Reliability Measures of a Computer System with Arrival Time of the Server and Priority to H/W Repair over S/W Up-gradation

J.K. Sureria
Department of Statistics
M.D. University, Rohtak Haryana (India)

S.C. Malik
Department of Statistics
M.D. University, Rohtak Haryana (India)

ABSTRACT

The main emphasis of this paper is on the evaluation of some important reliability measures of a computer system considering the concepts of redundancy, priority in repair disciplines and arrival time of the server. Two identical units of a computer system are taken up-one unit is initially operative and the other is kept as spare in cold standby. There is a direct independent failure of h/w and s/w from normal mode. The system is repaired at its h/w failure while s/w is up-graded as per requirements. Priority is given to the h/w repair over the s/w up-gradation. Failure time of the system is exponentially distributed while the distributions of h/w repair time, s/w up-gradation time and arrival time of the server are assumed as arbitrary. The system has been analysed stochastically in detail using semi-Markov process and regenerative point technique. The graphical study of the results has also been made.

Keywords: Computer System, H/w Repair, S/w Up-gradation, Arrival Time, Priority and Reliability Measures.

Subject Classification: Primary 90B25 and Secondary 60K110.

1. INTRODUCTION

The requirements of computer systems have been increased many folds in day to day activities of the society. The applications of these systems can also be visualized in many other important technical fields such as air traffic control, nuclear reactors, aircraft, automotive mechanical and safety control, telephone switching and hospital patient monitoring systems. When the requirements and dependencies on computers increase, the possibility of their failure also increases. And, the impact of these failures ranges from inconvenience to economic damage to loss of life. Therefore, it becomes necessary to operate such systems with proper care and high reliability. Thus an overall assessment of the reliability of computer systems is necessary to provide better services to the customers.

The method of redundancy has been suggested by various researchers including Naidu and Gopalan [1984] as one of the best technique in improving the reliability and performance of the operating systems. However, a little work has been reported by the researchers so far on reliability modeling of computer systems with redundancy. Recently, Malik and Anand [2010, 11, and 2012] and Malik and Sureria [2012] developed reliability models for computer systems by taking one more unit (computer system) in cold standby redundancy. They assumed that server attends the system immediately when required. But, practically, this assumption seems to be unrealistic because sometimes it would be very difficult for a server to reach at the system immediately may because of his pre-occupations. And, in such situations, server may take some time to arrive at the system.

Keeping in view of the practical applications, here reliability measures of a computer system are obtained considering the concepts of redundancy, priority in repair disciplines and arrival time of the server. Two identical units of a computer system are taken up in which one unit is initially operative and the other is kept as spare in cold standby. In each unit h/w and s/w work together and may fail independently from normal mode. There is a single server who takes some time to arrive at the system. Server repairs the system at its h/w failure while up-gradation of the s/w is made as per requirements. Priority to the h/w repair is given over the s/w up-gradation. All random variables are statistically independent. Repair and switch devices are perfect. Time to system failure is exponentially distributed while the distributions of h/w repair time, s/w up-gradation time and arrival time of the server are taken as arbitrary with different probability density functions (pdf). The expressions for some measures of system effectiveness such as mean sojourn times, mean time to system failure (MTSF), availability, busy period of the server due to h/w repair and s/w up-gradation, expected number of s/w up-gradations and expected number of visits by the server are derived using semi-Markov process and regenerative point technique. The graphical study of MTSF, availability and profit function has also been made giving arbitrary values to various parameters and costs.

2. NOTATIONS

E : The set of regenerative states

O : The unit is operative and in normal mode

Cs : The unit is cold standby

a/b : Probability that the system has h/w / s/w failure

λ_1/λ_2: Constant h/w/s/w failure rate

FHUr/FHUR: The unit is failed due to h/w and is under repair/under repair continuously from previous state

FHWr / FHWR: The unit is failed due to h/w & is waiting for repair/waiting for repair continuously from previous state

FSURp/FSURP: The unit is failed due to the s/w and is under up-gradation / under up-gradation continuously from previous state

FSWRp/FSWRP: The unit is failed due to the software and is waiting for up-gradation/waiting for up-gradation continuously from previous state

w(t) / W(t) : pdf / cdf of arrival time of the server due to h/w and s/w failure

f(t) / F(t) : pdf / cdf of s/w up-gradation time g(t) / G(t) : pdf / cdf of repair time of the unit due to hardware failure

q_{ij} (t)/ Q_{ij}(t): pdf / cdf of passage time from regenerative state i to a regenerative state j or to a failed state j without visiting any other regenerative state in (0, t]

