COMPUTER REDUNDANCY: DESIGN, PERFORMANCE, AND FUTURE

Ralph E. Kuehn

International Business Machines Corporation
Electronics Systems Center
Owego, N.Y.

INTRODUCTION

The theoretical basis for the use of redundancy in electronic systems was established in the papers by J. von Neuman[1] and E. F. Moore[2] and C. E. Shannon. These papers considered intermittent malfunctions of elements whose probability of failure was time invariant. W. E. Dickinson and R. M. Walker[3] extended this concept to permanent failures of elements and known failure probabilities as a function of time.

Intensive system engineering[4,5] efforts on the use of majority voting redundancy in digital computers has determined the relative merit of input and output voting, the optimimum size of circuitry between voters, reliability requirements of voters, and general design ground rules. Utilization of M. Cohn's[6] proposal to triplicate the majority voter has in the practical sense relieved the limiting restriction of voter reliability on computer reliability. The treatment of a hypothetical digital computer[4] capable of a useful computing function and implemented in majority voting redundancy indicated the excessive complexity of the exact analytical approach. A Monte Carlo model for simulating the statistical failure structure of a TMR (Triple Modular Redundancy) computer was developed as an extremely useful design tool.

The papers previously noted as well as many other mathematical and system engineering oriented papers have established an excellent background for the application of redundancy in the design of digital computers requiring high reliability. It is the purpose of this paper to describe the design of digital equipment utilizing TMR and QUAD redundancy. Design problems, reliability predictions, reliability achievement, and future implications will be covered.

TMR DESIGN

In 1961 the Electronics Systems Center of the Federal Systems Division initiated research and development under contract to Marshall Space Flight Center on a highly reliable flight computer system for the guidance and control of the uprated Saturn I and the Saturn V launch vehicles. The reliability goal for the computer was established at 99% for a 250 hour space mission including the launch and thrust portions. The computational requirements anticipated called for a maximum of 32K words of memory of 28 bits and an electronic part complement of approximately 12,000. Considering the initial requirements for production deliveries in the period 1964-8 and the projected part failure rates the reliability goal was significantly beyond the capability of the design and reliability art for a non-redundant (simplex) computer.

Trade off studies of attainable reliability versus cost, weight, and power resulted in the decision to use TMR in the logic and arithmetic sections and a modified form of duplex memories. The memory uses conventional toroidal cores in a self-correcting duplex[8] arrangement. Up to eight identical 4096 word memory modules may be operated in simplex for additional storage capability in ground operation or in duplex pairs for the high flight reliability. Figure 1 is a simplified block diagram of the memory. The error detect logic consists of a parity check and a memory drive current monitor. When both memories are operating without failure, each memory is controlled by its buffer register. Both memories are simultaneously read and updated, 14 bits in parallel. While a single cycle is required to read instructions (13 bits plus a parity bit per instruction), two cycles are required for reading and updating data (26 bits plus 2 parity bits). The parallel outputs of the buffer registers are serialized at the 512 K bit rate under control of the memory select logic. Initially only one memory output is transferred to the central computer but both are active. When an error is detected in the memory being used the error detect logic causes the memory select logic to switch immediately to data from the other memory. Both memories are then regenerated by the buffer register of the "good" memory thus correcting transient errors. After the error detection and parity checking circuits have indicated the correction of transient errors by lack of an error signal, each memory is again controlled by its own buffer register. The previously erroneous memory is not used by the computer until the "good" memory develops an error. Hence, instantaneous switching from one memory to the other results in continuous correct computer operation until simultaneous failure at the same storage location in both memories.

The duplex memory design presented substantial reliability design problems as well as reliability prediction problems. The reliability prediction model for a single memory module is fairly simple and basically dependent on part failure rates and redundancy of the associated electronics. The duplex model must include the

effectiveness of the error detection. The detailed model and the associated reliability equations have been treated elsewhere[9]. Sensitivity analysis of the reliability model of the proposed memory system indicated the need for close liaison between the designer and the reliability analyst.

The determination of the effectiveness of the error detection required answers to questions such as the following:

What is the probability of multiple failures resulting in correct parity?

What is the time distribution of errors compared to the distribution of times between successive readouts of memory addresses?

What is the reliability of the parity and memory drive current monitors?

Which memory part failures are non-detectable by the error detection and what is their probability of occurrence?

Will failures in the memory decoupling and power sequencing be detected?

