Intelligent Agents: Are they feasible in Swarm-Array Computing?

BLESSON VARGHESE AND GERARD MCKEE
School of Systems Engineering
University of Reading
Reading, Berkshire
UNITED KINGDOM
{hx019035, g.t.mckee}@reading.ac.uk

Abstract: - The work reported in this paper proposes ‘Intelligent Agents’, a Swarm-Array computing approach focused to apply autonomic computing concepts to parallel computing systems and build reliable systems for space applications. Swarm-array computing is a robotics a swarm robotics inspired novel computing approach considered as a path to achieve autonomy in parallel computing systems. In the intelligent agent approach, a task to be executed on parallel computing cores is considered as a swarm of autonomous agents. A task is carried to a computing core by carrier agents and can be seamlessly transferred between cores in the event of a predicted failure, thereby achieving self-* objectives of autonomic computing. The approach is validated on a multi-agent simulator.

Key-Words: - Autonomic computing, Swarm-array computing, Intelligent agents, Carrier agents.

1 Introduction

Autonomic computing has recently emerged as a domain of interest to computing researchers worldwide. What is autonomic computing, and what are its inspiration and vision? What are its distinct perspectives? What are the autonomic approaches? What needs to be focused ahead? These are the few questions answered in this section, before commencing discussions on Intelligent Agents and their feasibility in Swarm-Array Computing, the primary focus of this paper.

What is autonomic computing, and what are its inspiration and vision? With the advancements of computing techniques, biologically inspired computing has emerged as a major domain in computing. Many computing paradigms, namely amorphous computing, evolutionary computing and organic computing have emerged as a result of being inspired from natural phenomenon. Autonomic computing is one such biologically inspired computing paradigm based on the autonomic human nervous system [1].

Autonomic computing is a visionary paradigm for developing large scale distributed systems, reducing cost of ownership and reallocating management responsibilities from administrators to the computing system itself [2] – [9]. Autonomic computing paves the necessary foundation autonomic computing principles have paved necessary foundations towards self-managing systems.

Self-managing [10] systems are characterized by four objectives and four attributes. The objectives and attributes that contribute to self-management are not independent functions. The objectives considered are [1] [11] [12]: (a) Self-configuration – the capability of a computing system to automatically adapt to changes in the existing physical topology and software environment. The system must be also capable to seamlessly integrate new system components. Self-configuring systems are expected to increase resource availability. (b) Self-healing - the capability of a computing system to recuperate from faults and loss. Constant and consistent monitoring of the computing system is required to detect faults and loss. (c) Self-optimizing – the capability of a computing system to automatically tune resources and balance workloads to improve operational efficiency. (d) Self-protecting – the capability of a computing system to protect itself from malicious attacks originating from within and without the system. Self-protection safeguards the system from damages due to uncorrected cascading failures.

The attributes considered are [1] [11] [12]: (a) Self-awareness – the capability of a computing system to be aware of its internal state and knowledge of the possible states the system can transform to from the current state. (b) Self-situated – the capability of a computing system to be aware of the external operating conditions. (c) Self-monitoring – the capability of a computing system to detect the change of internal and external circumstances consistently. (d) Self-adjusting – the capability of a computing system to adapt to internal and external changes reflexively.

What are the perspectives of autonomic computing? There are mainly two perspectives, namely business and research oriented perspectives that provide a bird’s-eye view of the paradigm. Firstly, from a business oriented perspective, autonomic computing was proposed by IBM for better management of increasingly complex computing systems and reduce the total cost of ownership of systems today [5] [6], hence aiming to reallocate management responsibilities from administrators to the computing systems itself based on high-level policies [7].
[8]. With the aim to implement autonomic principles in personal computing environments, personal autonomic computing a subset of autonomic computing has also emerged [9].

