

Association for Information Systems AIS Electronic Library (AISeL)

PACIS 1993 Proceedings

Pacific Asia Conference on Information Systems
(PACIS)

December 1993

Optimal Capacity of a Firm's Computer Backup Center

Hsing Cheng
College of William and Mary

Follow this and additional works at: <http://aisel.aisnet.org/pacis1993>

Recommended Citation

Cheng, Hsing, "Optimal Capacity of a Firm's Computer Backup Center" (1993). *PACIS 1993 Proceedings*. 51.
<http://aisel.aisnet.org/pacis1993/51>

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 1993 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

OPTIMAL CAPACITY OF A FIRM'S COMPUTER BACKUP CENTER

Hsing K. Cheng, School of Business Administration, The College of William and Mary, Williamsburg, Virginia 23187-8795, Phone (804) 221-2879 Bitnet: hcheng@vmvm1

ABSTRACT

Due to great dependence on computers, today's business firms will suffer tremendous losses when their computer systems break down. A recent report shows that computer downtime cost U.S. business \$4 billion in 1991, primarily through lost productivity and lost revenues. Having a computer backup center (the hot site) immediately ready to take care of the information processing needs is the best solution for those industries for which continuous functioning of computers is critical. If a firm decides to install a computer backup center, a critical concern to the top management is to determine the optimal capacity of the backup center that minimizes the total cost to the firm as a whole. This paper develops analytical models and solution procedure for finding the optimal capacity of a firm's computer backup center.

Keywords: Computer breakdowns, Computer backup center, Optimal backup capacity.

INTRODUCTION

Due to great dependence on computers, today's business firms will suffer tremendous losses when their mainframe computer systems break down. According to a 1978 study by the University of Minnesota, the normal business activity of a typical financial institution will decline to 87% with a one-half day data processing center failure. A ten-day computer outage will entail a 97% loss of business activity. The same study shows that the average maximum computer downtime allowed for all industries is 4.8 days (Aasgard, et al., 1978). Although the University of Minnesota study is more than a decade old, it can be used as a lower bound estimate for the cost of computer breakdowns in today's computing environment.

A recent report shows that computer downtime cost U.S. business \$4 billion in 1991, primarily through lost productivity and lost revenues (Ballou, 1992). In the same report, mainframe computers are found to break down nine times per year in the average company. Computer downtime becomes even more of an issue for those companies that depend heavily on computers. As the vice president of an investment firm put it, "Downtime is one of those things in the financial industry that just cannot be tolerated." (Ballou, 1992).

Among a number of alternatives available to manage the computer downtime, having a computer backup center (the hot site) immediately ready to take care of the information processing needs is the best solution for those industries for which continuous functioning of computers is critical (*Data Management*, 1980). Banking is a typical example of an industry highly vulnerable to the failures of computer systems. If a firm decides to install a computer backup center, a critical question to the top management arises; what will be the optimal capacity of the backup center that minimizes the total cost to the firm as a whole? The purpose of this paper is to address this critical question. Analytical models and solution procedure are developed for finding the optimal capacity of a firm's computer backup center.

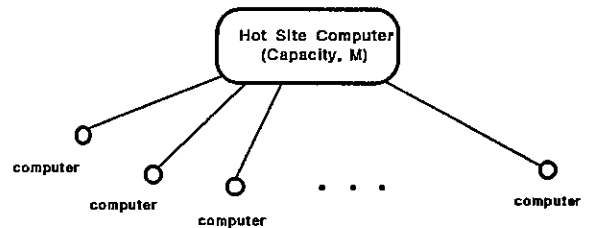


Figure 1. Schematic Diagram of the Computer Backup Center

THE MODEL

Assume that a firm under consideration has N identical users' computers to backup. The backup center, normally termed the hot site, is installed with a computer which can replace up to M users' computers at the same time. The schematic diagram of the computer backup center is shown in Figure 1. The objective of the firm is to determine the optimal capacity M^* that minimizes the total cost to the firm as a whole. The total costs include two components: the total opportunity costs of lost computing time even with a backup center, and the total capacity cost of the backup center.

To model the breakdown behavior of users' computers to be backed up, the computer systems are formulated as a two-state Markov chain with up (operational) and down states where the up time parameter is λ and down time parameter is μ . The N users' computers are identical and independent in the sense that they have the same up time and down time parameters but break down independent of each other. Since the hot site computer could also break down, the up time and down time of the hot site computer are assumed to be mutually independent and exponentially distributed with parameters of η and γ respectively. The hot site computer can be conceivably assumed to be more reliable than the users' computers. The service discipline of the hot site computer is First Come First Serve (FCFS).

