

A Survey of Autonomic Computing

Michelle Zhou

11/18/2005

Outline

- Definition of Autonomic Computing (AC)
- Four Fundamentals
- AC Architecture
- Industry Research Projects
- University Research Projects
- Major Research Challenges
- Appendix - Industry Standard Support
- References

Definition of Autonomic Computing

- “A type of computing model in which the system is self-healing, self-configured, self-protected and self-managed.” - *From webopedia.com based on IBM's definition*

Fundamentals of Autonomic computing

- **Self-configuring**
 - Systems adapt automatically to dynamically changing environments.
- **Self-healing**
 - Systems discover, diagnose, and react to disruptions.
- **Self-optimizing**
 - Systems monitor and tune resources automatically.
- **Self-protecting**
 - Systems anticipate, detect, identify, and protect themselves from attacks from anywhere.

Needs for Autonomic Computing

- Complexity
 - As the complexity of the computer system increases, the system becomes more prone to error and unmanageable. Finding a methodology to manage the ever-increasing complexity of computing environments has become a big challenge in the IT industry.

Autonomic Architecture

- The Managed Element
 - Any component in the autonomous system, such as a server, a database, or a file.
- The Autonomic Manager
 - Monitor
 - Analyze
 - Plan
 - Execute

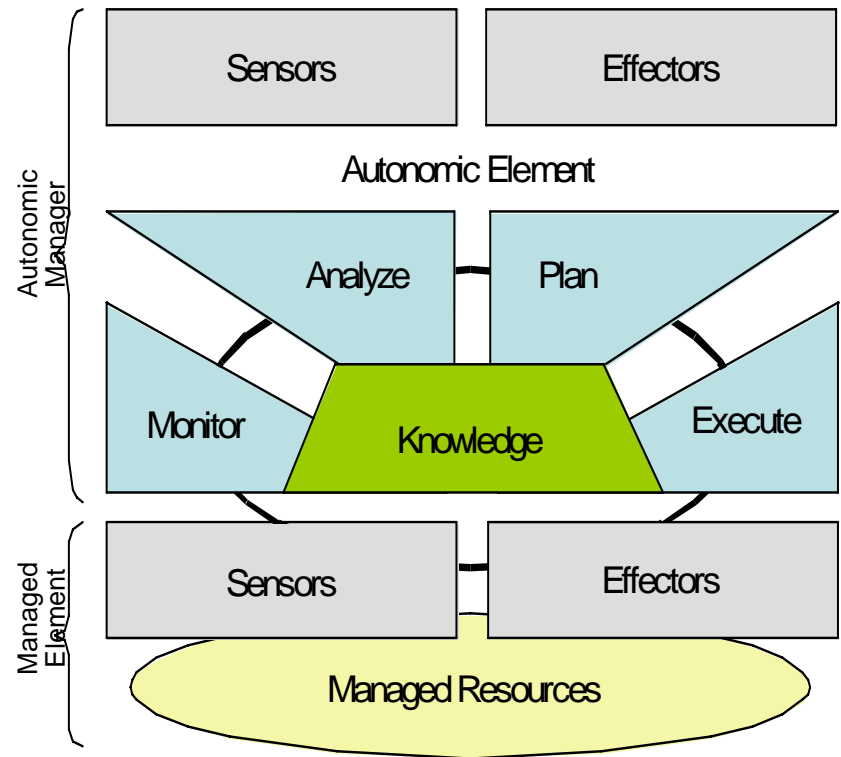


Figure 1 IBM's Autonomic Control Loop [2]

Industrial Research Projects

- IBM – Autonomic Computing
- SUN – N1
- Hewlett-Packard – The Adaptive Enterprise
- Microsoft – Dynamic Systems Initiative
- Intel – Proactive Computing

IBM Autonomic Computing Toolkit

Autonomic Computing

- The Autonomic Computing ToolKit (ATK), which is part of the Emerging Technologies ToolKit (ETTK), is available from the IBM alphaWorks website [4]. Contained within ETTK/ATK are several demonstrations of autonomic computing:
 - **Common Base Event (CBE)**
 - **Generic Log Adapter (GLA)**
 - **Agent Based Learning Environment (ABLE) [5]**
 - **Policy Management for Autonomic Computing (PMAC)**
 - **Autonomic Management Engine (AME)**

