Towards Cognitive B3G Networks: Autonomic Management of Access Points

K. Demestichas[1], E. Adamopoulou[1], M. Theologou[1], P. Demestichas[2], D. Boscovic[3], D. Bourse[4]

[1]National Technical University of Athens, Athens, Greece, cdemest@cn.ntua.gr
[2]University of Piraeus, Piraeus, Greece
[3]Motorola Inc., Corporate Technology, Schaumburg, USA
[4]Motorola Labs, Paris, France

***Abstract*—The heterogeneity and complexity of B3G wireless infrastructures dictates the need for novel functionality, in order to deliver enhanced services and meet user requirements. Cognitive networks comprising reconfigurable elements are an effective response towards this direction. The application of autonomic computing principles has the potential of tackling the complexity of managing heterogeneous environments. In this light, this paper introduces innovative functionality targeting at the autonomic management of access points. Four functional components are identified and detailed. A preliminary approach for the deployment of the platform on network infrastructures is also outlined.**

***Index Terms*— Cognitive networks, network management, autonomics, context awareness, profiles, policies**

I. Introduction

Over the last centuries, the economy has migrated from agriculture to industry, and from there, to the era of information and services. Technological advances occupy the centre-stage in our modern, digital life, affecting our everyday experience through a multitude of services and applications, varying from healthcare and education to finance, banking and public administration. Communication networks, and wireless networks in particular, have established their omnipresent nature in all residential, public and business environments. The number of end-users who benefit by the provision of the novel, feature-rich services is constantly increasing. However, this penetration of new technologies is accompanied by the position of more demanding requirements originating from both the end-users and the business community. While not extensive the following list summarizes the essence of these requirements: (*i) personalization; (ii) support of pervasive computing; (iii) context awareness; (iv) always-best connectivity; (v) ubiquitous application provision; (vi) seamless mobility.*

The B3G era is expected to provide the means for meeting the abovementioned requirements, by encompassing innovative functionality that lies not only in network infrastructure but in management principles as well. Currently the wireless landscape comprises several Radio Access Technologies (RATs), which can be roughly classified along the two major lines:

- Wireless wide area networking (WWAN) technologies, which include, among others, 2G/2.5G/3G mobile communications, the IEEE 802.16 suite, WiMAX and broadcasting technologies;
- Wireless short range networks (WShRNs), which include wireless local and personal area networks (WLANs/WPANs), as well as wireless sensor networks (WSNs).

This plethora, however, of access technologies combined with the constantly improving characteristics of mobile terminals, introduces inevitable heterogeneity and, consequently, increased complexity of the wireless networks. Each RAT has different capabilities, in terms of capacity, coverage, mobility support, cost, etc. Therefore, each RAT is best suited for handling certain, but not all, conditions. In this respect, a network operator will have to rely on different RATs for raising customer satisfaction, and achieving the required Quality of Service (QoS) levels, in a cost-efficient manner.

A direction for addressing complexity, without raising significantly the capital expenditure, is to attribute the wireless B3G infrastructures of network operators with “cognitive network” capabilities [1]-[2]. In general, cognitive systems determine their behaviour, in a reactive or proactive manner, based on the external stimuli (environment aspects), as well as their goals, principles, capabilities, experience and knowledge. In the case of cognitive networks, this definition can be translated as the ability to dynamically select the network’s configuration, through management functionality that takes into account the context of operation (environment requirements and characteristics), goals and policies [3] (corresponding to principles), profiles (capabilities), and machine learning [4]-[5].

As can be deduced from these definitions, cognitive wireless networks will consist of reconfigurable elements [6]-[7] and management functionality [1]. Regarding the management part, a modern research direction, in order to increase scalability and decrease complexity, is to comply with self-management paradigms, or in other words, to develop the management functionality in accordance with the autonomic computing principles [8]-[10].

It can be argued that autonomic management is most important at the access points of the infrastructure. These elements will be facing more frequently, and more drastically, changing situations. In this respect, they constitute the primary field, in which autonomic management can be applied. The autonomic manager will produce as output the best access point configurations for handling a given or anticipated situation. Such an autonomic management entity is specified in the next section.

This paper introduces innovative functionality, targeting at the autonomic management of access points (AMAP). It proposes the establishment of autonomic management entities (AMAP entities) for the optimal, efficient and autonomic management of the reconfigurable elements of the infrastructure.