$q_{ij.kr}$ (t)/$Q_{ij.kr}$(t): pdf/cdf of direct transition time from regenerative state i to a regenerative state j or to a failed state j visiting state k, r once in (0, t]

m_{ij} : Contribution to mean sojourn time (μ_i) in state S_i when system transits directly to state S_j so that $\mu_i = \sum_j m_{ij}$ and

$m_{ij} = \int t dQ_{ij}(t) = -q_{ij}^{*\prime}(0)$

Ⓢ/© : Symbol for Laplace-Stieltjes convolution/Laplace convolution

~ / * : Symbol for Laplace Steiltjes Transform (LST) / Laplace Transform (LT)

' (dash) : Used to represent alternative result

The following are the possible transition states of the system:

S_0= (O, Cs), S_1= (O, FHWr),
S_2= (O, FSWRp), S_3= (O, FSURp),
S_4= (O, FHUr), S_5= (FHWr, FHWR),
S_6= (FHWR, FHUr), S_7 = (FHWr, FHUR),
S_8 = (FSWRp, FHWR), S_9= (FSWRP, FHUr),
S_{10}= (FSWRp, FHUR), S_{11}= (FHUr, FSWRP),
S_{12}= (FHWr, FSWRP), S_{13}= (FSWRp, FSWRP),
S_{14}= (FSWRP, FSURp), S_{15}= (FHUr, FSWRp),
S_{16}= (FSWRp, FSURP),

The states S_0–S_4 & S_{15} are regenerative states while the states S_5–S_{14} and S_{16} are non-regenerative as shown in **fig. 1.**

3. RELIABILITY MEASURES

3.1 Transition Probabilities and Mean Sojourn Times

Simple probabilistic considerations yield the following expressions for the non-zero elements:

$$p_{ij} = Q_{ij}(\infty) = \int_0^\infty q_{ij}(t)dt \text{ as}$$

$p_{01}= \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2}$, $p_{02}= \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2}$,

$p_{14} = w^*(a\lambda_1 + b\lambda_2)$, $p_{18}= \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2}\left[1 - w^*(a\lambda_1 + b\lambda_2)\right]$,

$p_{15}= \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2}\left[1 - w^*(a\lambda_1 + b\lambda_2)\right]$, $p_{23} = w^*(a\lambda_1 + b\lambda_2)$,

$p_{2,13} = \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2}\left[1 - w^*(a\lambda_1 + b\lambda_2)\right]$, $p_{2,12} = \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2}\left[1 - w^*(a\lambda_1 + b\lambda_2)\right]$,

$p_{30} = f^*(a\lambda_1 + b\lambda_2)$,

$p_{3,15}= \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2}\left[1 - f^*(a\lambda_1 + b\lambda_2)\right]$, $p_{3,16}= \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2}\left[1 - f^*(a\lambda_1 + b\lambda_2)\right]$,

$p_{40}= g^*(a\lambda_1 + b\lambda_2)$, $p_{47}= \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2}\left[1 - g^*(a\lambda_1 + b\lambda_2)\right]$,

$p_{4,10}= \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2}\left[1 - g^*(a\lambda_1 + b\lambda_2)\right]$, $p_{56}= w^*(s)$,

$p_{64}= g^*(s)$, $p_{74}= g^*(s)$, $p_{89}= w^*(s)$,

$p_{93}= g^*(s)$, $p_{10,3}= g^*(s)$, $p_{11,4}= f^*(s)$,

$p_{12,11}= w^*(s)$, $p_{13,14}= w^*(s)$, $p_{14,3}= f^*(s)$,

$p_{15,4}= f^*(s)$, $p_{16,3}= f^*(s)$, $p_{15,3}= g^*(s)$ (1)

For f (t) = $\theta e^{-\theta t}$, g (t) = $\alpha e^{-\alpha t}$ & $w(t) = \beta e^{-\beta t}$ we have

$p_{13.89} = \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2 + \beta}$, $p_{14.56} = \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2 + \beta}$,

$p_{23.13,14} = \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2 + \beta}$, $p_{24.12,11} = \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2 + \beta}$,

$p_{33.16} = \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2 + \theta}$, $p_{34.15} = \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2 + \theta}$,

$p_{44.7} = \dfrac{a\lambda_1}{a\lambda_1 + b\lambda_2 + \alpha}$, $p_{43.10} = \dfrac{b\lambda_2}{a\lambda_1 + b\lambda_2 + \alpha}$ (2)

