1. Background and Challenges

- **Advantages of Cloud computing**
  - Elastic
  - Scalable
  - Cost-effective

- **Challenges of Cloud computing**
  - Hard to guarantee the user’s expectations of Quality of Service (e.g., reliability, timeliness)
  - Shared resources by a number of users
  - Compute-intensive processes
  - Network traffic

- **Research Challenges of Cloud Computing to Support DRE Systems**
  1. Need effective resource monitoring
  2. Real-time hypervisors and data-center networks
  3. Predictable dependability mechanisms

2. Literature Survey

- **Research criteria based on literature survey**
  1. Effective resource monitoring (common to all)
  2. Timeliness in data-center networks
  3. Real-time scheduling in hypervisors
  4. High availability via replication of virtual machines

- **Timeliness in data-center networks**
  1. DCTCP
     - TCP modified protocol
     - Better throughput than TCP
  2. D^3
     - Deadline aware control protocol
  3. D^2TCP
     - Reducing deadline miss ratio

- **Real-time scheduling in hypervisors**
  1. RT-Xen
     - 4 Fixed priority real-time schedulers used in Xen
  2. Scheduler S
     - The modified Xen scheduler for soft-real-time tasks

- **High availability via replication of virtual machines**
  1. Remus
     - Asynchronous replication and speculative execution
     - Supported with Xen
  2. Kemari
     - Lock-stepping and continuous check-pointing approach
     - Supported with KVM
  3. HydraVM
     - Storage-based and memory-efficient approach

3. Addressing Challenge 1 – Need for Effective Resource Monitoring

- **SQRT-C (A scalable and QoS-enabled cloud monitoring system)**
  - Using OMG Data Distribution Service (DDS) real-time publish/subscribe (pub/sub) middleware

- **SQRT-C software artifacts**
  - DDS-based pub/sub communication
  - Monitoring Manager

- **DDS-based pub/sub communication**
  - Disseminating monitoring information for virtual resources from the source (i.e., publishers) to the sinks (i.e., the subscribers)
  - Supporting the QoS requirements

- **Monitoring Manager**
  - Serving as the orchestrator for the deployment of data-writers and data-readers of the DDS pub/sub mechanism

- **Performance Evaluation**
  - SQRT-C outperforms RESTful services in terms of response time (Figure 2) and jitter for real-time applications

4. Ongoing and Future Work

- **Unresolved challenges**
  - Trade-off between timeliness and high-availability with strong consistency
  - Tradeoffs between response time and consistency
  - BASE (Basically Available replicated Soft state with Eventual consistency)
  - ACID (atomicity, consistency, isolation, and durability) database models

- **Redundancy-based fault recovery mechanisms for DRE systems**
  - Replication using primary-backup
  - A proactive, resource-aware fail-over strategy
  - A resource-aware allocation based on backup resource overbooking

- **Our proposed research**
  - Implementation of a fault-tolerant cloud architecture applying redundancy-based fault recovery mechanisms
  - Performance analysis for trade-off between strict timeliness and strong consistency
  - Integration of real-time hypervisors and deadline-aware data-center networks

- **Work in progress**
  - A framework for automated placement of virtual machine replicas for DRE systems
  - Bin-packing heuristics developed

5. References


Research supported in part by NSF CNS 0915976