Dependabilty Analysis for E-Commerce Site Business Decisions

Myron Hecht, Yutao He, Xuegao An
SoHaR Incorporated, Beverly Hills, CA
{myron, yutao, xuegao}@sohar.com

1. Introduction

Application of standard dependability analysis techniques to e-commerce sites can provide significant non-obvious answers to business questions related to e-commerce system architecture. This paper provides an example.

2. Example System

For an example system, we use a simplified representation of the eBay e-commerce web site shown in Figure 1 [1]. Uunet and Sprint Internet backbones are connected to redundant Cisco routers and then a set of (Compaq) front-end web servers running on Windows NT. The web servers in turn pass requests on auctions and/or bids to redundant Oracle Database Management System (DBMS) servers hosted on Sun Starfire servers These DBMSs in turn retrieve and update tables representing each of the auctions. The outcomes of the auctions are recorded in the Anaconda system that then participants by email.
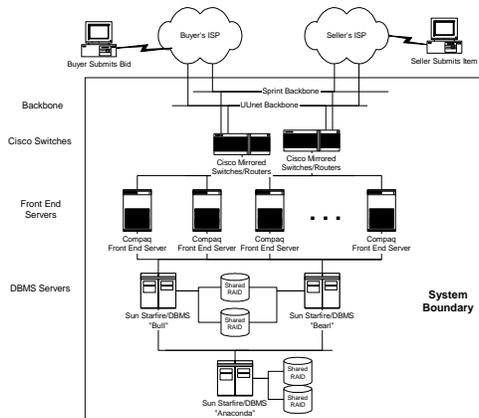

Figure 1   eBay System Configuration

3. Model Description

The system model was created using the MEADEP dependability analysis tool developed by SoHaR [3] Tool

allows a model to be composed as a hierarchy of Reliability Block and Markov Models representing each of the major elements. Hierarchical models allow simpler, easily verified models to be combined to represent a complex system, and represents the most important benefit of hierarchical modeling.
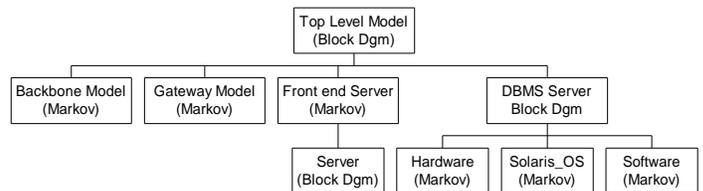

Figure 2. Organization of Hierarchical Model

Figure 3 shows an example of one of the submodels for the front-end web server farm. Loss of a single server results in a loss of capacity and is shown by the reward values in the state representations.
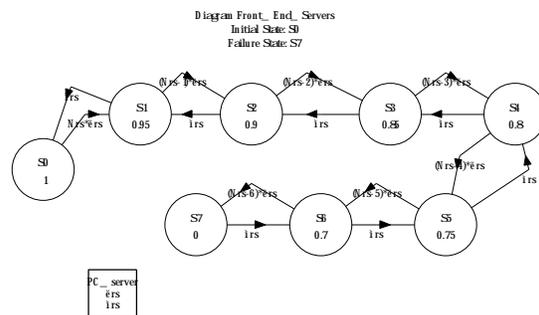

Figure 3. Front End Processor Model

Imperfect recovery and finite recovery times in active-standby systems such as database servers were modeled using coverage parameters Details of these models are described in [2] and [3].

4. Modeling Results

The results of the model using representative (but not actual eBay site) parameters include an overall system availability of 0.9957. The site is predicted to be off-line

every 61 hours with an annual downtime of about 38 hours. With 650 auction bids per minute (as of October, 1999 [1]), this downtime could amount to more than 1.46 million lost transactions. Reducing downtime has a significant financial impact. However, not all systems contribute equally. Figure 4 shows the failure rates by subsystem and shows that the highest failure rate is the front-end server PC subsystem. Although a natual conclusion is that this subsystem should have the highest priority for being addressed, the WAN backbone subsystem actually had the highest impact on site availability.
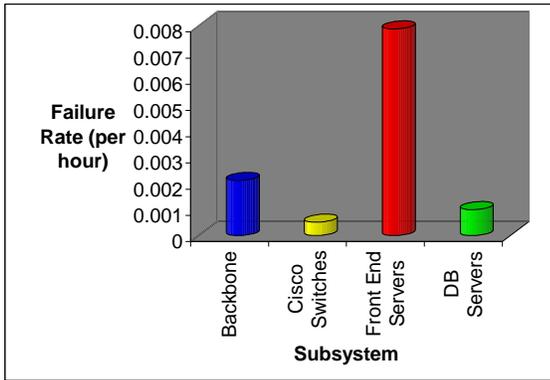