IBM Autonomic Computing Self-Healing Systems

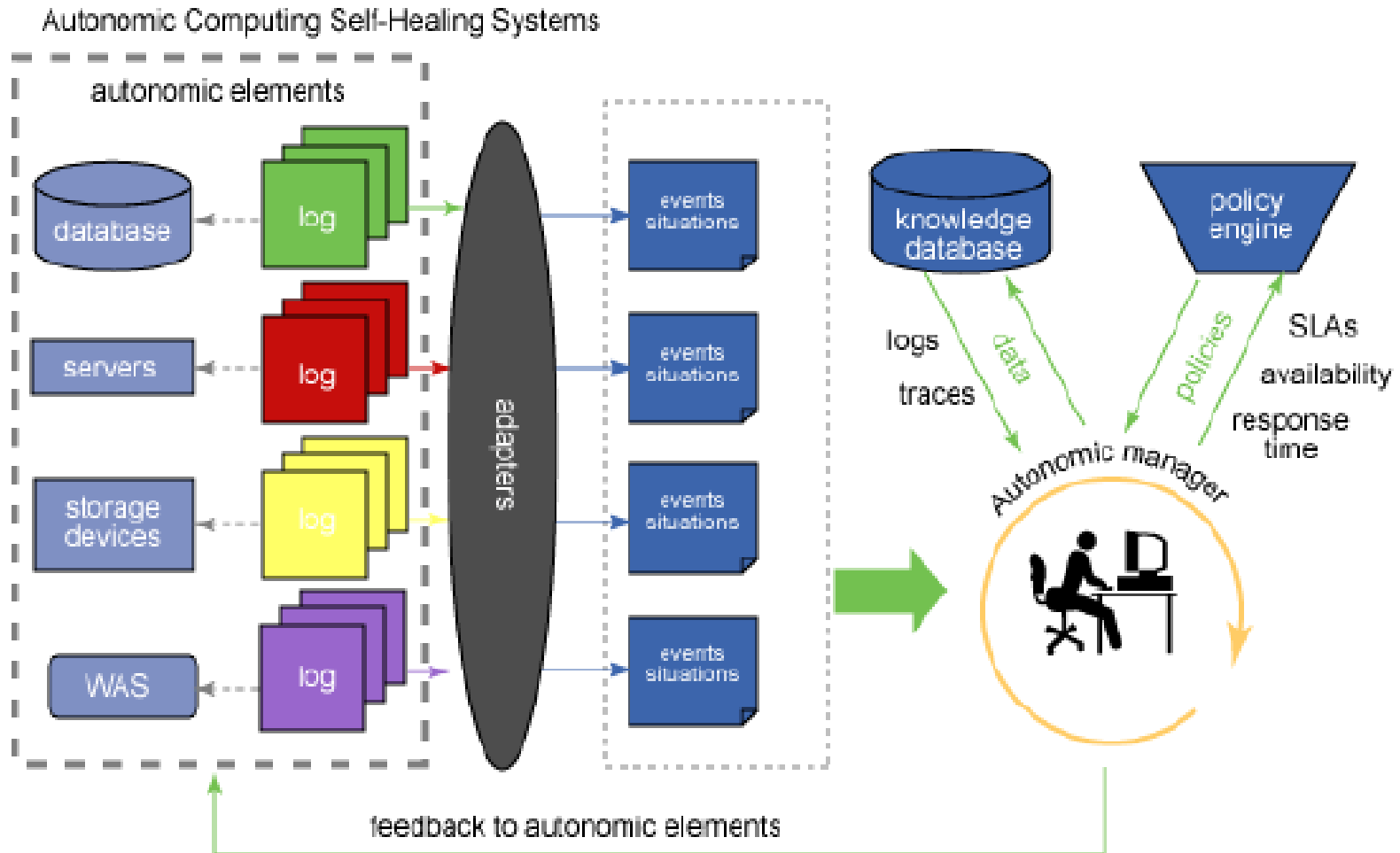


Figure 2 Autonomic Computing Systems

Other IBM Research Projects

- OptimalGrid
 - A research prototype of grid-enabled middleware designed to hide complexities of partitioning, distributing, and load balancing.
 - See <http://www.alphaworks.ibm.com/tech/optimalgrid> for additional information
- Tivoli
 - A suites of software products provide intelligent infrastructure management solutions that help customers understand and proactively manage the business value of their IT systems in an on demand world. It provides integrated views for managing and optimizing critical IT systems with policy-based resource allocation, security, storage and systems management solutions.
 - See <http://www-128.ibm.com/developerworks/tivoli> for additional information
- LEO: DB2's Learning Optimizer.
 - A comprehensive way to repair incorrect statistics and cardinality estimates from a query execution plan.
 - See <http://www.research.ibm.com/> for additional information
- UFiler
 - A web-based enterprise file system solution. It allows access to files anytime and anywhere, and files are protected through fine-grained access-control lists. UFiler desktop clients allow applications to access files stored in UFiler as if they were on a local disk.
 - See <http://www.research.ibm.com/sutonomic/academic/research.html> for additional information

