

SCALABILITY MANAGEMENT FOR E-BUSINESS SOLUTION: A RESOURCE-BASED VIEW

Dr. Binshan Lin, Louisiana State University in Shreveport, blin@pilot.lsus.edu

ABSTRACT

Scalability management is a key strategy that organizations are embracing to manage in e-business. This paper extends the resource-based view to the context of scalability management, and examines several essential ingredients to the success of scalability management for e-Business solution.

Keywords: Scalability, Availability, Performance, Adaptability, Workload, Scalability Management, Resource-Based View, Network Analysis, e-Business

INTRODUCTION

Scalability can be defined as the ability of a computer, product, or system to expand to serve a larger number of users without breaking down (12). In simple words, scalability is the capacity of a system to handle increasing workload. Another definition, in the context of e-business, is the capability of a server, application, or Web site not only to function well in the rescaled situation, but also to take full advantage of it (4). Scalability is a critical component of Web-enabled applications because the nature of the Web is such that workload cannot be predicted, but must be handled when it arises (23).

There are four dimensions of scalability: data size, speed, workload, and transaction cost (27). In analyzing each of these dimensions, one needs to evaluate system growth by the ideal of linear scaling. Dimensions related to result (such as total response time and cost) ought to increase in direct proportion with dimensions related to requirements (such as transaction and data volume).

SCALABILITY CHALLENGES IN E-BUSINESS

A recent empirical study shows that the size of the Web site was found to correlated with almost all content and design aspects, i.e., larger Web sites do not only contain more content bit also more different content features, more different forms of entertainment and are considered as more informative (9). On the Web “larger” does not only mean “more of the same”, larger sites are also richer sites. In general, scalability challenges exist in every larger Web site.

B2C Scalability Challenges

For B2C systems, transactions are generally simple and the greatest challenges are often in getting personalized content to a browser. Customerization and interactivity are key tool for profitable B2C sales (16). B2C scalability challenges include:

- The system must turn the understanding of customers’ needs into system design’s driving force.

The nature and magnitude of the scalability problem goes beyond even what the immense data volumes suggest. What really cranks up the scalability requirement is the in-depth analysis required to extract business value from the clickstream data.

- Much of what is to turn to B2C advantage comes from the details of the data. The data for data mining, for example, must be complete and comprehensive. An organization can have access to data on all customers, but what exactly does it know about each customer? Is sufficient history available to do effective trend analysis? Does the data represent a complete view of the business in terms of customers, competitors, the market, suppliers?
- Gathering detailed data clearly means dealing with very large volumes of data via the Web. Detailed data produces very large databases (VLDB), and data warehouse platform must provide the necessary scalability and performance at all levels of the architecture: hardware, operations systems, and DBMS.

B2B Scalability Challenges

For B2B systems, there may not be a human on the client side to guess what to do next, or to ask if something is unclear. A B2B transaction must be much more precise, while at the same time, adapting to many different business environments.

- The system must tie together information such as product catalogs from many suppliers and applications such as order management, scheduling, and auctions in real time.
- The site must implement complex business rules to customize business transactions between customers and suppliers.
- It must scale to serve thousands of concurrent customers initiating complex transactions with information drawn from hundreds of partners and suppliers.
- Competitive advantage comes from non-stop availability. The application cannot be taken down for updates, performance tuning, system reconfiguration or system failure.

A CONCEPTUAL FRAMEWORK

The prevailing paradigm of understanding how and why firms gain and sustain competitive advantage is the resource-based view of the firm (13, 21). Focusing on a firm level analysis, the resource-based view emphasizes the resources possessed, developed, and deployed by an organization and understanding the relationship of those internal resources with performance (26).

The theory primarily focuses on how and why some resources and capabilities are valuable, yet rare, difficult to imitate or nonsubstitutable by allowing those firms that have them to accrue superior business returns (4). The competitive advantage arises from the heterogeneity of valuable resources and capabilities (i.e., strategic assets) across firm. The sustainability of the advantage depends on isolating mechanisms, including causal ambiguity, social embeddedness, path dependence, complementarity, and assets specificity of resource and capabilities that make them immoral, nonimitatable, or rare (18).

Notice that resources and capabilities have several differences. Resources can be physical capital

resources (such as financial assets and information technology), human capital resources (such as marketing expertises), or organizational resources (such as team-based culture) (3). On the other hand, capabilities refer to a firm's capacity to deploy valued resources, usually in combination or in co-presence (21). Capabilities include such competences as trustworthiness, organizational flexibility, rapid response to new customer trends, and short product life cycles. The capabilities are firm specific and are developed over time (5).

