Towards Autonomic Management of Communications Networks

Brendan Jennings, Sven van der Meer, Sasitharan Balasubramaniam, Dmitri Botvich, Mícheál Ó Foghlú, and William Donnelly, Waterford Institute of Technology, Ireland
John Strassner, Motorola Labs

ABSTRACT

As communications networks become increasingly dynamic, heterogeneous, less reliable, and larger in scale, it becomes difficult, if not impossible, to effectively manage these networks using traditional approaches that rely on human monitoring and intervention to ensure they operate within desired bounds. Researchers and practitioners are pursuing the vision of *autonomic network management*, which we view as the capability of network entities to self-govern their behavior within the constraints of business goals that the network as a whole seeks to achieve. However, applying autonomic principles to network management is challenging for a number of reasons, including:

- A means is required to enable business rules to determine the set of resources and/or services to be provided.
- Contextual changes in the network must be sensed and interpreted, because new management policies may be required when context changes.
- As context changes, it may be necessary to adapt the management control loops that are used to ensure that system functionality adapts to meet changing user requirements, business goals, and environmental conditions.
- A means is required to verify modeled data and to add new data *dynamically* so that the system can *learn* and *reason* about itself and its environment.

This article provides an introduction to the FOCALE autonomic network management architecture, which is designed to address these challenges.

INTRODUCTION

The telecommunications industry has changed dramatically in recent years. Explosive growth of the Internet, the proliferation of mobile technologies, and fixed-mobile convergence has led to a progressively more complex, interconnected networking infrastructure. The ever increasing difficulty in managing multi-vendor environments and the services they provide has altered forever the dynamics of the industry, the expectations of its customers, and the business models with which it operates. In hardware, the impact of Moore's Law has had a profound effect across all sectors of the industry, encouraging equipment manufacturers, network operators, and service providers to continually strive to rapidly deploy the latest technology in order to gain competitive advantage. We believe that a downside of this rapid technology deployment is that current communications service offerings are inflexible in nature: they are rigidly defined; closely coupled to specific network technology; exhibit static functionality; and are often prone to security breaches. Critically, they are for the most part manually deployed and managed, requiring highly labor-intensive support structures, with consequent inflexibility and significant time to market constraints.

To address this problem, academic researchers and industrial implementers are moving away from traditional network management approaches based on centralized control of a (relatively) small number of managed entities. Observing that networks are becoming more dynamic, more heterogeneous, less reliable, and larger in scale; they instead are actively investigating the application of autonomic principles. Their aim is to simplify network management processes by automating and distributing the decision making processes involved in optimizing network operation. The goal is to enable expensive human attention to focus more on business logic and less on low-level device configuration processes.

In this article we introduce our approach to autonomic network management. We contend that the essence of autonomic management is the capability of a system to self-govern its behavior, but only within the constraints of the (human-specified) goals that the system as a whole seeks to achieve. We propose the use of information and ontological modeling to capture *knowledge* relating to network capabilities,

