

# Supervisory Control of a Database Unit

N. Eva Wu, *Senior Member, IEEE*, James M. Metzler, and Mark H. Linderman, *Member, IEEE*

**Abstract** — To effectively enhance service availability, this paper proposes a redundancy configuration for a database unit residing in a command and control (C2) system that supports air operations. The results of modeling, supervisory control, and performance analysis of the database unit are presented. The unit is modeled as a closed Markovian queuing network. State variable feedback is used to implement the functions of restoration and routing upon the identification of the failure of one of the database servers in the unit. Several control policies are evaluated in terms of the resulting mean time to unit failure, the steady state availability, the expected response time, and the service overhead of the database unit.

## I. INTRODUCTION

THE recent effort to install and test monitoring tools and to increase the level of redundancy in critical subsystems in air operation centers [1] has provided opportunities for vast performance improvement in its command and control (C2 hereafter) supporting systems. Our previous work on a controlled C2 processing unit [2] has demonstrated that reduced response time to service requests and shortened periods of system unavailability, as a result of automated monitoring and control, can raise significantly the probability to attain the desired outcome in an air operation. This paper shifts focus to one other critical C2 subsystem, a database unit. A simulation study [3] has been performed recently using Arena [4], [5] on a controlled database unit. The results indicate, however, that the architecture shown in Fig.1 is extremely inefficient, where the service burden rests almost entirely on the primary server, while the secondary server, though indispensable for the required system availability, is rarely utilized.

Fig.2 shows an alternative architecture for which the potential improvements in response time and in service availability are to be examined. The partition of the database into multiple sets of data (to be called data classes hereafter), and the simultaneous access to multiple servers allow the reduction of the response time to queries, whereas the presence of a secondary data class in every server leads to fault-tolerance and therefore higher service availability. The

performance improvement, however, cannot be achieved in a cost-effective manner without a reconfiguration scheme called a supervisory control that acts on the state information of the database system. This effort investigates several such schemes that differ by their control authorities. To assess the effectiveness of these schemes in a quantified manner, the model in Fig.2 (and that in Fig.1) is given the interpretation of a queuing network [6] with specific sets of operating policies and structural parameters. The control authorities considered include the ability to restore the lost data and/or the ability to route queries. In order to obtain an analytic model of manageable size for scrutinizing the effects of supervisory control, the archiving process is ignored, and the queuing network is of the closed type [7]. A simulation study is being conducted currently without these simplifications.

The paper is organized as follows. Section II of the paper models the database system in Fig.2 as a Markov chain [8] with supervisory control. Section III evaluates a set of performance measures under several supervisory control policies. Section IV concludes the paper. Section V acknowledges the contributions from our colleagues. Details of the database model are given in Appendix.

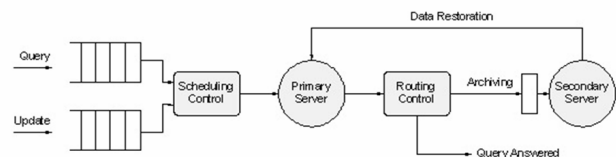


Fig.1 Redundant database unit

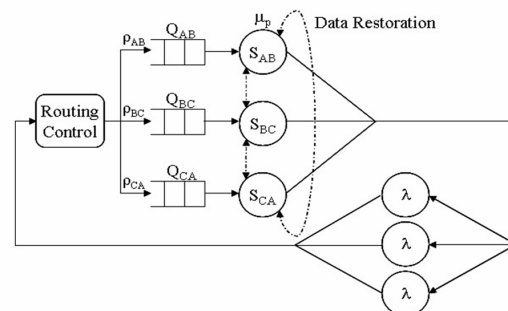


Fig.2 Partitioned database unit

## II. MODELING AND CONTROL

### A. Modeling

The database unit in Fig.2 contains three servers in parallel to answer three classes (A, B, C) of queries for which relevant information can be found in the partitioned

Manuscript received March 1, 2005. This work was supported in part by the U.S. Air Force Research Laboratory Contract F30602-020-C-0225.