It can be easily verified that

$p_{01}+p_{02} = p_{14}+p_{15}+p_{18} = p_{23}+p_{2,12}+p_{2,13} = p_{30}+p_{3,15}+p_{3,16} = p_{40}+p_{47}+p_{4,10} = p_{14} + p_{14.56} + p_{13.89} = p_{23} + p_{23.13,14} + p_{24.12,11} = p_{30}+p_{33.16}+p_{34.15} = p_{40}+p_{44.7}+p_{43.10} = 1$ (3)

The mean sojourn times (μ_i) in the state S_i are

$\mu_0 = \dfrac{1}{a\lambda_1 + b\lambda_2}$, $\mu_1 = \mu_2 = \dfrac{1}{a\lambda_1 + b\lambda_2 + \beta}$,

$\mu_3 = \dfrac{1}{a\lambda_1 + b\lambda_2 + \theta}$, $\mu_4 = \dfrac{1}{a\lambda_1 + b\lambda_2 + \alpha}$ (4)

$m_{01} + m_{02} = \mu_0$, $m_{14} + m_{15} + m_{18} = \mu_1$,
$m_{23} + m_{2,12} + m_{2,13} = \mu_2$, $m_{30} + m_{3,15} + m_{3,16} = \mu_3$
$m_{40} + m_{47} + m_{4,10} = \mu_4$ (5)

$m_{14} + m_{14.56} + m_{13.89} = \mu_1'$, $m_{23} + m_{23.13,14} + m_{23.12,11} = \mu_2'$,
$m_{30} + m_{33.16} + m_{3,15} = \mu_3'$, $m_{40} + m_{44.7} + m_{43.10} = \mu_4'$ (6)

for f (t) = $\theta e^{-\theta t}$, g(t)= $\alpha e^{-\alpha t}$ and $w(t) = \beta e^{-\beta t}$

we have

$$\mu_1' = \frac{\alpha(\alpha + a\lambda_1 + b\lambda_2)(\beta + 1) + \beta(a\lambda_1 + b\lambda_2)(\beta + b\lambda_2 + a\lambda_1)}{\alpha\beta(a\lambda_1 + b\lambda_2 + \alpha)(\beta + b\lambda_2 + a\lambda_1)}$$

$$, \mu_2' = \frac{\theta(\beta + a\lambda_1 + b\lambda_2) + \beta(a\lambda_1 + b\lambda_2)}{\theta\beta(\beta + b\lambda_2 + a\lambda_1)}$$

$$\mu_3^1 = \frac{\theta + b\lambda_2}{\theta(\theta + a\lambda_1 + b\lambda_2)}, \ \mu_4^1 = \mu_{15} = \frac{1}{\alpha} \quad (7)$$

3.2 Reliability and Mean Time to System Failure (MTSF)

Let ϕ_i (t) be the c.d.f. of first passage time from regenerative state i to a failed state. Regarding the failed state as absorbing state, we have the following recursive relations for ϕ_i(t):

$$\phi_i(t) = \sum_j Q_{i,j}(t)(S)\phi_j(t) + \sum_k Q_{i,k}(t) \tag{8}$$

where *j* is an un-failed regenerative state to which the given regenerative state *i* can transit and *k* is a failed state to which the state *i* can transit directly. Taking LST of above relation (8) and solving for $\tilde{\phi}_0(s)$. We have

$$R^*(s) = \frac{1-\tilde{\phi}_0(s)}{s} \tag{9}$$

The reliability of the system model can be obtained by taking Laplace inverse transform of (9). The mean time to system failure (MTSF) is given by

$$\text{MTSF} = \lim_{s\to o} \frac{1-\tilde{\phi}_0(s)}{s} = \frac{N_1}{D_1}\text{, where} \tag{10}$$

$N_1 = \mu_0 + p_{01}\mu_1 + p_{02}\mu_2 + p_{02}p_{23}\mu_3 + p_{01}p_{14}\mu_4$ and