The value of lost computing time to the i -th user is G_i in dollars per unit of time. Assume that the marginal capacity cost for the hot site computer is a constant c where the marginal capacity

Table 1 Summary of Notation

N :	no. of user's computers to backup
M :	capacity of the computer backup center, the hot site, in terms of the number of active users allowed at the same time
λ :	the failure rate of the user's computer systems
μ :	the repair rate of the user's computer systems
γ :	the repair rate of the hot site computer
η :	the failure rate of the hot site computer
W :	the expected time without replacement service from the hot site when user's computer systems break down
c :	the marginal capacity cost for the hot site computer
G_i :	the value of lost computing time for the i -th user

cost equals the cost per unit of time to increase the hot site capacity for replacing *one more* user's computer. The constant c reflects constant economies of scale of computing. Mendelson (1987) shows that the average cost per unit of capacity is constant, suggesting that the use of constant marginal capacity cost is appropriate. Let W be the part of user's computer down time for which no replacement is available from the backup center either because the hot site capacity has been exhausted by other users already down or because the hot site computer is also down. The notation used in this model is summarized in Table 1.

The dynamic interaction between individuals' computers and the hot site computer is very complicated. For instance, when a user's computer is down and needs replacement service from the hot site, more than M other computers could have been down already. Therefore, the request has to be placed on the queue. There are two ways of leaving the queue: either the individual user's computer is repaired before it receives replacement service from the hot site, or it finishes waiting and starts receiving replacement service from the hot site. The hot site computer could also break down when users are receiving replacement service. This complicated queuing phenomenon has been analyzed in Freimer, Sumita, and Cheng (1992). Given that the capacity of the hot site is M , the expected time without replacement service when a user's computer is down, $E[W|M]$, is readily available from Freimer, Sumita, and Cheng (1992). It is repeated here as Theorem 2.1.

Theorem 2.1 (Expected time without replacement service from the hot site)

$$(2.1) \quad E[W|M] = \sum_{n=0}^{N-1} e_n^{N-1} E[W_n|M]$$

where

$$(2.2) \quad e_n^{N-1} = C_n^{N-1} \left(\frac{\lambda}{\lambda + \mu} \right)^n \left(\frac{\mu}{\lambda + \mu} \right)^{N-1-n}$$

and

$$(2.3) \quad E[W_n|M] = \begin{cases} \frac{n-M+1}{(n+1)\mu} + \frac{M}{n+1} \left(\frac{\eta}{(\gamma+\eta)\mu} \right) & \text{for } M \leq n \leq N-1 \\ \frac{\eta}{(\gamma+\eta)\mu} & \text{for } n < M. \end{cases}$$

The total cost to the organization as a whole consists of two components, the first part being the total expected opportunity cost of the lost computing time across all users, and the second being the total capacity cost of the hot site computer. Let $g(M)$ be the total cost function given that the capacity of the hot site is M . Then one has

$$(2.4) \quad g(M) = \left(\frac{\lambda\mu}{\lambda + \mu} \right) \sum_{i=1}^N G_i E[W|M] + c \cdot M.$$

It is noted that the first term in Equation (2.4) equals the total expected costs of the lost computing time and the second term denotes the total capacity cost of the hot site. The total expected costs of the lost computing time equals the breakdown rate of

users' computers, $\frac{\lambda\mu}{\lambda + \mu}$, multiplied by the expected costs per

breakdown across all users. Both the total expected costs of the lost computing time and the total capacity cost are measured in dollars per unit of time. The following theorem shows that the total cost function $g(M)$ is a convex function of M .

Theorem 2.2

The total cost function $g(M)$ is convex in the sense that $\Delta^2 g(M) > 0$, where $\Delta g(M) = g(M+1) - g(M)$.