Sun – N1

- N1 is Sun's vision for the next-generation data center.
 - Sun's N1 manages widely distributed computing resources such as servers, storage, software, and networks, and enables them to operate as a single entity.
- Core concepts of N1
 - Virtualization
 - The process of modeling all the components in the network.
 - Application and service level provisioning
 - The process of configuring and deploying software on systems.
 - Dynamic policy management
 - Automations of the configuration and deployment of software in the system.

HP – The Adaptive Enterprise

- HP's approach to autonomic computing is embraced in its Adaptive Enterprise strategy.
 - Provide tools, services, and products that deliver IT service levels that match the flow of real-time business activities or changes whenever needed.
- HP's adaptive technologies fall into three main areas:
 - Dynamic resource optimization
 - Automated and intelligent management
 - Continuous and secure operations

Microsoft – Dynamic Systems Initiative

- To deliver self-managing dynamic systems
 - Help IT teams capture and use knowledge to design more manageable systems and automate ongoing operations, resulting in reduced costs and more time to proactively focus on what is most important to the organization.
- The System Definition Model (SDM) [7]
 - A key technology component of the DSI
 - Provide a common language, or meta-model, that is used to create models that capture the organizational knowledge relevant to entire distributed systems.

Intel – Proactive Computing

- A new era of computing
 - Focus on human-supervised operation, where the user stays out of the loop as much as possible until required to provide guidance in critical decisions.

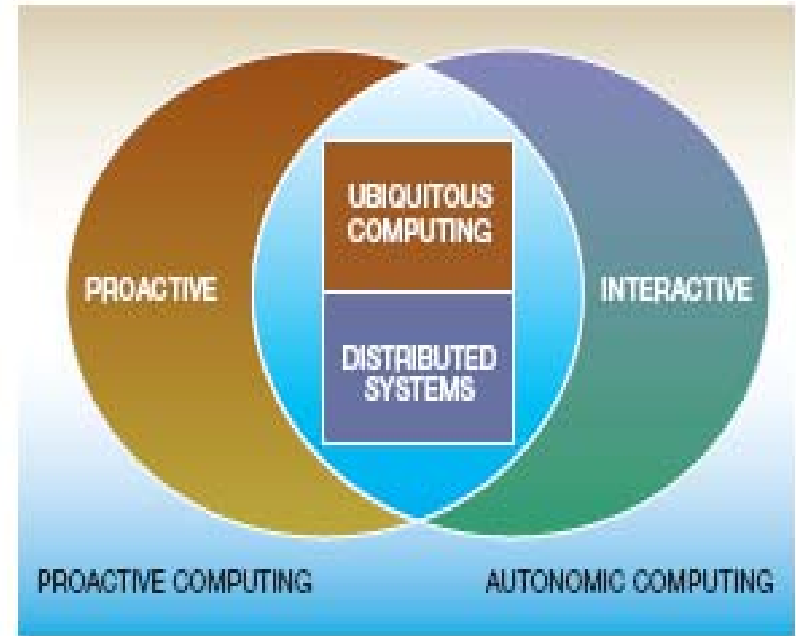


Figure 3: The relationship of computing paradigms [8]

University major research areas

- **Self-configuring**
 - Georgia Tech (self-configuring storage system)
 - George Mason University (self adjustment of computer system configuration parameters under dynamic loading)
 - University of Southern California at Los Angeles: network reconfiguration for applications optimization
 - Stanford University: stability aspects of system reconfiguration, ...
 - Università' di Modena e Reggio Emilia (Italy) – self organization, emergent behavior, and management of amorphous computing
- **Self-healing**
 - University of Maryland (remote repair of OS faults)
 - Brown University – transaction models for automated failure recovery in software
 - University of Minnesota – agent based network monitoring systems for faults and security
 - Duke University – proactive fault management at the applications level (cleaning & restart)
 - University of California at Berkeley - OceanStore, Recovery-oriented computing (ROC)
 - Campus Universitaire' de Beaulieu (France) - remote repair of OS faults
- **Self-optimizing:**
 - Michigan State University – adaptive middleware to support CORBA-based applications self-optimization, self-optimizing wireless network applications
 - University of Michigan Ann Arbor - self optimization of QoS in eCommerce systems
 - Vanderbilt University – self optimization of QoS in eCommerce systems
 - York University (UK) - optimization of database access

