Allocation of repairable spare parts inventories along ran automatic transfer line.

Sally Elizabeth Gager

Follow this and additional works at: https://preserve.lehigh.edu/etd

Part of the Industrial Engineering Commons

Recommended Citation

This Thesis is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.
ALLOCATION OF REPAIRABLE SPARE PARTS INVENTORIES ALONG AN AUTOMATIC TRANSFER LINE

by

Sally Elizabeth Gager

A Thesis
Presented to the Graduate Committee of Lehigh University in Candidacy for the Degree of Master of Science in Industrial Engineering

Lehigh University 1984
This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

June 14, 1984
(date)

Professor in Charge

Chairman of Department
Table of Contents

Abstract ................................................. 1
1.0 Introduction ......................................... 2
2.0 Model Description ..................................... 5
   2.1 Model Assumptions .................................. 16
   2.2 Mathematical Formulation of the Model ............ 19
   2.3 Previous Research .................................. 20
      2.3.1 Reliability Analysis ......................... 21
      2.3.2 Transfer Lines ............................... 24
3.0 Workstation Reliability Analysis ..................... 27
   3.1 Evaluation of Single Components
      Without Standbys .................................. 30
   3.2 Evaluation of Single Components
      With Standbys ..................................... 33
   3.3 Calculation of Workstation Performance
      Measures ............................................ 42
4.0 Analysis of the Transfer Line ......................... 45
   4.1 Evaluation of the Net Transfer Line
      Production Rate ................................... 46
5.0 Economic Considerations ............................... 50
   5.1 Derivation of the Total Relevant Cost
      Per Workpiece Produced ........................... 51
   5.2 Selection of the Buffer-Standby Policy ............ 57
5.2.1 Calculation of the Optimum Buffer Size .......................... 58
5.2.2 A Branch and Bound Method to Find the Optimum Standby Policy .................. 62
6.0 Example Problem ..................................................... 69
7.0 Recommendations for Further Study .................................. 79
8.0 Conclusion .............................................................. 81
References ................................................................. 82
Appendix 1 ................................................................. 85
Vita ......................................................................... 90
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Two Station Transfer Line with Buffer Storage</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Transfer Line with no Redundancy Modelled at Component Level</td>
<td>9</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Reconfigurable Redundancy</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Decomposition of an Optical Seam Tracker</td>
<td>12</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Transfer Line with Redundancy Modelled at Component Level</td>
<td>15</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Branch and Bound Decision Tree</td>
<td>65</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Decision Tree for the Example Problem</td>
<td>72</td>
</tr>
</tbody>
</table>
Abstract

This thesis considers a transfer line system consisting of two unreliable workstations separated by a finite buffer storage queue. Each workstation is modelled as a multicomponent series, subject to exponential failures and repairs, with redundancy in parallel. An analytical formula for the availability of the transfer line developed by Malathronas, Perkins, and Smith is used to derive the total relevant cost per unit produced. The optimum number of standbys and the optimum buffer storage capacity are obtained by using a branch and bound technique.
Allocation of Repairable
Spare Parts Inventories
Along an Automatic Transfer Line

1.0 Introduction

In recent years automatic transfer lines have become a key component in many high volume production systems. Because these lines typically involve high capital investment costs it is often crucial that they be designed in such a manner that production rates reach their full potential.

In order to achieve this goal systems designers must keep in mind one of the major problems associated with transfer lines: the increased impact of downtime. Most transfer lines used today are relatively inflexible. Not only are workpieces processed through a fixed sequence of machines but, also, the machines themselves must be arranged in a permanent configuration. The result of such inflexibility is that when any machine along a fixed transfer line breaks down other machines may subsequently be forced down.

There are several methods by which the impact of downtime can be minimized. The most common method is the insertion of one or more buffer storage zones between successive stages along the line. These zones decouple the stages from one another and serve to partially isolate
the workstations in any particular stage from the workstations in all other stages. Thus, the effects of individual workstation failures on the performance of the entire line are reduced and the proportion of time the line is operating is increased.

Another method that can be used to further improve flow line performance is to increase the reliability of workstations within each stage. This can be accomplished by following two approaches: fault avoidance and fault tolerance.

The goal of fault avoidance is to reduce the possibility of failures through preventive maintenance. Parts which deteriorate with age can be replaced at periodic intervals. Also, inspection procedures can be enforced where employees regularly spot check components for significant wear.

The goal of fault tolerance, on the other hand, is to establish a strategy for lessening the impact of failures. Often it is impossible to achieve high reliability regardless of the amount of maintenance involved. Random workstation failures may occur and if these occur frequently or if the resultant repair times are long a significant lowering of the average flow line rate can result. In this situation the only way for a production system to reach acceptable levels of performance may be through the adoption of a fault tolerant system design. Such a design would reduce the effects of component failures by providing
strategies for lessening their impact. An example of this is the use of redundancy among strategic components.

The purpose of this thesis is to formulate an analytical model of a transfer line which possesses both buffer storage capacity and a fault tolerant system design. The fault tolerance will be provided by the insertion of standbys for strategic workstation components. Both buffer storage and component standbys will be allocated by calculating the benefits which can be derived from their implementation and comparing these to the associated costs. In this manner, the most economical means of improving the transfer line can be chosen.
2.0 Model Description

A simplified model of the system under consideration is shown in Figure 1. In this system, the transfer line consists of two automated workstations separated by an inventory storage buffer. Workpieces enter the line at station 1 where one or more processing operations take place. They then advance to the storage buffer, if space to transfer them is available, and wait in a finite capacity queue for station 2. After the processing operations at station 2 take place the completed workpieces leave the system. It is assumed that all workpieces entering the line eventually leave the line as completed workpieces.

At any given moment each station is either in an operational up state or in a down state under continuous repair. When a station is operational it may process workpieces only if it is not starved or blocked. Station 1 will be blocked when there is no room in the storage buffer to put processed workpieces. Station 2 will be starved if there are no workpieces in the storage buffer available for processing. Because it is assumed that there is an infinite supply of unmachined workpieces in front of station 1, station 1 can never be starved. Also, due to the unlimited space for machined workpieces leaving the system station 2 can never be blocked.
I unlimited supply of unprocessed workpieces

Figure 1
Two Station Transfer Line with Buffer Storage
The controller performs two major functions within the system. The first major function is it provides computer control for the processing operations within each workstation. Depending upon the computational requirements, numerical control technology may be needed to operate one or more workstation components. Computer control may also be needed within workstations to store programs, adaptively control machine tools, and collect data on tool changes, breakdowns, and the net workstation production rate.

The second major function provided by computer control is the sequencing and synchronization of workpieces between stations. If a variety of workpieces are to be processed by the transfer line these workpieces must be launched in a sequence which maximizes workstation utilization while meeting product demands. Also, in order for the transfer line to achieve the highest net production rate possible, the workpieces must be able to move more or less independently of each other. The degree of workpiece independence, of course, will depend upon the size of the buffer and its success in isolating the individual workstations.

The function of the control system becomes more crucial when the model given in Figure 1 is expanded to differentiate between individual component failures.
As shown in Figure 2, each workstation can be treated as a collection of several components. Station 1, for example, is divided into \( n_1 \) components where \( r_c(t) \) is the reliability of the \( c^{th} \) component. When these components operate independently of one another, i.e. when the components have mutually independent failure and repair rates, they can be modelled as a probabilistic \( n \) component series. For an \( n \) component series system to work all \( n \) components must be functioning. Therefore, the overall workstation reliability, \( R(t) \), can be found by computing the product of the individual component reliabilities \( r_c(t) \).

The choice of the components within a workstation and the resulting degree to which workstations can be decomposed will depend upon the physical configuration of the workstation parts and their relationships with one another. Components must be physically separable units which fail and are repaired independently from all other components within the workstation. Also, if redundant standbys are to be incorporated into the model additional factors such as the space requirements of standby parts and reconfiguration constraints must be taken into account.

Redundancy can be added to the model by assuming that redundant spares will only be available for those components which are reconfigurable. A component is
Figure 2
Transfer Line with no Redundancy
Modelled at Component Level
reconfigurable if the control system is able to disconnect it whenever it fails and automatically reconfigure the workstation with a spare, if one is available.

The process of reconfiguration is illustrated in Figure 3. At any given moment only one of the redundant components is connected to the system outputs; all of the other components are on standby. Reconfiguration is triggered by an internal detection of faults in the active component or by a detection of faults in the output. The action taken following a fault detection can take several different forms. Occasional erroneous results can often be ignored. In many cases a second attempt or a retry of an operation may be successful, particularly if the failure was due to a transient or intermittent fault. Finally, if the fault is diagnosed as permanent the system may be able to reconfigure the active component, replacing it with one of the redundant spares. In this situation, the defective component is disconnected from the system and put into repair.

An example of a typical workstation decomposition is shown in Figure 4. In this figure, an optical seam tracker is divided into seven physically distinct components with mutually independent failures and repairs. Only those components which are reconfigurable will be considered for redundant standbys. In many instances, however,
Reconfigurable Redundancy. A detected mismatch during comparison of characteristic signals triggers reconfiguration. (Siewiorek and Swarfz, 1982)
Figure 4

Decomposition of an Optical Seam Tracker

(From Villiers, 1982, Figure 2)
those components which are not reconfigurable can be subdivided further into components which can be reconfigured. Also, it should be noted that depending upon the resultant economic gains attainable entire workstations or individual components can be redesigned to make reconfiguration possible.

Some components may not be reconfigurable simply because there is not enough space available at the workstation for standby components to be stored. If, for example, the robot shown in Figure 4 is extremely bulky the presence of a standby may intrude on the functions of the workstation. In this situation, it may be possible to design a reconfiguration system whereby the spare robot is stored several feet away from the workstation. Thus, when a failure occurs in the on-line component it can be automatically routed away from the workstation while the standby assumes its former position.