Proof:

Define $G = \left(\frac{\lambda\mu}{\lambda + \mu} \right) \sum_{i=1}^N G_i$. Let $E[Y|M]$ be the expected

time the user receives the replacement service from the hot site when the user's computer is down. Since the sum of $E[W|M]$ and $E[Y|M]$ equals the expected breakdown duration of the user's computer, i.e., $E[W|M] + E[Y|M] = 1/\mu$, one can rewrite $g(M)$ as

$$(2.5) \quad g(M) = G \cdot \left(\frac{1}{\mu} - E[Y|M] \right) + c \cdot M.$$

From Theorem 2.1 and after some algebra, one has

$$(2.6) \quad E[Y|M] = \frac{1}{\mu} \frac{\gamma}{\gamma + \eta} \sum_{n=0}^{M-1} e_n^{N-1} + M \sum_{n=M}^{N-1} \frac{1}{n+1} e_n^{N-1}.$$

Substituting (2.6) into (2.5) leads to

$$(2.7) \quad g(M) = G \cdot \left\{ \frac{1}{\mu} - \frac{1}{\mu} \frac{\gamma}{\gamma + \eta} \left(\sum_{n=0}^{M-1} e_n^{N-1} + M \sum_{n=M}^{N-1} \frac{1}{n+1} e_n^{N-1} \right) \right\} + c \cdot M.$$

Define $\Delta g(M) = g(M+1) - g(M)$. By noting the fact that

$$(2.8) \quad \sum_{n=0}^M e_n^{N-1} = \sum_{n=0}^{M-1} e_n^{N-1} + (M+1) \frac{1}{(M+1)} e_M^{N-1},$$

one has

$$(2.9) \quad \Delta g(M) = -G \frac{1}{\mu} \frac{\gamma}{\gamma + \eta} \sum_{n=M}^{N-1} \frac{1}{n+1} e_n^{N-1} + c.$$

Since the first term of (2.9) is an increasing function of M , $\Delta^2 g(M) > 0$ and the theorem follows.

Theorem 2.2 shows that the total cost function $g(M)$ is a discrete convex function of the hot site capacity M . Hence, there exists an optimal capacity M^* such that the total cost to the firm is minimized. A numerical example presented in the next section demonstrates how to find the optimal capacity of the computer backup center.

A NUMERICAL EXAMPLE

The actual optimal capacity M^* depends on the specification of parameters of $\lambda, \mu, \gamma, \eta, G_i$'s and c . Suppose a firm has the following computer configurations.

- $\eta = 1/240$ (day⁻¹)
- $\gamma = 1/2$ (day⁻¹)
- $\lambda = 1/180$ (day⁻¹)
- $\mu = 1/3$ (day⁻¹)
- $N = 80$ users' computers
- $c = 1.19$ in thousands per day
- $\sum_{i=1}^N G_i$: total opportunity costs across all users
- are \$450,000 per day.

In order to determine the minimum total costs, $g(M)$ is plotted against the capacity of the hot site, M . The behavior of total costs, $g(M)$, is shown in Figure 2. As predicted by Theorem 2.2, $g(M)$ is a discrete convex function of M . From the figure, one knows that the total cost is minimized when the hot site replaces two users' computers at the same time, i.e., $M^* = 2$.

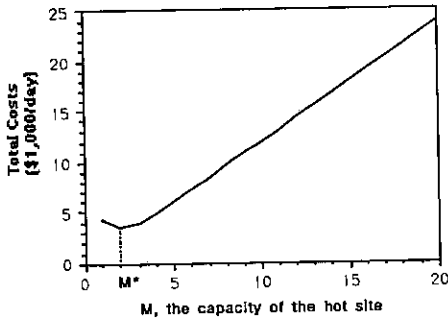


Figure 2. The total cost function, $g(M)$, vs. M

In fact, Figure 2 shows the trade-off between the two components of the total costs. After a closer examination of Figure 2, one finds that it is a straight line for $M \gg 3$ portion of the graph and the slope of the line equals 1.19. This implies that expanding the hot site capacity all the way for $M \gg 3$ will not reduce the total expected costs of users' lost computing time at all. The total expected costs of users' lost computing time reaches the asymptotic value and the increase in total cost function $g(M)$ is solely due to the marginal capacity cost of the hot site computer for the range $M \gg 3$. This fact can be verified when $E[W|M]$ is plotted against M in Figure 3.