Secondly, the research oriented perspective primarily focuses on the worms-eye view, laying necessary foundations for the newly emerging computing paradigm. There are two categories of ongoing research in the area of autonomic computing. Firstly, research describing approaches and technologies related to autonomic computing [10]. The aim of the approaches is to achieve autonomy without specifying the technology to be implemented [11]. Any existing technology capable of achieving autonomy (in any degree) can be used in the approaches. Secondly, research attempting to develop autonomic computing as a unified project [10]. The research lays emphasis on the means to achieve autonomy and initiatives are taken to define a set of standard practices and methods as the path towards autonomy.

What are the autonomic computing approaches? Autonomic computing researchers have adopted six different approaches, namely emergence-based, component/service-based, control theoretic based, artificial intelligence, swarm intelligence and agent based approaches to achieve self-managing systems.

The emergence based approach for distributed systems considers complex behaviours of simple entities with simple behaviours without global knowledge [12]. Intelligent behaviour is thus repercussions of interactions and coordination between entities. One major challenge in emergence based approaches is on how to achieve global coherent behaviour [13]. Autonomic computing research on emergence based approaches is reported in [12] - [15].

The component/service based approach in distributed systems employ service-oriented architectures. With advancements in software engineering practices, component/service based approaches are also implemented in many web based services. The autonomic element of the autonomic system is a component whose interfaces, behaviours and design patterns aim to achieve self-management. These approaches are being developed for large scale networked systems including grids. Autonomic computing research on component/service based approaches is reported in [16] - [19].

The control theoretic based approach aims to apply control theory for developing autonomic computing systems. The building blocks of control theory such as reference input, control input, control error, controller, disturbance input, measured output, noise input, target system and transducer are used to model computing systems and further used to study properties like stability, short settling times, and accurate regulation. Using a defined set of control theory methodologies, the objectives of a control system namely regulatory control, disturbance rejection and optimization can be achieved. These objectives are closely associated with the objectives of autonomic computing. Research on control theoretic based approaches applied to autonomic computing is reported in [20] - [22].

The artificial intelligence based approaches aim for automated decision making and the design of rational agents. The concept of autonomy is realized by maximizing an agent’s objective based on perception and action in the agent’s environment with the aid of information from sensors and in-built knowledge. Work on artificial intelligence approaches for autonomic computing is reported in [23] [24].

The swarm intelligence based approaches focuses on designing algorithms and distributed problem solving devices inspired by collective behaviour of swarm units that arise from local interactions with their environment [25] [26]. The algorithms considered are population-based stochastic methods executed on distributed processors. Autonomic computing research on swarm intelligence approaches is reported in [27] - [29].

The agent based approaches for distributed systems is a generic technique adopted to implement emergence, component/service, artificial intelligence or swarm intelligence based approaches. The agents act as autonomic elements or entities that perform distributed task. The domain of software engineering considers agents to facilitate autonomy and hence have a profound impact on achieving the objectives of autonomic computing. Research work based on multi-agents supporting autonomic computing are reported in [5] [30] - [35].

What needs to be focused ahead? The focus of researchers in autonomic computing should be towards two directions. Firstly, researchers ought to aim towards applying autonomic computing concepts to parallel computing systems. This focus is essential since most distributed computing systems are closely associated with the parallel computing paradigm. The benefits of autonomy in computing systems, namely reducing cost of ownership and reallocating management responsibilities to the system itself are also relevant to parallel computing systems. It is surprising that only few researchers have applied autonomic computing concepts to parallel computing systems in the approaches above.

Secondly, autonomic computing researchers ought to focus towards implementing the approaches for building reliable systems. One potential area of application that demands reliable systems is space applications. Space crafts employ FPGAs, a special purpose parallel computing system that are subject to malfunctioning or failures of hardware due to ‘Single Event Upsets’ (SEUs), caused by radiation on moving out of the
protection of the atmosphere [43] - [45]. One solution to overcome this problem is to employ reconfigurable FPGAs. However, there are many overheads in using such technology and hardware reconfiguration is challenging in space environments. In other words, replacement or servicing of hardware is an extremely limited option in space environments. On the other hand software changes can be accomplished. In such cases, autonomic computing approaches can come to play.