Another instance where components may not be reconfigurable occurs when components are permanently attached to workstations in fixed on-line positions. Reconfiguration requires the control system to switch off on-line components which fail and replace them with functioning standbys. There are two ways to modify a component which cannot be switched off-line. The first method is to subdivide the component into smaller components. The welding
equipment in Figure 4, for example, may be divisible into several subcomponents. If any of these subcomponents are reconfigurable they become candidates for standby redundancy. The second method is to redesign the component and/or the workstation to accommodate reconfiguration. However, it should be noted that alterations can be extremely expensive and, therefore, their costs should be weighed against the productivity gains which result from the added redundancy.

In summary, the system model is as shown in Figure 5. Workpieces enter the line at workstation 1 where they are processed, and then move on to a queue where they await processing at workstation 2. Each station is decomposed into several components some of which possess redundant standbys. It is assumed that the components have mutually independent failures and repairs. Also, only those components which are reconfigurable are eligible for redundant standbys.
Controller

unlimited
supply of
unprocessed
workpieces

reconfigurable
component

component

n1

standby for
component

1

queue

unlimited
space for
processed
workpieces

reconfigurable
component

component

n2

standby for
component

n2

Workstation 1

Workstation 2

Figure 5
Transfer Line with Redundancy
Modelled at Component Level
2.1 Model Assumptions

In order to build a mathematical model of the system, several assumptions will be made concerning the failure and repair distributions of workstations 1 and 2. Over the course of time, the components within each workstation will inevitably break down. Factors which cause these failures may include the type of workstation, its age, quality of manufacture, and working conditions such as maintenance, work load, temperature, and humidity. The time which it takes to repair the workstations may also vary depending upon the component which breaks down, its difficulty to repair, the availability of spare parts and special tools, and, in general, the motivation and skill of the repairman. Because of the significant number of variables which contribute to workstation failures and repairs it is not difficult to see why the choice of statistical distributions is still largely an art.

For the purpose of this discussion it is assumed that the failure distribution of each workstation has "no memory". In other words, the probability that a workstation fails during a certain time period is independent of the age of the workstation and the time which has passed since its last failure. Several precautionary measures may be necessary to insure the memoryless property...
of the failure distribution. First, because new workstations tend to experience decreasing failure rates, a certain amount of time following installation may be necessary for "debugging". The model should not be applied until after most factory defects have been detected and initial testing and adjustments have been made. A second precautionary measure to insure the memoryless property of the failure distribution is to keep the workstations in good condition. Poorly maintained equipment may experience increasing failure rates if machinery is allowed to wear out. In order for workstation failure rates to remain constant over time a program of preventive maintenance should be established whereby those workstation components which deteriorate with age are replaced or overhauled during weekends or other off-shift periods.

It will also be assumed that the repair distribution has no memory. This will only be true if the overall repairman skill level remains relatively constant over time. Repairmen, in essence, may not learn from their experience and work at faster rates as time passes by. Therefore, it is assumed that the repairmen are already well trained and are not in the process of learning their jobs.

The principle advantage with the previous two assumptions is that the failure and repair rates can be exponentially
distributed. Although exponential distributions may not result in a completely accurate description of the real world they can, under controlled conditions, approximate actual systems. The time between workstation failures, for example, will be exponential if the number of components is large and the repair time is negligible. Exponential distributions are necessary because they possess unique quantitative properties which greatly simplify calculations by allowing mathematical techniques such as the Markov chain approach and renewal theory to be used.
2.2 Mathematical Formulation of the Model

The system model shown in Figure 5 will be described mathematically by following three major steps. First, the workstation failure and repair rates will be found in terms of the number and type of standbys which are present. To do this, workstation components will be modelled as one-unit repairable systems, with or without standbys, whose performance measures can be found by using renewal theory and regeneration techniques.

The second step will be to determine the output of the two station transfer line in terms of the failure and repair rates of each workstation (found in step 1) and, also, the buffer storage capacity. Because of the considerable complexity involved in creating such a mathematical model, theories previously developed by Malathronas, Perkins, and Smith will be used.

Finally, the third step involves substituting the transfer line data calculated in step 2 into a branch and bound algorithm which minimizes the total relevant cost per workpiece produced. In this way, the most economical allotment of standbys and buffer storage capacity can be made.
2.3 Previous Research

The previous research which was used in formulating a mathematical model of the system falls into two major categories. The first major category is literature pertaining to the workstation performance measures. Each workstation is composed of several components which may or may not possess standbys. Therefore, in order to mathematically model the workstations we are primarily interested in literature which concerns multicomponent series systems with redundancy in parallel. In addition, complications such as imperfect switchover and repairable standbys will be added so that the workstation model will be as realistic as possible.

The second major category of literature is research which has been published concerning two workstation transfer lines with buffer storage. Although there has been a significant amount of research published in this area, most of the present models are of little use in the present investigation because they contain assumptions which are too restrictive. We are specifically interested in models which do not require workstations 1 and 2 to possess identical failure and/or repair rates. Hence, a transfer line system with two different workstations can be accurately modelled.
2.3.1 Reliability Analysis

A vast quantity of mathematical models and methods are available to analyze the reliability of complex systems. In this section we are specifically interested in literature which pertains to the modelling of the workstations. We wish to look at research concerning variations to the workstation model presented in section 2.0. These variations will be chosen with two goals in mind. First, the resultant workstation model should resemble a real life workstation as closely as possible. Second, the variations should remain simple enough so that the resultant computational complexity of mathematically analyzing the workstation model lies within reasonable bounds.

As mentioned in section 2.0, each workstation may be modelled as a series of several subsystems. Barlow and Proschan (2) provide simple formulas for computing the performance measures of a multicomponent series system if the failure and repair rates of the components are known. Therefore, in order to find the availability, downtime, and uptime of the workstations we must first analyze the component subsystems.

There are two different types of component subsystems: those which possess redundant standbys and those which do not possess redundant standbys. Because renewal
systems are typically described by Markov chains, Markov processes, renewal processes, or semi-Markov processes. Most redundancy models in literature are limited to only one redundant standby. If there is more than one standby the number of system states which must be modeled can be quite large. Hence, the analysis can become complicated as seen in Kulshrestha (9,10) and Chow (5).

In most circumstances, it is reasonable to assume that operating and standby components can have unequal failure rates. Such models are described by Chow (5,6), and Subramanian and Venkatakrishnan (17). When components fail while they are in the standby position at the same rate as the operating components the analysis, although often unrealistic, is much simpler as shown by Kulshrestha (9).

When an operating component fails a switching mechanism is required to put a standby component, if one is available, into operation. Many models do not take into account the possibility of switch failure (Kulshrestha (10), Kodama (7)). Other models allow imperfect switchover, but assume that when the switch fails the result is 'catastrophic', and the system is not allowed to recover (Nakagawa and Osaki (13), Subramanian and Venkatakrishnan (17)). For our workstations, we are most interested in imperfect switchover with
repair (Chow (5,6)) because it is reasonable to assume that the switching device may fail, and, also, that the repairman is capable of repairing the switch.
2.3.2 Transfer Lines

A significant amount of literature concerning transfer lines has been published in recent years. This section will focus specifically on research which provides analytical models for transfer line systems consisting of two workstations separated by a finite capacity storage buffer. The usefulness of the various analytical models will be discussed by comparing the model assumptions and noting how consistent these assumptions are with real world behavior.

One of the first analytical models of a two stage transfer line system was developed by Vladzievskii (19) and Sevastyanov (16) via a probabilistic approach. Their analysis assumes that the mean time to repair (MTTR) of the system is negative exponentially distributed, and the workstation failure rates are constant when both workstations are running. Vladzievskii and Sevastyanov also assume that only one workstation can be down at a time. This assumption is unrealistic because it ignores the possibility that when one workstation is down the other workstation may fail.

Chronologically, the next significant transfer line model was developed by Finch. As described by Koenigsberg (8), Finch's model assumes that both the downtimes and
uptimes of each workstation have negative exponential distributions. Unlike the model presented by Vladzievskii and Sevastyanov, Finch's model assumes that both workstations may fail simultaneously. Finch, however, adds an additional assumption that a workstation may fail while it is undergoing repair. This assumption does not depict most real life situations.

Perhaps the most renowned and widely applied transfer line model is that described by Buzacott (3). Buzacott's model assumes that workstation uptimes and downtimes are either geometrically distributed or constant. It has several reasonable assumptions: workstations may fail only when they are operating and both workstations may fail simultaneously. Also, it is assumed that the probability of two repairs or two failures within the same production cycle is negligible. Buzacott's formulas for the output of the transfer line have been found to give accurate predictions when failures are infrequent and repairs are short. The major drawback of the model is that both workstations must be identical.

Murphy (12) and Wijngaard (20) use renewal theory to discuss transfer lines where the workstations need not be identical. Both of these models assume that the downtimes and uptimes of the workstations are negative exponentially distributed. Murphy assumes that only one
repairman is available, so repair on workstation 2 stops when workstation 1 is down. Wijngaard assumes that the failure rates of the workstations are the same when they are undergoing repair as when they are operating. Murphy does not arrive at an analytical expression for the production rate of the transfer line. Wijngaard, on the other hand, does come up with production rate formulas but they require lengthy numerical calculations and, hence, do not provide insight on the effect of different parameters on the line performance without extensive numerical exploration.

The most complete model to date is that presented by Malathronas, Perkins, and Smith (11). This model assumes that workstation uptimes and downtimes are negative exponentially distributed. Like Buzacott's model workstations fail only when they are operating and both workstations may fail simultaneously. It is superior to Buzacott's model, however, because the workstations need not be identical. The only restriction is that the workstation production rates must be equal.
3.0 Workstation Reliability Analysis

The problem of finding the overall reliability of each workstation given the assumption of one repairman per component will be solved by following three major steps. First, the components within the workstations which operate without the presence of standbys will be analyzed. These components can be modelled as simple one-unit repairable systems whose performance measures will be evaluated by using renewal theory and regeneration techniques.