The remainder of this paper is structured as follows: Section II presents the functional architecture of an AMAP entity, and specifies the role, data and knowledge layer of each component. In Section III, a preliminary approach for the deployment of AMAP entities on network infrastructures is outlined. Finally, concluding remarks are drawn and future work directions are provided in Section IV.

II. Overall Architecture and High-Level Operation

The functional architecture of an AMAP entity is depicted in Figure 1. Each entity is divided into components, with each component comprising two functional layers: the *data layer* and the *knowledge layer*. The first one is concerned with the retrieval of contextual data, while the second one refers to the knowledge and experience developed on the basis of the extracted data. The exact role of each component, as well as the interactions among them, is described in detail in what follows.

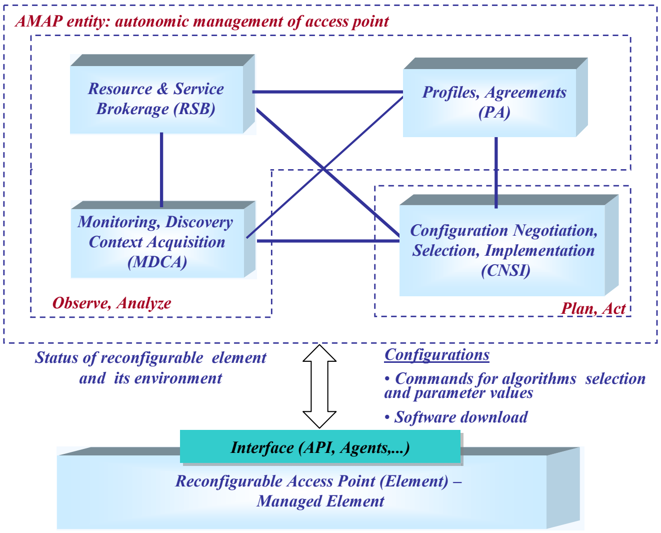
AMAP entity: autonomic management of access point
Resource & Service Brokerage (RSB)
Profiles, Agreements (PA)
Monitoring, Discovery Context Acquisition (MDCA)
Configuration Negotiation, Selection, Implementation (CNSI)
Observe, Analyze
Plan, Act
Status of reconfigurable element and its environment
Configurations
• Commands for algorithms selection and parameter values
• Software download
Interface (API, Agents,...)
Reconfigurable Access Point (Element) – Managed Element

Figure 1. Functional analysis (components) of an AMAP entity

A. Profiles and Agreements (PA)

1) Data Layer

This component is mainly motivated by the need of supporting personalization and pervasive computing. Therefore, the corresponding data layer should provide information on the following: *(a)* preferences and agreements of the end-users, and *(b)* equipment capabilities of both the managed element (reconfigurable element) and other equipment elements and devices within a particular service area. Pertaining to the first set, the information should specify: *(i)* the applications that can be used; *(ii)* for each application, the permissible QoS levels and the respective typical traffic and mobility behaviour; *(iii)* for each application and QoS level, the importance (utility), and cost information (e.g., maximum allowable or indicative tolerable). Regarding the equipment capabilities, what should be specified is: *(i)* the set of wireless transceivers used for serving the equipment/users in the service area; *(ii)* for each transceiver, the set of RATs that can be supported; *(iii)* for each transceiver and RAT, the sets of frequency carriers (spectrum) that can be used. Therefore, this part describes all the potential configurations of each transceiver (focusing on the primary parts of each configuration, which are the RAT and frequency). The configuration of all the transceivers yields a (main part of the) configuration of the access point. Consequently, this part of the component provides information on the access point capabilities. Information concerning other equipment in the service area is analogous to that of the managed element.

2) Knowledge Layer

The functionality of this layer is focused on the development of knowledge on the end-user behaviour and preferences. The following issues should be addressed:

- What is the behaviour of a certain user, regarding applications, traffic and mobility, in a given time-epoch/user-role, location, etc.?
- What are the user preferences for QoS levels in a given time-epoch/user-role, location, etc.?

The user preference for an application and QoS level is expressed through a utility value. The traffic behaviour per application covers the frequency of application usage, the duration of the usage, and the information volume generated in each usage. The mobility behaviour covers aspects like speed, direction, and location.