0163-6804/07/$20.00 © 2007 IEEE

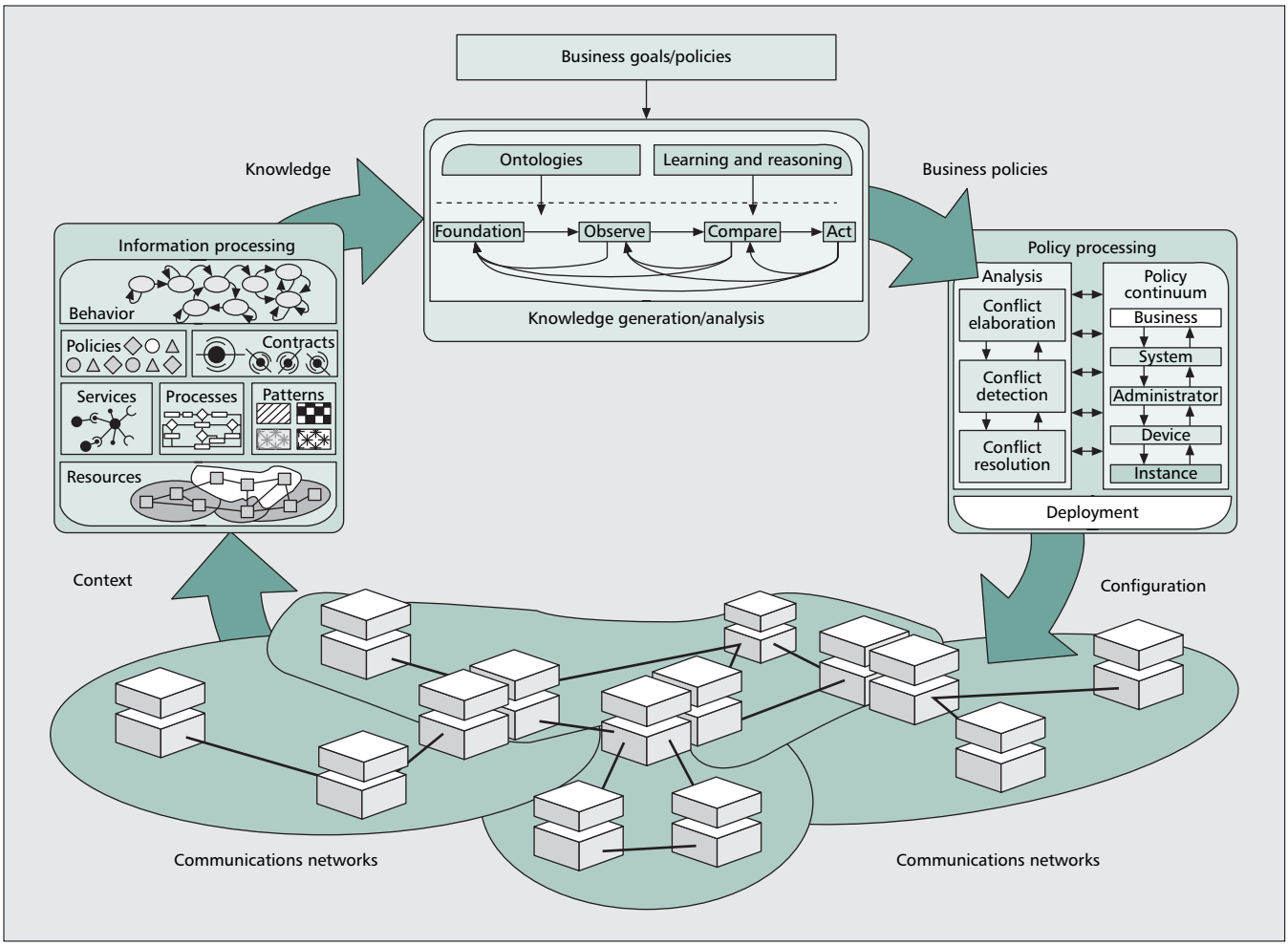

Figure 1. *Conceptual autonomic control loop for network management.*

environmental constraints, and business goals and policies, together with reasoning and learning techniques, to enhance and evolve this knowledge. Knowledge embedded within system models is used by policy-based network management systems [1], incorporating translation/code generation and policy enforcement processes that automatically configure network elements in response to changing business goals and/or the environmental context. This realizes an autonomic control loop in which, as depicted in Fig. 1, the system senses changes in itself and its environment. It analyzes this information to ensure that business goals and objectives are met, expedites changes should these goals and objectives be threatened, and closing the loop, observes the result.

We believe an information and ontological modeling-based approach delivers considerable improvements over existing manually-configured network management systems, because it supports context-driven reconfiguration of networks with minimal human intervention at all but the high-level business view. Nevertheless, to deliver full autonomic network management capabilities, we believe it also is necessary to introduce decentralized processes and algorithms into the network infrastructure to maintain optimal or near-optimal behavior in terms of global stability, improved performance and adaptability, robustness, and security. As described by Babaoglu et al. [2], many of these processes and algorithms can be profitably modeled on various biological processes found in the natural world. However, to ensure that they act in accordance with business goals, we argue that such processes and algorithms should themselves be modeled, so that their operation can be configured automatically by appropriate management policies.

This article is structured as follows: we briefly summarize current research in the areas of autonomic computing and autonomic networking, contrasting our approach with other ongoing efforts. We introduce our conceptual model of an autonomic network management system, highlighting the main components that are required to deliver effective management. We describe our work on embodying these concepts in Foundation Observation Comparison Action Learn rEason (FOCALE), our architecture for autonomic network management, and we discuss an initial prototypical realization of FOCALE. Finally, we summarize the article and discusses future work and open issues relating to the development and standardization of our work.

Put simply, the autonomic paradigm seeks to reduce the requirement for human intervention in the management process through the use of one or more control loops that continuously re-configure the system to keep its behavior within desired bounds.

AUTONOMIC COMPUTING AND AUTONOMIC NETWORKING

For many years, researchers have been aware that the structural and operational complexity of communications networks — indeed, distributed computing systems in general — has been increasing to the point where it is having a negative impact on the economic viability of introducing new products and services to the market. In recent years the paradigms that created the most interest as methods of addressing this problem are those of *autonomic computing* [3] and later, *autonomic networking* [4]. Put simply, the autonomic paradigm seeks to reduce the requirement for human intervention in the management process through the use of one or more control loops that continuously re-configure the system to keep its behavior within desired bounds.

The term autonomic computing was coined by IBM as an analogy to the autonomic nervous system, which maintains homeostasis (essentially maintaining equilibrium of various biological processes) in our bodies without the need for conscious direction. Autonomic computing attempts to manage the operation of individual pieces of IT infrastructure (such as servers in a data center) through the introduction of an *autonomic manager* that implements an autonomic control loop in which the managed element and the environment in which it operates is monitored. Collected data is analyzed, and actions are taken if the managed element is deemed to be in an undesired (sub-optimal) state [3]. The autonomic control loop is made up of monitor, analyze, plan, and execute components, all of which rely on a common knowledge repository. The monitor component gathers data, filters and collates it as required, and then presents it to the analyze component, which seeks to understand the data and determine if the managed element is acting as desired. The plan component takes the data and determines if action should be taken to reconfigure the managed element. The execute component translates the planned actions into a set of configuration commands that can be applied to the managed element.

The vision of autonomic computing can be summarized as that of a *self-managing* IT infrastructure in which equipment has software deployed on it that enables it to self-configure, self-optimize, self-heal, and self-protect. That is, it will exhibit what has come to be known as *self-** behavior. Clearly this is a powerful vision, albeit one that is acknowledged by Kephart [5] as requiring significant advances in the state of the art over a number of years. Therefore, it was natural that the networking community would extend this vision from autonomic management of individual elements to autonomic networking — the collective (self-) management of networks of communicating computing elements. Of course, some of the network management work of the 1980s and 1990s could be retrospectively termed *autonomic networking*, as some of the self-* issues were addressed; however, in practice, the term is a twenty-first-century one. Autonomic networking is currently a burgeoning research area that seeks to integrate results from disciplines ranging from telecommunications network management to artificial intelligence and from biology to sociology. For a summary of the state of the art in this fast-moving area, see Dobson et al. [6].

Much of the focus of research in autonomic network management is on the development of highly distributed algorithms that seek to optimize one or more aspects of network operation and/or performance, in essence aiming to provide various self-management capabilities. In this context, many researchers are investigating the potential use of biologically-inspired algorithms and processes. As noted in [6], complex biological systems "*tend to exploit decentralized and uncoupled coordination models, relying primarily on environment-mediated local information transfer*." Examples include homeostasis (the tendency towards stable equilibrium between interdependent elements), chemotaxis (the movement of a cell in a particular direction corresponding to a chemical gradient), and stigmergy (indirect communications between organisms through modification of their local environment), all of which have been used as inspiration for algorithm development. For descriptions of specific examples, see [2].

Although work on the development of decentralized, self-management algorithms is crucial — and noting that significant advances have been made — we believe that deployment of these algorithms, although necessary, will not be sufficient. Equally important will be the flexible specification and enforcement of the goals these algorithms collectively seek to achieve — goals designed to ensure that the network successfully delivers services to users. Policy-based network management [1], is widely seen as an appropriate management paradigm to facilitate higher-level, human-specified cognitive decision making; therefore, many researchers are examining how policy-based management can be leveraged to help realize the autonomic vision (a good example is the work of Agrawal et al. [7]). However, to our knowledge, little work has been done to date on integrating distributed self-management algorithms with policy-based management — for example, by enabling policies to re-parameterize such algorithms to change their behavior to adapt to changing business goals. We believe this step is essential to provide a solution that balances the need for explicit control over network operation with the benefits of highly efficient and robust self-management algorithms and processes. Given this, a key goal of our work is to develop an architecture for autonomic network management that can profitably integrate the *top-down* explicit control model of traditional network management with the *bottom-up* emergent behavior associated with biologically-inspired, distributed algorithms. This requires numerous advances in the state of the art, chiefly with regard to:

