in running many-task workloads on Cloud Computing infrastructure
- **Storage architectures and implementations**
  - Distributed file systems
  - Parallel file systems
  - Distributed metadata management
  - Content distribution systems for large data
  - Data caching frameworks and techniques
  - Data management within and across data centers
  - Data-aware scheduling
  - Data-intensive computing applications
  - Eventual-consistency storage usage and management
- **Programming models and tools**
  - MapReduce, its generalizations, and implementations
  - Many-task computing middleware and applications
  - Parallel programming frameworks
  - Ensemble MPI
  - Service-oriented science applications
- **Large-scale workflow systems**
  - Workflow system performance and scalability analysis
  - Scalability of workflow systems
  - Workflow infrastructure and e-Science middleware
  - Programming paradigms and models
- **Large-scale many-task applications**
  - High-throughput computing (HTC) applications
  - Data-intensive applications
  - Quasi-supercomputing applications, deployments, and experiences
  - Application coupling, integration, and composition
  - Algorithms for many-task applications- Monte Carlo, parameter sweep/search, uncertainty quantification
  - Performance evaluation
- **Performance evaluation**
  - Theoretical vs. real systems
  - Simulations
  - Reliability and fault tolerance of large systems

## Important Dates

- Papers Due: March 13th, 2015
- First round decisions: June 1st, 2015
- Major Revisions if needed: August 1st, 2015
- Final decisions: September 1st, 2015
- Publication Date: Fall 2015 (may vary depending on production queue)

## Paper Submission

Authors are invited to submit unpublished and original work to the IEEE Transactions on Cloud Computing (TCC), Special Issue on Many-Task Computing in the Cloud. If the paper is extended from an initial work, the submission must contain at least 50% new material that can be qualified as "brand" new ideas and results. The paper must be in the IEEE TCC format, namely 14 double-column pages or 30 single-column pages (Note: All regular paper page limits include references and author biographies). These length limits are taking into account reasonably-sized figures and references. For details on paper formatting, please visit TCC website. Papers must be submitted using the submission system: <https://mc.manuscriptcentral.com/tcc-cs>, by selecting the special issue option "SI-MTC".

## Guest Editors:

- Dr. Ioan Raicu (Illinois Institute of Technology & Argonne National Laboratory)
- Dr. Justin M. Wozniak (Argonne National Laboratory & University of Chicago)
- Dr. Ian Foster (University of Chicago)
- Dr. Yong Zhao (University of Electronic Science and Technology of China, China)

For any queries on this special issue, please contact Guest editors by email: [mtc15-tcc-editors@datasys.cs.iit.edu](mailto:mtc15-tcc-editors@datasys.cs.iit.edu).