In more detail, the probabilistic relations that should be managed can be described as:

- Probability that a specific, traffic or mobility, parameter will take a certain value, in a given context, which is associated with the user identity, location, time/role, and application.
- Probability that the utility will take a certain reference value, in a given context, which is associated with the user identity, location, time/role, application used and the QoS level of the application provision.

B. Monitoring, Discovery and Context Acquisition(MDCA)

1) Data Layer

The component should provide the means for: *(a) acquiring the context* in which the managed access point operates; *(b) monitoring* the efficiency with which each contextual situation is handled; *(c) discovering* alternate capabilities that exist in the environment and can be used for context handling.

Context acquisition. The context in which the managed element ("who") operates, consists of the: *(i)* location ("where"); *(ii)* time epoch ("when"); *(iii)* traffic, mobility and interference conditions encountered ("what"), caused by the equipment/users in the service area. Therefore, there should be information on: *(i)* the set of equipment and users in the service area; *(ii)* their locations and the time epoch; *(iii)* the configuration of devices and the applications requested; *(iv)* the traffic and mobility behaviour.

The overall traffic, mobility and interference conditions faced by the managed element, are random, but typically, remain stationary for a certain time length. This essentially defines the duration of a contextual situation.

Monitoring. This procedure collects basic data from the transceivers of the managed element, specifically: *(i)* the selected configuration; *(ii)* the QoS levels offered, in response to a certain contextual situation (traffic, mobility, and interference conditions), encountered in a certain time. The basic data gathered from transceivers can be processed to produce aggregate data at various levels of abstractions.

Discovery. The functionality of this part lies in the estimation of the performance of alternate configurations in a certain context, e.g., time. Therefore, this part of the component should have information on the achievable performance, e.g., interference levels, bit-rate, coverage, associated with each alternate transceiver configuration. Discovery may require the cooperation between the managed access point and the devices in the service area.

2) Knowledge Layer

This part of the component can add robustness (reliability, stability) to the monitoring, discovery and context acquisition procedures. Moreover, it can provide the means for reasoning, not only on what currently goes on, but also on what is likely to happen in the future, in the managed element and in its environment. The following essential questions can be addressed:

- What is the set of equipment and users that should be anticipated in a given time epoch?
- What are the traffic and mobility conditions that should be anticipated in a given time epoch?
- What are the QoS levels that can be anticipated to be achieved by a given configuration in a given context, i.e., time epoch, location, and respective traffic, mobility and interference conditions?
- What are the anticipated capabilities, e.g., in terms of bit-rate, of a given alternate configuration, in a given time and location?
- Given a certain situation in the current time epoch (e.g., traffic, interference, mobility conditions), what are the situations that should be anticipated in a subsequent time epoch?

The first two bullets fall in the realm of context acquisition. The third bullet is related to monitoring, while the fourth is related to discovery. The fifth is targeted to context predictions.

Traffic requirements, mobility conditions, and configurations' capabilities, in certain time epochs, have a stochastic nature. Regarding discovery, the configuration capabilities will not be constant, especially, in an environment that is not centrally controlled, as the one considered. Capabilities depend on the configurations used in the environment, by other network elements, and the resulting interference conditions. Therefore, learning mechanisms can yield the typical capabilities of alternate, candidate configurations, in certain contexts (e.g., time-epochs). Finally, context predictions are necessary for the proactive handling of situations. This is necessary for meeting the seamless mobility and ubiquitous provision requirements.

This part can be based on monitoring, discovery and context acquisition measurements. Based on these measurements, probabilistic relations can be managed. Context predictions can be based on the monitoring of situations, in each time epoch. In this respect, the transitions that occur, in successive time epochs, can also be recorded. The frequency of the transitions is a basic way for computing probabilities, and predicting future situations.

C. Resource and Service Brokerage (RSB)

1) Data Layer

This component enables an AMAP entity to interact with other entities in its environment. This can assist in the satisfaction of ubiquitous provision, in conjunction with seamless mobility. Moreover, it can be used for enhancing always-best connectivity. The goal of the interactions is to acquire data on the status of the "neighbouring" elements. The status of a neighbouring element may provide some information on QoS levels that it can offer. Then, the behaviour of the managed access point can be influenced by this information, in order to improve QoS (therefore, enhance always best connectivity) and avoid QoS fluctuations (therefore, assist in seamless mobility and ubiquitous provision). Eventually (as addressed in the corresponding knowledge layer paragraph), the need for these interactions can be reduced, and each access point can develop knowledge on the capabilities of the near-by elements. However, the existence of such an entity can ensure smooth migration from current management models to the autonomics.