$D_1 = 1 - p_{01}p_{14}p_{40} - p_{02}p_{23}p_{30}$

3.3 System Availability

Let A_i (t) be the probability that the system is in up-state at instant‘t’ given that the system entered regenerative state i at t = 0. The recursive relations for A_i (t) are given as

$$A_i(t) = M_i(t) + \sum_j q_{i,j}^{(n)}(t) © A_j(t) \tag{11}$$

where *j* is any successive regenerative state to which the regenerative state *i* can transit through n transitions and M_i (t) is the probability that the system is up initially in state $S_i \in E$ is up at time t without visiting to any other regenerative state, we have

$$M_0(t) = e^{-(a\lambda_1+b\lambda_2)t}, M_1(t) = M_2(t) = e^{-(a\lambda_1+b\lambda_2)t}\overline{W}(t),$$
$$M_3(t) = e^{-(a\lambda_1+b\lambda_2)t}\bar{F}(t), M_4(t) = e^{-(a\lambda_1+b\lambda_2)t}\bar{G}(t) \tag{12}$$

Taking LT of above relations (11) and solving for $A_0^*(s)$, the steady state availability is given by

$$A_0(\infty) = \lim_{s\to 0} sA_0^*(s) = \frac{N_2}{D_2}\text{, where} \tag{13}$$

$N_2 = p_{30}(1-p_{44.7})\,[\mu_0 + p_{01}\mu_1 + p_{02}\mu_2] + [1-p_{47.7}-p_{01}p_{40}(1-p_{13.89})]\,\mu_3 + p_{01}p_{30}(1-p_{13.89})\,\mu_4$

$D_2 = p_{30}(1-p_{44.7})\,[\mu_0 + p_{01}\mu_1' + p_{02}\mu_2'] + [1-p_{47.7}-p_{01}p_{40}(1-p_{13.89})]\,[\mu_3' + \mu_{15}'] + p_{01}p_{30}(1-p_{13.89})\,\mu_4'$

3.4 Busy Period Analysis for Server

3.4.1 Due to Hardware Repair

Let $B_i^H(t)$ be the probability that the server is busy in repairing the unit due to hardware failure at an instant ‘t’ given that the system entered state i at t = 0. The recursive relations $B_i^H(t)$ for are as follows:

$$B_i^H(t) = W_i^H(t) + \sum_j q_{i,j}^{(n)}(t) © B_j^H(t) \tag{14}$$

where *j* is any successive regenerative state to which the regenerative state *i* can transit through n transitions and W_i^H (t) be the probability that the server is busy in state S_i due to hardware failure up to time t without making any transition to any other regenerative state or returning to the same via one or more non-regenerative states and so

$$W_4^H(t) = e^{-(a\lambda_1+b\lambda_2)t}\bar{G}(t) + \left[a\lambda_1 e^{-(a\lambda_1+b\lambda_2)t} © 1\right]\bar{G}(t)$$
$$+\left[b\lambda_2 e^{-(a\lambda_1+b\lambda_2)t} © 1\right]\bar{G}(t), W_{15}^H(t) = \bar{G}(t) \tag{15}$$

3.4.2 Due to Software Up-gradation

Let B_i^S (t)be the probability that the server is busy due to up-gradation of the software at an instant ‘t’ given that the system entered the regenerative state i at t = 0. We have the following recursive relations for B_i^S (t):

$$B_i^S(t) = W_i^S(t) + \sum_j q_{i,j}^{(n)}(t) © B_j^S(t) \tag{16}$$

where *j* is any successive regenerative state to which the regenerative state *i* can transit through n transitions and W_i^S (t) be the probability that the server is busy in state S_i due to up-gradation of the software up to time t without making any transition to any other regenerative state or returning to the same via one or more non-regenerative states and so

$$W_3^S(t) = e^{-(a\lambda_1+b\lambda_2)t}\overline{F}(t) + (a\lambda_1 e^{-(a\lambda_1+b\lambda_2)t} © 1)\overline{F}(t)$$
$$+(b\lambda_2 e^{-(a\lambda_1+b\lambda_2)t} © 1)\overline{F}(t) \tag{17}$$

Taking LT of above relations (14) and (16) and solving for B_0^{*H} (s) and B_0^{*S} (s), the time for which server is busy due to repair and up-gradations respectively is given by

$$B_0^H = \lim_{s\to 0} sB_0^{*H}(s) = \frac{N_3^H}{D_2} \tag{18}$$

$$B_0^S = \lim_{s\to 0} sB_0^{*S}(s) = \frac{N_3^S}{D_2}\text{, where} \tag{19}$$

$$N_3^H = p_{01}p_{30}(1-p_{13.89})\tilde{W}_4^H(0) +$$
$$p_{3,15}[p_{43.10} + p_{02}p_{04} + p_{01}p_{40}p_{13.89}]\tilde{W}_{15}^H(0)$$
$$N_3^S = [p_{43.10} + p_{02}p_{04} + p_{01}p_{40}p_{13.89}]\tilde{W}_3^S(0)$$

and D_2 is already mentioned.