University major research areas (Contd.) *Autonomic Computing*

- Self-protecting
 - Texas A&M University - Cooperating Security Manager,
 - Event Monitoring Enabling Responses to Anomalous Live Disturbances (EMERALD), SRI International [8]
 - Purdue - Intrusion Detection Using Autonomous Agents, COAST
- Grid computing environments
 - University of Arizona – “Automania” applications management in grid computing environments, agent-based network attack framework
 - Rutgers University - application management
 - Carnegie Mellon - rule-based grid computing management
 - Ohio State University: agent based optimization of desktop grid computing,
- Routing
 - University of Central Florida (adaptive "smart" routing optimization of networks)
 - MIT (reinforcement learning in ad hoc networks for optimal routing decisions)
 - University of Texas at Austin (adaptive network routing and scheduling for optimization)
 - Imperial College (UK) - adaptive “smart” routing optimization of networks

Major AI and learning methods

Autonomic Computing

(By scanning papers of the First and Second International Conference on Autonomic Computing, 2004 and 2005)

- Data Mining, clustering method to discover classes of configurations [11]
- Decision tree to locate the source of a system fault [12], to learn to classify the properties of files. [13]
- Reinforcement learning to repair the network failure [14], to optimize the network routing and scheduling [15], to dynamic allocate resources. [19]
- Automated planning for goal-type policies.[16]
- Multi-dimensional utility function to optimize the task execution [17].
- Bayesian networks to construct the dependency model of software component [18].

University Research Projects

- **OceanStore** (University of California, Berkeley) [9]
 - A global persistent data store designed to scale to billions of users. It provides a consistent, highly-available, and durable storage utility atop an infrastructure comprised of untrusted servers.
 - Any computer can join the infrastructure, contributing storage or providing local user access in exchange for economic compensation. Users need only subscribe to a single OceanStore service provider, although they may consume storage and bandwidth from many different providers.
- **Recovery-Oriented Computing** (University of California, Berkeley and Stanford University) [10]
 - A project to investigate novel techniques for building highly-dependable Internet services.
 - Emphasizes recovery from failures rather than failure-avoidance.

(From [20])

- **AUTONOMIC ELEMENT CHALLENGES**

- Specific autonomic elements
 - Research directed towards improving the self-managing capability of specific components such as databases, storage systems, servers, etc.
- Generic autonomic element technologies
 - Research on technologies that are generally applicable to autonomic elements, including planning, modeling, forecasting, optimization, etc.
- Generic autonomic element architectures, tools, and prototypes
 - Research on the internal structure of autonomic elements, tools to help create autonomic elements, and reference implementations of autonomic elements built with these tools.

Research Challenges (Contd.)

- AUTONOMIC SYSTEM CHALLENGES
 - Autonomic system technologies
 - Research on generic technologies that entail interactions among multiple autonomic elements to achieve system-level goals, including problem determination and remediation, automated provisioning, workload management, automated installation and configuration, integrity management, etc.
 - Autonomic system architectures and prototypes
 - Research on system-level architectures that effectively govern interactions among autonomic elements, and prototypes that assemble the many other pieces described in this framework to demonstrate systems with improved self-management relative to existing systems.
 - Autonomic system science
 - Research on fundamental science of large-scale autonomic computing systems, addressing questions of learning, stability, control and emergent behavior in multiagent systems, and also addressing questions of how to quantify the degree of self-management in systems.

Research Challenges (Contd.)

- HUMAN-COMPUTER CHALLENGES

- Human studies

- Research on present and future interactions between human administrators and other users and self-managing systems, to determine what interfaces and other modes of interaction are most effective.

- Policy

- Research on methods for eliciting high-level policies from people, representing and appropriately transforming those policies within autonomic systems, and managing behavior with respect to those policies.