2) Knowledge Layer

The interest in this component is to develop knowledge regarding the capabilities of neighbouring elements. The key question is: "what are the most likely capabilities of neighbouring elements in each time epoch". The motivation is that the capabilities of neighbouring elements are uncertain. Knowledge should be developed in this respect, for facilitating seamless mobility and ubiquitous provision.

D. Configuration Negotiation, Selection and Implementation (CNSI)

1) Data Layer

This component has the following role: (i) To consider as input the *context* that has to be faced, the *profiles* of users and equipment, the capabilities of other elements, and the *goals* and *policies*, of the NO (in general) and managed access point. (ii) To produce as output the *optimal (cross-layer) configurations*, for the managed access point and the equipment in its service area. The configurations should lead to optimal (best-possible) context handling.

Taking into account the input considered, it can be argued that the component is context-aware, policy-based and goal-driven. In general, changes in context can signify changes in the goals and policies that should be used. As indicated in the MDCA description, contextual situations of the managed element can be delineated by changes in: *(i)* time (e.g., migration from one time epoch to another); *(ii)* the traffic requirements and mobility levels in the service area; *(iii)* the potential behaviour of the managed access point, due to alterations in the interference conditions.

Goals. The goals are expressed through an objective function that has to be optimized. Different objective functions can be used depending on the context. In general, the objective function is associated with the offered QoS levels (which have to maximised), and the cost (which has to be minimised) of providing these QoS levels.

Policies. In a sense, they specify rules (constraints) that should be respected, and functionality that should be followed for reaching the goals. These rules and functionality may differ depending on the context. Rules can refine the information specified in the managed element and equipment/user profiles, or introduce additional constraints. In this respect, each policy can have rules on the following: *(i)* regarding the managed element profile, the set of configurations allowed per transceiver of the managed access point (these are a subset of the overall access point capabilities, and implicitly affect also the equipment in the service area); *(ii)* regarding the user profile, the set of applications/services allowed, and the respective permissible QoS levels per application (this may also have a positive flavour, e.g., by allowing only the highest QoS levels, among those permissible ones, for certain classes); *(iii)* the set of applications that are allowed to be offered through each configuration. The functionality designated by policies can be optimisation algorithms, parameters values, and, potentially, a set of suggested configurations.

Selection. In the light of the above, the CNSI component can handle the different contexts according to the following general strategy. Firstly, there should be retrieval of the appropriate goals and policies, which correspond to the context encountered. Next, there should be investigation of whether there are suggested configurations (solutions for context handling). In case there are, the best (according to the goals), feasible (according to the rules), suggested configuration is selected. Otherwise, if no feasible, suggested solutions exist, the optimisation algorithms are applied. The configuration selection will be done from this set. The output may also involve agreements with other elements, which can be backbone elements or near-by access points.

Negotiation. As indicated above, there is selection from a set of "best" configurations per contextual situation. This increases the reliability regarding the achievement of the appropriate QoS levels, by offering several alternate solutions. Moreover, it enables the resolution of conflicts, i.e., the selection of a new configuration, in case the performance of the one in use is compromised, for example, due to contemporary use in the vicinity of the managed element. Finally, it opens several negotiation possibilities. It enables the element to potentially trade (change) its configuration, in order to allow other elements to achieve adequate QoS levels. Third, it enables the element to potentially trade (change) its configuration, in order to allow other elements to achieve adequate QoS levels.

Implementation. This last part involves interfacing with the managed element. The implementation of the reconfigurations results to an information flow, towards the element, that specifies parameter values, new algorithms, or new executable code that should be activated and used.

2) *Knowledge Layer*

The following question should be addressed: "what are the configurations (solutions) typically used, given a specific context of the element operation (time, location, traffic, mobility, interference conditions) and the respective policies used". In other words, the interest here is on knowing what are the typical solutions used for handling certain contexts. The corresponding probabilistic relation can be described as: "probability that a configuration will be used given the policy and contextual condition (traffic, mobility, interference conditions)". Probabilistic models can increase the confidence on the most likely behaviour of policies, in handing certain contexts. This knowledge will enable faster and more reliable selection of configurations. This part can rely on the mere recording of the outcome of the CNSI invocations. The probabilistic associations are between the context addressed, the policy applied and the configuration proposed.