N. Eva Wu is with the Department of Electrical and Computer Engineering, Binghamton University, Binghamton, New York, 13902-6000, USA. (telephone: 607-777-4375; fax: 607-777-4464; e-mail: [evawu@binghamton.edu](mailto:evawu@binghamton.edu))

James M. Metzler is with the Department of Electrical and Computer Engineering, Binghamton University, Binghamton, NY 13902-6000, USA. (e-mail: [james.metzler@binghamton.edu](mailto:james.metzler@binghamton.edu))

Mark H. Linderman is with the Information Directorate of the Air Force Research Laboratories, Rome Research Site, Rome, NY, 13441-4505, USA. (e-mail: [mark.linderman@rl.af.mil](mailto:mark.linderman@rl.af.mil))

sets A, B, C of the database, respectively. Server  $S_{AB}$  contains database class A as the primary class and database class B as the secondary class. Server  $S_{BC}$  contains database class B as the primary class and database class C as the secondary class. Server  $S_{CA}$  contains database class C as the primary class and database class A as the secondary class. The failure of a server implies the loss of two classes of data within the server. A system level failure is declared when two servers fail, in which case one class of data is said to be lost. The queues preceding servers  $S_{AB}$ ,  $S_{BC}$ , and  $S_{CA}$  are named  $Q_{AC}$ ,  $Q_{BC}$ , and  $Q_{CA}$ , respectively. All queues are of sufficient capacity. Service is provided on a FCFS basis at each server.

The three delay elements imply that there are always three customers present in the unit at any given time. A new query is generated at a delay element upon the completion of the service to a query at one of the servers. The delay elements are intended to be also reflective of the response time to the querying customers by other service nodes in the C2 supporting system, which are not explicitly modeled. Any new query is assumed to be equally likely to seek database class A or B or C. Therefore routing probabilities  $\rho_{AB}$ ,  $\rho_{BC}$ , and  $\rho_{CA}$  are assigned the same values under the normal operation condition.

The use of a queuing network model for the database is based on its suitability to involve control actions and our intention to capture their effects on the system performance. The model is built in this study with the premise that event life distributions have been established for the process of query generation ( $\exp(\lambda) \equiv 1 - e^{-\lambda t}$ ), the process of service completion ( $\exp(\mu)$ ), the process of server failure ( $\exp(\nu)$ ), the process of data restoration ( $\exp(\gamma)$ ), and the process of unit overhaul ( $\exp(a)$ ) when the failed database unit is repaired. All such processes are independent. Standard statistical methods that involve data collection, parameter estimation, and goodness of fit tests [9] exist for identifying event life distributions. Since all event lives are assumed to be exponentially distributed, the database unit can be conveniently modeled as a Markov chain specified by a state space  $\mathcal{X}$ , an initial state probability mass function (pmf)  $\pi_x(0)$ , and a set of state transition rates  $A$  [7], [8]. The reader uninterested in the details of model building can advance to the paragraph right above Equation (1).

### 1) State space $\mathcal{X}$

A state name is coded with a 6-digit number indicative of all queue lengths and server states in the unit. With some abuse of notations, a valid state representation is given by  $x = Q_{AB}Q_{BC}Q_{CA}S_{AB}S_{BC}S_{CA}$ , where queue length  $Q_{AB}$ ,  $Q_{BC}$ ,  $Q_{CA} \in \{0, 1, 2, 3\}$  with total length  $L \equiv Q_{AB} + Q_{BC} + Q_{CA} \leq 3$ , and server state  $S_{AB}$ ,  $S_{BC}$ ,  $S_{CA} \in \{0, 1, 2\}$ . Server state "2"  $\equiv$  data are lost in both the primary and the secondary classes in a server, "1"  $\equiv$  the data in the primary class have been restored and data in the secondary class have not been restored, and "0"  $\equiv$  data in both primary class and secondary class in a

server are intact. A server is said to be in the down state if it is either at state "1" or at state "2". For example, state 110020 indicates that server  $S_{AB}$  is up with one customer in its queue, server  $S_{BC}$  is down with both classes of data gone and one customer in its queue, and server  $S_{CA}$  is up and idle. Note that the queue length includes the customer being served. There are 540 valid states in the system. The total number of states is reduced to 147 when the states of system level failures are aggregated. The symmetry of the system permits the arrangement of customers in the queues at the time of system level failure to be captured in one of seven states, allowing the system to return to an equivalent state upon completion of the system overhaul. A set of alternative state names are assigned from  $\mathcal{X} = \{1, 2, \dots, 147\}$  with 000000 mapped to  $x = 1$  and the aggregated system failure states mapped to  $x \in \{141, 142, 143, 144, 145, 146, 147\}$ .

