Performance and Availability of Internet Data Centers

Daniel A. Menascé • *George Mason University*

I've used previous columns to discuss several quality-of-service (QoS) metrics, including response time, throughput, and availability, in the context of Web scalability. In most of my past discussions, though, I treated these metrics in isolation; here, I'll use an Internet data center (IDC) as a motivating example to discuss how performance and availability are interrelated. IDCs provide the means for geographically distributed Internet users to tap into other computers and applications.

IDC users pay for the services they obtain, so they want them to be delivered according to established service-level agreements (SLAs). These agreements indicate service performance levels and availability. An IDC must provision enough capacity and redundant resources to ensure that it can meet its performance and availability SLAs. Failing to do so can incur a loss of revenue or, in some cases, penalties.

An IDC Model

IDC-provided services must meet response-time, throughput, and availability constraints — for example, average response time should not exceed two seconds for a given type of application, the IDC must be able to process at least 8,000 requests per second on average, and a given application should be available at least 99.999 percent of the time.

To explain the relationship between throughput and availability, we can look at a simple performance and availability model for an IDC (see Figure 1). The IDC has M equivalent machines that can all satisfy the user requests submitted to it. Assuming that each of these machines has a processing capacity of X requests/sec, the IDC's maximum theoretical capacity is thus equal to $M \times X$ requests/sec. Assuming that machines fail at a rate of γ failures/sec, a failed machine joins a queue of machines waiting to be fixed by one of the N members of the repair staff. When a machine is repaired, it returns to the pool of operational machines.

Consider the following performance and availability trade-offs:

- As the number of machines increases, the cost of operating the data center increases along with the IDC's total throughput and availability.
- As the repair staff's size N and skill level increase, the cost of operating the IDC increases. However, failed machines also return to the pool of available machines faster, thus improving the IDC's throughput and availability.

The average throughput $\bar{X}$ of requests submitted to the IDC is a function of the number j of machines in operation. Thus,

$$\bar{X} = \sum_{j=1}^{M} (jX) \times p_j = X \sum_{j=1}^{M} j \times p_j = X \times \bar{M} \qquad (1)$$

where p_j is the probability that j machines are in operation, and $\bar{M}$ is the average number of machines in operation.

The probability p_j is a function of the following factors:

- machine failure rate γ ($1/\gamma$ is the average time it takes for a machine to go from the operational to the failed state, which we also call mean time to failure, or MTTF);
- number of people N on the repair staff;
- average time F it takes for a member of the repair staff to fix a machine (the average time to bring a machine from the failed to the oper-

 Published by the IEEE Computer Society 1089-7801/04/$20.00

ational state, also known as mean time to repair, or MTTR, is the sum of F plus the average time a machine spends in the queue waiting to be repaired).

The IDC's availability is, obviously, the fraction of time that it's available; we define it as the probability that at least one machine is operational, or $(1 - p_0)$. This very simple availability metric does not reflect the level of performance the IDC provides, however; the IDC would be considered 100-percent available if only one machine were up 100 percent of the time or if all M machines were up 100 percent of the time. A very clear performance difference exists between these two extreme scenarios. A more meaningful availability-related metric is the probability A_j — that at least j machines are in operation — which we can compute as

$$A_j = \sum_{k=j}^{M} p_j. \quad (2)$$

To complete the analysis of the performance and availability model, we need to find the values of the probabilities p_j ($j = 0, \ldots, M$). As indicated in my recent book,[1] we can find these probabilities by using a Markov chain[2] with $M + 1$ states numbered from 0 to M. A state k represents the fact that k machines are in the failed state. The transition rate from state k to state $k + 1$ — in other words, the overall failure rate — is given by $(M - k)\gamma$. This happens because when k machines are in the failed state, $(M - k)$ machines are in operation, and each one of them can fail at a rate of γ. The overall repair rate at state k, or the transition rate from state k to $k - 1$, depends on whether the number of failed machines exceeds the size N of the repair staff. If $k \leq N$, then the overall repair rate is k/F because k machines will be in the process of being repaired, and each one is repaired at a rate of $1/F$. If $k > N$, only

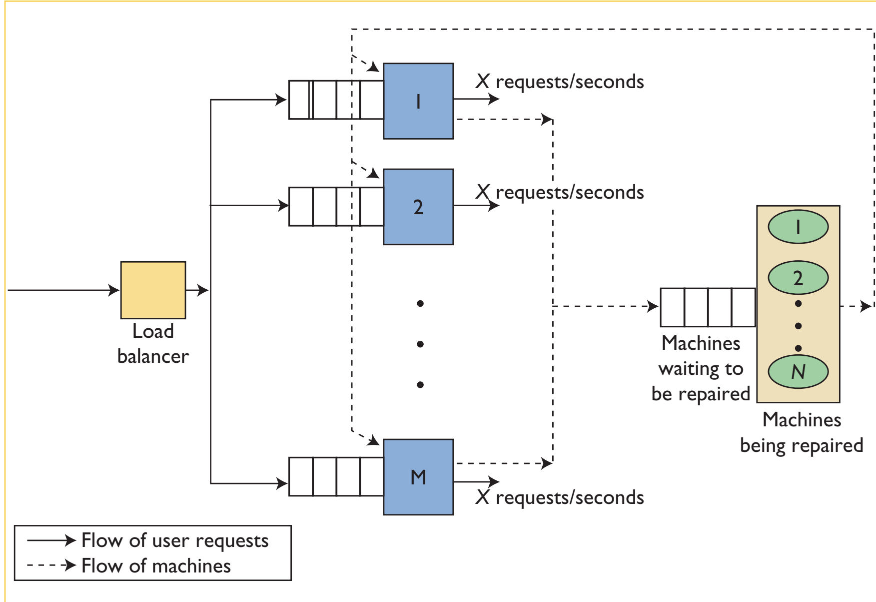

Figure 1. Performance and availability model for an Internet data center (IDC). User requests are submitted to one of M equivalent machines. When a machine fails, it joins the queue of machines waiting to be repaired.