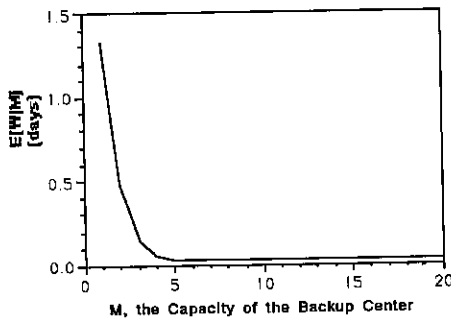


Figure 3. The time without replacement, $E[W|M]$, vs. M

EXTENSION OF THE MODEL TO HETEROGENEOUS COMPUTERS

In the previous sections, all users' computers are assumed to be identical in that they have the same reliability parameters, λ 's and μ 's. The model in Section 2 would be more useful if it can be extended to deal with the case where users' computers are heterogeneous in terms of different reliability parameters. The purpose of this section is to extend the previous model to allow users' computers to have different uptime and downtime parameters, thereby incorporating the previous model as a special case.

Assume that a firm under consideration wants to install a computer backup center to backup N different users' computers. Let S be the set of indices of all users' computers and $S = \{1, 2, \dots, N\}$. The i -th user's computer is formulated as a two-state Markov chain with the uptime parameter λ_i and down time parameter μ_i , $i \in S$. The λ_i and μ_i can be different from λ_j and μ_j when $i \neq j$. The failures and repairs of users' computers are independent of each other and of the hot site computer. The up time and down time of the hot site computer are assumed as before to be mutually independent and exponentially distributed with parameters of η and γ respectively. The hot site computer can replace M users' computers at the same time. The firm wants to find the optimal capacity M^* of the hot site so that the total costs are minimized.

Since users' computers have different uptime and downtime parameters, the expected time without replacement service from the hot site is no longer the same for every user. The expected time without replacement service has to be determined from each user's standpoint. Let k be the index of the computer under consideration, $k \in S$, and define $E[W^k|M]$ as the expected time without replacement service for the k -th user given that the hot site's capacity is M .

As a first step in determining $E[W^k|M]$, we analyze a Markov chain describing the k -th user in the queue waiting for replacement service when the number of other users' computers already down is more than the hot site can handle. Again, the k -th user leaves the queue either he obtains access to the hot site or his computer is repaired. Hence, we consider a Markov chain $N(t)$ on $\mathcal{N} = G \cup B$, where $G = \{n: M \leq n \leq N-1\}$ and $B = \{s_1, s_2\}$. The Markov chain $N(t)$ has two absorbing states, s_1 for replacement and s_2 for repair. The interpretation of the Markov chain $N(t)$ is that when the k -th user's computer fails at time $t=0$, $N(t) = n$ means he is still waiting for the replacement service and there are n other users ahead of him in the queue.

Let T_{nB}^k be the first passage time of $N(t)$ from $n \in G$ to the set B , and let D_n^k be the absorbing state reached. That is,

$$(4.1) \quad T_{nB}^k = \inf\{t: N(t) \in B \text{ given that } N(0)=n \in G\}$$

and

$$(4.2) \quad D_n^k = N(\infty) \text{ given } N(0)=n \in G.$$

Hence, $E[T_{nB}^k]$ equals the expected time spent in the queue for the k -th user waiting for replacement service from the hot site and $P\{D_n^k = s_1\}$ is the probability to receiving replacement service when the waiting in the queue ends.

Define S_n as the set of the indices of all other computers that are already down. Hence $|S_n| = n$ and $k \notin S_n$. Let $S_{n-1} = S_n \setminus \{i_n\}$, where $i_n \in S_n$; $S_{n-2} = S_{n-1} \setminus \{i_{n-1}\}$, where $i_{n-1} \in S_{n-1}$; ...; and $S_M = S_{M+1} \setminus \{i_{M+1}\}$, where $i_{M+1} \in S_{M+1}$. It is clear that the set G can be partitioned as

$$(4.3) \quad G = \bigcup_{j=M}^n G_j$$

where

$$(4.4) \quad G_j = \{<g_{i_n}> | g_{i_n} = 0 \text{ or } 1, i_n \in S_n, \text{ and } \sum_{i_n} g_{i_n} = j\},$$

and $<g_{i_n}>$ is an n -dimension vector.

Then, one has the following theorems.