Detailed analysis of the memory operation in the presence of part failures resulted in an estimate that 93% of expected first failures would be detected by the combined parity and drive current monitors. Analysis of memory module failures during factory acceptance test and field operation indicated 95% effectiveness in detecting first memory failures basically confirming the memory design.

Figure 2 gives a schematic representation of the TMR redundancy used in the computer logic. Each module is identical, receiving the same problem simultaneously. The outputs of each module feed into "majority elements" designated as voters. The output is determined by the majority of the voter input signals. The number and placement of voters presents an important design and reliability analysis problem. References[4,5] give a detailed approach to optimizing the system reliability considering the amount of circuitry between voters, the reliability of the voters, and certain basic assumptions. In actual design additional factors must be considered, such as:

1. Voter circuit delays and drive requirements.

2. Number of signal transfers out of a proposed signal module.

3. Second level package size limitations due to connector capabilities, expected fabrication yield, and test capabilities.

4. Maintenance, trouble shooting time, and logistic requirements.

Considerations such as the above resulted in a number of tradeoffs and a reasonably optimum design. The computer logic was divided into seven modules with an average of 13 voted outputs.

An exact analytical approach to predicting the reliability of a computer of such complexity is impractical. A program[10] for the IBM 7090 was developed using a Monte Carlo technique for failure generation and logic simulation for tracing the effects of a failure. The program generates a random set of component part failures considering the failure rate for each component part type, the number of each type component part used in the computer, and the time for a complete mission. The resulting numbers are related to a specific part in a logic block and when weighted by the conditional probability of an open appears as an open or shorted part. Logic block failure parameters supplied to the program then determine whether the logic block fails to a logical "1" or "0". The program then traces the failed signal through the simulated logic to the input to a voter. If the program finds two of the three inputs to a voter, failed in the same direction, a system failure has occurred. The steps above are continued for the duration of the mission or until a system failure occurs whichever happens first. This process was typically repeated for twenty thousand mission times. The reliability of the computer for the specified mission is

$$R = \frac{\text{number of successful simulated missions}}{\text{total number of simulated missions}}$$

The estimate for the reliability of the computer logic for a 250 hour mission obtained from the program for 20,000 simulated missions is 0.9994.

While the end utilization of the computer is the guidance and control of the Saturn launch vehicle, early studies indicated that all but a very few operating hours would occur on the ground because of the large number of acceptance and verification tests. Early field and flight experience indicates that less than 1% of total operating time will be in flight. Consequently maintenance and trouble shooting activities are of extreme importance. NASA personnel suggested the use of Disagreement Detectors (DD) shown in Figure 2. At selected voter locations the voter inputs are monitored by a DD which provides an error indication output[7] when there is a disagreement among the voter inputs. The DD outputs are stored in a register for use in ground trouble shooting and for telemetering computer status during flight.

TMR RELIABILITY VERIFICATION

The field measurement and/or analytical verification of a high reliability for a redundant computer is extremely difficult and may be infeasible depending on the expected usage. The confirmation of a computer reliability of 0.99 for a flight mission of 250 hours at the 90%

confidence level[11] requires approximately 57,500 computer operating hours without failure. Since the expected flight time in the entire program is less than a thousand hours, flight reliability verification is clearly infeasible.

Since all computers are under complete surveillance in the field and have individual elapsed time indicators the measurement of Mean Time Between Component Failure (MTBCF) is quite feasible with high accuracy and confidence. The provision of the DD and associated disagreement register ensures detection of all logic malfunctions (solids, intermittents, and one time occurrences). It is of course reasonable to proceed from a MTBCF measured at the 90% confidence level, assumed part failure distributions, and the probability of part open circuits assuming the part has failed to arrive at a deduced flight reliability. At this time however, the verification of part open probabilities and part failure distributions at an acceptable confidence level is not possible from field data. The most that can be said at this time is that the assumptions have not been disproved.

A summary of field failure surveillance as of 25 September 1967 is given in Table 1. Field data covers

Surveillance Data	Number	Predicted MTBCF
Incidents	30	
Significant Failures	13	
Flight Significant Failures	7	1,316

Surveillance Data	Required MTBCF	Observed MTBCF
Incidents		522
Significant Failures		1,205
Flight Significant Failures	691	2,235

Table 1. Surveillance Data Summary

22 computers operated in the field for 15,674 power on hours. Incidents are defined as the total number of reported malfunctions in usage. This includes most intermittents and single datum errors made reportable by means of the disagreement detectors and associated register. Significant failures are those component malfunctions remaining after evaluation has excluded malfunctions due to external causes, multiple part failures due to an initial part failure and human errors. Flight significant failures are those component malfunctions that could have occurred in flight. This classification is arrived at by censoring the significant failure classification for those failures unique to the ground operational environment.