The resource-based view assumes economic rationality in managerial choices for resources, notwithstanding the common cognitive biases, limitations, and organizational inertia that strategic decision making face (2). On the other hand, individual managers have little discretion over capabilities or the pace of their development although capabilities cannot be disentangled from the firm's past experience and the overall skills and knowledge of the management (2).

The information technology (IT) field has long linked IT with competitive advantage. IT is a resource that can generate competitive advantage value (8), but only if deployed so that it leverages or exploits pre-existing resources in the firm via co-presence and complementarity (19). In terms of sustainability, IT's competitive advantage resides in the organization's managerial skills related to IT, not in the technology itself (14).

In order to implement an effective scalability management, several important elements are needed to identify and examine.

Ingredient 1: Scalability Assessment

Perfect scalability - excluding cache initializations - is linear (1). Linear scalability, relative to load, means that with fixed resources, performance decreases at a constant rate relative to load increases. Linear scalability, relative to serve resources, means that with a constant load, performance improves at a constant rate relative to changes in resources (1).

In general, three major factors involves in assessing scalability of an e-commerce platform: performance, availability, and adaptability.

Availability is the ability to maintain service and performance all the time. A service breakdown, in fact, isn't an option for an e-commerce business, because customers like to shop at all hours (7). For a company like *buy.com*, which focuses on offering heavily discounted prices on many product categories, an outage of an hour or two can exact a heavy cost, both in disappointed customers and lost sales.

Availability can also refer to the probability that the system is operational according to its specification at a given point in time. If one looks at the limit of availability as time goes to infinity, it refers to the expected percentage of time that the system under consideration is available to perform functions as intended. Availability and reliability of a system are considered to be contradictory objectives (22). In the e-Business platform, it is usually accepted that it is easier to develop highly available systems as

opposed to highly reliable systems.

Many supposedly scalable systems lack build-in mechanisms to handle a rapidly changing workload (6).

As key ingredient of scalability management, adaptability must be addressed during the assessment procedure. This avoids the risk of finding out too late that the solution cannot meet real business demands.

Ingredient 2: An Effective Business Process Support

A business process engine must support high-performance process definition, redefinition, and process state update and interrogation. Like other element, process management is most scalable if distributed. Process definition can be replicated from a central repository, thereby enabling local and efficient lookups for each process instance. A slow process engine in e-Business will become a serious bottleneck and impede scalability.

Ingredient 3: Build Supportive Infrastructure

The third major ingredient is to build the supportive infrastructure. It is tempting to describe its technical characteristics by saying that “we are scalable” or “we are scalability-enabled.” More importantly, we need to understand the characteristics of the infrastructure that draw customers to the Web site and encourage them to continue to use the Web as a primary way of using products and services. One key question related to the third ingredient is: “Do customers find their online experience supported by the appropriate functionality?” (25).

A supportive infrastructure includes (1) a scalable architecture, and (2) supportive facilities. First, a scalable architecture is a technology infrastructure that can logically and physically increase in performance and capacity with continuity to meet reasonable growth and change over time. A scalable architecture contains a strategic migration plan for continuous growth and progress.

Ingredient 4: Determine Approaches By Workload Pattern

Web sites are distinguished primarily by workload pattern. All Web sites types are considered to have high-volume of dynamic transactions. After evaluation of the sites for such characteristics that pertain to transaction complexity, volume swings, data volatility, security, and other, one is able to categorize the five types: publish/subscribe, online shopping, self-service, trading, and B2B sites.

Several scaling approaches are available in current practices. These scaling approaches are related to the objectives of scalability management. For example, if the objective is to increase the speed of a component, one would consider using a faster or special machine and/or creating a machine cluster. If a large number of requests can be handled using a cache instead of the database, the overall load on the database is reduced, affording greater scalability for the entire system. In practice, the approaches that reduce load on one component actually make other components more efficient, thus compounding the scalability effect.

Ingredient 5: Scalability Can Be Realized Through Design

Scalability starts with technology, but it is realized through design. Certainly much of the scalability management must depend on the hardware platform and technology infrastructure, but effective scalability must be designed into the application. For example, some significant application design guidelines to emerge are extensive use of caching, including network caching and site specific caching, and data modeling techniques to reduce the number of database queries per customer visit.

Building quality into e-Business service design can increase the availability of resources to prioritized requests. This involves allocating limited shareable resources to the requests that need them most, by classifying and qualifying requests based on business policies. For example, one may classify users who buy or are paying ahead of users who browse or are searching. This deification is directly applicable to online shopping, self-service, and online trading sites. One may also distinguish between requests based on complexity, i.e., separating the typical simple transactions (such as product display) from the complex transaction (such as payment authorization).

CONCLUSION

Scalability will emerge as a major catalyst for e-business and a key differentiator for individual companies. This paper tried to synthesize important lessons learned from organizations in terms of the five major ingredients for making scalability management successful in an organization. Investments in scalability for e-business solution tend to be irreversible and highly path dependent, placing firms on the so-called trajectory of capability development (24).