Theorem 4.1 (Expected time to absorption)

$$E[T_{nB}^k] = \frac{1}{\mu_k + \sum_{i_k \in S_n} \mu_{i_k}} + \sum_{i_k \in S_n} \frac{\mu_{i_k}}{\mu_k + \sum_{i_k \in S_n} \mu_{i_k}} \left(\frac{1}{\mu_k + \sum_{i_{n-1} \in S_{n-1}} \mu_{i_{n-1}}} + \sum_{i_{n-1} \in S_{n-1}} \frac{\mu_{i_{n-1}}}{\mu_k + \sum_{i_{n-1} \in S_{n-1}} \mu_{i_{n-1}}} \left[\frac{1}{\mu_k + \sum_{i_{n-2} \in S_{n-2}} \mu_{i_{n-2}}} + \dots \right] \right) + \sum_{i_{M+2} \in S_{M+2}} \frac{\mu_{i_{M+2}}}{\mu_k + \sum_{i_{M+2} \in S_{M+2}} \mu_{i_{M+2}}} \left(\frac{1}{\mu_k + \sum_{i_{M+1} \in S_{M+1}} \mu_{i_{M+1}}} + \sum_{i_{M+1} \in S_{M+1}} \frac{\mu_{i_{M+1}}}{\mu_k + \sum_{i_{M+1} \in S_{M+1}} \mu_{i_{M+1}}} \left[\frac{1}{\mu_k + \sum_{i_M \in S_M} \mu_{i_M}} + \dots \right] \right)$$

Proof:

When $n = M$, the expected time to absorption of the chain $N(t)$ equals the inverse of $\mu_k + \sum_{i_M \in S_M} \mu_{i_M}$, and the theorem holds true.

Suppose the theorem holds true for the case $n > M$. Then, $E[T_{n+1,B}^k]$ consists of two components. The first is the expected time of the chain $N(t)$ to make the first transition, which equals the inverse of $\mu_k + \sum_{i_{n+1} \in S_{n+1}} \mu_{i_{n+1}}$. Since the expected time to absorption

from state s_2 is zero, the second component of $E[T_{n+1,B}^k]$ equals the sum of the probability of reaching a state in G_n multiplied by the expected time to absorption from that state. That is,

$$(4.5) \quad E[T_{n+1,B}^k] = \frac{1}{\mu_k + \sum_{i_{n+1} \in S_{n+1}} \mu_{i_{n+1}}} + \sum_{i_{n+1} \in S_{n+1}} \frac{\mu_{i_{n+1}}}{\mu_k + \sum_{i_{n+1} \in S_{n+1}} \mu_{i_{n+1}}} E[T_{n,B}^k]$$

The mathematical induction proceeds and the theorem follows.

Theorem 4.2

$$P[D_n^k = s_1] = \sum_{i_1 \in S_0} \frac{\mu_{i_1}}{\mu_k + \sum_{i_1 \in S_0} \mu_{i_1}} \left(\sum_{i_{n-1} \in S_{n-1}} \frac{\mu_{i_{n-1}}}{\mu_k + \sum_{i_{n-1} \in S_{n-1}} \mu_{i_{n-1}}} \left(\dots \sum_{i_{n-2} \in S_{n-2}} \frac{\mu_{i_{n-2}}}{\mu_k + \sum_{i_{n-2} \in S_{n-2}} \mu_{i_{n-2}}} \left(\dots \sum_{i_k \in S_M} \frac{\mu_{i_k}}{\mu_k + \sum_{i_k \in S_M} \mu_{i_k}} \right) \right) \right)$$

Proof: The theorem is easily seen from mathematical induction.

The expected time without replacement given that there are n other computers already down, $E[W_n^k|M]$, consists of two parts. The first part is the expected time, if any, spent in the queue as described in Theorem 4.1. The second part is the expected cumulative down time of the hot site computer while the k -th user is receiving replacement service, which equals $\frac{\eta}{(\gamma + \eta)\mu_k}$ from the analysis in Freimer, Sumita, and Cheng (1992). The following theorem is thus immediate.

Theorem 4.3

$$(4.6) \quad E[W_n^k|M] = \begin{cases} E[T_{n,B}^k] + P[D_n^k = s_1] \frac{\eta}{(\gamma + \eta)\mu_k} & M \leq n \leq N-1, \\ \frac{\eta}{(\gamma + \eta)\mu_k} & n < M. \end{cases}$$

To complete the specification of the expected time without replacement service for the k -th user, let S be the set of the indices of all users' computers and define the generating function $K(z)$ as in the following

$$(4.7) \quad K(z) = \prod_{i \in S \setminus \{k\}} (d_i z + u_i)$$

$$\text{where } d_i = \frac{\lambda_i}{\lambda_i + \mu_i} \text{ and } u_i = \frac{\mu_i}{\lambda_i + \mu_i}$$

Then $K(z)$ can be rearranged as

$$(4.8) \quad K(z) = \sum_{n=0}^{N-1} p_n z^n$$

where p_n equals the ergodic probability that n other computers are down.

