Towards a Compositional Service Architecture for Real-Time Cloud Robotics

Shahina Begum*, Moris Behnam*, Thomas Larsson*, Thomas Nolte*†, Kristian Sandström*†
*MRTC / Mälardalen University, Västerås, Sweden
†Software Systems / ABB Corporate Research, Västerås, Sweden

***Abstract*—This paper presents our ongoing work towards a compositional service architecture that integrates cloud technology for computational capacity targeting real-time robotics applications. In particular, the paper focuses at the challenges inherent within the data center where the services are executing. Here, we outline characteristics of the services applied in the real-time cloud robotics application, along with the service management and corresponding task model used to execute the services. Moreover, several key challenges that must be addressed towards integrating cloud technology in real-time robotics are identified.**

I. Introduction

Collaborative robotics [1] is a promising solution to the next generation of flexible automation systems. This research field has emerged to provide safe collaboration among robots and humans, where the robot by construction is not able to harm its surrounding environment. Increasing autonomy of such robots will enhance their performance in terms of unsupervised action, which will simplify coordination between systems of humans and robots. Such autonomous yet collaborative embedded systems will require sophisticated and predictable sensing and actuation, along with complex computation, all with predictable ultra-short delay times.

However, to keep the per-unit cost low it is desirable to avoid expensive embedded hardware. Instead, by connecting and integrating robots and sensors in the cloud it will be possible to increase autonomy and intelligent functionality for future real-time collaborative cloud robotics systems [2], resorting to relatively low cost and light weight embedded sensors and processing capacity while utilizing the vast computational and data resources available in the cloud [3]. To make this happen we need time predictable algorithms running on cloud enabled system architectures providing both high performance and predictable computational capacity.

In this paper, we look closer at three areas of research with corresponding research challenges that must be addressed towards enabling the application of real-time could robotics: (a) Flexible architectures, integrating robots and sensors with local cloud (aka fog) technology, allowing for adaptive runtime operation of resources and services, (b) Adaptive and predictive algorithms, applied on cloud services for robot motion planning, intelligent multi-sensor fusion for 3D awareness, and decision support, and (c) Predictable platform technology, providing parallel computational capacity on nodes and communication on networks.

II. Architecture

The system architecture used in this paper includes several robots and sensors (cameras) connected through wired and wireless communication networks to a private cloud (fog). The architecture must be time predictable, supporting a correct integration of robots with sensors, other robots and the cloud. Real-time operation also requires ultra-fast (and time predictable) adaptive and predictive algorithms implementing machine learning, decision making and motion planning that will rely on new cloud services. This level of predictability required for sustainable correct real-time operation prevents a straightforward technology transfer, adoption and implementation from current cloud systems. Instead we need an enhanced predictable platform technology to guarantee the level of Quality-of-Service (QoS) required for real-time cloud robotics.

III. Services

Two central services of the real-time cloud robotics application are the 3D environment service and the motion control service.

The 3D environment service utilizes sensor data collected from multiple sensor sources. Intelligent sensor data fusion together with smart sensor technologies ensures recognition of environmental features, e.g., shapes and properties enabling continuous sensing of the 3D environment. Thus, it enhances the detection and positioning of objects. To create a 3D model of the environment, we will combine images with other sensor data to generate environment textures and geometry of the 3D model. Sensor data fusion using a Bayesian framework will be used to detect dynamic range measurements in a dynamic environment. Moreover, fusing supplement data obtained from 2D imagery will generate high precision 3D point cloud data. Temporal analysis for remote sensing will be performed based on entropy analysis. Over the last few years several technological advancements enriches autonomous 3D modeling [4]. However, reliable 3D modeling is still a complex task. In particular, when the environment is changing the number of uncertainties increase in primitive shape detection and higher-level knowledge inference. In the proposed work, we will devise novel strategies to fuse sensor signals at feature level and compare the result in terms of recognition accuracy. This will be done based on our existing model on sensor fusion [5].

In the targeted application there can be several parallel instances of this service, depending on how many environment models that should be provided. Individual sensors can contribute with data to one or more of these environment services.

The motion control service uses the 3D environment model together with a 3D model of the robot. Each robot arm has an associated 3D robot model. By leveraging on cloud computational performance and availability of vast amount of data, we can provide robots with real-time decision support allowing for collision free motion of robotic arms. For such operation, the robots must have a notion of the environment and they need to know how to move within this environment.

Determining robot motions in dynamic collaborative environments requires sophisticated solutions. Algorithms will be designed to enable real-time adaptive collision avoidance and motion planning, which rely on seamless integration with the 3D environment construction module and cloud robotics services. By using a smarter combination of collision query types, an improved awareness of the geometric situation can be derived, which can be exploited to determine a richer set of collision-free trajectories. To support reliable on-the-fly re-planning in a predictable way as a response to unforeseen obstacles or events, real-time dynamic motion planning is needed. This includes robot kinematics, preferred error-bounded path computation, minimization of estimated trajectory costs, and real-time scheduling of operations. Parallel solutions which exploit modern heterogeneous platforms will be designed and evaluated together with a framework for physical simulation, algorithm visualization, and auto-tuning. Algorithms can operate in off-line and on-line contexts, depending on the nature of the environments and obstacles. The dynamic case is particularly challenging, since on-line methods must deal with unpredictable obstacles and re-planning of routes and motion trajectories continuously. Solutions can use, e.g., spatial space decomposition, occupancy grid maps, graph data structures, physical simulation, and evolutionary methods [6].