### 2) Initial state pmf $\{\pi_x(0), x = 1, 2, \dots, 147\}$

It is assumed that the database unit starts operation from state  $x = 1$ , i.e., the initial state probability is given by vector  $\pi(0) = [1 \ 0 \ \dots \ 0]$ . When overhaul is considered at the occurrence of a system level failure, the system returns to a state with an equivalent arrangement of customers in the queues once the database unit is renewed [8] and ready for operation again.

### 3) Set of state transition rates $A$

A transition rate table containing all transition rates is created following a similar procedure as that described in [10], however with a more compact representation. The state transition table is given in Appendix. The list of current states occupies the first column of the table. In the row corresponding to each state, the set of all feasible next states are listed with each next state followed by the rate at which the next state is reached. Events that trigger the transitions and the corresponding transition rates are given as follows. A newly generated query enters one of the servers with rate  $\rho^{u_2} (3-L) \times \lambda$  where  $\rho^{u_2}$  is a controlled routing probability by control variable  $u_2$ . A query is answered at a server with rate  $\mu$ . A complete data loss occurs at a server with rate  $\nu$ . Data in the primary data class of a server are restored with rate  $\gamma_p u_1$  where  $u_1$  authorizes whether to restore the lost data. Data in the secondary data class of a server are restored with rate  $\gamma_s u_1$ . Finally, the failed database unit is renewed with rate  $\omega u_3$ , where  $u_3$  decides whether to repair the failed system. All rates are relative, for their net effects depend on the time unit specified.

Let  $X \in \mathcal{X}$  denote the random state variable at time  $t$ . The set of state transition functions

$$P_{i,j}(t) \equiv P[X(t) = j | X(0) = i], i, j = 1, 2, \dots, 147 \quad (1)$$

for the continuous-time Markov chain can be solved from the forward Chapman-Kolmogorov equation [7]

$$\dot{P}(t) = P(t)Q, P(0) = I, P(t) = [P_{i,j}(t)], \quad (2)$$

where  $Q$  is called an infinitesimal generator or a rate transition matrix whose  $(i,j)^{\text{th}}$  entry is given by the rate associated with the transition from current state  $i$  to next

state  $j$  in the rate transition table. State probability mass function at time  $t$

$$\pi(t) = [\pi_1(t) \ \pi_2(t) \ \cdots \ \pi_{147}(t)], t \geq 0 \quad (3)$$

is computed by

$$\pi(t) = \pi(0)P(t). \quad (4)$$

At this point a Markov model for the database unit of Fig.2 has been established. The state probabilities are the basis for evaluating the performance of the database unit, which is conducted in Section III.

### B. Control policies

Our ultimate goal is to eliminate all single point failures, and to mitigate the effects of a single server failure on the performance of the database unit. Our approach is to base the supervisory control actions on the state information, which effectively alter the transition rates when loss of data occurs in a single server.

Taking into consideration the symmetry of the model, the control policy is described only for the case of a failed server  $S_{AB}$ . When routing control is effective, the routing probabilities are determined by the state of  $S_{AB}$  and by whether the lost data can be restored. Thus,  $\rho^{u_2} = \rho_{AB}(S_{AB}, u_1)$ ,  $\rho_{BC}(S_{AB}, u_1)$ ,  $\rho_{CA}(S_{AB}, u_1)$  with  $\rho_{AB} + \rho_{BC} + \rho_{CA} = 1$ . The control policies considered for this study are summarized as follows.