Let I_j be an indicator function, where $j \in S \setminus \{k\}$, and define

$$(4.9) \quad p(I_j) = \begin{cases} d_j & \text{if } I_j = 1, \\ u_j & \text{if } I_j = 0. \end{cases}$$

Then, one sees that

$$(4.10) \quad p_n = \sum_{I: \sum_{i \in S \setminus \{k\}} I_i = n} \prod_{i \in S \setminus \{k\}} p(I_i)$$

The expected time without replacement service for the k -th user, $E[W^k|M]$, is thus derived as

$$(4.11) \quad E[W^k|M] = \sum_{n=0}^{N-1} \sum_{I: \sum_{i \in S \setminus \{k\}} I_i = n} \left[\left(\prod_{i \in S \setminus \{k\}} p(I_i) \right) E[W_n^k|M] \right]$$

where the $E[W_n^k|M]$ in (4.11) is computed according to Theorem 4.3, and the corresponding S_n 's in Theorems 4.1 and 4.2 are the set of indices j where $j \in S_N$ and $I_j = 1$.

Hence, the total cost function $g(M)$ is specified by the following equation

$$(4.12) \quad g(M) = \sum_{k=1}^N \left(\frac{G_k \lambda_k \mu_k}{\lambda_k + \mu_k} \right) E[W^k|M] + c \cdot M,$$

where G_k is the opportunity cost of lost computing time for the k -th user in dollars per unit of time. The first term in Equation (4.12) amounts to the total opportunity costs of lost computing time summed across all users and the second term equals the total capacity cost of the hot site. One cannot derive a theorem showing the convexity of $g(M)$ as in Theorem 2.2 for identical users case because of the very complicated form of $g(M)$ in (4.12). The behavior of $g(M)$ will be explored numerically in the next section.

Numerical Exploration of the Heterogeneous Computers Model

The computation complexity of the total cost function $g(M)$ for the heterogeneous computers case is enormous. It is first noted that the complexity of evaluating the expected time without replacement service for the k -th user, $E[W^k|M]$, in Equation (4.11) is of order $O(2^{N-1})$. When the number of other computers already down is n , one has to consider all C_n^{N-1} possible combinations of the breakdown patterns of other computers and calculate the $(\prod_{i \in S \setminus \{k\}} p(I_i)) E[W_n^k|M]$ accordingly. Therefore, the complexity of the algorithm for evaluating Equation (4.11) is of order $O(C_0^{N-1} + C_1^{N-1} + \dots + C_{N-1}^{N-1}) = O(2^{N-1})$. Thus, the computation complexity for finding the total cost function $g(M)$ is of order $O(N \cdot 2^{N-1})$.

It is impractical to calculate the total cost function $g(M)$ in Equation (4.12) for large N because of the combinatorial explosion of the complexity. However, the behavior of $g(M)$ is explored through the following numerical example. Assume that the firm wants to install a computer backup center to backup seven departmental computers. The physical parameters of the hot site computer is still the same as the ones used in Section 3, i.e., $\eta = 1/240$ (day⁻¹) and $\gamma = 1/2$ (day⁻¹). The physical parameters of the departmental computers and opportunity cost of the lost computing time for each department are shown in Table 2.

Table 2. Physical parameters of the departmental computers

Department k	λ_k (day ⁻¹)	μ_k (day ⁻¹)	G_k (\$1,000/day)
1	1/60	1/1.8	220
2	1/80	1/2.0	240
3	1/100	1/2.2	260
4	1/120	1/2.4	280
5	1/140	1/2.6	300
6	1/160	1/2.8	320
7	1/180	1/3.0	240

The firm's objective is to find the optimal capacity for the hot site such that the total costs are minimized. Figure 4 shows the graph of $g(M)$ as a function of M , the hot site capacity. Once again, the convexity of the total cost function is observed and the optimal backup capacity equals two. The tail of the graph still shows a straight line with the slope equal to 1.19, the marginal capacity cost of the hot site. The trade-off between the total capacity cost and the total opportunity costs of the lost computing time across all departments is still evident by observing the behavior of $E[W^k|M]$ as shown in Figure 5.