Management of heterogeneous functionality: One of the problems in applying autonomic principles to networks is that networks are

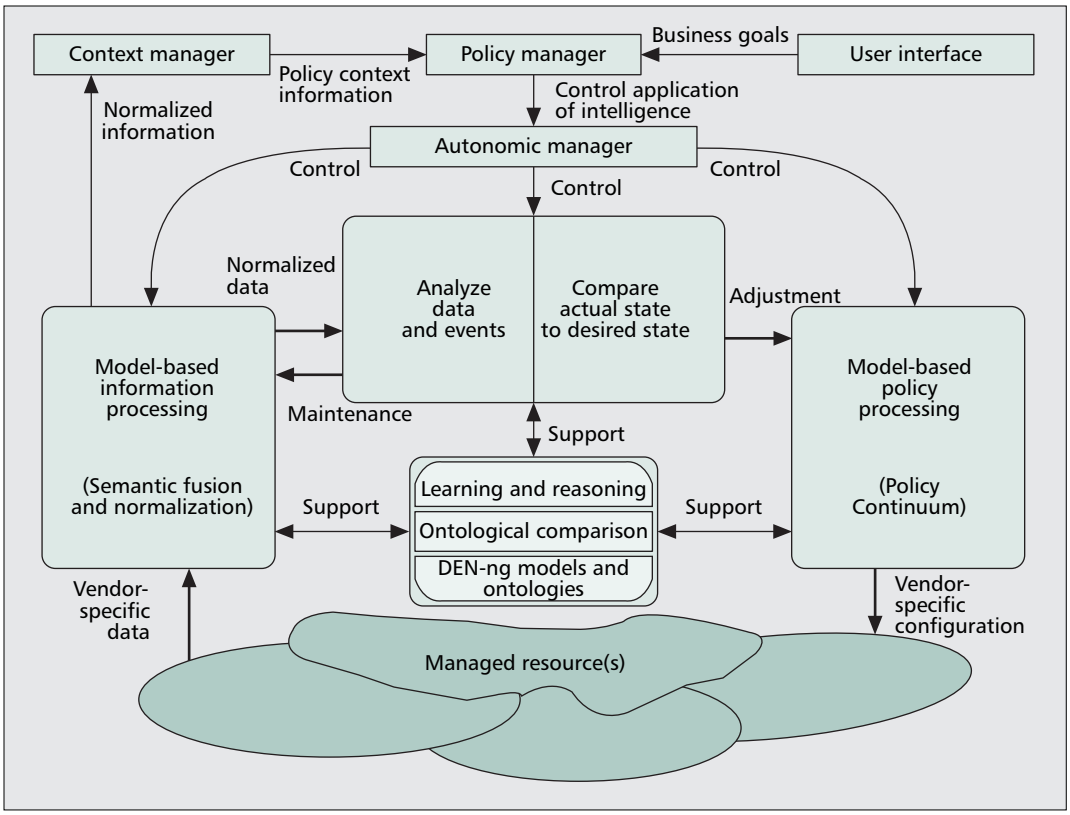

Figure 2. *Conceptual representation of an autonomic network management system.*

One of the promises of autonomic operation is the capability to adapt the functionality of the system in response to changes in user requirements, business rules, and/or environmental conditions. This requires a more flexible governance approach than is provided to date.

made up of many different devices. These devices have different programming models and provide different management data, describing the same or similar concepts. This makes it imperative to harness information models and ontologies to abstract away vendor-specific functionality to facilitate a standard way of reconfiguring that functionality. Achieving this will enable legacy resources with no inherent autonomic capabilities to be managed by autonomic systems.

Adaptability: One of the promises of autonomic operation is the capability to *adapt* the functionality of the system in response to changes in user requirements, business rules, and/or environmental conditions. This requires a more flexible governance approach than is provided to date. In particular, the system must sense context changes and use policies specific to the new context to effect the required re-configuration of network devices.

Application of learning and reasoning techniques to support intelligent interaction: Current examples of network device configuration and management rely on vendor-specific snapshots of static data. For example, statistics can be gathered and analyzed to determine if a given device interface is experiencing congestion. However, existing management data does not tell the user why congestion is occurring. This information must be inferred using these and other data and retained for future reference. Hence, there is a need to incorporate sophisticated, state-of-the-art learning and reasoning algorithms into autonomic network management systems.

Conceptual Representation of an Autonomic Network Management System

Figure 2 presents a conceptual representation of an autonomic network management system that incorporates an autonomic control loop enabled by the presence of one or more system models and ontologies. These abstract the static structure, functionality, and dynamic behavior of the underlying network infrastructure, management functionality, and offered services. These models and ontologies are continuously updated by the model-based information processing component in response to the changing operational context of the network. This component gathers raw context information from managed resources (e.g., Simple Network Management Protocol (SNMP) alarms). Using various analysis techniques, the component infers the impact, or potential impact, of this information (e.g., a network failure means that customer X is not receiving the quality of service (QoS) level indicated in their service level agreement (SLA) for service Y). It then passes normalized data relating to current operational context to the event analysis component, which employs ontological engineering, in conjunction with learning and reasoning techniques, to analyze whether the actual state of the system corresponds to the desired state (as indicated by a currently deployed set of policies).

If there is a mismatch between the detected system state and the desired state, two courses of action are possible. First, if a deployed policy

The combination of a set of enhanced DEN-ng information and data models, combined with domain-specific ontologies for augmenting those models with required semantics, enables information gathered from the network to be analyzed and used to ensure the models accurately reflect the current operational status.

exists that specifies what should be done in this particular scenario, that policy is triggered by the policy processing component. The component utilizes knowledge embodied within system models to automatically generate and apply updated network device configurations that should return the network to the desired state. Alternatively, if no such policy exists (as happens occasionally, as it is never possible to model all possible operational scenarios for a complex network), information models and ontologies are analyzed to determine what actions are required to return the network to the desired state. This will be codified in a new policy that will be passed to the policy processing component, where as described previously, it is triggered to appropriately reconfigure the network.

The control loop described previously is controlled by an autonomic manager that influences the deployment of the policies that effect decision making within the loop. The autonomic manager receives up-to-date business-level policies from the policy manager, which in turn manages policies that are created or modified by humans via a user interface (e.g., business analysts may create policies indicating the services a new customer may access), or which are modified by the system itself, based on information supplied by the context manager (e.g., in cases of network failures, policies relating to certain customers could be modified to deny them access to services in order to give preferential access to other, more important, customers).