How can a bridge be built between autonomic computing approaches and parallel computing systems? How can autonomic computing approaches be extended towards building reliable systems for space applications? The work reported in this paper is motivated towards bridging this gap by proposing swarm-array computing, a novel technique to achieve autonomy for distributed parallel computing systems and experimenting the feasibility of a proposed approach on FPGAs that can be useful for space applications.

The remainder of the paper is organized as follows. Section 2 introduces swarm-array computing. Section 3 investigates the feasibility of the proposed approach by simulations. Section 4 concludes the paper and considers future work.

2 Swarm-Array Computing

Swarm-array computing is a swarm robotics inspired approach and is proposed as a path to achieve autonomy. The development of the swarm-array computing approach is from the foundations of parallel and autonomic computing paradigms. The constitution of the swarm-array computing approach can be separated into four constituents. Three approaches are proposed that bind the swarm-array computing constituents together. The four constituents and the three approaches are considered in the following sub sections.

2.1 Constituents

There are four prime constituents that make up the constitution of swarm-array computing. They are the computing system, the problem/task, the swarms and the landscape considered in this section.

Firstly, the computing systems which are available for parallel computing are multi-core processors, clusters, grids, field programmable gate arrays (FPGA), general purpose graphics processing units (GPGPU), application-specific integrated circuit (ASIC) and vector processors. With the objective of exploring swarm-array computing, FPGAs are selected as an experimental platform for the proposed approaches.

FPGAs are a technology under investigation in which the cores of the computing system are not geographically distributed. The cores in close proximity can be configured to achieve a regular grid or a two dimensional lattice structure. Another reason of choice to look into FPGAs is its flexibility for implementing reconfigurable computing.

The cores of the computing system can be considered as a set of autonomous agents, interacting with each other and coordinating the execution of tasks. In this case, a processing core is similar to an organism whose function is to execute a task. The focus towards autonomy is laid on the parallel computing cores abstracted onto intelligent cores. The set of intelligent cores hence transform the parallel computing system into an intelligent swarm. The intelligent cores hence form a swarm-array. A parallel task to be executed resides within a queue and is scheduled onto different cores by the scheduler. The swarm of cores collectively executes the task.

The intelligent cores described above are an abstract view of the hardware cores. But then the question on what intelligence can be achieved on the set of cores needs to be addressed. Intelligence of the cores is achieved in two different ways. Firstly, by monitoring local neighbours. Independent of what the cores are executing, the cores can monitor each other. Each core can ask the question of ‘are you alive’ to its neighbours and gain information. Secondly, by adjusting to core failures. If a core fails, the process which was executed on the core needs to be shifted to another core where resources previously accessed can be utilized. Once a process has been shifted, all data dependencies need to be re-established.

To shift a process from one core to another, there is a requirement of storing data associated and state of the executing process, referred to as checkpointing. This can be achieved by a process monitoring each core or by swarm carrier agents that can store the state of an executing process. The checkpointing method suggested is decentralized and distributed across the computing system. Hence, though a core failure may occur, a process can seamlessly be transferred onto another core. In effect, awareness and optimizing features of the self-ware properties are achieved.

Secondly, the problem/task to be executed on the parallel computing cores that can be considered as a swarm of autonomous agents. To achieve this, a single task needs to be decomposed and the sub tasks need to be mapped onto swarm agents. The agent and the sub-problems are independent of each other; in other words, the swarm agents are only carriers of the sub-tasks or are a wrapper around the sub-tasks.

The swarm displaces itself across the parallel computing cores or the environment. The goal would be to find an area accessible to resources required for executing the sub tasks within the environment. In this case, a swarm agent is similar to an organism whose
function is to execute on a core. The focus towards autonomy is laid on the executing task abstracted onto intelligent agents. The intelligent agents hence form a swarm-array.

