Chapter II
Economics of Downtime

CHAPTER OVERVIEW

After introducing some basic facts on how today’s businesses are faced with several types of business risks, the second chapter tends to explain one of the major problems that a contemporary business may face in regard to the impact of its computing infrastructure to business results. System downtime is briefly explained and the resulting “economics of downtime” is elaborated in order to demonstrate the direct financial impacts of the unavailability of so-called “business-critical applications” to business results.

DOWNTIME AND UPTIME

This section introduces basic terms of continuous computing and business continuity: downtime and uptime. It introduces the main framework for so-called “economics of downtime” by providing several empirical data on negative financial effects of system downtime.

A serious glitch on computer system, human error, or any kind of natural disaster (fire, flood, earthquake) can halt or shutdown company’s application environment operated and supported by that computer system, in most cases enterprise server (servers). Such an event can cause partial or full unavailability of enterprise data and applications. When system is not operational for any reason, it is said to be down and the time in which it does not operate is called “downtime.”

Downtime refers to a period of time or a percentage of timespan that a machine or system (usually a computer server) is offline or not functioning, usually as a
result of either system failure (such as a crash) or routine maintenance (Downtime, Wikipedia). The opposite is uptime.

Even in the very beginning of the e-business era, more than a decade ago, Datamation (1995) emphasized how important the system downtime is. In August 1995, Datamation quoted the results of a survey of 400 large companies, which unveiled that, “…downtime costs a company $1400 per minute, on average. Based on these figures, 43 hours of downtime per year would cost $3.6 million. One hour of downtime per year amounts to $84,000 per year.”

Uptime is a measure of the time a computer system has been “up” and running. It is often used as a measure of computer operating system reliability and stability, in that this time represents the time a computer can be left unattended without crashing, or needing to be rebooted for many administrative or maintenance purposes. Long uptime can also indicate negligence as many critical updates require reboot on same operating systems (Downtime, Wikipedia).

In modern business, even a few minutes of system downtime may cause thousands or even millions in lost revenues. In addition, such situations may result in bad decisions, unsatisfied customers, broken image of the company. Simply put, when mission-critical applications are considered, system downtime (both planned and unplanned) should be avoided or minimized. This fact emphasizes the need for system’s reliability, availability and scalability. IDC (2006) underscores the fact that a true high availability model expands the concept of availability beyond an infrastructure perspective in terms of system or servers.

There are many definitions for availability. Some consider availability of the data only, while others speak on applications availability or availability of the server or storage subsystem. Many “buzzwords” are in use, such as:

- high availability
- business continuity
- business continuance
- business resilience
- always-on computing
- 24x7x365 computing
- fault tolerant computing
- disaster tolerant computing
- disaster recovery
- disaster resiliency
- real-time computing
- zero latency enterprise, and so forth.
Customer Satisfaction in the Context of Online Gaming Service: The Hedonic Experience Factor
www.igi-global.com/article/customer-satisfaction-in-the-context-of-online-gaming-service/117549?camid=4v1a