The entire system, but particularly the model-based information processing component, relies on the presence of information and data models that embody the knowledge required to represent managed resources (routers and other network devices) and their control, using autonomic principles. In our work, we use the DEN-ng information model (outlined in [8]), which we are currently enhancing with finite state machines to model behavior and augmenting with ontological models that embody semantic information that cannot be represented in the Unified Modeling Language (UML). DEN-ng is a comprehensive information model for telecommunications, capturing everything from business concepts (e.g., products, service level agreements, and customers) to low-level device functionality (e.g., packet marking, forwarding, and queuing). It is designed to be readily augmented with vendor specific information and data models; it thereby provides a highly flexible and extensible modeling solution. Probably the best known application of DEN-ng is in the TM Forum standards, where parts of the model are used as the basis of the shared information and data (SID) modeling effort.

The combination of a set of enhanced DEN-ng information and data models, combined with domain-specific ontologies for augmenting those models with required semantics, enables information gathered from the network to be analyzed and used to ensure the models accurately reflect the current operational status. Considering the previous examples: when devices in the network fail, resulting in localized lack of connectivity or decreases in available bandwidth, the myriad alarms that are raised can be collectively analyzed to ascertain which services and therefore, which customers, are affected. Forwarding knowledge expressed in these terms facilitates decisions regarding which customers should be given preferential access to the network during the period in which the network is congested and therefore incapable of supporting all of its customers. Moreover, the DEN-ng models and associated ontologies provide a knowledge base that can be used by machine learning and reasoning algorithms to both analyze collected data and automatically generate new knowledge that then can be leveraged to autonomically manage the underlying network infrastructure. In particular, this allows data gathered from the network to be analyzed and used to ensure the models accurately reflect its current operational status.

Another significant benefit of DEN-ng is its comprehensive policy model, which facilitates the specification of policy representation languages that are tightly coupled to the information model of the network, which policies authored in that language will govern (as discussed later, this characteristic is particularly useful for policy analysis). Furthermore, we can apply the DEN-ng Policy Continuum [8], in which policies at different levels of abstraction are organized in stratified business, system, network, device, and device instance views — mirroring the different constituencies of people who will work together — to define and deploy the policies that provide a product or service. Implementation of the Policy Continuum enables these constituencies, who understand different concepts and use different terminologies, to manipulate sets of policy representations at a view appropriate to them, and to have those view-specific representations mapped to equivalent representations at views appropriate for manipulation by other constituencies.

The task of automating the refinement of business-level policies (specified in terms of entities, such as products, services, and customers) through the continuum into corresponding device instance policies (specified in terms of entities such as packet-marking rules or firewall configuration rules) is very challenging, as information must be added (and removed) as the policies become more specific in nature. Our approach is to harness the expressive power of ontologies to detail the nature of the relationships between concepts at different continuum levels. Policy refinement processes can access this knowledge from the ontologies and applying ontological engineering techniques, propose candidate refinements that in certain cases must be ratified by humans; thus, we envisage policy refinement as a (semi-)automated process, where the level of human intervention decreases over time as the system learns from the outcome of previous refinements.

The model-based policy processing component also must incorporate policy conflict analysis algorithms that:

- Elaborate newly defined/modified policies (e.g., by adding conditions relating to system constraints not evident to the policy author) so that conflicts are easier to detect.

- Detect sets of policies that will or potentially could conflict, given certain network context.
- Resolve conflicts by modifying or removing polices based on separate resolution policies or by referring back to the appropriate policy author for a decision.

Policy conflict analysis must be done at each level of the continuum, with high level policies only "deployed" if they, and all the policies associated with them at lower levels of the continuum, are detected as being conflict-free. Policy conflict analysis is widely researched and is acknowledged as an extremely difficult challenge; however, we believe significant advances can be made by harnessing the semantic information available in DEN-ng and associated ontologies to facilitate more powerful conflict analysis algorithms than those currently available. Our initial work on this approach is described later and in more detail in [9].

Finally, we note that the model-centered approach (previously outlined) primarily provides for explicit control of network behavior and as such, can be viewed as an evolution of traditional network management approaches. However, this approach is limited by the capabilities (and configurability) of the network devices and the capability of maintaining up-to-date information and ontological models of very complex and highly dynamic network topologies. Additionally, we believe that true autonomic network management will require the deployment of processes and algorithms within network devices themselves. These would act in a highly distributed manner, serving to optimize network behavior with respect to stability, performance, robustness, and security — in effect, providing the kind of self-management functionality discussed earlier. We argue that these processes and algorithms must be incorporated into the overall model-centered management process so that their operation can be re-parameterized to modify their behavior to satisfy high-level policies. This provides the required integration point between the (top-down) model-centered management approach and the (bottom-up) self-management approach in which highly distributed algorithms applying local rules give rise to the desired emergent global behavior.

The FOCALE Autonomic Network Management Architecture

In this section, we describe how the concepts for autonomic management described earlier are realized in the context of a concrete architectural model for distributed autonomic network management systems. The FOCALE architecture is based on the observation that business objectives, user requirements, and environmental context all change dynamically. Therefore, a single, statically defined, management control loop is insufficient — we need the ability to adapt the behavior of the control loop so that it can effectively manage the network to react appropriately to observed or hypothesized changes. FOCALE implements two control loops: *a maintenance control loop* is used when no anomalies are found (i.e., when either the current state is equal to the actual state, or when the state of the managed element is moving towards its intended goal); an *adjustment control loop* is used when one or more policy reconfiguration actions must be performed, and/or new policies must be codified and deployed.