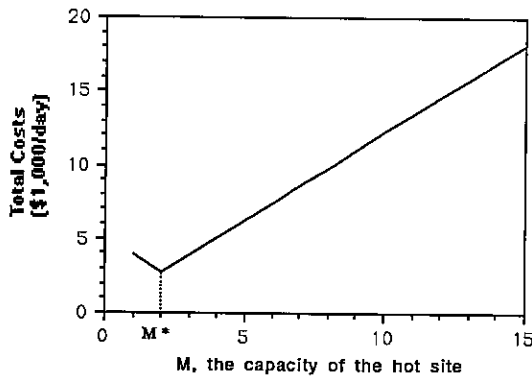


Figure 4. Total cost function (heterogeneous computers) vs. Hot Site Capacity, M

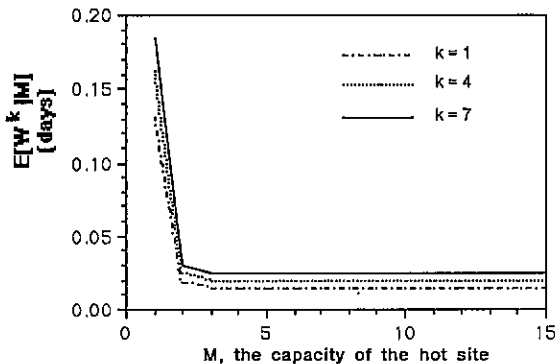


Figure 5. $E[W^k|M]$ vs. M graph

REFERENCES

Aasgard, D. O. et al., "An Evaluation of Data Processing 'Machine Room' Loss and Recovery Strategies," MISRC Working Papers, University of Minnesota, Minneapolis, MN, 1978.

Ballou, Melinda-Carol, "Survey Pegs Computer Downtime Costs at \$4 Billion," *Computerworld*, August 10, 1992, pp. 53 and 56.

Bronson, Gail, "Blue-Chip Backup", *Forbes*, January 26, 1987, pp. 80 - 81.

Business Week, "Another Crunch for Computer Makers", January 8, 1990, pp. 97.

Cheng, Hsing K., "Optimal Internal Pricing and Backup Capacity of Computer Systems Subject to Breakdowns," Submitted for publication.

Data Management, "Developing A Contingency Plan", Vol. 18, 1, 1980, pp. 10-12, p 14, and p 48.

Data Management, "Disaster Recovery Operations Approach", Vol. 18, 1, 1980, pp. 18-19.

Freimer, Marshall, Ushio Sumita, and Hsing K. Cheng, "Analysis of Economics of Computer Backup Service," *IEICE Trans. Commun.*, Vol. E75-B, No. 5, May 1992, pp. 385-400.

Gray, J. N., "Why do computers stop and what can be done about it?," Presentation at the Fifth Symposium on Reliability in Distributed Software and Database Systems, Los Angeles, California, January, 1986.

Haerder Theo, and Andreas Reuter, "Principles of Transaction-Oriented Database Recovery," *ACM Computing Surveys*, Vol. 15, No. 4, December 1983, pp. 287-317.

Lucisano, Rocco, Manager of Sales and Marketing, Agway Data Service Inc., personal communication, March, 1992.

Mason, Janet (1990), "Hot Sites Turn Up the Heat on America", *Computerworld*, April 23, 1990, pp. 88-89.

Mendelson, Haim, "Economics of Scale in Computing: Grosch's Law Revisited," *Communications of the ACM*, Vol. 30, No. 12, 1987, pp. 1066-1072.

New York Times, "Just Waiting for a Great Big Blackout," November 7, 1990, pp. D-1 and D-4.

Sullivan, W. M. (1980), "A 'First' Recovery Site Corporation Evolves", *Data Management*, Vol. 18, 1, 1980, pp. 15-16, and p 28.

Toigo, Jon William, *Disaster Recovery Planning - Managing Risk & Catastrophe in Information Systems*, Yourdon Press, 1989.

Wallace John J. and Walter W. Barnes, "Designing for Ultrahigh Availability: The Unix RTR Operating System," *IEEE Computer*, August 1984, pp. 31 - 39.

Wass, W. and J. S. Keen, "Disaster's Top Ten," *DG Review*, Vol. 9, No. 6, Dec. 1988, pp. 43 - 50.