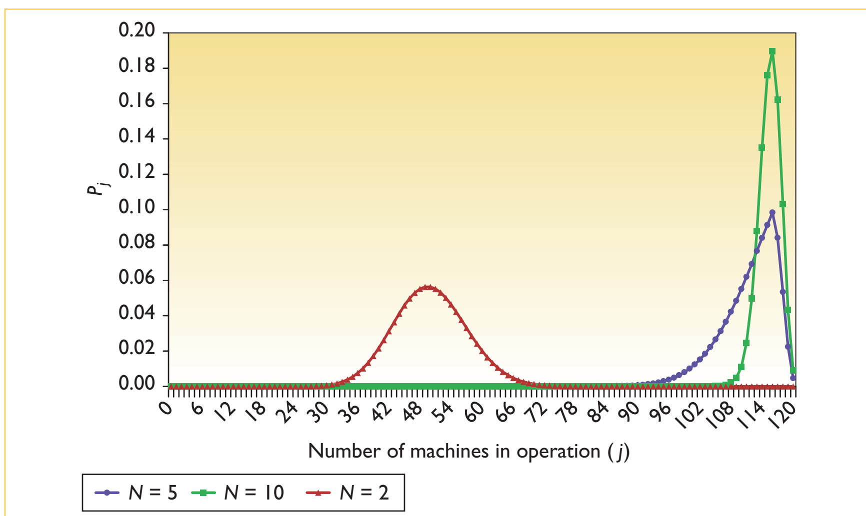

Figure 2. Probability distribution. The lines show the probability that j machines are in operation for three different values of the size N of the repair staff. As N increases, the probability distribution peak moves toward the total number of machines.

N machines are being repaired, so the overall repair rate is N/F.

Using the Performance and Availability Model

Let's look at an IDC with 120 machines, each with a capacity to process 50 requests/sec. The IDC's maximum possible throughput is therefore 6,000 (or 120 × 50) requests/sec. Each of the 120 machines has an MTTF equal to 500 minutes, and each member of the repair staff takes 20 minutes, on average, to repair a machine.

Figure 2 shows the distribution of the probability p_j that exactly j

Table 1. Average throughput of the IDC as a function of repair-staff size.

Size of repair staff	Average throughput (requests/sec)	Maximum throughput (percent)
2	2,500	41.7
5	5,575	92.9
10	5,769	96.1

machines are operational for three values (2, 5, and 10) of the number of people on the repair staff. As the figure shows, with a staff of two people, the probability distribution peaks for 50 machines at about 5.6 percent. Also, we can see that for $N = 2$, the probability of 70 or more machines being operational is negligible. When the repair staff's size increases to five people, the distribution peaks for 116 machines at a probability value close to 10 percent. If five more people join the repair staff, the distribution also peaks at 116 machines, but at a value close to 19 percent. The average number of machines in operation is equal to 50, 111.5, and 115.4 for $N = 2$, 5, and 10, respectively.

We can compute the IDC's average throughput by using Equation 1 and the probabilities shown in Figure 2. Table 1 shows these values for $N = 2$, 5, and 10. As the table indicates, a small repair staff ($N = 2$) yields a throughput of 2,500 requests/sec, which is only 41.7 percent of the maximum throughput of 6,000 requests/sec achievable if all machines are operational all the time. A five-person repair staff significantly improves the average throughput to 92.9 percent of the maximum theoretical throughput, but doubling the size of the repair staff at this point brings very little performance advantage and significantly increases maintenance costs. This is because, given the MTTR and average repair time F values, the probability that a repair person will be available to start working on a machine as soon as it fails is already very high when $N = 5$.

Final Remarks

The availability of the IDC we just considered is extremely close to 100 percent for any of the three values of N. However, as we just saw, the performance levels for these three configurations differ vastly. Moreover, the probability A_j — that at least j machines are in operation — varies significantly with N. The probability that at least 100 out of the 120 machines are in operation is virtually zero for $N = 2$, 0.965 for $N = 5$, and virtually one for $N = 10$, for example. These considerations indicate that availability alone is not a very meaningful metric unless it is coupled with the performance level a computer system provides.

References

1. D.A. Menascé, V.A.F. Almeida, and L.W. Dowdy, *Performance by Design: Computer Capacity Planning by Example*, Prentice Hall, 2004.
2. L. Kleinrock, *Queuing Systems, Vol. I: Theory*, John Wiley & Sons, 1995.

Daniel A. Menascé is a professor of computer science, the codirector of the E-Center for E-Business, and the director of the MS in E-Commerce program at George Mason University. He received a PhD in computer science from UCLA and published the books *Performance by Design*, *Capacity Planning for Web Services*, and *Scaling for E-Business* (Prentice Hall, 2004, 2002, and 2000). He is a fellow of the ACM and a recipient of the 2001 A.A. Michelson Award from the Computer Measurement Group. Contact him at menasce@cs.gmu.edu.