E. *AMAP Entity Operation*

The high-level operation of an AMAP entity is summarized in Figure 2.

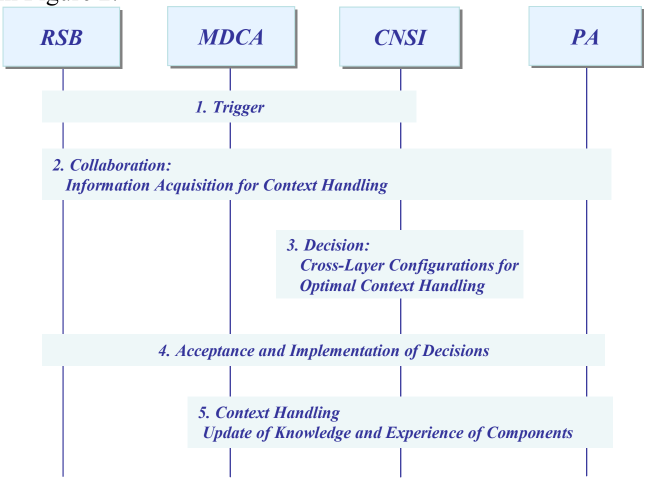
RSB
MDCA
CNSI
PA
1. Trigger
2. Collaboration:
Information Acquisition for Context Handling
3. Decision:
Cross-Layer Configurations for
Optimal Context Handling
4. Acceptance and Implementation of Decisions
5. Context Handling
Update of Knowledge and Experience of Components

Figure 2. High-level operation of an AMAP entity

In *phase 1*, a trigger is received. Triggers may originate from the MDCA (new context), the CNSI directly (new policies and goals activated by the NO), or the RSB (preparation for handling of future contexts). In *phase 2*, the trigger is forwarded to the CNSI. The CNSI collaborates with the PA, MDCA and RSB for acquiring information on the profiles (of managed element, equipment/users), the situation encountered, as well as capabilities of other elements (AMAP entities). In *phase 3*, a decision is derived by the CNSI, on the optimal (cross-layer) configurations, taking also into account the policies and goals, apart from the profiles, context, and capabilities of collaborating entities. These configurations should lead to the best possible handling of the overall resulting situation (new context). In *phase 4*, it is assumed that the CNSI decisions become known, accepted and applied. *Phase 5* is targeted to new context handling (handling of new situation that caused the trigger). During this phase, all the components collaborate for enhancing their knowledge and experience.

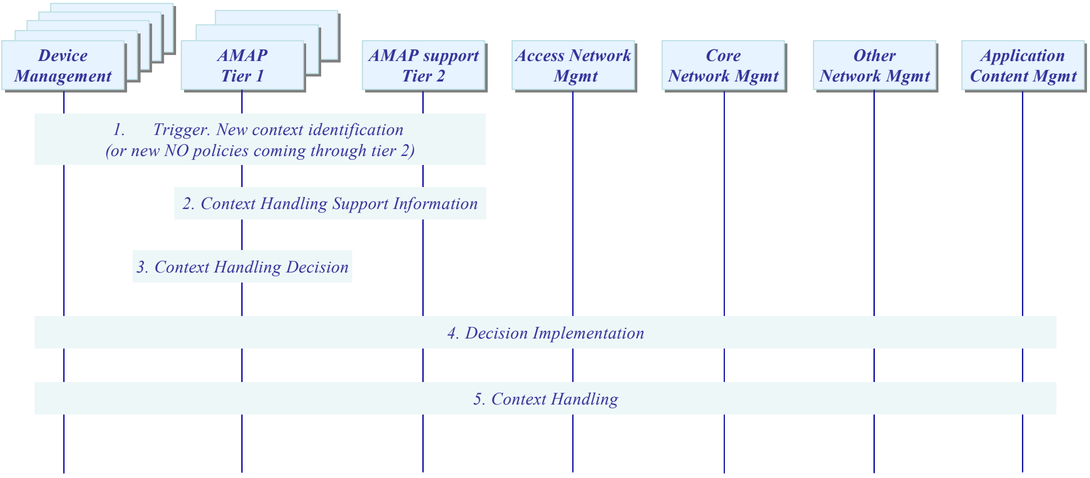
Device Management
AMAP Tier 1
AMAP support Tier 2
Access Network Mgmt
Core Network Mgmt
Other Network Mgmt
Application Content Mgmt
1. Trigger. New context identification (or new NO policies coming through tier 2)
2. Context Handling Support Information
3. Context Handling Decision
4. Decision Implementation
5. Context Handling

Figure 3. Design Analysis – Deployment of AMAP entities: operation scneario

III. Design Analysis

This section briefly addresses the deployment of the AMAP management entities in the B3G world. It can be envisaged that the deployment will require two new types of entities, which are categorised in two tiers. The first tier has the AMAP entities of the previous section. The second tier has AMAP-support entities.

