Large-scale Systems Performability

Thilo Kielmann

- Open M.Sc. Projects (with)
  - Ana-Maria Oprescu (presenter)
  - Stefania Costache, Kaveh Razavi, Alexandru Uta
  - Guillaume Pierre (U Rennes)
  - QuoVide (Heerhugowaard)
Large-scale Systems Performability

- Large-scale systems:
  - hundreds or thousands of nodes
  - geographically distributed
  - “everything cloud”

- Performability:
  - performance, fault-tolerance, cost, energy
  - trading these properties in system design
Thilo's group

Kaveh Razavi
Ana Oprescu
Stefania Costache
Alexandru Uta
Ana, Stefania, Kaveh: Efficient resource management in clouds
Kaveh: Control the VM boot process

- Controlling the VM boot process
  - In cloud infrastructures: make booting virtual machines (VMs) faster.
  - Resume from pre-booted snapshots, without support from the VM itself (for any VM image)
  - Project: Observe the I/O traffic of a booting VM and decide when it is booted, then take a snapshot
Managing efficiently rented resources is challenging
- cloud providers typically offer static sets of VM types
- workload resource demands change over time

Analyze prediction algorithms for (re)computing optimal mappings of applications to VMs.
- Scalability is key (requests are processed on-line)
- Validation via simulations using real workload traces
Ana, Cosmin(UvA): Data-intensive portfolio scheduling

- Bags-of-tasks applications have different data transfer to compute ratios
- Network characteristics affect the set of optimal schedules
- What is the relationship between this ratio and various scheduling policies?
- What is the trade-off between amount of apriori information and decision accuracy?
Ana: Heterogeneous pricing models in clouds

- Scheduling applications under heterogeneous pricing models:
  - Google: first 10 min fixed rate, then per minute, Amazon: per hour, Azure: per minute
  - Impact on set of optimal schedules (total execution time and monetary cost)
  - Analyze algorithms for trade-off time-to-solution versus accuracy
Alex, Stefania: Many Task Computing
Alex: Fault-tolerance

- Fault-tolerance for many-task computing applications
  - MemFS, our in-memory file system stores temporary files in the main memory of an MTC application
  - Study existing MTC applications and their file communication patterns
  - Devise an FT-storage scheme that combines file redundancy with the ability to recompute certain files, while minimizing both types of overhead
Alex, Stefania: Data locality scheduling

- Data locality scheduling of MTC applications on an in-memory file system
  - MemFS, our in-memory file system stores temporary files in the main memory of an MTC application
  - Build a scheduler that starts tasks on those nodes where files (file stripes) are stored
  - Resolve conflicting goals when multiple files are needed per task
Stefania: Contention-adaptive scheduling

- Contention-adaptive scheduling of MTC applications on multi-core infrastructures
  - In a (MTC) workflow, different types of tasks have different resource requirements (CPU, memory, network, I/O)
  - Build a scheduler for heterogeneous MTC applications where different types of tasks together utilize the capabilities of a server without creating contention
Integrated blackbox+whitebox performance modeling of Cloud applications

- Clouds provide vast amounts of virtual machines of various types
- Users can hardly make good choices to fit their applications' needs
- Project: model performance needs for selecting cloud resources
- (some funds available for staying in Rennes)
QuoVide (Heerhugowaard)

- Small Dutch company providing social media platforms for large companies like KPN or RTL, three projects:
  - Apache Spark for real-time social intranet analytics
  - A Change Data Capture (CDC) toolkit for keeping system layers in sync with the database
  - CockroachDB (Google Spanner) as a basis for a SAAS solution for social networking software
Wiki links

- A Change Data Capture (CDC) toolkit for keeping system layers in sync with the database
- Apache Spark for real-time social intranet analytics
- CockroachDB (Google Spanner) as a basis for a SAAS solution for social networking software
- Contention-adaptive scheduling of MTC applications on multi-core infrastructures
- Controlling the VM boot process
- Data locality scheduling of MTC applications on an in-memory file system
- Dealing with heterogeneous pricing models in clouds
- Fault-tolerance for many-task computing applications
- Integrated blackbox+whitebox performance modeling of Cloud applications
- Optimizing IaaS resource management through workload pattern prediction
- Various projects in Virtual Reality at Personal Space Technologies
- Various projects in game development at Sticky Studios