IV. Platform and service management

We target an execution platform covering multi-, many-core and distributed systems. The services must be managed online, and the services will utilize a parallel task model for computational performance.

The service manager handles the services of the cloud robotics application online with respect to their resource requirements needed to satisfy a desired level of QoS. The resources are managed holisticly, handling joint reservations for sets of heterogeneous resources across the platform. This resource management includes reallocating of bandwidth among reservations according to actual usage, thus reducing bandwidth waste and boosting efficiency. Central challenges include end-to-end delay analysis, reservations definition to enforce end-to-end properties, and reservations bandwidth monitoring and reallocation. Addressing these challenges will allow for robots to execute cloud services with bounded latency in an open setting with efficient resource usage.

The parallel task model is central for the execution of services in a data center consisting of a large number of computing cores. Hence, we need a suitable parallel task model for cloud computing, allowing for time predictable execution of services. The task model and its associated run-time mechanism should be compatible with dynamic addition and removal of services online in a predictable manner, as robots and sensors are added and removed during run-time. Hence it is highly desirable that the tasks of the task model are compositional with respect to timing and resource usage.

V. Work in Progress

This work outlined in this paper addresses the challenging issues of virtualization, scheduling and timing analysis for QoS guarantees, and parallel models for software execution. As a first step the models and their implementations are developed for parallel execution of static systems, where the architecture and components are fully known beforehand, i.e., focus is on the task model and its parallel execution. Later we will allow for dynamic behavior supporting adaptivity, adding and removing both sensors and robots. Here the service manager becomes important, as well as protocols for predictable service discovery and distributed resource management.

The targeted execution platform consists of virtualized many-core computers in the cloud. Compositional and parallel task models suitable for virtualized environments in the cloud has not been well investigated. Targeting soft real-time constraints, some recent solutions allow for allocation and scheduling of parallel tasks modeled by fork/join or DAG models, e.g., [7]. For the real-time cloud robotics application, we will develop compositional solutions targeting hard real-time systems. In such a setting it should be possible to add resource users, i.e., sensors and robots, to the execution platform, without creating unpredictable interference among resource users.

The current work-in-progress builds upon our experience with flexible network reservations [8] and other contemporary projects, such as IRMOS [9], to open the way to adaptive resource and service management in a heterogeneous resource setting.

References

[1] A. Shikany, "Collaborative robots end user industry insights," in *RIA Robotic Industries Association*, 2015.

[2] G. Mohanarajah, V. Usenko, M. Singh, R. D'Andrea, and M. Waibel, "Cloud-based collaborative 3d mapping in real-time with low-cost robots," in *IEEE Transactions on Automation Science and Engineering, 12(2):423-431*, 2015.

[3] B. Kehoe, S. Patil, P. Abbeel, and K. Goldberg, "A survey of research on cloud robotics and automation," in *IEEE Transactions on Automation Science and Engineering, 12(3):398-409*.

[4] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "Rgb-d mapping: Using kinect-style depth cameras for dense 3d modeling of indoor environments," in *Intl. Journal of Robotics Research 0(0) 117*, 2012.

[5] S. Begum, S. Barua, R. Filla, and M. Ahmed, "Classification of physiological signals for wheel loader operators using multi-scale entropy analysis and case-based reasoning," in *Expert Systems with Applications 41:295-305*.

[6] P. Raja and S. Pugazhenthi, "Optimal path planning of mobile robots: A review," in *Intl. Journal of Physical Sciences 7.9: 1314-1320*, 2012.

[7] S. Wu, L. Zhou, H. Sun, H. Jin, and X. Shi, "Poris: A scheduler for parallel soft real-time applications in virtualized environment," in *IEEE Transactions on Parallel and Distributed Systems, vol.PP, no.99*, 2015.

[8] J. Silvestre, R. Marau, P. Pedreiras, and L. Almeida, "On-line qos management for multimedia real-time transmission in industrial networks," in *IEEE Transactions on Industrial Electronics, 58(3):1061-1071*, 2011.

[9] T. Cucinotta, F. Checconi, G. Kousiouris, D. Kyriazis, T. Varvarigou, A. Mazzetti, J. P. Z. Zlatev, M. Boniface, S. Berger, D. Lamp, T. Voith, and M. Stein, "Virtualised e-learning with real-time guarantees on the irmos platform," in *IEEE Intl. Conf. on Service-Oriented Computing and Applications (SOCA 2010)*, 2010.