Since the end use is flight, comparisons among predicted, required, and observed MTBCF for flight significant failures are of much interest. The initial prediction for the computer of 1,316 hours is significantly above the requirement of 691 hours. Considering that a demonstration was included in the contract with a fee incentive, the above condition was necessary to reduce the producer's risk to an acceptable value. The observed MTBCF for flight significant failures is 170% of the predicted value. Since the observed value is calculated based on only 7 failures the lower 90% confidence value may be more appropriate for comparison. The lower 90% confidence limit on the observed MTBCF for flight significant failures is 1,327 hours which compares favorably with the predicted value. To date three flights have been completed with perfect performance by the computer. While the large majority of flights are still in the future, the flight and ground operation results to date confirm the decision to accept a flight incentive of maximum reward for no flight failures of the computer and maximum penalty for one or more computer flight failures.

QUAD DESIGN

In 1960 the Electronics Systems Center of the Federal Systems Division initiated research and development under contract to Grumman Aircraft Engineering Corporation on the Primary Processor and Data Storage (PPDS) for the NASA Orbital Astronomical Observatory. The IBM designed equipment incorporates the central timing source, command storage, data storage, electronics for routing commands to control various spacecraft functions, and electronics for routing data from the experiment packages to the data storage. Initial analysis indicated that the functional requirements of the PPDS could be accomplished with a 15,000 component part machine and 100,000 cores for memory.

The initial reliability requirement was 95% for a year operation in space. Considering the scheduled delivery date of late 1962 for flight hardware and the projected state of the reliability art with respect to component parts the reliability of a simplex machine was projected as 1% for a year in space. The use of duplex and TMR implementation as well as a Minuteman type parts improvement program were considered and rejected on the basis of failing to meet the reliability requirement or presenting too high a risk. Study of redundant component circuits (QUAD) resulted in the conclusion that this method would result in a satisfactory reliability within the schedule and cost constraints.

The concept of QUAD redundancy[12] at the passive component part level is illustrated in Figure 3 with the dashed line indicating an option dependent on the probability of opens and shorts for the passive parts under consideration. Considering only opens and shorts two or more parts must fail to cause total network failure provided the source and load can perform satisfactorily over the range of impedance provided by the total network in the presence of part failures. The reliability equation for the net-

work of Figure 3 without the dashed connection is given in the literature[12,13] as

$$R_{AB} = (1 - Q_{NS}^2)^2 - (1 - [1 - Q_{NO}]^2)^2$$

where Q_{NS} is the probability that N shorts

and Q_{NO} is the probability that N opens

If the probability of failure in the short mode is zero, all elements N can be made parallel-redundant and the reliability is $R = 1 - Q_{NO}^K$ where K is number of parallel elements.

A physical and mathematical analysis of the basic QUAD network clearly reveals the importance of short and open probabilities at the component part level to the circuit designer and reliability analyst. In the PPDS these probabilities were established for over 20 basic part types for each of three environments namely ground, powered flight, and orbit. Historical failure data, physics of failure analysis, overstress testing, process analysis, and engineering judgment were used in establishing open and short failure probabilities.

The inclusion of active component parts into QUAD networks has received much attention[13,14]. Extensive design efforts on digital circuits using discrete transistors have disclosed a number of basic problems[14] including the following

- Transistor parameters are subject to more demanding requirements than in simplex design.
- A QUAD configuration circuit must be applied so as to drive significantly less load than a simplex circuit.
- The QUAD design will dissipate significantly more power if maximum speed is required.
- The QUAD approach is inherently slower with respect to signal propagation time.
- Part failure within a QUAD materially increases semi-conductor dissipation in remaining operational portions of QUAD.

Detailed circuit design of the 22 distinct circuits used in the PPDS confirmed the above design problems.

Figure 4 shows a simplex Memory Address Register circuit and the QUAD counterpart used in the PPDS design for comparison purposes. In the QUAD design the capacitors, diodes, and resistors associated with a transistor quadrant are simplex rather than QUAD. This results from the particular resistors and capacitors having a negligible probability of shorting, the effect of opens on the resultant quadrant operation, the predominant contribution of the transistor to the quadrant failure probability, and the reliability versus power, weight, and cost trade-offs. It should be noted that resistors capable of causing the entire circuit to fail are in QUAD parallel.