Renewal theory will also be used in the second step of the analysis, determining the performance measures of those components which have a redundant standby. In this case, however, the analysis is complicated by the need to define additional system states. Not only may a component be operating or undergoing repair (the only system states which are possible for the components evaluated in step one), but they may also be on standby or waiting for a repairman. Also, since the added complexity of imperfect switchover is included in the model, the states of the switches which are used to replace failed components with functioning standbys must be interpreted. These switches can be in an operational up state, undergoing repair, or waiting for a repairman.
The third step involved in evaluating the reliability of each workstation is to combine the reliability performance measures of the individual components within the workstations. As mentioned previously, each workstation is assumed to be a series system consisting of several components. Thus, workstation failure coincides with the failure of any one of the station components. Because the failure distributions used in this model are continuous, only one component can fail at any given moment. Also, in order to make this model as realistic as possible it will be assumed that whenever a component is undergoing repair all other components will remain in a state of suspended animation. In other words, these components will stop operating until the failed component has been repaired. At that time they will resume operation in the same condition they were in before the failure occurred.

The reliability improvement of the workstations due to the inclusion of repairable standby components will be measured in several different ways. The mean time to failure (MTTF) and the mean time to repair (MTTR) for each workstation must be found because these values are necessary to compute the corresponding increase in the production rate of the transfer line. Another performance measure that is useful is the long run fraction of time a workstation is down due to the failure of a
particular component. This measure can be used to pinpoint the effectiveness and possible future applications of redundant standbys.

Finally, it is often desirable to know the probability that a workstation is operating at a specific time $t$. This measure, called availability, is expressed symbolically as $A(t)$. Recall that the reliability $R(t)$ of a workstation is the probability that no failure has occurred prior to time $t$. We are not concerned with measuring workstation reliability, however, because many component failures may occur within a workstation and yet it may still be operating. Due to the presence of redundancy, we are more interested in knowing how effective the standbys are at keeping the workstations operating and, hence, available.
3.1 Evaluation of Single Components Without Standbys

In this section the performance measures of single components which do not have redundant standbys will be evaluated. Consider a system consisting of one component and one repairman. The system is said to be up if the component is running. However, at any instant the operational unit may fail and, thus, be forced into an inoperative state. When this occurs, repair of the failed component commences immediately. After the component has been repaired it is assumed to be as good as new and is returned to operation.

In practice, the system passes through a number of operative and inoperative periods. Assuming that the failure times and repair times are all independent random variables, the instant when the system comes up after having been down is a regeneration point. At such an instant no part of the life of the operating unit has expired and repair has just been completed.

As time passes by, the system will go through many regeneration points. The intervals between these successive regeneration points form a sequence of cycles which are probabilistically equivalent to one another. Let $V_i$ be the length of the $i^{th}$ cycle (the interval between the $i^{th}$ and the $(i+1)^{th}$ regeneration points.) The $V_i$, $i=1,$
2, . . are a sequence of independent and identically distributed random variables and as such form a renewal process. Also, because each cycle is composed of uptime and downtime the system can be described by two renewal processes. Let:

\[ V_i = U_i + D_i, \quad i = 1, 2, \ldots \]

where \( U_i, \quad i = 1, 2, \ldots \) and \( D_i, \quad i = 1, 2, \ldots \) are each renewal processes representing the uptime and downtime during the \( i \)th cycle respectively. Under these conditions, \( V_i \) is said to be an alternating renewal process.

If the system becomes operational at time \( t_i \), breaks down at time \( x \) and becomes operational once again at time \( t_{i+1} \), the cycle \( V_i \) will be as follows:

\[ \begin{align*}
  & t_i \quad x \quad t_{i+1} \\
  & V_i \quad \uparrow \quad U_i \quad \rightarrow \quad D_i \quad \rightarrow 
\end{align*} \]

The interval \((t_i, x)\) is the uptime, \( U_i \), and the interval \((x, t_{i+1})\) is the downtime \( D_i \). The regeneration points occur at times \( t_i \) and \( t_{i+1} \).

In Appendix 1 it is shown that the steady-state availability of the system, \( A \), is given by:

\[ A = \frac{u_c}{\lambda_c + u_c} \]

Another way of expressing this availability is to
put it in terms of the expected uptime of the system $E(U)$ and the expected downtime $E(D)$. Since $1/\lambda_c$ is equal to the mean time to failure and $1/u_c$ is equal to the mean time to repair, $E(U)$ and $E(D)$ are as follows:

$E(U) = 1/\lambda_c$ \hspace{1cm} 3.1.1

$E(D) = 1/u_c$ \hspace{1cm} 3.1.2

Using this notation the system availability may be written as:

$$A = \frac{E(U)}{E(U) + E(D)}$$
3.2 Evaluation of Single Components with Standbys

In this section the performance measures of single components which have redundant standbys will be evaluated. Consider a system consisting of two identical components, a switch, and a repairman. The system is said to be operational if the on-line component is running. However, at any instant the on-line component may fail. When this occurs the switch is used to interchange the failed component with a standby, if one is available, whereby the system remains operational. Repair of the failed component commences immediately and after the component has been repaired it is assumed to be as good as new.

The system will fail if the switch fails or if the standby is undergoing repair when the on-line component fails. If the system goes down because of switch failure then the standby component must be operating (or the switch would not have been attempted.) Hence, the repairman is available and can follow either of two policies. The first policy is for the repairman to fix the switch while it is on-line. In this case the entire workstation would remain down for the amount of time that it takes the repairman to successfully repair the switch. The second policy is for the repairman to manually perform the switching operation and send the switch elsewhere to be repaired.
off-line. Under most circumstances manually interchanging the components would be much faster than fixing the switch. Therefore, the second policy will be used as it will result in greater workstation uptime.

If the system goes down because the standby component is not operating at the time when the on-line component fails then the repairman will be occupied fixing the standby. If the switch is broken then the repairman will manually perform the switching operation at which time the system will come up.

Assuming that the failure times and the repair times are all independent random variables, the instant when the system comes up after having been down is a regeneration point. At such an instant, one of the components has just begun operating, and the repairman has just started fixing the other component. In other words, at a regeneration point no part of the life of the operating component has expired, and no part of the repair time of the failed component has been completed.

In the course of time, the system will go through many regeneration points. The length of the interval between the \( i \)th and the \( i+1 \)th regeneration points, \( V_i \), will be composed of two renewal processes representing the uptime during the \( i \)th cycle, \( U_i \), and the downtime, \( D_i \). Thus, the length of the \( i \)th cycle, \( V_i = U_i + D_i \),
is said to be an alternating renewal process.

The long run availability of the system, \( A \), may be calculated by using the key theorem of renewal theory. This theorem states that if the sequence of \( V_i \), \( i=1,2, \ldots \) are mutually independent the system availability may be expressed as:

\[
A = \frac{E(U)}{E(U) + E(D)}
\]

where \( E(U) \) is the expected system uptime and \( E(D) \) is the expected downtime.

As mentioned in the previous section when no standbys are present and the system consists of only one component, \( E(U) \) is equal to \( \frac{1}{\lambda_c} \), and \( E(D) \) is equal to \( \frac{1}{u_c} \). When standbys are added to the system, however, considerable complexity results, as the following analysis will reveal.

The uptime in a cycle ends after the \( N^{th} \) component failure if switching of the \( N^{th} \) component is unsuccessful or if the repairman has not finished work on the previously failed component. Let:

- \( X_1 = \) The life of the component which started running at the beginning of the cycle.
- \( X_j = \) The life of the component installed after the \( (j-1)^{st} \) component failed, \( j=2, \ldots N \).

and
\[ Y_1 = \text{The repair time which started at the beginning of the cycle.} \]
\[ Y_j = \text{The repair time of the (j-1)st component which began after the (j-1)st component failed,} \]
\[ j=2, 3, \ldots N. \]

The uptime during the cycle is, therefore, \( U = x_1 + x_2 + \ldots x_N. \) Uptime ends at the time of the \( N^{th} \) failure if one of the following two conditions holds:

**Condition 1:** An operating component is not available on standby.
\( (Y_N \geq X_N, \text{repair on the (N-1)st component is not finished at the time of the N}^{th} \text{failure}) \)

**Condition 2:** An operating component is available but switching is unsuccessful.
\( (Y_N < X_N, \text{but the switch is broken}) \)

Assume that the life of a component has the cumulative distribution function \( F(x) \) and the probability density function \( f(x) \). Assume also that the repair time of a failed component has the cumulative distribution function \( G(x) \) and the probability density function \( g(x) \). The steady-state probability, say \( q \), that a repair time is less than the corresponding failure time (for renewal intervals other than the initial interval when both components are known to be in working condition) is:

\[ q = P(Y_j < X_j) = 1 - P(Y_j \geq X_j) \]
\[
\int_0^\infty G(x) dF(x)
\]

Given a probability of switch failure equal to \( p \), the probability of uptime ending at the time of the \( j \)th component failure can be derived as follows:

\[ P(N = 1) = P(\text{uptime ending at the time of the first component failure}) \]
\[ = P(\text{condition 1 doesn't hold or condition 2 doesn't hold}) \]
\[ = 1 - q + q \cdot p \]
\[ = 1 - (1-p) \cdot q \]

\[ P(N = 2) = P(\text{uptime ending at the time of the second component failure}) \]
\[ = P((\text{condition 1 holds and condition 2 holds (for the first component failure)}) \]
\[ \text{and (condition 1 doesn't hold or condition 2 doesn't hold (for the second component failure))}) \]
\[ = ((1-p) \cdot q) \cdot (1 - (1-p) \cdot q) \]

\[ P(N = j) = P(\text{uptime ending at the time of the } j^{\text{th}} \text{ component failure}) \]
\[ = P((\text{condition 1 holds and condition 2 holds (for the first } j-1 \text{ component failures)}) \text{ and (condition 1 doesn't hold or condition 2 doesn't hold (for the } j^{\text{th}} \text{ component failure}))} \]
To find the availability of the system it is first necessary to evaluate the expectation of the uptime; i.e.:

\[ E(U) = E(x_1 + x_2 + \ldots + x_n) \]

Each \( x_c \) has the expectation \( 1/\lambda_c \) (the mean time to failure of a component), and \( N \) is a random variable having the geometric distribution with parameter \((1-p).q\). Moreover, \( N \) is a stopping time and it follows that:

\[ E(U) = E(x_c) \cdot E(N) \]

\[ = \frac{1/\lambda_c}{1 - (1-p).q} \]

Next, it is necessary to evaluate the expectation of the downtime. Downtime results if the repairman has not finished work on the previously failed component or if switching is unsuccessful (conditions 1 or 2 mentioned above). Distinguishing between these two cases yields:

\[ P(\text{lack of an operating system} | \text{downtime}) = \frac{1-q}{1-(1-p).q} \]

\[ P(\text{switching is unsuccessful} | \text{downtime}) = \frac{pq}{1-(1-p).q} \]

If the downtime begins with the lack of an operating unit, that portion of the expected downtime is:
If the downtime begins with a switching failure the expected downtime is:

$$E(D)_2 = \eta$$

where $\eta$ is the expected amount of time it takes the repairman to manually interchange the components.