The intelligent agents described above are an abstract view of the sub-tasks to be executed on the hardware cores. Intelligence of the carrier agents is demonstrated in two ways. Firstly, the capabilities of the carrier swarm agents to identify and move to the right location to execute a task. In this case, the agents need to be aware of their environments and which cores can execute the task. Secondly, the prediction of some type of core failures can be inferred by consistent monitoring of power consumption and heat dissipation of the cores. If the core on which a sub-task being executed is predicted to fail, then the carrier agents shift from one core to another gracefully without causing an interruption to execution, hence making the system more fault-tolerant and reliable. An agent can shift from one core to another by being aware of which cores in the nearest vicinity of the currently executing core are available.

Thirdly, a combination of the intelligent cores and intelligent swarm agents leads to intelligent swarms. The intelligent cores and intelligent agents form a multi-dimensional swarm-array. The arena in which the swarms interact with each other is termed as landscape.

Fourthly, the landscape that is a representation of the arena of cores and agents that are interacting with each other in the parallel computing system. At any given instance, the landscape can define the current state of the computing system. Computing cores that have failed and are predicted to fail are holes in the environment and obstacles to be avoided by the swarms.

A landscape is modelled from three different perspectives which is the basis for the swarm-array computing approaches discussed in the next section. Firstly, a landscape comprising dynamic cores (are autonomous) and static agents (are not autonomous) can be considered. In this case, the landscape is affected by the intelligent cores. Secondly, a landscape comprising of static cores and dynamic agents can be considered. In this case, the landscape is affected by the mobility of the intelligent agents. Thirdly, a landscape comprising of dynamic cores and dynamic agents can be considered. In this case, the landscape is affected by the intelligent cores and mobility of the carrier agents.

2.2 Approaches
At this point it is appropriate to consider how the constitution of swarm-array computing fits together? To answer this question, three approaches that combine the constituents of swarm-array computing are proposed.

In the first approach, only the intelligent cores are considered to be autonomous swarm agents and form the landscape. A parallel task to be executed resides within a queue and is scheduled onto the cores by a scheduler. The intelligent cores interact with each other as considered in section 2.1 to transfer tasks from one core to another at the event of a hardware failure.

In the second approach, only the intelligent swarm agents are considered to be autonomous and form the landscape. A parallel task to be executed resides in a queue, which is mapped onto carrier swarm agents by the scheduler. The carrier swarm displace through the cores to find an appropriate area to cluster and execute the task. The intelligent agents interact with each other as considered in Section 2.1 to achieve mobility and successful execution of a task. Figure 1 describes the approach diagrammatically.

In the third approach, both the intelligent cores and intelligent agents are considered to form the landscape. Hence, the approach is called a combinative approach. A parallel task to be executed resides in a queue, which is mapped onto swarm agents by a scheduler. The swarm agents can shift through the landscape utilizing their own intelligence, or the swarm of cores could transfer tasks from core to core in the landscape. The landscape is affected by the mobility of intelligent agents on the cores and intelligent cores collectively executing a task by accommodating the intelligent agent.
However, in this paper the major focus is the second approach and is only considered for experimental studies.

3 Simulation Studies

Simulation studies were pursued to validate and visualize the proposed approach in swarm-array Computing. Since FPGA cores are considered in this paper and the approach proposed in this paper considers executing cores as agents; hence a multi-agent simulator is employed. This section is organized into describing the simulation environment, experimental platform and model and simulation results.

3.1 Simulation Environment

The feasibility of the proposed swarm-array computing approach was validated on the SeSAm (Shell for Simulated Agent Systems) simulator. The SeSAm simulator environment supports the modelling of complex agent-based models and their visualization [41] [42].