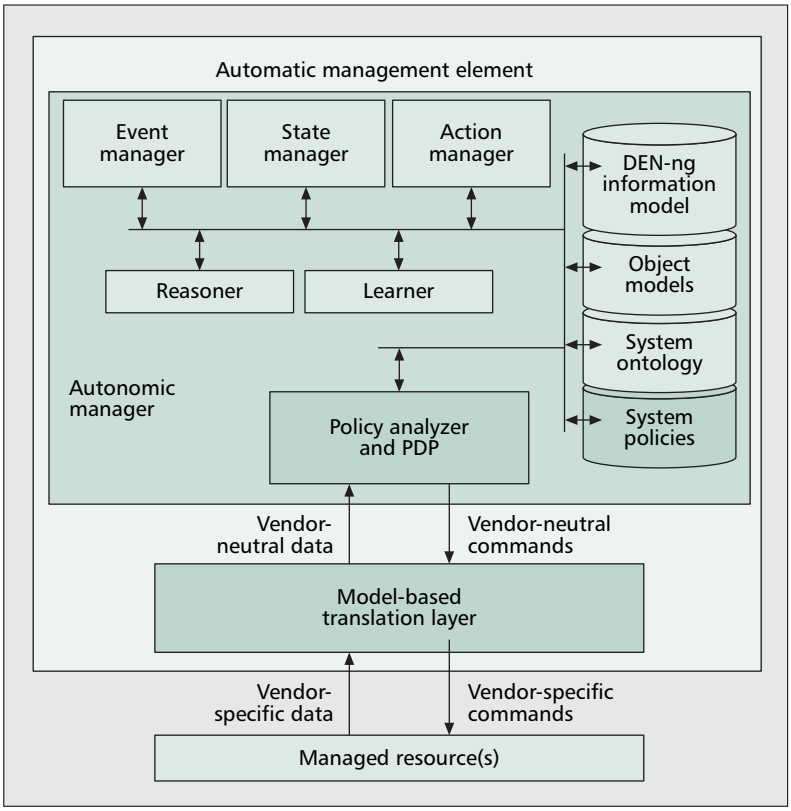

Figure 3. *FOCALE autonomic management element functional architecture.*

Of course, it is unreasonable to assume that a single entity can maintain all the information required to realize the FOCALE control loops for large scale networks containing large numbers of heterogeneous (in terms of available functionality, vendor-specific programming model, and specific configuration) devices. Therefore, FOCALE must be a distributed architecture, to the degree that even individual network devices may incorporate autonomic management software, implementing the maintenance and adjustment control loops. To this end, FOCALE assumes that any managed resource (which can be as simple as a device interface or as complex as an entire system or network) can be associated with an autonomic management element (AME), by interfacing the functionality of the managed resource to the functionality of an autonomic manager (AM) using a model-based translation layer (MBTL), as shown in Fig. 3. As Fig. 4 shows, AMEs can be modularized to first form a uniform autonomic management domain (AMD) and then to an autonomic management environment; with each level containing policy, security, discovery, context, and analysis services that serve to harmonize the operation of the AMEs/AMDs.

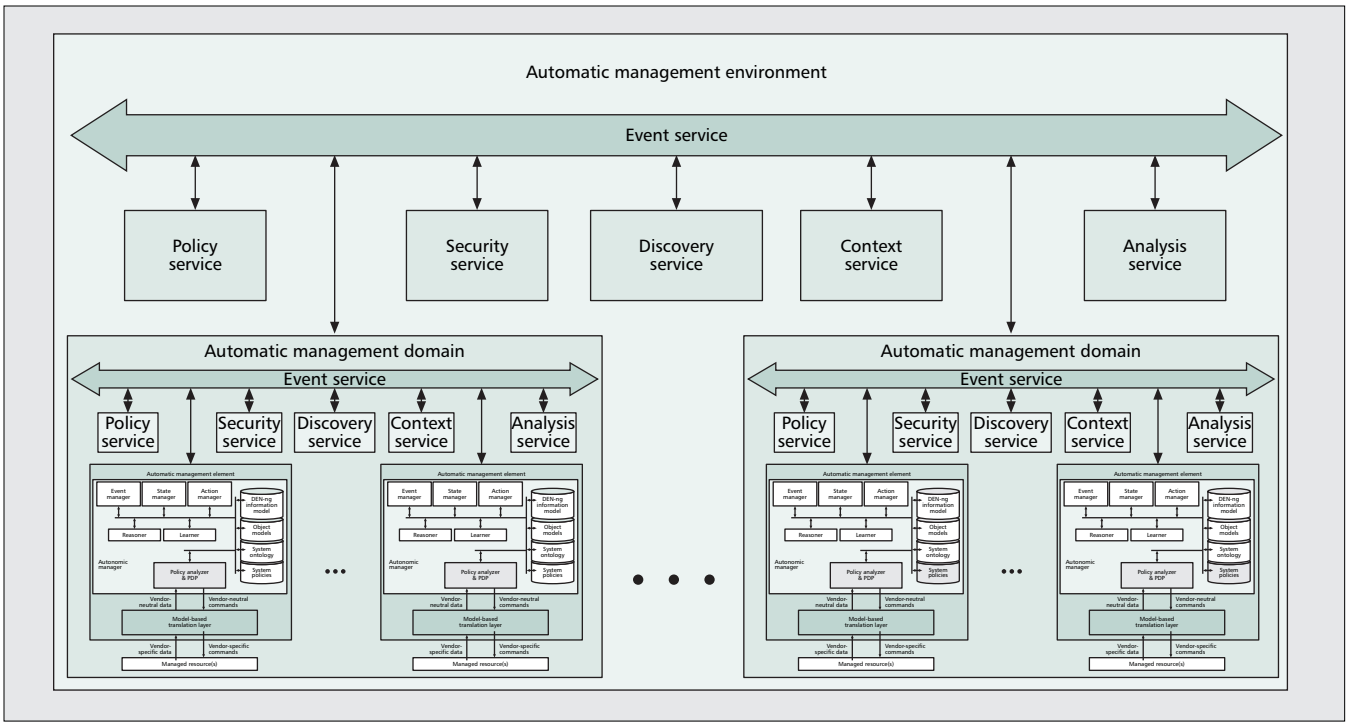

Figure 4. *FOCALE autonomic management environment functional architecture.*

The autonomic management architecture contains two main functional components: the AM and the MBTL. The AM is independent of the vendor-specific functionality/data of the underlying managed resource(s), which facilitates easier communication between AMEs for coordination of management decision making. Each AM realizes the autonomic management functionality described in the previous section via an event manager, a state manager, an action manager, a reasoner, a learner, and a policy analyzer/policy decision point (PDP). All these subcomponents can communicate with each other and have access to the DEN-ng information model, an object model reflecting the current state of the AME's managed resource(s), the system ontology, and the set of deployed policies governing the AME's managed resource(s). When the AM receives context information via the MBTL, the policy analyzer/PDP ascertains if the conditions of any deployed policies are satisfied; if they are, then the corresponding actions are applied via the MBTL. If the policy analyzer/PDP does not recognize the context, it contacts the event and state managers, which use the models/ontologies to ascertain if the system is in a desired state. If it is not, the state manager employs the reasoner to identify actions that will lead the system back toward its desired state. Once identified, the action manager coordinates the enforcement of these actions by the policy analyzer/PDP. Subsequently, the learner monitors the effectiveness of actions identified in this manner; if successful, these actions are codified as one or more policies that then are added to the set of system policies. Of course AMs also have the ability to communicate with other AMs to coordinate activities such as analysis of global network state or introduction of new policies.

