Preface

The goal of autonomic computing is to reduce the increasing complexity of managing large computing systems. As computing systems evolve, they are subject to the effect of continuous growth in the number of degrees of freedom that must be well-managed in order to maintain their efficiency. Two major factors contribute to the increase in the number of degrees of freedom. One is the different rates at which the capabilities of computing elements, such as the CPU, memory, disks, and networks, have historically increased. The disparity between the capabilities of various elements provides opportunities to use different strategies for a task, depending upon the environment. In turn, this calls for a dynamic approach in order to make judicious choices for achieving targeted efficiency. The other factor is the tendency of current systems to exhibit a global range in the demand for their services and the resources they employ for rendering the services. Changes in the demands or resources in one part of a system can have a significant effect on other parts of the system.

On the human side, users, developers, and systems administrators must become more sophisticated in detecting and solving problems. The addition of layer upon layer of system software brings the promise of simplifying environments, yet inevitably requires new levels of expertise. Where will this spiral take us? Autonomic computing is the result of the realization that unless we begin to build computing systems that reduce the complexity for those who use and manage them, we will not have the time or the expertise to unravel problems arising in newer systems.

The information technology industry, which has been an important contributor to the world's economy by increasing productivity, could instead become an inhibitor to advances in both the developed and emerging economies around the world. A worst-case scenario can be imagined. The plot of an early *Star Trek* television program comes to mind ("The Ultimate Computer," 1968) in which the M-5 multitronic computer, which can run the United Star Ship Enterprise without any human intervention, goes awry in its mission and threatens the survival of the star ship. After considerable effort, two characters in the program, Spock (the First and Science Officer) and the ever-capable Scottie (the finest engineer, ever), are able to disconnect M-5 and gain back control. Thus, autonomic computing must focus on actually reducing complexity, not simply hiding it.

Computing systems must become more capable of detecting and correcting problems by recognizing impending situations that will likely cause trouble. But these actions must be taken in a way that allows the systems to be tracked and manually overridden. Autonomic actions that are obscure or that appear to be unpredictable or untraceable will produce distrust and lack of confidence among users and limit general acceptance of this initiative.

How can the benefits of an autonomic system be measured and quantified? This question must be analytically approached and answered as autonomic computing matures. For industrial firms, such as IBM, each investment into technologies, including those that are self-regulating and responsive to external situations, must be justified by the expected return on investment. Can this return be crisply measured in terms of shorter development cycles, increased employee productivity, greater revenue, or increased sales? Will customers see the value in this approach and believe it brings them additional advantage in their markets? How will customers measure this

0018-8670/03/$5.00

value? In the abstract, it is simple to believe in and understand the value of autonomic computing. Downtime of systems, problems that are difficult to diagnose, Web site failures, and poor user response times requiring skilled experts to repair or improve them are all well-known conditions and reported on regularly. Security vulnerabilities and malicious virus attacks are another dimension of the complexity problem that can produce catastrophic effects.

On the Internet, where the cost to a consumer of switching to a competing business is nil, the loss of an exclusive customer is a major event for a business. One of the most widely publicized examples of how the failure of a complex, fragile system can have disastrous effects on a business is eBay's original nonscalable, failure-prone architecture, which produced a series of failures that affected customers between 1998–2001. According to CNET news (June 14, 1999), in a group of consumers surveyed, 53 percent made no change in their behavior after experiencing technical problems at a Web site, and only 9 percent ceased to use the site. Nearly one-fourth, or *24 percent*, found a new site and used both old and new, whereas 13 percent found a new site but only used it once. These percentages indicate that about one-fourth of consumers begin to shop around after encountering a problem.

It is not enough to provide verbal assurances or white papers outlining the value of autonomic computing to solve these long-standing problems. Rather, we must begin to build a body of literature that demonstrates its value. In this issue, the first dedicated to autonomic computing, we begin that process.

As plans for a body of literature on autonomic computing developed, we felt that the *IBM Systems Journal* would provide an excellent first venue for a set of papers on the subject. We then had to decide on a set of appropriate topics. The final list, refined after discussions with many colleagues, reflects the current thinking about autonomic computing, as well as approaches being taken in research and development. The topics included in this issue are: infrastructure, storage, systems management, middleware, tools, clients, and services and applications. Present in the background of all these topics is the application of theoretical principles in algorithms and optimization.

Systems have traditionally been designed in a layered, building-block fashion, in which boundaries indicate a change in function, speed, access level, and so on. IBM's view of autonomic computing follows this pattern, which is reflected in the content and flow of the papers in this issue. The server infrastructure, which acts as the fundamental system base, must provide a solid foundation for autonomic computing and must be capable of responding to varying workload demands in a timely way. The ability to seamlessly allocate resources to adjust to these needs and balance need against demand is a challenge that is addressed via the concept of dynamic reconfiguration. It is one of the first steps toward the autonomic environment and is addressed in the paper, "Dynamic Reconfiguration: Basic Building Blocks for Autonomic Computing on IBM pSeries Servers." At the opposite end of the research being described, papers focus on how autonomic personal computing and user environments will be affected. The challenge here is to respond to user needs, allow individuals to use their machines for work or play, provide them with a robust and versatile environment, and yet not impose constraints. Users want full function with ease of use—not a return to the world of remote terminals.

This issue contains 16 papers and a Technical Forum article. Its content represents a first, but incomplete, step toward the examination of autonomic computing. We believe that this issue will begin a long and interesting discourse on a subject that will remain at the core of our industry for many years to come.

In the first paper of the issue, "The Dawning of the Autonomic Computing Era," A. G. Ganek and T. A. Corbi provide an overview of autonomic computing. They discuss why it is needed, what it is, and how it might be implemented, while focusing on IBM's initiative in this area. They include synopses of the other papers in this issue.

The next issue of the *Journal* is devoted to storage systems, with a focus on work being done in IBM.

Lorraine Herger, Issue Coordinator
Kazuo Iwano, Issue Coordinator
Pratap Pattnaik, Issue Coordinator
Alfred G. Davis, Associate Editor
John J. Ritsko, Editor-in-Chief