The environment has provisions for modelling agents, the world and simulation runs. Agents are characterized by a reasoning engine and a set of state variables. The reasoning engine defines the behaviour of the agent, and is implemented in the form of an activity diagram, similar to a UML-based activity diagram. The state variables of the agent specify the state of an agent. Rules that define activities and conditions can be visually modelled without the knowledge of a programming language. The building block of such rules is primitives that are pre-defined. Complex constructs such as functions and data-types can be user-defined.

The world provides knowledge about the surroundings the agent is thriving. A world is also characterized by variables and behaviours. The modelling of the world defines the external influences that can affect the agent. Hence, variables associated with a world class can be used as parameters that define global behaviours. This in turn leads to the control over agent generation, distribution and destruction.

Simulation runs are defined by simulation elements that contribute to the agent-based model being constructed. The simulation elements include situations, analysis lists, simulations and experiments. Situations are configurations of the world with pre-positioned agents to start a simulation run. Analysis lists define means to study agents and their behaviour with respect to time. Simulations are combinations of a situation, a set of analysis items and a simulation run; or in other words a complete definition of a single simulation run. Experiments are used when a combination of single simulation runs are required to be defined.

3.2 Experimental Platform & Model

As considered in section 2.1, the swarm-array computing approach needs to consider the computing platform, the problem/task and the landscapes. The parallel computing platform considered in the studies reported in this paper is FPGAs and is modelled in SeSAm. The hardware cores are arranged in a 5 X 5 regular grid structure. The model assumes serial bus connectivity between individual cores. Hence, a task scheduled on a core can be transferred onto any other core in the regular grid.

The breakdown of any given task to subtasks is not considered within the problem domain of swarm-array computing. The simulation is initialized with sub-tasks scheduled to a few cores in the grid. Each subtask carrying agent consistently monitors the hardware cores. This is possible by sensory information (in our model, temperature is sensed consistently) passed onto the carrier agent. In the event of a predicted failure, the carrier agent displaces itself to another core in the computing system. The behaviour of the individual cores varies randomly in the simulation. For example, the temperature of the FPGA core changes during simulation.

If the temperature of a core exceeds a predefined threshold, the subtask being executed on the core is transferred by the carrier agent to another available core that is not predicted to fail. During the event of a transfer or reassignment, a record of the status of execution of the subtask maintained by the carrier agent also gets transferred to the new core. If more than one sub-task is executed on a core predicted to fail, each sub-task may be transferred to different cores.

3.3 Simulation Results

Figure 2 is a series of screenshots of a random simulation run developed on SeSAm for eight consecutive time steps from initialization. The figure shows the executing cores as rectangular blocks in pale blue colour. When a core is predicted to fail, i.e., temperature increases beyond a threshold, the core is displayed in red. The subtasks wrapped by the carrier agents are shown as blue filled circles that occupy a random position on a core. As discussed above, when a core is predicted to fail, the subtask executing on the core predicted to fail gets seamlessly transferred to a core capable of processing at that instant.

The simulation studies are in accordance with the expectation and hence are a preliminary confirmation of the feasibility of the proposed approach in swarm-array computing. Though some assumptions and minor approximations are made, the approach is an opening for applying autonomic concepts to parallel computing platforms.
Conclusions

In this paper, a swarm-array computing approach based on intelligent agents that act as carriers of tasks has been explored. Foundational concepts that define swarm-array computing are introduced. The feasibility of the proposed approach is validated on a multi-agent simulator. Though only preliminary results are produced in this paper, the approach gives ground for expectation that autonomic computing concepts can be applied to parallel computing systems and build reliable systems for space applications.

Future work will aim to study the third proposed approach or the combinative approach in swarm-array computing. Efforts will be made towards implementing the approaches in real time and exploring in depth the fundamental concepts associated with the constituents of swarm-array computing.

References:


[31] T. De Wolf and T. Holovet, “Towards Autonomic Computing: Agent-Based Modelling, Dynamical Systems Analysis, and Decentralised Control” in the


[42] SeSAm website: http://www.simsesam.de