Combining the previous two expected downtimes and their probabilities of occurring yields the net expected downtime:

$$E(D) = \frac{1-q}{1-(1-p)q} \int_0^\infty \left( \int_0^\infty \frac{(1-G(t+x))}{(1-G(x))} \cdot dt \right) dF(x)$$

$$+ \frac{pq}{1-(1-p)q} \cdot \eta$$

In order to evaluate the expected uptime $E(U)$ and the expected downtime $E(D)$ it is first necessary to define the cumulative distribution functions $F(x)$ and $G(x)$ for the life of a component and the repair time of a component.
Let:
\[ F(x) = 1 - e^{-\lambda_c x}, \quad x \geq 0 \quad \text{and} \]
\[ G(x) = 1 - e^{-\mu_c x}, \quad x \geq 0. \]

The probability \( q \) that a failed component is repaired before the operating component fails may then be evaluated as follows:
\[
q = \int_0^\infty G(x) dF(x)
\]

\[
= \int_0^\infty (1 - e^{-\mu_c x})(\lambda_c e^{-\lambda_c x}) dx
\]

\[
= \left\{ \int_0^\infty e^{-\lambda_c x} dx - \int_0^\infty e^{-(\lambda_c + \mu_c) x} dx \right\} \cdot \lambda_c
\]

\[ = \lambda_c \left( \frac{1}{\lambda_c} - \frac{1}{\lambda_c + \mu_c} \right) \]

\[ = \frac{\mu_c}{\lambda_c + \mu_c} \]

Substituting this value of \( q \) into the formula for \( E(U) \) yields:
\[
E(U) = \frac{1/\lambda_c}{1 - (1-p)q} = \frac{1/\lambda_c}{1 - (1-p)(\mu_c/\lambda_c + \mu_c)}
\]

The expected downtime if downtime begins with the lack of an operating component may be evaluated as:
\[
E(D)_1 = \int_0^\infty \left( \int_0^\infty \frac{1 - G(t+x)}{1 - G(x)} \cdot dP(x) \right) \, dt
\]
\[
= \int_0^\infty \left( \int_0^\infty \frac{e^{-u_c(t+x)} e^{-\lambda_c x}}{e^{-u_c(x)}} \cdot dx \right) \, dt
\]
\[
= \lambda_c \int_0^\infty \left( \int_0^\infty e^{-u_c x - \lambda_c x} \cdot dx \right) \, dt
\]
\[
= \frac{1}{u_c}
\]

Substituting this value of \(E(D)_1\) and the value obtained for \(q\) into the formula obtained for \(E(D)\) yields:

\[
E(D) = \frac{1 - q}{1 - (1-p) \cdot q} \left( \frac{1}{u_c} \right) + \frac{pq}{1 - (1-p) \cdot q} \left( \eta \right)
\]
3.3 Calculation of Workstation Performance Measures

Each workstation \( w, w=1,2 \), can be modelled as a series system consisting of \( n_w \) subsystems arranged in a series. The subsystems are one of two types: those with a standby and those without a standby. In order to distinguish between the two types of subsystems the following notation will now be introduced, let:

\[ h_w(s) = \text{The standby policy for subsystem } s, s=1,2, \ldots, n_w, \text{ of workstation } w, w=1,2. \]

where

\[ h_w(s) = 0, \text{ if the subsystem } s \text{ does not have a standby and } h_w(s) = 1, \text{ otherwise.} \]

If \( h_w(s) = 0 \) the subsystem \( s \) consists of a single component. According to equations 3.1.1 and 3.1.2 the expected uptime and the expected downtime of the subsystem are as follows:

\[ E(U)_s = 1/\lambda_s \]
\[ E(D)_s = 1/\mu_s \]

If \( h_w(s) = 1 \) the subsystem \( s \) consists of a single component with a standby. The expected uptime and the expected downtime of the subsystem are given by equations 3.2.1 and 3.2.2:
\[ E(U)_s = \frac{1/\lambda_c}{1 - (1-p)(q)} \]
\[ E(D)_s = \frac{1-q}{1 - (1-p)(q)} \left( \frac{1}{u_c} \right) + \frac{pq}{1 - (1-p)(q)} \left( \eta \right) \]

where

\[ q = 1 - \lambda_c (\lambda_c + u_c) \]

The performance measures of the workstations may now be evaluated. Because redundancy is incorporated in the transfer line, the failure rate \( \lambda_w \), repair rate \( u_w \), and availability \( A_w \) for each workstation \( w, w=1,2 \), will depend upon the standby policy \( h_w(s) \). Therefore, the formulas for \( \lambda_w, u_w \), and \( A_w \) derived by Barlow and Proschan (1975) will be given by:

\[ \lambda_w(h_w(s)) = \sum_{s=1}^{n_w} \frac{1}{E(U)_s} \quad \text{3.3.1} \]

\[ u_w(h_w(s)) = \lambda_w(h_w(s)) \left( \frac{\sum_{s=1}^{n_w} E(D)_s / E(U)_s}{E(U)_s} \right)^{-1} \]

\[ A_w(h_w(s)) = \frac{u_w(h_w(s))}{(u_w(h_w(s)) + \lambda_w(h_w(s)))} \quad \text{3.3.3} \]

The impact of individual subsystems on the performance of a workstation may be determined by evaluating \( D_{sw} \), the long run fraction of time that the workstation \( w \) is down due to the failure of subsystem \( s \). \( D_{sw} \) may be computed as:

\[ - 43 - \]
\[ D_{sw} = \left( \frac{E(D)}{E(U)} \right) A_w \]
4.0 Analysis of the Transfer Line

As seen in Figure 5, the transfer line consists of two workstations with an interstation storage queue. By using the techniques presented in the previous section the net failure rate, repair rate, and availability of each workstation can be calculated for any component redundancy level. In order to determine how much redundancy, if any, should be incorporated into the system a method is needed whereby the production rate of the transfer line is found in terms of the performance of the workstations and the buffer storage capacity.
4.1 Evaluation of the Net Transfer Line Production Rate

The model proposed by Malathronas, Perkins, and Smith assumes that the transfer line is "balanced". In other words, the nominal production rate of workstation 1, \( R_{c1} \), is equal to the nominal production rate of workstation 2, \( R_{c2} \).

Malathronas, Perkins, and Smith also assume that the formula for the net production rate of the transfer line is a function of only the buffer capacity, \( b \). Because the model presented in this thesis takes into account both the buffer storage capacity and the standby policy, \( h_w(s) \), the notation for the net production rate of the transfer line will be changed from \( R(b) \) to \( R(b, h_w(s)) \).

The formula for \( \hat{R}(b, h_w(s)) \) presented by Malathronas, Perkins, and Smith is:

\[
\hat{R}(b, h_w(s)) = \left( \frac{\lambda_1}{u_1} - \frac{\lambda_2}{u_2} e^{-k} \right) \frac{A_1 A_2 R_{cw}}{\frac{\lambda_1}{u_1} A_2 - \frac{\lambda_2}{u_2} A_1 e^{-k}}
\]

where

\[
k = \frac{(u_1 + u_2 + \lambda_1 + \lambda_2)(\lambda_1 u_2 - \lambda_2 u_1) b}{(u_1 + u_2)(\lambda_1 + \lambda_2) R_{cw}}
\]

In this formula, \( A_1 \) and \( A_2 \) represent the availabilities...
of workstations 1 and 2 respectively. The availability of workstation \( w \), \( A_w \), is a function of the standby policy \( h_w(s) \). As shown in section 3.1, \( A_w \) may be written as:

\[
A_w = A_w(h_w(s)) = \frac{u_w(h_w(s))}{u_w(h_w(s)) + \lambda_w(h_w(s))}
\]

where \( u_w(h_w(s)) \) is the repair rate of workstation \( w \) and \( \lambda_w(h_w(s)) \) is the failure rate of workstation \( w \). The values for \( u_w(h_w(s)) \) and \( \lambda_w(h_w(s)) \) can be calculated by using 3.3.1 and 3.3.2.

The validity of equation 4.1 can be partially checked by computing its limiting behavior for zero buffer capacity \( (b=0) \) and infinite buffer capacity \( (b=\infty) \).

When the buffer has zero capacity, \( k=0 \) and \( e^{-k}=1 \). Substituting \( e^{-k}=1 \) into equation 4.1 yields:

\[
R(0,h_w(s)) = \left( \frac{\lambda_1 - \lambda_2}{u_1} \frac{e^{-k}}{u_2} \right) A_1A_2 R_{cw}
\]

\[
= \frac{R_{cw}}{\lambda_1 + \lambda_2 + 1}
\]

In order to completely understand equation 4.1.1 we shall introduce some new notation. Let:

\( T_{cw} = \) The nominal time required to transfer plus the processing time at workstation \( w \) (when
\[ F_w = \text{The frequency which workstation } w \text{ fails per cycle.} \]
\[ T_{dw} = \text{The average downtime required to diagnose a problem and make repairs when a failure occurs at workstation } w. \]

Using the variables \( T_{cw}, F_w, \) and \( T_{dw}, \) we may redefine \( R_{cw}, \lambda_w, \) and \( u_w \) as follows:

\[ R_{cw} = \frac{1}{T_{cw}} \]
\[ \lambda_w = \frac{F_w}{T_{cw}} \]
\[ u_w = \frac{1}{T_{dw}} \]

Substitution of the three previous equations into equation 4.1.1 yields:

\[ R(0,h_w(s)) = \frac{1}{T_{cw} + F_{1}T_{d1} + F_{2}T_{d2}} \]

Because there is no buffer storage, the transfer line will be forced down if either workstation 1 or workstation 2 fails. Therefore, the production rate can be found by taking the reciprocal of the time required to process a workpiece when the line is up and running, \( T_{cw}, \) plus the expected downtimes from workstations 1 and 2, \( F_{1}T_{d1} \) and \( F_{2}T_{d2}. \) The equation for \( R(0,h_w(s)) \) is, thus, correct.