$$u_1 = \begin{cases} 0, & S_{AB} = 2, S_{BC} \text{ serves}, S_{CA} \text{ serves (no restoration)} \\ I, & \begin{cases} S_{AB} = 2, S_{BC} \text{ serves}, S_{CA} \text{ restores class A data} \\ S_{AB} = 1, S_{CA} \text{ serves}, S_{BC} \text{ restores class B data} \end{cases} \end{cases} \quad (5)$$

$$u_2 = \begin{cases} 0, & S_{AB} = 2, \rho_{AB} = \rho_{BC} = \rho_{CA} = \frac{1}{3} \\ I, & \begin{cases} S_{AB} = 2, \rho_{AB}(2, u_1), \rho_{BC}(2, u_1), \rho_{CA}(2, u_1) \\ S_{AB} = 1, \rho_{AB}(1, u_1), \rho_{BC}(1, u_1), \rho_{CA}(1, u_1) \end{cases} \end{cases} \quad (6)$$

Four sets of routing probabilities are shown in the following table as examples, where  $S_{BC}=0$  and  $S_{CA}=0$  are assumed.

Table 1 Examples of routing probabilities

$u_1$	$u_2$	$S_{AB}$	$\rho_{AB}$	$\rho_{BC}$	$\rho_{CA}$
0	1	2	0	1/2	1/2
1	0	2 (1)	1/3 (1/3)	1/3 (1/3)	1/3 (1/3)
1	1	2 (1)	0 (1/6)	2/3 (1/6)	1/3 (2/3)
1	1	2 (1)	0 (0)	1 (0)	0 (1)

The composition of  $u_1$  and  $u_2$  gives rise to four different control policies. The case of  $(u_1, u_2) = (0, 0)$  corresponds to the case of a single point failure, and is therefore not considered in the performance analysis. The control policies in the other three cases are named

- Policy 1:  $(u_1, u_2) = (0, 1)$  when a server is down,  
 Policy 2:  $(u_1, u_2) = (1, 0)$  when a server is down, (7)  
 Policy 3:  $(u_1, u_2) = (1, 1)$  when a server is down.

Note that policy 2 does not permit routing, whereas policy 1 does not permit restoring. As can be seen, policy 3 allows variations in the routing probabilities to the intact servers. A special consideration with the case  $u_1=0$  is the rerouting of

the customers who have arrived at a server before the server fails to the delay elements.

The presence of supervisory control in the transition rate table is seen via  $u_1, u_2, u_3, n_1 = 1-u_1, n_2 = 1-u_2$ , and  $n_3 = 1-u_3$ . The values of  $u_1, u_2, u_3$  represent specific control actions associated with data restoration, query routing, and unit overhaul, respectively.

## III. PERFORMANCE ANALYSIS

### A. Time to system failure

When  $u_3 = 0$ , the Markov chain model for the database unit contains seven absorbing states  $x \in \{141, 142, 143, 144, 145, 146, 147\}$  at which the chain remains forever once it is entered. These are the states of system level failure. The rest of the 140 states are transient states. Decompose the state probability vector

$$\pi(t) \equiv \underbrace{[\pi_\tau(t)]}_{1 \times 140} \quad \underbrace{[\pi_\alpha(t)]}_{1 \times 7}, \quad (8)$$

where vector  $\pi_\tau(t)$  contains the transient state probabilities, and  $\pi_\alpha(t)$  are the absorbing state probabilities. Decomposing the rate transition matrix  $Q$  and the state transition function matrix  $P(t)$  solved from (2) accordingly yields

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ 0 & 0 \end{bmatrix}, P(t) = \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ 0 & I \end{bmatrix}. \quad (9)$$

From (2), (4), and (9), it can be determined that the probability density function of time to system failure, or time to absorption, is given by

$$\dot{\pi}_\alpha(t) = \pi_\tau(0)P_{11}(t)Q_{12}, \pi_\alpha(0) = 0, \quad (10)$$

where

$$\pi_\tau(0) = [1 \ 0 \ \cdots], P_{11}(t) = e^{Q_{11}t}. \quad (11)$$

In addition, the mean time to failure of the database unit can be shown to be [8].

$$MTTF = -\pi_\tau(0)Q_{11}^{-1}1_\tau, 1_\tau = [1 \ \cdots \ 1]^T \quad (12)$$

Fig.3 below shows the dependence of mean time to failure of the database unit on the restoration rate.