Unlike the AM, the MBTL must have in-depth knowledge of the managed resource(s) to enable it to translate normalized vendor-specific data gathered from the managed resource(s) into DEN-ng compliant vendor-neutral data (context information) to pass to the policy analyzer/PDP and vice versa for configuration commands. As alluded to in the previous section, DEN-ng can be readily extended with vendor-specific information and data models (e.g., relating to new releases of command line interface (CLI) command sets for a family of network devices). Assuming that the DEN-ng information model is extended in this manner for all the managed resource(s), and furthermore, that the system ontology is extended to incorporate semantic information detailing the meaning of various vendor specific data/commands, the MBTL can employ ontological engineering techniques, including semantic similarity matching (for a description, see [10]), to map between DEN-ng vendor neutral representations and vendor specific representations.

The basis of the MBTL approach is depicted in Fig. 5, which shows a typical network scenario in which different devices having different data models are managed using different tools. This creates cognitive dissonance between data in the two data models. Because there is no common vocabulary with established meanings defining the data and relationships, it is impossible to directly compare data from different sources, which in turn means that it is impossible to see if those data are related to each other. By using an ontology to augment the facts represented in these data models, each fact can be mapped into a common vocabulary, which enables each fact to be augmented with appropriate semantics — enabling cognitive similarity between these different facts to be established. Conceptually, the

association between a set of nodes in a model and the set of nodes in an ontology creates a new set of associations, bridging the gap between how knowledge is represented between these different approaches.

Prototype Implementation

We now provide a brief overview of our ongoing work in building a prototype realization of the FOCALE architecture; for a fuller description of this prototype, see [9]. Figure 6 depicts our current implementation of a single FOCALE AME that targets aspects of traffic conditioning in a simulated IP-based network of an ISP, over which customers are offered a small number of communications services. Note that the simulated network is very loosely coupled to the AME implementation — in the next phase of development, we plan to replace the simulation with real routers that will be configured by CLI commands generated by the AME and that will provide context information to the AME via SNMP. As it stands, our OPNET™ based simulation is configured to emit information relating to network events and read and apply new router configurations generated by the AME.

Essential to the prototype are the object model and system ontology, which provide synchronized models representing the current state of the simulated network, the customers and services it supports, and the policies deployed in it. To initially set up these models, we created a configuration domain specific language (DSL) and editor that enables the creation of information model instances (i.e., object models) to represent the structure of the managed system. We use the textual DSL framework called xtext, created by open Architecture Ware (available through Eclipse's generative modeling technologies (GMT) initiative). The configuration DSL is generated from the model elements in the DEN-ng model that are marked as relevant for describing the structure of our particular managed network. This offline process is depicted at the top of Fig. 6. The same technique used to create the configuration DSL and its editor were used to generate an event-condition-action (ECA) policy DSL that is based on the policy representation entities in the DEN-ng model. The ECA policy DSL has the following semantics: on the occurrence of a set of events, if the condition clause evaluates to true, then execute

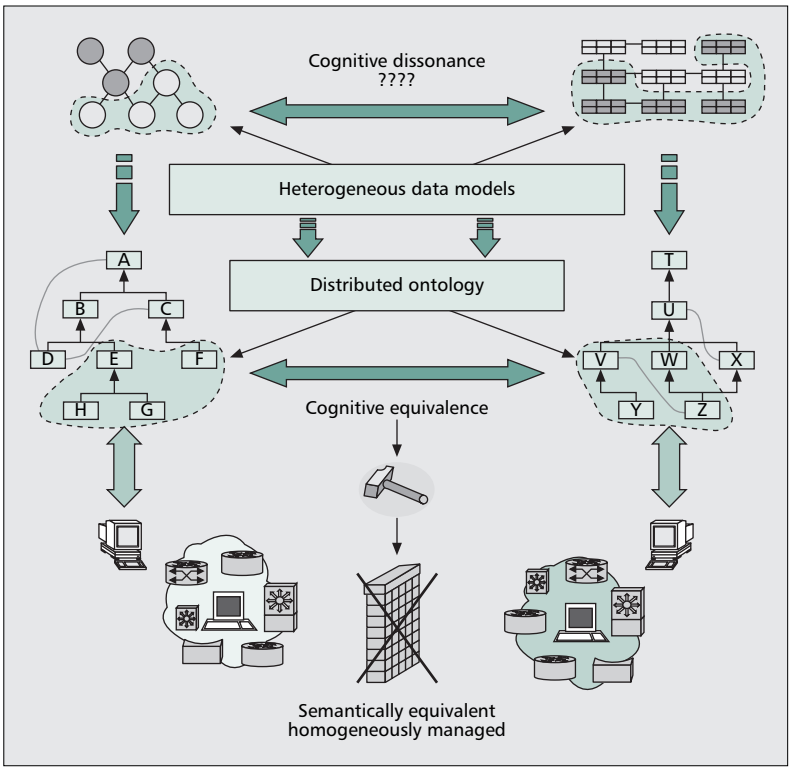

Figure 5. *Use of ontologies to identify cognitive equivalence across heterogeneous data models.*