When the buffer capacity becomes infinite the value of \( k \) depends on the availabilities of workstations 1 and 2
2, \( A_1 \) and \( A_2 \). This can be seen by rewriting the expression for \( k \):

\[
k = \frac{(u_1 + u_2 + \lambda_1 + \lambda_2)(u_2/u_1)(1/A_1 - 1/A_2)b}{(u_1 + u_2)(\lambda_1 + \lambda_2)R_{cw}}
\]

If \( A_2 > A_1 \) then \( k > 0 \) and the \( \lim_{b \to \infty} e^{-k} = 0 \). Therefore, the production rate of the line, \( R(\infty, h_w(s)) \), becomes:

\[
R(\infty, h_w(s)) = \frac{(\lambda_1/u_1)A_1A_2R_{cw}}{(\lambda_1/u_1)A_2} = A_1R_{cw}
\]

If \( A_1 > A_2 \) then \( k > 0 \) and the \( \lim_{b \to \infty} e^k = 0 \). Multiplying the numerator and denominator of equation 4.1 by \( e^k \) yields:

\[
R(\infty, h_w(s)) = -\frac{(\lambda_2/u_2)A_1A_2R_{cw}}{(\lambda_2/u_2)A_1} = A_2R_{cw}
\]

When the previous two expressions are combined the formula for \( R(\infty, h_w(s)) \) becomes:

\[
R(\infty, h_w(s)) = \min(A_1R_{c1}, A_2R_{c2})
\]

This result is as expected. When the buffer storage capacity approaches infinity the workstations become decoupled. Hence, the workstation with the lowest availability is a bottleneck limiting the production rate of the line.
5.0 Economic Considerations

When a system designer is presented with the option of using component standbys and/or buffer storage to increase the net production rate of a transfer line the most important decision criterion will undoubtedly be economics. Obviously, component redundancy reduces workstation downtime and, thus, increases the transfer line production rate. Likewise, buffer storage capacity also improves the output of the system. What is needed, then, is a method of comparing the relative effects of each of these two devices along with the costs associated with their implementation. In this manner, the most economical means of increasing the production rate of the transfer line can be chosen.
5.1 Derivation of the Total Relevant Cost per Workpiece Produced

In order to accurately describe the various costs within the transfer line system some additional definitions and notation will now be introduced. Let:

\[ R(b, h_w(s)) = \text{The net production rate of the transfer line when the buffer storage capacity is } b \text{ and the standby policy for subsystems, } s=1,2, \ldots n, \text{ in workstations } w=1,2 \text{ is } h_w(s). \text{ The dimensions are workpieces/unit time.} \]

\[ r_1 = \text{The cost of having one dollar tied up in inventory for a unit time interval. The dimensions are } \$/\text{unit time.} \]

\[ v = \text{The value of a workpiece in terms of raw materials and the value added through the processing operations at workstations 1 and 2. The dimensions are } \$/\text{workpiece.} \]

\[ b = \text{The buffer storage capacity.} \]

\[ b = \text{The average number of workpieces stored in the buffer.} \]

\[ r_2 = \text{The cost of allocating a workpiece space in buffer storage for one unit time interval. The dimensions are } \$/\text{workpiece/unit time.} \]

\[ C_{sw} = \text{The cost incurred per unit time interval from installing a standby in subsystem} \]
The sum of the costs incurred by installing standbys along the transfer line when the standby policy for subsystems $s=1,2,\ldots,n$ in workstations $w=1,2$ is $h_w(s)$. The dimensions are $\$/unit time.

$F =$ The sum of the fixed costs incurred per unit time interval from operating the transfer line. The dimensions are $\$/unit time.

The total relevant cost per workpiece produced is the sum of the costs incurred from operating the transfer line (per unit time) divided by the number of workpieces produced (per unit time). There are several different types of transfer line costs which must be taken into account: standby, inventory, and fixed.

The standby cost, $C(h_w(s))$, is the total cost of installing the standbys specified by the standby policy $h_w(s)$. Of course, if no standbys are used in either workstation, this cost will be equal to zero.

The inventory costs are those costs which are incurred as a result of storing workpieces in the buffer. There are two types of inventory costs: space and carrying. The space cost is that cost which results from allocating workpieces space in the buffer storage zone and is equal to $r_2b$. This cost can be quite large, especially if the workpieces are large or if the buffer storage device
is expensive to build and/or to operate.

The second type of inventory cost is the expense which results from having the dollar value of the workpieces tied up in inventory. This cost, called the carrying cost, depends upon the average number of workpieces stored in the buffer and is equal to $r_1 b v$.

Finally, the fixed cost, $F$, measures all of the transfer line costs which are not a function of either the buffer capacity or the standby policy. The fixed cost should take into account the maintenance, power, and equipment expenditures required to keep the transfer line system operational.

Dividing the sum of the transfer line costs by the number of workpieces produced yields the total relevant cost per unit produced. This can be written as:

$$ TRC(b, h_w(s)) = \frac{1}{R(b, h_w(s))} \left( r_2 b + r_1 b v + C(h_w(s)) + F \right) $$

where $R(b, h_w(s))$ is calculated by using equation 4.1.

The average number of workpieces stored in the buffer must be a fraction of the buffer capacity which is between zero and one. Letting $f$ be this fraction, $\bar{b} = fb$, and:

$$ TRC(b, h_w(s)) = \frac{1}{R(b, h_w(s))} \left( r_2 b + r_1 fb v + C(h_w(s)) + F \right) $$
In order to find an approximation for the value of \( f \) the following notation will now be introduced:

\( Y_t = \) A random variable which denotes the fraction of the buffer which is filled at any time \( t \).

\( \pi_1(y) = \) The steady state probability that both workstations are down and \( 0 < Y_t < 1 \).

\( \pi_2(y) = \) The steady state probability that workstation 1 is down, workstation 2 is up, and \( 0 < Y_t < 1 \).

\( \pi_3(y) = \) The steady state probability that workstation 1 is up, workstation 2 is down, and \( 0 < Y_t < 1 \).

\( \pi_4(y) = \) The steady state probability that both workstations are up and \( 0 < Y_t < 1 \).

\( \pi_5(y) = \) The steady state probability that workstation 1 is down, workstation 2 is up, and \( Y_t = 0 \).

\( \pi_6(y) = \) The steady state probability that both workstations are up and \( Y_t = 0 \).

\( \pi_7(y) = \) The steady state probability that workstation 1 is up, workstation 2 is down, and \( Y_t = 1 \).

\( \pi_8(y) = \) The steady state probability that both workstations are up and \( Y_t = 1 \).

The expected value of \( Y_t \), \( E(Y_t) \), is equal to the mean value of the fraction of workpieces stored in the buffer at any given time. \( E(Y_t) \) can be written as:

\[
E(Y_t) = \int_0^1 y \left[ \sum_{i=1}^4 \pi_i(y) \right] dy + 0(\pi_5(y) + \pi_6(y)) + 1(\pi_7(y) + \pi_8(y))
\]

- 54 -
As derived by Malathronas, Perkins, and Smith ( ), the steady state probabilities $\pi_i(y), i=1,2, \ldots 8$, are as follows:

$$
\pi_1(y) = \frac{(\lambda_1 + \lambda_2) \pi_0 e^{-ky}}{(u_1 + u_2)} \\
\pi_2(y) = \pi_0 \\
\pi_3(y) = \pi_0 e^{-k} \\
\pi_4(y) = \frac{(u_1 + u_2) \pi_0 e^{-ky}}{(\lambda_1 + \lambda_2)} \\
\pi_5(y) = \frac{(\lambda_1 + \lambda_2) R_c \pi_0}{u_1 \lambda_2} \\
\pi_6(y) = \frac{R_c \pi_0}{b \lambda_2} \\
\pi_7(y) = \frac{(\lambda_1 + \lambda_2) R_c \pi_0 e^{-k}}{b \lambda_1 u_2} \\
\pi_8(y) = \frac{R_c \pi_0 e^{-k}}{b \lambda_1}
$$

where

$$
\pi_0 = \left[ \frac{(1-e^{-k}) \beta \delta}{k} + \frac{R_c}{b} \left( \frac{1}{u_1} + \frac{e^{-k}}{u_2} \right) \right] \\
+ \frac{R_c}{b} \left( \frac{\lambda_1}{u_1 \lambda_2} + \frac{\lambda_2 e^{-k}}{u_2 \lambda_1} \right)
$$
Substituting these equations into equation 5.1.1, letting \( f \) equal \( E(Y_t) \), and simplifying yields:

\[
f = E(Y_t) = e^{-k} \pi_0 \left( \frac{-k}{k^2} \left( \frac{(\lambda_1 + \lambda_2) + (u_1 + u_2)}{(u_1 + u_2)(\lambda_1 + \lambda_2)} \right) \right)
\]

\[
+ \frac{1}{2} + \frac{R_c}{b \lambda_1} \left( \frac{(\lambda_1 + \lambda_2)}{u_2} + 1 \right)
\]

\[
+ \frac{\pi_0}{2}
\]

5.1.2

It should be noted that because equation 5.1.2 is based upon the continuous system analysis of Malathronas, Perkins, and Smith it will yield an approximate value for \( f \). The system which we are modelling produces individual workpieces and is, hence, discrete.
5.2 Selection of the Buffer-Standby Policy

In this section, a buffer-standby policy will be chosen such that the total relevant cost per workpiece produced is minimized. As shown in equation 4.1, the total relevant cost per unit produced is a function of two decision variables: the buffer storage capacity \( b \), and the standby policy \( h_w(s) \). The first step towards minimizing \( TRC(b, h_w(s)) \) will be to optimize \( b \) for fixed values of \( h_w(s) \). Thus, the problem is reduced to one of a single decision variable, the standby policy \( h_w(s) \).