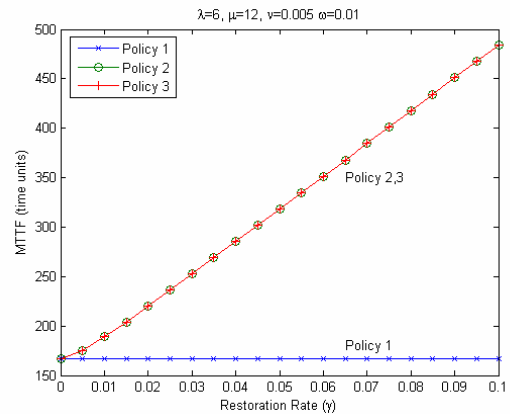


Fig.3 Database unit mean time to failure versus restoration rate

### B. Steady-state availability

Suppose as soon as the database unit reaches a system level failure, an overhaul process starts. Suppose with a rate

$\omega$  the unit is repaired, and at the completion of the repair, the unit immediately starts to operate again. In this case  $u_3$  is set to 1 in the model, whereas it is set to 0 in the case of an absorbing chain. The existence of a unique steady-state distribution of the Markov chain when  $u_3=1$  is guaranteed if the chain is irreducible (or ergodic) [7]. Ergodicity is satisfied under policy 2 and policy 3. Although ergodicity is not met under policy 1 without eliminating the few unreachable states in this case, a unique steady state distribution is obtained nevertheless in our computation. The steady state availability, which can be roughly thought of as the fraction of time the database unit is up, is given by

$$A_{sys} = I - \pi_F(\infty), \quad (13)$$

where  $\pi_F(\infty)$  is the sum of the system level failure state probabilities, determined by solving

$$\pi(\infty)Q = 0, \text{ and } \sum_{x=1}^{147} \pi_x(\infty) = I. \quad (14)$$

Fig.4 shows the steady-state availability as a function of restoration rate at a fixed overhaul rate. Fig.6 demonstrates the benefit of the success in supervisory control to steady-state availability.

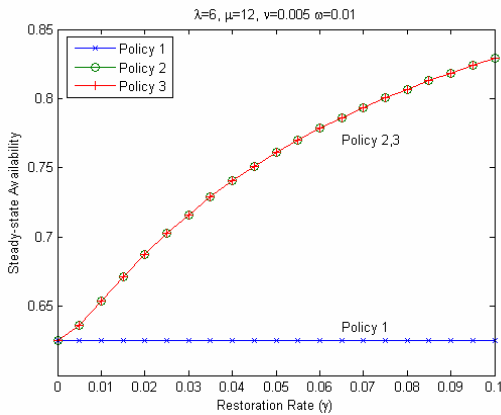


Fig.4 Steady-state availability of the database unit versus restoration rate

### C. Response time

The average response time  $E[R]$  is the expectation of the ratio of total amount of time that all customers spend in the upper portion of the system to the number of customers that are serviced. A loose argument is given below to justify the way  $E[R]$  is computed in this paper. Define the vector  $C$  where  $c(i)$  is the number of customers in the system at state  $i$ . The numerator of  $E[R]$  is then  $\pi(\infty)Ct$ . Computing the number of customers that are serviced requires counting the number of transitions from one state to another that have occurred that have introduced a new customer to the system. Define a matrix  $N$  such that  $n(i,j)$  is equal to the number of customers introduced into the system when the system transitions from state  $i$  to state  $j$ . The total number of transitions for a given  $i$  and  $j$  is then

$$T(i, j) = tN(i, j)\pi_i(\infty)Q(i, j). \quad (15)$$

Therefore, the average response time  $E[R]$  of the system is

taken as

$$\frac{\pi(\infty)Ct}{\sum_{i=1}^{147} \sum_{j=1}^{147} T(i, j)} = \frac{\pi(\infty)C}{\sum_{i=1}^{147} \sum_{j=1}^{147} N(i, j)\pi_i(\infty)Q(i, j)}. \quad (16)$$

Fig.5a and Fig.6 show the average response time as a function of restoration rate with the overhaul rate fixed, and a function of overhaul rate with the restoration rate fixed, respectively, for all three policies. The routing probabilities in rows 1 through 3 in Table 1 are in fact used for calculating all performance measures resulting from Policies 1 through 3, respectively. Policy 1 enjoys a lower response time because the intact servers need not deny customers in order to restore the failed server. Also, customers present at the time of server failure in policy 1 are emptied into the delay elements and incur no response time gains.