Figure 4. Results by Subsystem

This result was found using parametric analysis and demonstrates its importance in defining an overall availability strategy. The Model Evaluator (ME) module of MEADEP facilitates such analyses for design and operations tradeoffs. Figure 5 shows the annual downtime as a function of the service provider as the repair rate is being varied from 0.5 per hour (corresponding to a repair time of 2 hours) to a repair rate of 2 per hour (corresponding to a repair time of 0.5 hours). The ordinate shows the expected downtime decreases from approximately 38 hours to approximately 30. This result may be quite significant when negotiating a service level agreement.
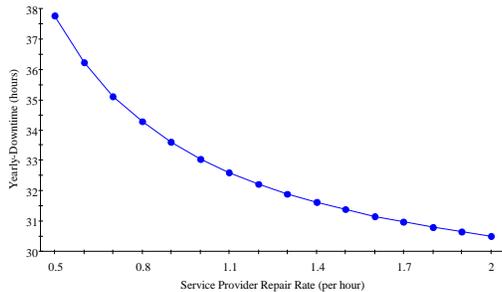

Figure 5. Yearly System Downtime as a Function of Service Provider Repair Rates

The management benefit of such modeling becomes apparent when downtime is translated into lost transactions such is as shown in Figure 6. With a restoration time of 30 minutes, the expected number of lost transactions per year is approximately 1.2 million. However, as the restoration time approaches 5 hours, the number of lost transactions increases to 2.8 million. If the value of each transaction is $0.10, the value of a service level agreement that guarantees a restoration time of 0.5 hours or less is more than $160,000 per year.
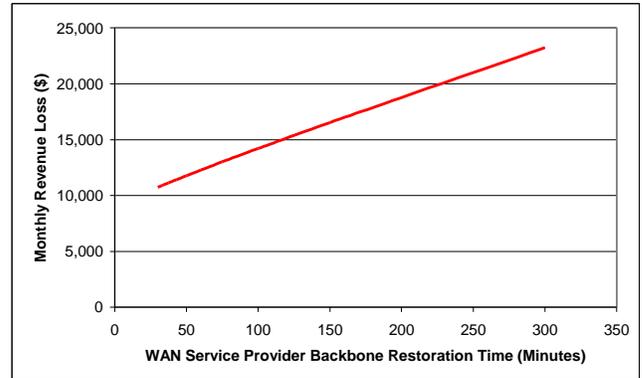

Figure 6. Impact of Restoration Time

The model can be used to evaluate the value the additional capacity. Increasing the capacity of both backbone WANs so that either could handle the full load (as opposed to either handling half the load) could decrease the monthly revenue loss by as much as $30,000 or as little as $9,000 depending on the reliability of the service providers. The lower the reliability (MTBO), the greater the benefit of increasing the capacity. With these results, it is also possible to assess the value of a service level agreement with the additional dimension of capacity. For example, it may be that provisioning each of the backbones with double the capacity is a lower cost option than securing service level agreements with both WAN backbone service providers. With the MEADEP reward function, it is also possible to consider other options such as providing full capacity in the first backbone and half capacity in the second one, or providing 75% capacity in each. The optimum configuration depends on the traffic profile and the MTBO and restoration times of the WAN backbones, the load profile of the server, the value of the traffic, and other system-specific factors.

References

[1] Joseph Menn, "Prevention of Online Crashes Is No Easy Fix", Los Angeles Times, October 16, 1999, Section C, Page 1

[2] D. Tang, et. al., *Experience in Using MEADEP,* Proc. 1999 Reliability & Maint. Symp. Wash., DC, 1999

[3] MEADEP description, www.sohar.com/meadep, 2000