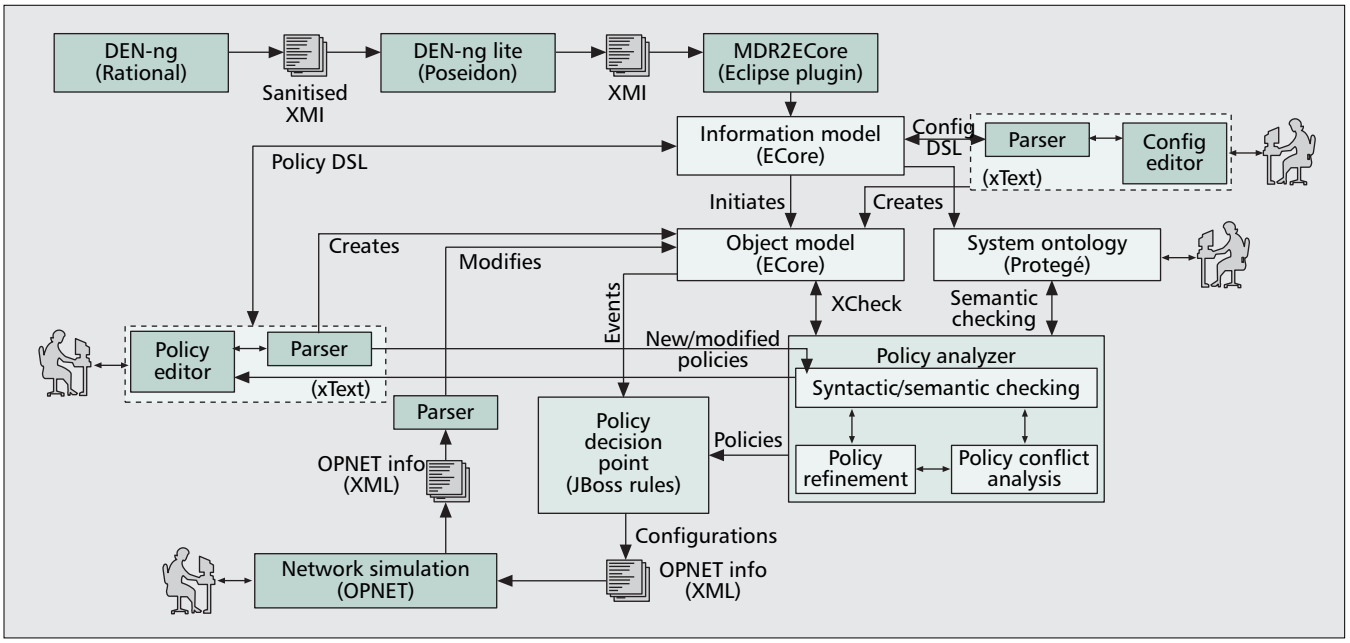

Figure 6. *FOCALE autonomic management element prototype.*

The Rete algorithm efficiently stores rules in memory in the form of a network so that it can take advantage of rule patterns to reduce the number of conditions that must be evaluated. Policies are particularly amenable to Rete-based rule engines.

the action clause. Separation of the configuration and policy DSLs and their editors allows the policy editor to be used during system operation to create, modify, or withdraw policies.

A formal representation of the information model subsets used for generating the DSLs (configuration and policy) is required for automated reasoning in policy analysis. The ontology Web language (OWL)-based system ontology provides this ability. We use IBM's Eclipse modeling framework (EMF) ontology definition metamodel (EODM) to produce an OWL representation of the identified subset of the DEN-ng information model. The resulting baseline ontology can be viewed within Eclipse with the integrated ontology development toolkit (IODT) plug-in or saved and opened with the Protégé OWL editing tool. Our baseline system ontology has been edited to embody semantic information useful for analysis processes such as policy conflict detection.

The policy analyzer uses the object model and system ontology models to build a more complete understanding of the characteristics and behavior of policies and how they affect managed resources. To help provide this understanding, we translate policies into a format suitable for deployment on a rule inference engine-based policy decision point. We use JBoss Rules (based on the Drools rule engine), which uses a tailored object-oriented form of the Rete algorithm called Rete-OO for evaluating the rules. The Rete algorithm efficiently stores rules in memory in the form of a network so that it can take advantage of rule patterns to reduce the number of conditions that must be evaluated. Policies are particularly amenable to Rete-based rule engines, as sets of deployed policies typically share event, condition, and even action parts.

As depicted in Fig. 6, these components cooperate to realize FOCALE maintenance and adjustment control loops. For the maintenance loop, a parser detects changes in the operational context of the network simulation (e.g., a bandwidth utilization threshold on a link being exceeded) and updates the object model accordingly. This object model change triggers evaluation of policies deployed in the JBoss rules engine; if the conditions of one or more policies are satisfied, the appropriate actions are invoked on the simulated network via reconfiguration of simulation parameters (e.g., a low priority customer is denied access to a service, thereby reducing link bandwidth utilization to below the threshold value). For the adjustment loop, human users can use the policy DSL editor to create/modify/withdraw policies via an iterative process in which policies proposed by the user are analyzed by the policy analyzer for syntactic and semantic correctness and potential for conflict with other deployed policies. This occurs at each step of their refinement in the policy continuum. If at any stage, the problem cannot be resolved by the policy analyzer, the user is informed of the problem and prompted to re-edit the policy. After the policy analysis is complete, the created/modified policy is converted to JBoss format and deployed on the JBoss PDP. In this manner, the operation of the maintenance control loop is adjusted.

SUMMARY AND OUTLOOK

This article advocates the principle of *self-governance* as the basis for realizing communications networks that operate and are managed autonomically. We introduced the FOCALE autonomic network management architecture, which has the following distinctive characteristics:

- Emphasis on the use of business goals (codified as policy rules) to determine how resources in the network should be collectively utilized to best deliver services to users.
- Context-aware policy management processes to adapt the management control loops used to ensure that system functionality adapts to meet changing user requirements, business goals, and environmental conditions.
- A novel combination of information and data modeling, augmented by ontological data, to enable the system to *learn* and *reason* about itself and its environment.

We emphasized the difficulties imposed by the inherently heterogeneous and interconnected nature of networked communications systems and described how FOCALE architectural components are designed to overcome these difficulties.

The grand vision of autonomic computing and autonomic networking is that of a completely self-managing infrastructure that itself can access, or generate, the knowledge it requires to enable it to optimally react to changing operational context. Although this vision is very attractive, it is unlikely ever to be fully realized, especially in the context of the extremely complex and dynamic global communications infrastructure. In developing FOCALE, we take a more pragmatic approach: rather then seeking to entirely eliminate human intervention from the management process, we seek to minimize it and to focus it more on business concerns than on low-level device configuration. In particular, we acknowledge that human intervention sometimes will be required for the refinement of policies in the policy continuum and also to resolve policy conflicts never before encountered by the system. Of course, the degree to which this happens is highly dependent on the completeness, correctness, and timeliness of the knowledge embodied within the system information, data and object models, and associated ontologies. We believe that DEN-ng is the most exhaustive and well-structured information model currently available and as such, provides a solid basis for realization of FOCALE.

In the article, we also briefly described our prototype realization of FOCALE. Currently, we are using this prototype to examine the tradeoffs between management sophistication and derived benefits as a function of network size, number of users, mix of traffic, and other factors. In parallel, we are working on replacing the simulated target environment with a small number of real routers. This process should provide valuable insights into the practical implications of using information models and ontologies to automate the generation of CLI commands. Subsequently, we plan to explore integration of

bio-inspired algorithms into FOCALE and to demonstrate their added value via implementation in our prototype. After sufficient experimental results are obtained, we plan to submit our results to the Autonomic Communications Forum standards body.