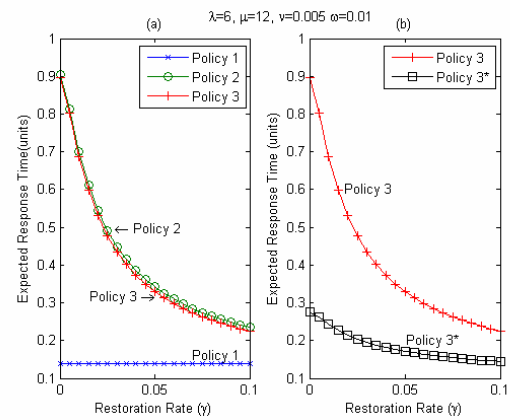


Fig.5 Average query response time versus restoration rate

Fig.5b shows the effect of applying Policy 3\*: routing all customers to the intact server that is not restoring the failed server, an alternative to Policy 3. The reduced response time in policy 3\* results from customers not waiting at a failed server. This policy may not be as advantageous in a system of higher traffic intensity.

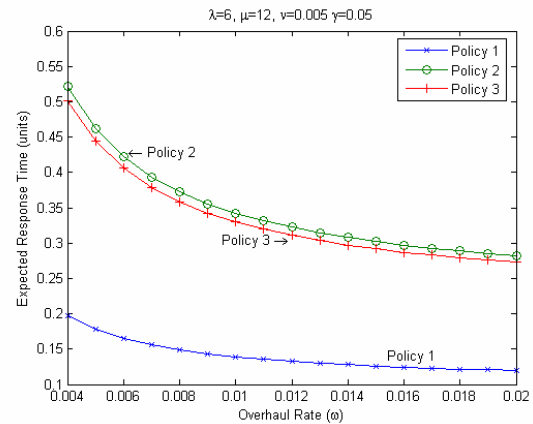


Fig.6 Average query response time versus overhaul rate

### D. Overhead

Overhead is a quantity introduced to reflect the ratio of the time invested on helping the database unit to survive longer

to its overall busy time. It is a measure of the cost of supervisory control. More specifically,

$$\theta \equiv \frac{\Pr[S_{AB} \text{ restores or fails | unit is not failed}]}{\Pr[S_{AB} \text{ restores or fails or serves | unit is not failed}]} \quad (17)$$

Overhead  $\theta$  is calculated for both the absorbing chain ( $u_3 = 0$ ) as a function of time, and the irreducible chain ( $u_3 = 1$ ) as a function of server failure rate. These are shown in Fig.7 and Fig.8. In Fig.7, it is seen that restoration incurs a higher overhead in the early life of the unit. As the database unit ages, its server becomes more likely to fail. A control policy that permits restoration becomes advantageous. There is a reduction in overhead across all policies with an increase in the arrival rate because of the resulting increased utilization. In Fig.8, for sufficiently low server failure rate, overhead is always lower with restoration. When server failure rate passes some threshold, however, restoration becomes expensive. Overhead is expected to gain more significance as a function of time and a function of server failure rate when the server life distributions have an increasing failure rate, such as in the case of Weibull distribution.