Each AMAP entity of the first tier has to exist in an environment that has: *(i)* equipment management entities; *(ii)* AMAP entities of the two tiers; *(iii)* legacy management platforms that are targeted to the access and core network, other networks (owned by other NOs), and the application and content segment.

The second tier can have a less decentralized role: *(i)* to complement the first tier, in case of missing knowledge, e.g., on profiles, policies on how to handle certain contexts; *(ii)* to act as an intermediary between the first tier, and the NO (business level, in general). New policies and goals introduced by the NO can be first fed to the entities of the second tier, prior to being distributed to the first tier.

Figure 3 shows a use case that shows the collaboration between the different entities of the physical architecture. In phase one, the scenario is triggered, specifically, from the identification of a new contextual situation, by a tier-1 AMAP entity, potentially through the collaboration with the devices in its service area. In the second phase, the AMAP entities collaborate and exchange supporting information and knowledge that will lead to optimal context handling. In the third phase, there is the decision on context handling. The fourth phase is targeted to the implementation of the decision. In the fifth phase there is the context handling.

IV. Conclusions

This paper specified an autonomic management platform targeted to the access points of a wireless B3G infrastructure. Our work showed how typical requirements that should be served by wireless B3G infrastructures justify the need for cognitive network technologies, in general. Next, our focus was on the functional analysis (components) of each autonomic entity, which provides the means for monitoring, discovery, context acquisition, profile and agreements management, resource and service brokerage, and configuration negotiation, selection and implementation. For each component, its functionality, data and knowledge were presented. Finally, the paper briefly addressed issues related to the deployment of the platform on network infrastructures.

Our future work involves two challenging phases: First, the completion of the development of the platform, the realization of the necessary integration in B3G environments, and the pursuit of standardization activities. Second, the introduction of further cognitive technologies, especially at the fixed network, and the realization of integrated experiments.

Acknowledgment

The work presented is partially funded by the European Commission, under the Sixth Framework Program, in the context of the IST Integrated Project E2R2.

References

[1] P.Demestichas, D.Boscovic, V.Stavroulaki, A.Lee, J.Strassner, "m@ANGEL: autonomic management platform for seamless wireless cognitive connectivity to the mobile Internet", *IEEE Commun. Mag.*, Vol. 44, No.6, June 2006

[2] R. Thomas, L. DaSilva, A. MacKenzie, "Cognitive networks", *In Proc. 1st IEEE Symposium on Dynamic Spectrum Access Networks 2005 (DySPAN 2005)*, Baltimore, USA, pp. 352-360, Nov. 2005

[3] J.Strassner, "Policy-based network management: solutions for the next generation", Morgan Kaufmann (series in networking), 2005

[4] T. Mitchel, "Machine learning", McGraw-Hill, 1997

[5] R. E. Neapolitan, "Learning Bayesian Networks", Prentice Hall (series in artificial intelligence), 2002

[6] P.Demestichas, G.Vivier, K.El-Khazen, M.Theologou, "Evolution in wireless systems management concepts: from composite radio to reconfigurability", *IEEE Commun. Mag.*, Vol. 42, No. 5, pp. 90-98, May 2004

[7] Software Defined Radio Forum Web site, www.sdrforum.org, 2007

[8] J. Strassner, "Autonomic networking – theory and practice", *In Proc. 9th IFIP/IEEE International Symposium on Network Management (IM'2005)*, Nice, France, May 2005

[9] J. Kephart, D. Chess, "The vision of autonomic computing", *IEEE Computer*, Vol. 36, No.1, pp. 41-50, January 2003

[10] M. Hinchey, R. Sterritt, "Self-managing software", *IEEE Computer*, Vol. 39, No. 2, pp. 107-109, February 2006