Future research could concentrate on the development of a general framework to study scalability management. The academic community needs such a general framework to facilitate the comparison of the results of scalability management to e-business solutions. We invite future researchers to build on and/or challenge our extensions to the resource-based view on scalability management for e-business solution. The use of normative studies by different researches can speed up the development of a solid e-business solution knowledge base. Managers in practice are very eager to make use of such a knowledge base.

REFERENECES

1. Allaire Corporation. (1999). A Scalable Web Application Development with ColdFusion 4.0, Technical Report, www.allaire.com.
2. Amit, R. and Schoemaker, P.J.H. (1993). A Strategic Assets and Organizational Rent, *Strategic Management Journal*, (14), 33-46.
3. Barney, J.B. (1991). Firm Resources and Sustained Competitive Advantage, *Journal of Management*, (17), 99-120.

4. Barney, J.B. (1997). *Gaining and Sustaining Competitive Advantage*, Reading, MA: Addison-Wesley.
5. Collis, D.J. and Montgomery, C.A. (1997). Competing on Resources: Strategy in the 1990s, *Harvard Business Review*, (73), 118-128.
6. Columbus, L. (2000). *Realizing e-Business with Application Service Providers*, Sams Publishing.
7. Davis, J.R. and Winter, R.A. (2000). Business Intelligence Scalability for Real Workloads, Technical Report, www.winter.com.
8. Deitel, H.M., Deitel, P.J. and Steinbuhler, K. (2001). *e-Business and e-Commerce for Managers*, Upper Saddle River, NJ: Prentice Hall.
9. Ginzberg, A and Venkatraman, N. (1992). Investing in New Information Technology: The Role of Competitive Posture and Issue Diagnosis, *Strategic Management Journal*, (13), 353-377.
10. Huizingh, E.K.R.E. (2000). The Content and Design of Web Sites: An Empirical Study, *Information & Management*, (37), 123-134.
11. Keyes, J. (1998). *Datacasting: How to Stream Databases Over the Internet*, New York, NY: McGraw-Hill.
12. McGoveran, D. (2000). Architected Scalability, *eAI Journal*, (2), 10.
13. Laudon, K. and Laudon, J.P. (2001). *Essentials of Management Information Systems*, 4th Ed, Upper Saddle River: NJ, 2001.
14. Mahoney, J.T. and Pandian, J.R. (1992). The Resource-Based View within the Conversation of Strategic Management, *Strategic Management Journal*, (13), 363-380.
15. Mata, F.J., Fuerst, W.L. and Barney, J.B. (1992). The Resource-Based View Within the Conversation of Strategic Management, *Strategic Management Journal*, (13), 363-380.
16. Mencnarowski, D. and Zielinski, K. (2000). Improving Scalability of Event-Driven Distributed Objects Architectures, *Software: Practice and Experience*, (30), 1509-1529.
17. Mulvey, B. (2001). Inside the B2C Challenge, @ *eDirections*, Winter 2001, pp.16-21.
18. Perlmutter, B. *Virtual Private Networking*, Upper Saddle River, NJ: Prentice Hall, 2000.

19. Peteraf, M. (1993). The Cornerstones of Competitive Advantage: A Resource-Based View, *Strategic Management Journal*, (14), 179-191.
20. Ross, J.W., Beath, C.M. and Goodhuie, D.L. (1996). Develop Long-Term Competitiveness Through IT Assets, *Sloan Management Review*, 31-42.
21. Rorvig, M.E. (1990). The Simple Scalability of Documents, *Journal of the American Society for Information Science*, (41), 590-598.
22. Schendel, D. (1994). Introduction to Competitive Organizational Behavior: Toward and Organizationally Based Theory of Competitive Advnatage, *Strategic Management Journal*, (15), 1-4.
23. Siewiorek,D.P. and Swarz, R. S. (1982). *The Theory and Practice of Reliable System Design*, Bedford, MA: Digital Press.
24. Skinner, J. (2001). Extended Scalability on the Data Tier, Technical Report, 2001, Microsoft Company.
25. Teece, D., Pisano, G., and Shuen, A. (1997). Dynamic Capabilities and Strategic Management, *Strategic Management Journal*, (18), 509-533.
26. Venkatraman, N. (2000). Five Steps to a Dot-Com Strategy: How To Find Your Footing on the Web, *Sloan Management Review*, (41), 15-28.
27. Wernerfelt, B. (1984). A Resource-Based View of the Firm, *Strategic Management Journal*, (5), 171-180.
28. Winter, R. (1999). Lexicology of Scale, *Intelligent Enterprise*, (2), 70-72.