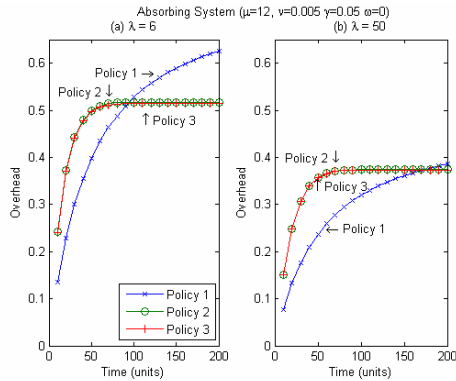


Fig.7 Overhead versus time in the absorbing chain

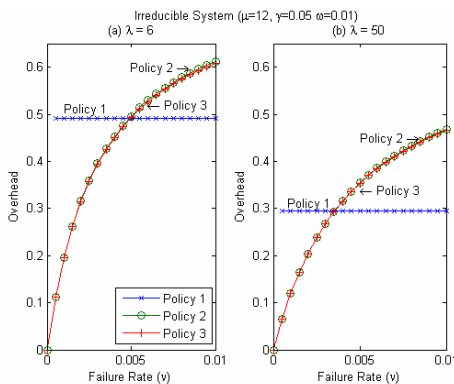


Fig.8 Overhead versus failure rate in the irreducible chain

#### IV. CONCLUSIONS

This paper modeled a redundant database unit in C2 for investigation of fault-tolerance and responsiveness afforded by a set of supervisory control policies. In all the performance measures examined, restoration ( $u_1$ ) is more effective than routing ( $u_2$ ). It is expected that when the number of queries increase, or the traffic becomes more intensive, the effectiveness of routing will be more apparent.

The study presented in this paper is limited by our ability to deal with complex problems analytically. Most restrictive is the size of the state space. The closed-queuing network model shown in Fig.2 presents perhaps the smallest possible state space for which the investigation on control policies is nontrivial. Besides answering queries, the database unit also must be updated from time to time. In that case, two types of service requests exist and the state space must be expanded. Almost equally restrictive is the assumption that times to event occurrence are exponentially distributed. Since there is only one parameter in an exponential distribution, it is likely to be unsuitable to truthfully describe some of the processes. Discrete event simulations are being carried out where the simplifying assumptions are removed to substantiate our claims on the benefit of supervisory control under more general settings in terms of the types of services, the number of customers, and the types of distributions of event lives.

Also ongoing is the extension of this study to incorporate the effect of decision and control under uncertainty and time delay due to, for example, incomplete state information and the time required for state estimation, respectively. The results will be reported in a future paper.

#### ACKNOWLEDGMENT

Both authors thank Ms. Sudha Thavamani, a student at Binghamton University pursuing her Ph.D. degree under N. Eva Wu, for her assistance in the simulations with Arena [3] that supported the preliminary study [2] of the database unit shown in Fig.1. N. Eva Wu also thanks Dr. Timothy Busch at the AFRL Rome Research Site for his insights in many sessions of discussion on the issue of architecture of the command and control supporting systems.

#### REFERENCES

- [1] N. E. Wu, and T. Busch, "Operational reconfigurability in command and control," *Proc. American Control Conference*, 2004.
- [2] N. E. Wu, and T. Busch, "An example of supervisory control in C2," *Proc. IEEE Conference on Decision and Control*, 2004.
- [3] N.E. Wu, and J. M. Metzler, "Reconfigurability in command and control systems: data loss prevention in a redundant database unit," *Report to AFRL RRS*, 2005.
- [4] Rockwell Software, Inc. *Arena*, Academic Version 7.01.00, 2004.
- [5] W. D. Kelton, R. P. Sadowski and D. T. Sturrock, *Simulation with Arena*, 3rd Ed., McGraw-Hill, 2004.
- [6] K.S. Trivedi, *Probability and Statistics with Reliability, Queuing and Computer Science Applications*, Prentice-Hall, 1982.
- [7] C.G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*, Kluwer, 1999.
- [8] E.P.C., Kao, *An Introduction to Stochastic Processes*, Duxbury Press, 1997.
- [9] S. Zacks, *Introduction to Reliability Analysis: Probability Models and Statistics Methods*, Springer-Verlag, 1992.
- [10] N. E. Wu, "Coverage in fault tolerant control," *Automatica*, vol.40, 2004, pp.537-548.

APPENDIX

Table 2 Transitions and transition rates of the database unit model with all rates valid for all policies, based on which matrix Q is formed

Table with 24 columns: States (Q, S, R, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R), Arrivals (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R), Completions (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R), Failures (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R), System Failures (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R), and Restorations (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R). Rows represent transitions between states, with numerical values for rates.