Appendix: Industry standard support

	Solution install	Common system administration	Problem determination	Autonomic monitoring	Policy-based management	Complex analysis	Transaction measurements
Distributed Management Task Force, Inc. (DMTF) Common Information Model (CIM)			•	•			•
Internet Engineering Task Force (IETF) Policy Core Information Model (RFC3060) Simple Network Management Protocol (SNMP)			•	•	•		
Organization for the Advancement of Structured Information Standards (OASIS) Web Services Security (WS-Security) Web Services Distributed Management (WS-DM)			•		•		•
Java Community Process Java Management Extensions (JSR3, JMX) Logging API Specification (JSR47) Java Agent Services (JSR87) Portlet Specification (JSR168)		•	•	•	•	•	•
Storage Networking Industry Association (SNIA) BlueFin			•	•			
Global Grid Forum (GGF) Open Grid Services Architecture (OGSA) Open Grid Services Infrastructure (OGSI) Web Services Common Resource Model (WS-CRM)			•	•	•	•	•
The Open Group Application Response Measurement (ARM)			•	•			•
New Autonomic Computing Standards (to be developed)	•	•	•	•	•	•	•

Figure 3: Mapping of industry standards to autonomic computing core capabilities [2]

Reference

1. P. Horn, *Autonomic Computing: IBM's Perspective on the State of Information Technology*, IBM Corporation (October 15, 2001); available at http://www.research.ibm.com/autonomic/manifesto/autonomic_computing.pdf
2. Murch, R.: *Autonomic Computing*. Prentice Hall PTR, 2004.
3. IBM, "An architectural blueprint for autonomic computing", April 2003, available at <http://www-306.ibm.com/autonomic/pdfs/ACwpFinal.pdf>
4. IBM alphaWorks, <http://www.alphaworks.ibm.com/autonomic/all>
5. IBM developerWorks, Meyer, M., "The features and facets of the Agent Building and Learning Environment (ABLE)", <http://www-128.ibm.com/developerworks/views/autonomic/library.jsp>
6. HP OpenView services, <http://www.managementsoftware.hp.com/service/index.html>
7. System Definition Model Overview White Paper, <http://www.microsoft.com/windowsserversystem/dsi/sdmwp.msp>
8. R. Want, T. Pering, D. Tennenhouse, "Comparing Autonomic and Proactive Computing", *IBM Systems Journal*, vol. 42, no. 1, pp. 129-135, January 2003
9. OceanStore, <http://oceanstore.cs.berkeley.edu/>
10. Recovery-Oriented Computing, <http://roc.cs.berkeley.edu/>
11. E. Kiciman and Y.-M. Wang, "Discovering correctness constraints for self-management of system configuration", *Proceedings of the International Conference on Autonomic Computing*, May 2004.
12. M. Chen, A. Zheng, J. Lloyd, M. Jordan, and E. Brewer. "Failure diagnosis using decision trees", *Proceedings of the International Conference on Autonomic Computing*, May 2004
13. M. Mesnier, E. Thereska, g. Ganger, D. Ellard, and M. Seltzer, "File classification in self-* storage system", May 2004
14. M. Littman, N. Ravi, E. Fenson, and R. Howard, "Reinforcement Learning for Autonomic Network Repair", *Proceedings of the International Conference on Autonomic Computing*, May 2004
15. Shimon Whiteson and Peter Stone, "Towards Autonomic Computing: Adaptive Network Routing and Scheduling", *Proceedings of the International Conference on Autonomic Computing*, May 2004
16. B. Srivastava and S. Kambhampati, "[The Case for Automated Planning in Autonomic Computing](#)" *Proceedings of the International Conference on Autonomic Computing*, June 2005
17. A. Ranganathan and R. Campbell, "[Self-Optimization of Task Execution in Pervasive Computing Environments](#)", *Proceedings of the International Conference on Autonomic Computing*, June 2005
18. Y. Li, W. Li, M. Zhang and C. Hou, "[An Active Method to Building Dynamic Dependency Model for Distributed Components](#)", *Proceedings of the International Conference on Autonomic Computing*, June 2005
19. D. Vengerov, N. Iakovlev, "[A Reinforcement Learning Framework for Dynamic Resource Allocation: First Results](#).", *Proceedings of the International Conference on Autonomic Computing*, June 2005
20. J. Kephart, "Research Challenges of Autonomic Computing", *Proceedings of the 27th International Conference on Software Engineering (ICSE) 2005*

Q & A