3.5 Expected Number of Software Up-gradations

Let $R_i^S(t)$ be the expected number of up-gradations of the failed software by the server in (0, t] given that the system entered the regenerative state i at t = 0. The recursive relations for $R_i^S(t)$ are given as

$$R_i^S(t) = \sum_j Q_{i,j}^{(n)}(t)(S)\left[\delta_j + R_j^S(t)\right] \tag{20}$$

where j is any regenerative state to which the given regenerative state *i* transits and δj =1, if *j* is the regenerative state where the server does job afresh, otherwise δj = 0. Taking LST of relations (20) and solving for $\tilde{R}_0^S(s)$. The expected number of software up-gradations per unit time is given by

$$R_0(\infty) = \lim_{s\to 0} s\tilde{R}_0^S(s) = \frac{N_4}{D_2}\text{, where} \quad (21)$$

$$N_4 = (1 - p_{13.89})[p_{43.10} + p_{02}p_{04} + p_{01}p_{40}p_{13.89}]$$

and D_2 is already mentioned.

3.6 Expected Number of Visits by the Server

Let N_i (t) be the expected number of visits by the server in (0, t] given that the system entered the regenerative state i at t= 0. The recursive relations for N_i (t) are given as

$$N_i\ (t) = \sum_j Q_{i,j}^{(n)}(t)(S)\left[\delta_j + N_j\ (t)\right] \quad (22)$$

where j is any regenerative state to which the given regenerative state *i* transits and δj =1, if *j* is the regenerative state where the server does job afresh otherwise δj = 0.

Taking LST of relation (22) and solving for $\tilde{N}_0(s)$. The expected numbers of visits per unit time by the server are given by

$$N_0(\infty) = \lim_{s\to 0} s\tilde{N}_0(s) = \frac{N_5}{D_2}\text{, where} \quad (23)$$

N_5 = (1- $p_{44.7}$) p_{30} and D_2 is already specified.

4. PROFIT FUNCTION

The profit incurred to the system model in steady state can be obtained as

$$P = K_0A_0 - K_1B_0^H - K_2B_0^S - K_3R_0 - K_4N_0 \quad (24)$$

where

K_0 = Revenue per unit up-time of the system

K_1 = Cost per unit time for which server is busy due to h/w repair

K_2 = Cost per unit time for which server is busy due to s/w up-gradation

K_3 = Cost per unit time s/w up-gradation

K_4 = Cost per unit time visit by the server

and $A_0, B_0^H, B_0^S, R_0, N_0$ are already defined.