ACKNOWLEDGEMENTS

We wish to acknowledge the valuable insights provided by Nazim Agoulmine in the development of FOCALE and the work on prototype design and implementation performed by Keara Barrett, Alan Davy, Steven Davy, and Elyes Lehtihet. This work received support from Science Foundation Ireland under the Autonomic Management of Communications Networks and Services award (grant no. 04/IN3/I404C).

REFERENCES

[1] M. Sloman, "Policy Driven Management for Distributed Systems," *J. Net. Sys. Mgmt.*, vol. 2, no. 4, Dec. 1994, pp. 333–60.

[2] O. Babaoglu *et al.*, "Design Patterns from Biology for Distributed Computing," *ACM Trans. Autonomous Adapt. Sys.*, vol. 1, no. 1, Sept. 2006, pp. 26–66.

[3] J. O. Kephart and D. M. Chess, "The Vision of Autonomic Computing," *Computer*, vol. 36, no. 1, Jan. 2003, pp. 41–50.

[4] J. Strassner, "Autonomic Networking — Theory and Practice," Tutorial, *Proc. 2004 IEEE/IFIP NOMS '04*, Apr. 2004, p. 927.

[5] J. O. Kephart, "Research Challenges of Autonomic Computing," *Proc. 27th Int'l. Conf. Software Eng.*, ACM Press, 2005, pp. 15-22.

[6] S. Dobson *et al.*, "A Survey of Autonomic Communications," *ACM Trans. Autonomous Adapt. Sys.*, vol. 1, no. 2, Dec. 2006, pp. 223–59.

[7] D. Agrawal, K. W. Lee, and J. Lobo, "Policy-Based Management of Networked Computing Systems," *IEEE Commun. Mag.*, Oct. 2005, vol. 43, no. 10, pp. 69–75.

[8] J. Strassner, "DEN-ng: Achieving Business Driven Network Management," *Proc. 8th IEEE/IFIP NOMS '02*, Apr. 2002, pp. 753–66.

[9] K. Barrett *et al.*, "A Model Based Approach for Policy Tool Generation and Policy Analysis," *Proc. 1st IEEE Int'l. Global Info. Infrastructure Symp.*, 2007, pp. 99–106.

[10] Y. Kalfoglou and M. Schorlemmer, "Ontology Mapping: the State of the Art," *Knowl. Eng. Rev.*, vol. 18, no. 1, 2003, pp. 1–31.

BIOGRAPHIES

BRENDAN JENNINGS [M] (bjennings@tssg.org) received B.Eng. in electronic engineering and Ph.D. degrees from Dublin City University, Ireland, in 1993 and 2001, respectively. He is a senior investigator at Waterford Institute of Technology, Ireland, working in the areas of policy-based network management, management of composed communications services, network monitoring and planning, and accountability processes in digital ecosystems. He has published 45 papers in international journals and conference proceedings and has participated in the work of standards bodies ETSI, FIPA, TM Forum, and ACF. For 2007 he is co-chair of the 2nd IEEE International Workshop on Modeling Autonomic Communications Environments (MACE 2007) and finance co-chair of the 3rd International Week on Management of Networks and Services (Manweek 2007).

SVEN VAN DER MEER (vdmeer@ieee.org) [M] received his Diploma and Ph.D. degrees in 1996 and 2002, respectively, from Technical University Berlin, Germany. He joined Waterford Institute of Technology in 2002, where he is currently a senior investigator for network and service management. He works mainly in the field of autonomic management as a technical leader of Irish and European research programs, developing strong links with industrial partners (HP, Motorola Labs, and IBM). He is a standing member of the organization/steering committees of MANWEEK, MACE, and MUCS, and is active in standardization (TM Forum, ACF).

SASITHARAN BALASUBRAMANIAM [M] received his B.Eng. (electrical and electronic) degree in 1998, his M.E.Sc. (computer and communication engineering) degree in 1999, and his Ph.D. (computer science) in 2005, all from the University of Queensland, Australia. He is currently a senior investigator at the Telecommunication Software and Systems Group, Waterford Institute of Technology. His research interests include bio-inspired autonomic network management, sensor and ad hoc networking, and pervasive computing.

DMITRI BOTVICH received his Bachelor's (mathematics) degree and Ph.D. (mathematics) from Moscow State University, Faculty of Mechanics and Mathematics, Russia, in 1980 and 1984, respectively. He is currently a principal investigator at the Telecommunication Software and Systems Group, Waterford Institute of Technology. His research interests include bio-inspired autonomic network management, security, trust management, sensor and ad hoc networking, queuing theory, and mathematical physics.

MÍCHEÁL Ó FOGHLÚ [M] is co-founder and research director of the Telecommunications Software and Systems Group at Waterford Institute of Technology, Ireland. His background is interdisciplinary: computer science, linguistics, English, education, and artificial intelligence (as a student at the U.K. universities of Keele, Cambridge, and Central Lancashire). For 10 years he has worked on converged IP-based communications management and services in over 40 Irish and EU funded projects. He leads the Irish IPv6 Task Force and the Irish National IPv6 Centre.

WILLIAM DONNELLY is director of the Telecommunications Software and Systems Group at Waterford Institute of Technology, Ireland. He has over 15 years of experience in research and development of telecommunications network management systems, in both industry and academia. He is a Science Foundation Ireland principal investigator, researching autonomic network management. His research interests are in the areas of bio-inspired network management solutions and management solutions for next-generation Internet-based electronic media.

JOHN STRASSNER [M] is a fellow of the technical staff and directs Motorola's autonomic networking research. He is also an adjunct professor at Waterford Institute of Technology. He was a past fellow of Cisco Systems and the chief strategy officer of Intelliden. His research interests include policy management and knowledge engineering, including ontologies, machine learning, and reasoning. He has been awarded the Daniel Stokesbury memorial award for excellence in network management. He is a Distinguished Fellow of the TeleManagement Forum. He is the chairman of the Autonomic Communications Forum and also vice chairman of WG6 (Reconfigurability and 233 Autonomics) in the WWRF. He has published two books and over 140 papers.

we plan to explore integration of bio-inspired algorithms into FOCALE and to demonstrate their added value via implementation in our prototype. After sufficient experimental results are obtained, we plan to submit our results to the Autonomic Communications Forum standards body.