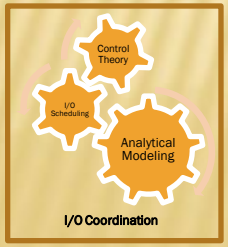
Patricia J Teller and Sarala Arunagiri

The University of Texas-El Paso

# I/O COORDINATION TO IMPROVE HEC SYSTEM PERFORMANCE:

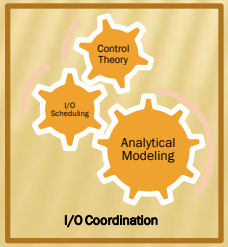
# A MARRIAGE OF ANALYTICAL MODELING, CONTROL THEORY, AND DIFFERENTIATED I/O PERFORMANCE




ER25948

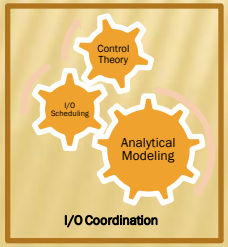
# RESEARCH GOALS

1. Extend scalability of checkpoint/restart
2. Reduce I/O system stress & resultant failures
3. Increase meaningful utilization of allocated HEC resources → enhance HEC system performance



# SIGNIFICANT PROGRESS TO DATE – MODELING

- **Problem description:**
  - Via analytical and parametric models, which used parameter sets that describe a few high-end computer systems and a hypothetical application, we demonstrated that *there exist potential opportunities for periodic-checkpointing applications to improve their performance by intelligent selection of the checkpoint interval*, i.e., the frequency with which checkpointing is performed.
  - Selection of the checkpoint interval involves a *trade-off between execution time and the number of checkpoint I/O operations performed*.
  - *Challenge: Prove that such opportunities exist for all values of the models' parameters*
- Using our analytical models, in collaboration with an ORNL scientist (Dr. Nagaswara S. V. Rao) *mathematically proved that "for all values of the parameters of our analytical models of a periodic-checkpointing application's execution time and the number of checkpoint I/O operations, regions of opportunity exist that potentially enhance system performance".*

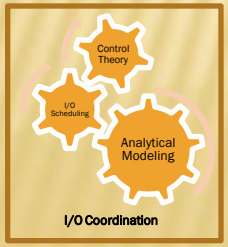



Solved a problem identified in our proposal
Related paper (for publication) in progress

ER25948

# SIGNIFICANT PROGRESS TO DATE – I/O SCHEDULING

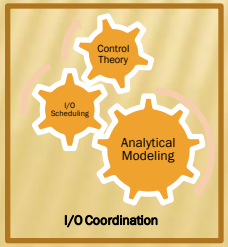
- Performed *bottleneck analysis of RAID storage systems*
- Designed, developed, and implemented *FAIRIO, an I/O scheduling algorithm that provides differentiated service with respect to disk time utilization of RAID storage utilities*
- *Demonstrated the provisioning of RAID I/O differentiated service* under various conditions by performing a set of experiments using a modified version of DiskSim 3.0
- *Submitting a related paper for publication*




Designed FAIRIO, an I/O scheduling algorithm that provides differentiated service w.r.t. disk-time utilization of RAID storage utilities
Paper being submitted (September) for publication

# SIGNIFICANT PROGRESS TO DATE – CHECKPOINTING

- *Established a collaborative working relationship with Texas Advanced Computing Center (TACC), Austin, TX – Drs. Jim Browne, John Hammond, and Dan Stanzione – and John Phillips, NAMD developer*
- *Designing experiments regarding periodic-checkpointing applications*, the first two of which will be NAMD and RAxML
  - NAMD, recipient of a 2002 Gordon Bell Award, is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems. Based on Charm++ parallel objects, NAMD scales to hundreds of processors on high-end parallel platforms and tens of processors on commodity clusters using gigabit ethernet.
  - RAxML (Randomized Axelerated Maximum Likelihood) is a program for sequential and parallel Maximum Likelihood-based inference of large phylogenetic trees – it is highly parallelizable.
- Analyzing *data regarding the experimental system*, in particular, MTTI and statistics re: I/O contention
- *Instrumenting NAMD and RAxMLto record information that can be used to characterize solution time, productive I/O, and defensive I/O*
- Once the experimental procedure has been tested, we will collaborate with the bigger users of TACC systems to conduct experiments using their codes (we will target codes of interest to DoE) and then will establish collaboration with DoE Labs (likely ORNL, who we have already visited, and LLNL)




**Almost ready to commence experiments using NAMD and RAxML that, hopefully, will validate the results of our analytical models**

# SCHOLARLY ACTIVITIES

- **Posters**
  - "Using Mathematical Modeling to Enhance the Scalability of Checkpointing Applications," Sarala Arunagiri, John Daly, Patricia J. Teller, and Ricardo Portillo, poster presented by Sarala Arunagiri at the 2010 DOE Applied Mathematics Program Meeting, 3-4 May 2010, Berkeley, CA.
- **Presentations**
  - "HEC I/O Coordination based on Judicious Checkpointing and I/O QoS," Patricia J. Teller (PI), DoE Joint Mathematics/Computer Science Institute Kickoff Meeting, SC09, Portland, Oregon, 17 November 2009.
  - "HEC I/O Coordination based on Judicious Checkpointing and I/O QoS (Extended version)," Patricia J. Teller (PI), The National Center for Computational Sciences (NCCS) Seminar Series, Oak Ridge National Laboratory, Oak Ridge, TN, 10 October 2009. Note that this visit to Oak Ridge National Laboratory, which was made by both Drs. Teller and Arunagiri (Co-PI) included meetings with many different research groups. We are near a point in our research where it makes sense to follow-up on this visit and suggest possible collaborative activities.
- **Publications in preparation**
  - "FAIRIO: An Algorithm for I/O Performance Differentiation," Sarala Arunagiri, Yipkei Kwok, Seetharami R. Seelam, Patricia J. Teller, and Ricardo Portillo, submitting in September 2010.