STATE TRANSITION DIAGRAM

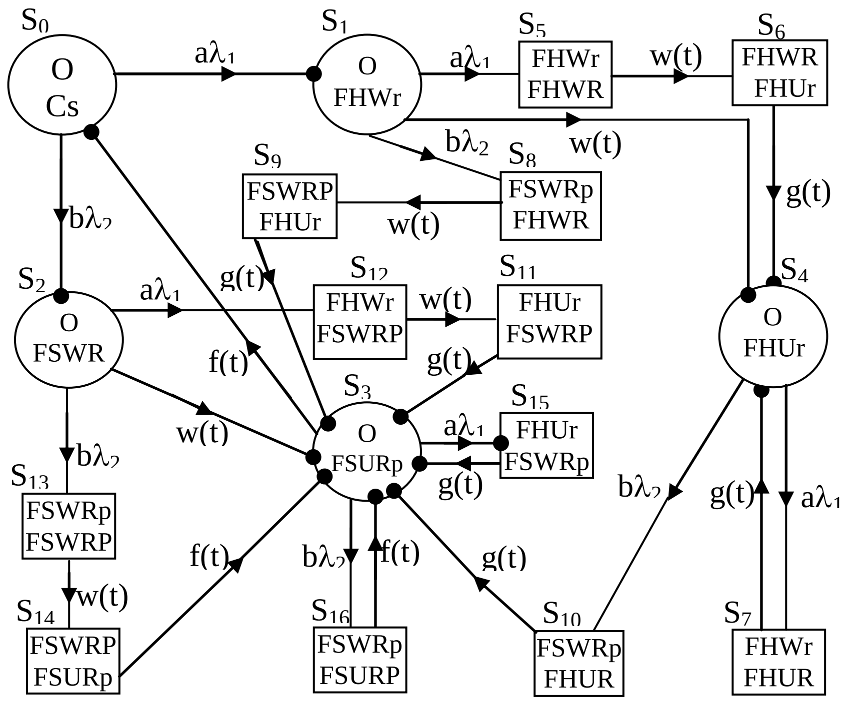

Fig. 1

○Up-state □ Failed state ●Regenerative point

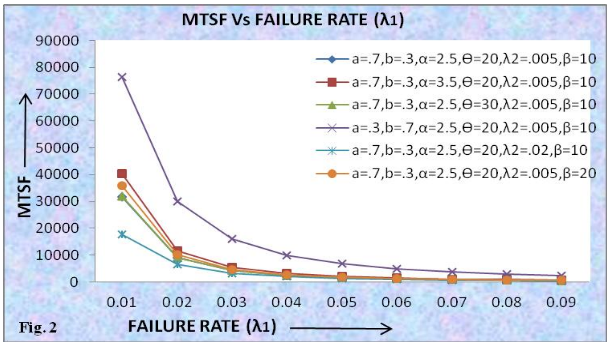

Fig. 2

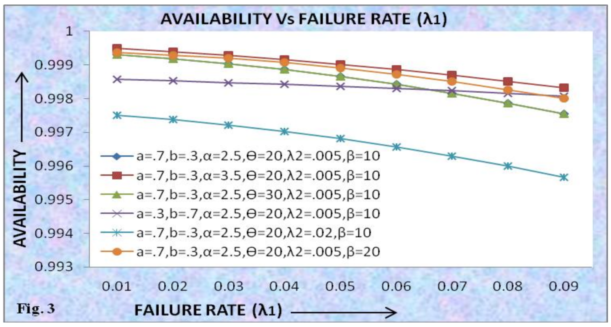

Fig. 3

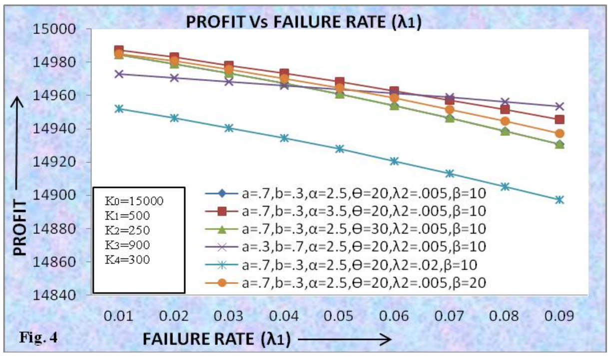

Fig. 4

5. CONCLUSION

The results for some important reliability measures of a computer system are obtained giving particular values to various parameters and costs. It is observed that mean time to system failure (MTSF), availability and profit go on decreasing with the increase of h/w and s/w failure rates (λ_1 and λ_2) as shown in figures 2, 3 and 4 respectively. However, the values of these measures increase with the increase of repair rate (α) and arrival rate (β) of the server.

Hence it is found that the reliability and availability of a computer system can be improved by the concept of redundancy in case server takes more time to arrive at the system. On the other hand, if priority is given to h/w repair over s/w up-gradation in a redundant computer system then system can be made more available to use either by increasing h/w repair rate or by reducing the arrival time of the server.

6. REFERENCES

[1]. Naidu, R.S. and Gopalan, M.N. (1984): Cost-benefit analysis of one-server two-unit system subject to arbitrary failure, inspection and repair. **Journal of Reliability Engg., Vol. 22,** *pp. 8-11*

[2]. Malik, S.C and Anand, Jyoti (2010): Reliability and Economic Analysis of a Computer System with Independent Hardware and Software Failures**.** *Bulletin of Pure and Applied Sciences*.Vol.29E (No.1), pp.141-153.

[3]. Malik, S.C and Anand, Jyoti (2011): Reliability Modeling of a Computer System with Priority for Replacement at Software Failure over Repair Activities at H/W Failure. *International Journal of Statistics and System*, ISSN 0973-2675, Vol. 6,No. 3,pp.315-325.

[4]. Malik, S.C. and Anand, J. (2012): Probabilistic Analysis of a Computer System with Inspection and Priority for Repair Activities of H/W over Replacement of S/W. *International Journal of Computer Applications*, Vol.44 (1), pp. 13-21.

[5]. Malik, S.C and Sureria, J.K. (2012): Profit Analysis of a Computer System with H/W Repair and S/W replacement. *International Journal of Computer Applications*, Vol. 47, No.1, pp.19-26