

Thesis Title: Efficient Techniques for Scheduling DAG Applications in Distributed Environments

Student: Debabrata Senapati

Roll No: 176101003

Abstract:

A Cyber-Physical System (CPS) integrates two sub-systems - a cyber sub-system and a physical sub-system. The cyber sub-system is often a heterogeneous distributed computing system that executes applications for regulating mechanisms associated with the physical sub-system, typically consisting of electromechanical components. Real-Time Cyber-Physical Systems (RT-CPSs) are characterized by their ability to respond to events that may happen in their operating environment within stipulated time bounds. The accuracy of these systems depends not only on the delivered results but also on their completion times. Applications in today's RT-CPSs are often represented by Directed Acyclic Graphs (DAGs) due to their distributed nature and complex interactions among component functionalities. In such DAGs, nodes represent tasks associated with the application, while edges denote interdependencies among tasks. To meet functionality-specific high-performance demands, these DAGs are often implemented on heterogeneous RT-CPS platforms, where (i) the same task may exhibit different execution time requirements on different processors, and (ii) inter-task messages containing the same amount of data may incur distinct transmission times on the different communication channels due to variations in channel bandwidths. The RT-CPS applications may be aperiodically triggered by an external event or may execute in infinite loops, periodically acquiring data from the environment through sensors at a particular frequency, processing the same, and then producing processed data via actuators.

*This dissertation deals with the design of resource allocation mechanisms for DAG-structured applications on heterogeneous distributed RT-CPSs. The thesis which unfolds through the dissertation is as follows: For the mentioned system scenarios, the list-based design philosophy is effective towards obtaining low-overhead but efficient scheduling mechanisms for satisfying diverse objectives/constraints related to resource usage efficiency, timeliness, energy, security, temperature, etc. All list scheduling heuristics typically consist of two phases, (i) *Task prioritization*: for listing tasks in a specific priority order, and (ii) *Task-to-processor mapping*: for allocating the tasks in the order of their priorities on suitable processors and associating with them appropriate execution start times. The contributions of this thesis are categorized into five phases as follows: (i) The first phase focuses on the development of an efficient real-time DAG-scheduling framework which attempts to minimize a generic penalty function. The designed penalty function can be amicably adopted towards its deployment in various application domains such as real-time cyber-physical systems like automotive and avionic systems, cloud computing, smart grids, etc. (ii) In the second phase, we develop a state-space search guided heuristic scheduling algorithm called *HMDS*, whose objective is to minimize schedule length. By controlling the nature and extent of state-space exploration, *HMDS* can adapt itself to deliver the best possible solution within a given time bound. (iii) A mechanism for co-scheduling multiple independent periodic DAG applications has been devised in the third phase. The objective of the scheduling algorithm is to minimize dissipated energy. (iv) Subsequently, in the fourth phase, a security-aware real-time DAG scheduling strategy has been designed. The scheme maximizes total security utility for a given application having known minimum security strength specifications for its messages. (v) Finally, in the last phase of the dissertation, we have developed a mechanism to construct minimum makespan schedules for precedence-constrained task graph applications with*

known thermal characteristics on a heterogeneous processing platform. The efficacy of the developed scheduling schemes has been extensively evaluated through simulation-based experiments using randomly generated DAGs and/or real-world benchmarks. Prototype real-platform implementations as well as real-world case studies have also been presented to exhibit the practical applicability of the proposed algorithms.