The second and final step towards minimizing \( TRC(b, h_w(s)) \) will be to find the optimum standby policy \( h_w^*(s) \). A branch and bound method is presented whereby only those standby policies which have the potential of lowering the total relevant cost are considered as candidates for the optimum standby policy.
5.2.1 Calculation of the Optimum Buffer Size

In this section a method for calculating the optimum buffer size, \( b^* \), for a fixed standby policy, \( h_w(s) \), will be derived.

The optimum value of the buffer size, \( b^* \), can be found by setting the partial derivative of the cost function \( TRC(b, h_w(s)) \) equal to zero and then solving for \( b^* \). The partial derivative of equation 5.1 with respect to the buffer storage capacity \( b \) is given by:

\[
\frac{\partial TRC(b, h_w(s))}{\partial b} = \frac{\partial (1/R(b, h_w(s)))(r_2 b + r_1 f b v + C(h_w(s)) + P)}{\partial b} + \frac{(1/R(b, h_w(s)))(r_2 + r_1 \frac{\partial f}{\partial b})}{\partial b}
\]

where

\[
\frac{\partial (1/R(b, h_w(s)))}{\partial b} = \frac{(\lambda_1/u_1)A_2 - (\lambda_2/u_2)A_1 e^{-k}}{((\lambda_1/u_1) - (\lambda_2/u_2) e^{-k})A_1 A_2 R_c}
\]

and

\[
\frac{\partial k}{\partial b} = \frac{(u_1 + u_2 + \lambda_1 + \lambda_2)(\lambda_1 u_2 - \lambda_2 u_1)}{(u_1 + u_2)(\lambda_1 + \lambda_2) R_{cw}}
\]

Simplifying the expression for \( TRC(b, h_w(s))/\partial b \) yields:

- 58 -
\[
\frac{\partial \text{TRC}(b, h_w(s))}{\partial b} = \frac{1}{\left(\frac{\lambda_1}{u_1} - \frac{\lambda_2}{u_2} e^{-k}\right)^{A_1 A_2 R_c w} + \left(\frac{\lambda_1}{u_1} - \frac{\lambda_2}{u_2} e^{-k}\right)^{A_1 A_2 R_c w}} \left\{ \begin{array}{c}
- \frac{\lambda_2}{u_2} e^{-k} \frac{\partial}{\partial b} A_1 A_2 R_c \frac{\lambda_1}{u_1} (A_1 - A_2) \\
+ \left(\frac{\lambda_1}{u_1} A_2 e^{-k} - \frac{\lambda_2}{u_2} A_1 e^{-k}\right) \left(r_2 + r_1 v_b \frac{\partial f}{\partial b} + r_1 f_v\right) \end{array} \right. \\
\left(\frac{\lambda_1}{u_1} - \frac{\lambda_2}{u_2} e^{-k}\right)^{A_1 A_2 R_c w}
\right\}
\]

where

\[
\frac{\partial f}{\partial b} = - \frac{\partial}{\partial b} e^{-k} \pi_0 \left\{ - (k+1) \left(\frac{\lambda_1 + \lambda_2}{k^2} \left(\frac{u_1 + u_2}{\lambda_1 + \lambda_2}\right) \right) \left(\frac{\lambda_1 + \lambda_2}{u_2} \right) + \right. \\
+ \frac{1}{2} + \frac{R_c w}{b} \left(\frac{\lambda_1 + \lambda_2}{u_2} + 1\right) \left(\frac{\lambda_1 + \lambda_2}{u_2} \right) \left(\frac{\lambda_1 + \lambda_2}{u_2} \right) + \\
\left. e^{-k} \pi_0 \left(2 \frac{\partial k}{\partial b} \left(\frac{k+1}{k^3}\right) - \frac{\partial k}{\partial b} \frac{k}{k^2}\left(\frac{\lambda_1 + \lambda_2}{u_1 + u_2}\right) \right) \right.
\]

- 59 -
\[ + \frac{(u_1 + u_2)}{(\lambda_1 + \lambda_2)} - \frac{R_c}{b^2 \lambda_1} \left( \frac{(\lambda_1 + \lambda_2)}{u_2} + 1 \right) \]

\[ = e^{-k} \sum_{0} \left\{ \frac{\partial k}{\partial b} \left( \frac{(k+1)}{k^2} + \frac{2}{k^3} \right) \left( \frac{1}{k^2} \right) \left( \frac{(\lambda_1 + \lambda_2)}{(u_1 + u_2)} + \frac{(u_1 + u_2)}{(\lambda_1 + \lambda_2)} + \frac{1}{2} + \frac{R_c}{b \lambda_1} \left( \frac{(1-1/b)}{u_2} \right) \right) \right\} \]

Next, it is necessary to set the partial derivative of \( TRC(b, h_w(s)) \) equal to zero:

\[ \frac{\partial TRC(b, h_w(s))}{\partial b} = 0 \]

\[ \frac{-\lambda_2 e^{-k} \frac{\partial k}{\partial b} \lambda_1 (A_1 = A_2)}{u_2 \frac{\partial b}{u_1}} (r_2 b^* + r_1 b^* f v + C(h_w(s)) + F) \]

\[ = \frac{\lambda_1 - \lambda_2 e^{-k}}{u_1 u_2} \left( r_2 + r_1 v b \frac{\partial f}{\partial b} + r_1 f v \right) \]

Using equation 5.2.1 the optimum value of the buffer size, \( b^* \), can be found by substituting the values of
\( \lambda_1, \lambda_2, A_1, A_2, \) and \( C(h_w(s)) \) which correspond to a fixed standby policy \( h_w(s) \) and then using a root finding technique such as the Newton-Raphson procedure to solve for \( b^* \). Therefore, the task of minimizing the total relevant cost per workpiece produced is reduced to choosing the optimum standby policy \( h_w^*(s) \).
5.2.2 A Branch and Bound Method to Find the Optimum Standby Policy

It was shown in section 5.2.1 that for any fixed standby policy, \( h_w(s) \), the optimum buffer storage size, \( b^* \), for that particular standby policy can be calculated. We now want to isolate the optimum standby policy \( h_w^*(s) \). This standby policy and its corresponding optimum buffer storage size constitute the overall optimum buffer-standby policy for the transfer line.

One method of isolating the optimum standby policy is to use exhaustive enumeration. The optimum buffer size for every possible standby policy can be calculated. Then, using equation 5.1.1 the total relevant cost associated with each buffer-standby policy can be found. Finally, the optimal standby policy, \( h_w^*(s) \), and the best buffer storage size, \( b^* \), correspond to the buffer-standby policy which has the minimum total relevant cost per unit produced.

The method of exhaustive enumeration, however, may be extremely tedious, especially if the number of standby options is large. If workstation 1 has \( n_1 \) subsystems and workstation 2 has \( n_2 \) subsystems then the maximum number of ways of placing \( r \) standbys along the transfer line, disregarding order, is given by:
\[ \sum_{r=0}^{n_1+n_2} \binom{n_1+n_2}{r} \]

For example, if workstation 1 has 3 subsystems and workstation 2 has 5 subsystems then the maximum number of standby options is:

\[ \sum_{r=0}^{3+5} \binom{3+5}{r} = \binom{8}{0} + \binom{8}{1} + \binom{8}{2} + \binom{8}{3} + \binom{8}{4} + \binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8} \]

\[ = 1 + 8 + 28 + 56 + 70 + 56 + 28 + 8 + 1 \]

\[ = 256 \]

That is, if all 8 subsystems along the transfer line are reconfigurable then 256 different standbys policies must be investigated.

Essentially, exhaustive enumeration requires that every standby policy must be investigated regardless of whether or not it has any potential for improving the cost equation. Therefore, although this method is theoretically simple and will always find the optimal solution, it may result in considerable wasted effort.

Another method of finding the optimal buffer-standby policy is to use a branch and bound approach. The branch and bound approach makes use of information contained in the cost equation to create bounds on the solution.
Bounds are useful because they can trim a large problem down to a manageable size. Bounds tell us which policies cannot possibly improve the cost of producing a unit, and, therefore, shouldn't be investigated.

Figure 6 shows a schematic representation of the branch and bound decision tree for isolating the optimum standby policy. At each branch of the tree a decision is made whether or not to add a standby to the transfer line. For example, at the top of the tree there are two initial choices: don't add a standby to subsystem 1 of workstation 1 ($h_1(1)=0$), or do add a standby to subsystem 1 of workstation 1 ($h_1(1)=1$). As one descends down the tree additional decisions are made until, finally, at the root of the tree all $h_w(s)$, $w=1,2$, $s=1,2 \ldots n_w$, have been set. Hence, a complete standby policy results when a branch of the tree has been fathomed. If every branch were fathomed, all standby policies would be investigated, as in the previous method of exhaustive enumeration.

The bounds on the tree will be defined as follows:

$$\text{Bound}(h_w(s)) = \text{The minimum total cost per unit produced with the standby policy } h_w(s), w=1,2, s=1,2 \ldots n_w.$$ 

In other words, the bound will represent the lowest possible cost, $\text{TRC}(b,h_w(s))$, that can be obtained when
Figure 6

Branch and Bound Decision Tree
the standby policy \( h_w(s) \) is used.

The total relevant cost per unit produced was derived in section 5.1 and is as follows:

\[
TRC(b, h_w(s)) = \frac{1}{R(b, h_w(s))} (\sigma T_2 b + r_1 fbv + C(h_w(s)) + F)
\]

It may be observed that \( TRC(b, h_w(s)) \) will be minimized if these conditions are satisfied:

1) The production rate \( R(b, h_w(s)) \) is maximized.
2) The inventory and storage costs associated with the buffer capacity \( r_2 b \) and \( r_1 fbv \) are minimized.
3) The sum of the standby costs, \( C(h_w(s)) \), is minimized.

In order to maximize the production rate we note that the more standbys which are present along the transfer line the better the line will operate. Therefore, when calculating the production rate we should assume a standby policy whereby standbys are added to all of the subsystems which have not already been set (i.e., all those located below the bound on the decision tree). The optimum buffer capacity which maximizes the production rate can be found by substituting the standby policy with the most standbys possible into equation 5.2.1 and then solving for \( b^* \).

In order to minimize the cost associated with the buffer capacity we note that if there is no buffer capacity then the buffer cost is equal to zero. Therefore, when
computing the inventory and storage costs we should assume that the buffer capacity is zero. At first, it may seem inconsistent to assume that \( b \) is equal to zero when calculating the buffer costs while assuming a different buffer capacity when maximizing the production rate. After all, in real life the buffer capacity must be fixed. We must remember, however, that when calculating bounds we need not create actual buffer-standby policies. The bound \( \text{Bound}(h_w(s)) \) simply represents the lowest total cost per unit produced that can hypothetically be achieved with the standby policy \( h_w(s) \).

In order to minimize the sum of the standby costs, \( C(h_w(s)) \), we note that the standby costs will be lowest if as few standbys are allocated along the transfer line as possible. Therefore, when calculating the sum of the standby costs we should assume a standby policy whereby no standbys are added to any of the subsystems which have not already been set. Only the costs from those standbys which have already been assigned will be summed. It should be noted that the standby policy \( h_w(s) \) used in calculating \( C(h_w(s)) \) is not the same standby policy as that used in calculating \( R(b,h_w(s)) \). Again, this is okay because we are only attempting to find a bound.

The procedure for calculating bounds is summarized as follows:
\[ \text{Bound}(h_w(s)) = \frac{1}{R(b, h_w(s))_{\text{max}}} \left( (r_2 b + r_1 fbv)_{\text{min}} + C(h_w(s))_{\text{min}} + F \right) \]

\[ = \frac{1}{R(b^*, h_w(s)_1)} \left( 0 + C(h_w(s)_2) + F \right) \]

where

- \( h_w(s)_1 \) = The standby policy whereby standbys are added to all of the subsystems located below the bound on the decision tree.

and

- \( h_w(s)_2 \) = The standby policy whereby standbys are not added to any of the subsystems located below the bound on the decision tree.
6.0 Example Problem

The analytical techniques presented in this investigation will now be illustrated by an example problem.

Consider a transfer line which has two workstations separated by a buffer storage queue. Workstation 1 consists of four components and workstation 2 consists of two components. The failure rate and the repair rate of each component is as follows:

<table>
<thead>
<tr>
<th>Workstation</th>
<th>Component</th>
<th>Failures per hr.</th>
<th>Repairs per hr.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\lambda_c$</td>
<td>$u_c$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>.008</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>.002</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>.010</td>
<td>.25</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>.010</td>
<td>.5</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>.008</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>.010</td>
<td>.5</td>
</tr>
</tbody>
</table>

Suppose that components 1, 3, and 4 within workstation 1 are reconfigurable. Suppose, also, that component 2 within workstation 2 is reconfigurable. The cost incurred per hour from installing a standby for each of these reconfigurable components is as follows:
<table>
<thead>
<tr>
<th>Workstation</th>
<th>Component Subsystem</th>
<th>Standby Cost per hr. ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>s</td>
<td>C_{sw}</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>(not reconfigurable)</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>(not reconfigurable)</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>.5</td>
</tr>
</tbody>
</table>

The probability that the switch which interchanges an on-line component and its standby fails is .02.
If the switch is broken the expected time it will take a repairman to manually interchange the components is .05 hr.

In this system, the cost of having one dollar tied up in inventory for a year is .20 $ (.0001 $/hr assuming 40 hrs/wk and 52 wks/yr). Likewise, the cost of allocating a workpiece space in buffer storage is 3.0 $/yr (.00144 $/hr). The nominal production rate of each workstation is 240 units/hr and the value added to each workpiece at workstations 1 and 2 is 1 $. The fixed cost incurred from running the transfer line is 30 $/hr.

We want to determine the buffer storage capacity b, and the standby policy h_w(s) which minimizes the total relevant cost per workpiece produced.
As mentioned previously, there are two methods of finding the most cost-effective buffer-standby policy: exhaustive enumeration and the branch and bound technique. We will not use the exhaustive enumeration method because it may not be necessary to explore every possible standby policy. By using the branch and bound technique only those standby policies which have the potential of lowering the total relevant cost will be explored.

The decision tree for the problem under consideration is shown in Figure 7. At each node of the tree a decision is made whether or not to add a standby to a component subsystem. If a standby is added to subsystem $s$ in workstation $w$ then $h_w(s)=1$. Otherwise, $h_w(s)=0$. It should be noted that there are 16 possible standby policies. Thus, if exhaustive enumeration were used 16 different standby policies would have to be explored.

As a first step in applying the branch and bound technique to the decision tree we will completely fathom one branch of the tree. Any branch may be chosen, but it is a good idea to make an "educated guess" and pick the branch which we think will yield the lowest total relevant cost per workpiece produced. The reason for making such a choice is that the total relevant cost per workpiece produced corresponding to the first fully fathomed branch is our initial "best" standby policy.
Figure 7
Decision Tree for the Example Problem
Only those branches of the tree which have bounds that are less than the cost associated with the initial guess have the potential of yielding a superior standby policy. Hence, the better our first fully fathomed solution is, the fewer branches of the tree we will need to explore.

For the first branch we will choose a standby policy where no standbys are added to the transfer line system. This is usually a suitable first guess especially when the cost of standbys is high. Using equations 3.3.1, 3.3.2, and 3.3.3 the failure rates, repair rates, and availabilities of workstations 1 and 2 can be calculated for the standby policy $h_w(s)=0$, $w=1,2$, $s=1,2, \ldots n_w$:

\[
\begin{align*}
\lambda_1 &= 0.008 + 0.002 + 0.010 + 0.010 = 0.030 \\
\lambda_2 &= 0.008 + 0.010 = 0.018 \\
u_1 &= 0.03(0.5/125 + 0.2/500 + 4/100 + 2/100)^{-1} = 0.4658 \\
u_2 &= 0.018(0.5/125 + 0.2/100)^{-1} = 0.750 \\
A_1 &= 0.4658/0.4958 = 0.9395 \\
A_2 &= 0.750/0.768 = 0.9766
\end{align*}
\]

The net production rate can be found by substituting the previous values into equation 4.1:

\[
R(b, h_w(s)) = \frac{0.0644 - 0.024e^{-k}}{0.06289 - 0.0225e^{-k}} \frac{(240)(.9395)(.9766)}{}
\]
where

\[
k = \frac{(0.4658 + 0.75 + 0.03 + 0.018)(0.03(0.75) - 0.018(0.4658))b}{(0.4658 + 0.75)(0.03 + 0.018)(240)}
\]

We now wish to find the value of \( b \) which minimizes the total relevant cost per workpiece produced. As shown in equation 5.1 the total relevant cost is as follows:

\[
TRC(b, h_w(s)) = \frac{1}{R(b, h_w(s))} (0.00144b + 0.001fb + 30)
\]

Because the inventory cost associated with the average number of workpieces stored in the buffer, 0.0001fb, is negligible we can ignore this cost, and, hence, spare ourselves the task of using equation 5.1.2 to calculate \( f \). Substituting the value of \( R(b, h_w(s)) \) into the previous equation and minimizing it to find the optimum value of \( b \) yields:

\[
b^* = 10 \text{ workpieces} \quad \text{and,}
\]

\[
TRC(b^*, h_w(s) = 0, w = 1, 2, s = 1, 2, \ldots n_w) = 0.136032
\]

The next step of the branch and bound technique is to backtrack up the decision tree and explore other standby policies. Looking at Figure 7, we see that the next branch to be explored is the standby policy \( h_1(s) = 0, s = 1, 2, 3, 4, h_2(1) = 0, h_2(2) = 1 \). The failure rate, repair rate, and availability of workstation 2 will change as a result of the addition of a standby for component 2.
Using equations 3.2.1 and 3.2.2 the expected uptime and the expected downtime of the component subsystem 2 are as follows:

\[ q = \frac{0.5}{0.5} = 0.998 \]

\[ E(U) = \frac{1/0.01}{1-(1-0.02)(0.998)} = 4553.734 \]

\[ E(D) = \frac{1-0.998 + 0.02(0.998)(0.05)}{0.02196 \cdot 0.02196} = 0.22759 \]

These values can then be substituted into equations 3.3.1, 3.3.2, and 3.3.3 to yield:

\[ \lambda_2 = \frac{1}{125} + \frac{1}{4553.7} = 0.0082196 \]

\[ u_2 = 0.0082196(0.228/4553.7 + 0.5/125)^{-1} = 2.0295 \]

\[ A_2 = 2.0295/2.0377 = 0.99597 = 0.9960 \]

The net production rate results from substituting the \( \lambda_1, u_1, \) and \( A_1 \) found for the case when no standbys are in workstation 1, along with the previous three values into equation 4.1:

\[ R(b, h_w(s)) = \frac{0.0644 - 0.00405e^{-k}}{0.06414 - 0.003805e^{-k}} (224.5713) \]
The total relevant cost per workpiece produced is given by equation 5.1:

$$TRC(b, h_w(s)) = \frac{1}{R(b, h_w(s))} (0.00144b + 0.00001fb + 30 + 0.5)$$

Substituting the value of $R(b, h_w(s))$ into the previous equation and minimizing it to find the optimum value of $b$ yields:

$b = 17$ workpieces and,

$$TRC(b, h_1(s) = 0, s=1,2,3,4, h_2(1) = 0, h_2(2) = 1) = .13583$$

We now have a new lowest total relevant cost per workpiece produced. Therefore we should temporarily add a standby to component 2 of workstation 2 and proceed to backtrack further up the decision tree.

The next step of the analysis is to calculate the bound, $Bound(h_1(1) = 0, . . h_1(3) = 0, h_1(4) = 1, h_2(1), h_2(2)$ determined by the bounding procedure). In order to find this bound we must find the maximum production rate that can be achieved given the bound's standby policy. Using equation 4.1 we get:

$$R(b^*, h_1(1) = 0, . . h_1(3) = 0, h_1(4) = 1, h_2(1) = 0, h_2(2) = 1) = 226.732 \text{ workpieces/hr}$$
Now, we wish to minimize the costs associated with the bound's standby policy. Setting the inventory costs $r_2 b$ and $r_1 f b v$ equal to zero, and assuming the only standby cost is that associated with component 4 of workstation 1 yields the bound:

$$\text{Bound}(h_1(1)=0, \ldots h_1(3)=0, h_1(4)=1, h_2(1), h_2(2) \text{ determined by the bounding procedure})$$

$$= (1/226.732)(0 + 1 + 30)$$

$$= .13673$$

Therefore, the lowest cost per workpiece that could be achieved by proceeding down from the $h_1(4)=1$ is .13673. Because this cost is greater than the lowest cost obtained thus far (.13583) we should not explore this branch further.

The next bounds obtained are $\text{Bound}(h_1(1)=0, h_1(2)=0, h_1(3)=1, h_1(4), \ldots h_2(2) \text{ determined by the bounding procedure})$ which equals .13857 and $\text{Bound}(h_1(1)=1, h_1(2), \ldots h_2(2) \text{ determined by the bounding procedure})$ which equals .13628. Because both of the previous two bounds are greater than .13583 we can stop. The optimum buffer-standby policy and the lowest cost has been obtained. They are as follows:

$$b^* = 17 \text{ workpieces}$$

$$h_w^*(s) = h_1(s)=0, s=1, 2, 3, 4, h_2(1)=0, h_2(2)=1$$
TRC* = .13583 $/workpiece
7.0 Recommendations for Further Study

Two areas are recommended for future research. The first area concerns the transfer line model. In this investigation we focused on a transfer line system consisting of two workstations separated by a single buffer storage zone. By expanding the model so that it can accommodate a variety of workstation and buffer layouts additional transfer line systems could be depicted. Malathronas, Perkins, and Smith (11) are currently working on an extension of their system which incorporates more than one buffer storage zone.

The second area recommended for future research is the relaxation of several assumptions contained in the reliability analysis. In this investigation it was assumed that at most one standby could be available to provide redundancy for an on-line subsystem. If the model were expanded to include two or more standbys, however, it would be possible to see if subsystems which have extremely high failure rates and/or long repair times require more than one standby.

The reliability analysis assumes that one repairman is available for each component subsystem. It would be more realistic, though, to assume that there is one (or more) repairman available per workstation. In order
to keep the workstation up and running as much as possible when failures occur in this system repair work should wait in a priority queue for a repairman. Because the input source of failures is finite and the number of priorities is greater than two the analysis of such a workstation would be quite complicated.
8.0 Conclusion

In this investigation the most cost effective buffer-standby policy was obtained for a transfer line system consisting of two unreliable workstations separated by a finite buffer storage queue. By using renewal theory to derive the overall reliability performance measures of each workstation, the research of Malathronas, Perkins, and Smith was extended to include the presence of standbys. Therefore, the production rate of the transfer line could be improved not only by increasing the buffer size but, also, by enacting a policy whereby strategic workstation components are candidates for redundancy. A branch and bound technique was formulated whereby only those buffer-standby policies which have the potential of minimizing the total relevant cost per workpiece produced must be considered as candidates for the optimum buffer-standby policy.
References


Appendix 1. Derivation of the Steady State Availability of Single Components without Standbys

The uptime and downtime distributions will now be analyzed in order to determine the availability of a system consisting of a single component with no standby. For a more general discussion of this analysis see Barlow and Proschan (1). Assume that the time to failure of a component is a random variable \( X \) with the cumulative distribution function \( F(x) \). Assume also that the repair time of a failed component is a random variable with the cdf \( G(x) \). The convolution, \( H \), of \( F \) and \( G \) is defined as:

\[
H(t) = \int_0^t G(t-x) dF(x)
\]

This function will be used to determine \( M_{ij}(t) \), the expected number of visits to state \( j \) in \((0,t)\) such that at time 0 the system enters state \( i \). If the up state is denoted by 1 and the failed state by 0, then the renewal functions may be derived by using the following logic. When the component is operational at time 0, the expected number of visits to the failed state, given that the first failure occurs at time \( x \), is \( 1 + M_{00}(t-x) \). Therefore,

\[
M_{10} = \int_0^t (1+M_{00}(t-x)) dF(x)
\]
Moreover, if the unit is operational at time 0 the expected number of visits to the operational "up" state given that the first failure occurs at time \( x \) is \( M_{01}(t-x) \). Thus,

\[
M_{11}(t) = \int_{0}^{t} M_{01}(t-x) dF(x)
\]

In a similar manner \( M_{01}(t) \) and \( M_{00}(t) \) are found to be:

\[
M_{00}(t) = \int_{0}^{t} M_{10}(t-x) dG(x)
\]

and

\[
M_{01}(t) = \int_{0}^{t} (1+M_{11}(t-x)) dG(x)
\]

Laplace transforms will now be used to solve for the renewal functions \( M_{ij}(t) \) in terms of \( F \) and \( G \). This technique is used widely in renewal theory because it deals almost exclusively with positive random variables and also because it greatly simplifies calculations dealing with convolutions. The Laplace-Stieltjes transform of the function \( M_{ij}(t) \) will be denoted by \( M_{ij}^{*}(s) \) and is defined by the integral
\[ M_{ij}^*(s) = L(M_{ij}(t)) = \int_{0}^{\infty} e^{-st} dM_{ij}(t). \]

Therefore, taking the transform of the convolution \( H(t) \) we have:

\[ H^*(s) = G^*(s)F^*(s) \]

Similarly, the transforms of the renewal functions are:

\[ M_{10}^*(s) = F^*(s) + M_{00}^*(s)F^*(s) \quad (1) \]
\[ M_{11}^*(s) = M_{01}^*(s)F^*(s) \quad (2) \]
\[ M_{00}^*(s) = M_{10}^*(s)G^*(s) \quad (3) \]
\[ M_{01}^*(s) = G^*(s) + M_{11}^*(s)G^*(s) \quad (4) \]

\( M_{10}(s) \) and \( M_{11}^*(s) \) can be found by algebraically solving equations (1), (2), (3), and (4) to get:

\[ M_{10}^*(s) = \frac{F^*(s)}{1 - F^*(s)G^*(s)} \]

and \[ M_{11}^*(s) = \frac{F^*(s)G^*(s)}{1 - F^*(s)G^*(s)} \]

Then, \( M_{10}(t) \) and \( M_{11}(t) \) are derived by taking the inverse transforms \( L^{-1} \) of \( M_{10}^*(s) \) and \( M_{11}^*(s) \).

The probability that the system is operating at a specified time \( t \) will now be found. Let \( P_{ij}(t) \) equal
the probability that at time $t$ the component is in state $j$ if at time $t=0$ the component was in state $i$. Thus,

$$P_{10}(t) = M_{10}(t) - M_{11}(t)$$

and

$$P_{11}(t) = 1 - P_{10}(t) = 1 - M_{10}(t) + M_{11}(t)$$

As defined by Barlow and Proschan (1975) the availability of a system is equal to the probability that it is operational at any given time. Therefore, assuming that the component is up at time $t=0$ the availability is given by:

$$A(t) = P_{11}(t)$$

Similarly, the limiting availability, $A$, and the limiting average availability, $A_{av}$, are as follows:

$$A = \lim_{t \to \infty} A(t) = \lim_{t \to \infty} P_{11}(t)$$

and,

$$A_{av} = \lim_{t \to \infty} \frac{1}{T} \int_0^T A(t) dt = \lim_{t \to \infty} \int_0^T P_{11}(t) dt$$

In order to demonstrate how this technique can be used, the availabilities of an example component will now be determined. According to the assumptions previously mentioned, each component within workstations one and two will possess exponential failure and repair rates. Therefore, suppose that a certain component $c$ fails exponentially with a failure rate $\lambda_c$. Suppose also
that the repair is exponential with the repair rate $u_c$.

The cumulative distribution functions are then $F(t) = 1 - e^{-\lambda_c t}$ and $G(t) = 1 - e^{-u_c t}$. Taking the Laplace transforms of $F$ and $G$ yields:

$$F^*(s) = \frac{\lambda_c}{s + \lambda_c}, \quad G^*(s) = \frac{u_c}{s + u_c}$$

To calculate $P_{11}(t)$, $M_{10}^*(s)$, and $M_{11}^*(s)$ must first be determined:

$$M_{10}^*(s) = \frac{F^*(s)}{1 - F^*(s)G^*(s)} = \frac{\lambda_c(s + u_c)}{s^2 + (\lambda_c + u_c)s}$$

$$M_{11}^*(s) = \frac{F^*(s)G^*(s)}{1 - F^*(s)G^*(s)} = \frac{\lambda_c u_c}{(s + \lambda_c)(s + u_c) - \lambda_c u_c}$$

Then, by taking the inverse transforms of $M_{10}^*(s)$ and $M_{11}^*(s)$ and setting $P_{11}(t)$ equal to $1 - M_{10}(t) + M_{11}(t)$, we obtain:

$$P_{11}(t) = \frac{u_c}{\lambda_c + u_c} + \frac{\lambda_c e^{-(\lambda_c + u_c)t}}{\lambda_c + u_c}$$

This value is equal to the system availability $A(t)$. Moreover, as the value of $t$ increases the limiting availability, $A$, will be approached:

$$A = \lim_{t \to \infty} A(t) = \frac{u_c}{\lambda_c + u_c}$$
Vita

Elizabeth Gager, the daughter of Louise and early youth were spent in the Puna Valley, Hawaii. From 1982 through 1986, she was an instructor of science program in industrial engineering and operations research from the University of California at Berkeley. From 1982 through 1986, she was an instructor at Lehigh University, where she was a part of a master of science program in industrial engineering.