The relatively precise circuitry required for core memories such as current drivers and temperature compensators eliminates component part redundancy as a practical approach to high reliability memories. The command storage of 256 - 30 bit words uses a four quadrant array packaged in a single array with dimensions of 32x32x30. The schematic for the QUAD memory module with associated addressing is given in Figure 5. Each word is written into word locations in each of the four quadrants A1, A2, B1, and B2. The word is written into and read from two locations (A1 and A2 or B1 and B2) simultaneously. The basic assumption was made that all failed core locations yield a "zero" output. Based on this assumption voting on the quadrant outputs is accomplished with latches. The presence of a "one" in any of the four locations is recognized as a "one". Partial compliance with the "fail to zero" assumption can be assured by circuit design. Design ground rules were accordingly established in the memory electronic areas to discriminate against failure modes having a significant probability of "one" failures. The "fail to one" situation has been further allievated by providing the capability to bypass failed locations and by gating the four outputs to allow quadrants of the array which contain failures to the "one" state to be ignored.

The excellent masking of component part malfunctions by proper QUAD design presents a serious problem in maintainability design, acceptance test specifications, and operational policy. At the circuit level terminals can be brought out or special probes designed to make the necessary measurements to ensure that at initial acceptance testing all component parts are functional. There is of course an obvious resources penalty compared to circuits that can be completely tested by means of output measurements. Factory acceptance testing and field verification testing of higher levels of assembly involving hundreds or thousands of QUAD circuits is practical only on an input output basis. An operational policy of maintenance only on system failure naturally follows. Reliability degradation prior to mission initiation during storage, spacecraft integration, and system checkout must be included in mission success predictions.

The PPDS reliability requirement, based on operation in space for a year, was modified to include 3300 hours of ground operation prior to launch in order to include the effects of the unknown but statistically predictable reliability degradation existing at launch. For long life unmanned missions such as the Orbital Astronomical Observatory, the PPDS policy of maintenance only at system failure and inclusion of ground

peration degradation in reliability prediction ppears to be completely compatible with QUAD edundancy.

QUAD RELIABILITY VERIFICATION

While the number of planned OAO missions s small, the mission duration is sufficiently ong to permit the measurement of flight reliability at an acceptable level by the end of the light program. This of course is of little omfort when reliability assurance prior to first light is desired.

The ground operating experience on the PPDS oes provide some verification of QUAD redundancy. s of 25 September 1967, five PPDS have accumulated approximately 5000 operating hours in the ield without a system failure. While statistically inconclusive, this is certainly encouraging rom the engineering point of view.

A single PPDS used for system integration nd environmental qualification was returned for efurbishment and engineering changes after approximately 3000 field operating hours. Acceptance level tests at the system level, unit level, nd circuit level were progressively made during disassembly. No discrepancies were discovered at the system and unit level. Detailed failure analysis was conducted on all circuits failing to meet their initial acceptance conditions. Failure analysis revealed five component part failures chargeable to the 3000 hours of field operation.

The summation of the predicted component part failure rates for ground operation of the PPDS is 1785×10^{-6} per hour. At this rate, five failures would be expected in 2800 ground operating hours ($5/1785 \times 10^{-6}$). Field operation has then indicated a most probable part failure rate somewhat lower than that used in the prediction. Based on the fact that these five component part failures were masked at the system level by the QUAD redundancy and the careful detailed analysis of each circuit for part failure effects, it can be concluded that Quad redundancy has been verified in ground field operation at an engineering judgment level and that the reliability prediction is reasonable.

FUTURE MILITARY, SPACE, AND COMMERCIAL IMPLICATIONS

The future of computers utilizing redundancy can reasonably be projected for military and space applications. It is expected that the long term trend of complexity increasing more rapidly than reliability improvement will continue. Space and military requirements for ultra high reliability in real time and for high reliability on long unmanned missions will continue to need redundant computers as the answer to their mission reliability requirements. Figure 6 presents the capability of simplex computer design for 95% reliability with complexity, mission time, and mean part failure rates as the variables. As an example, if one assumes a mean part failure rate of 1×10^{-10} per hour being available in 1975, a simplex computer requiring 10,000 parts for computational needs and 95% reliability could be designed for a mission of 60,000 hours as indicated by point A on Figure 6. Similarly for a predictable computational requirements of 10^7 component parts, the availability of a mean part failure rate of 1×10^{-10} per hour would permit a simplex computer to meet a 95% reliability requirement for a mission of 80 hours. This operation point corresponds to B on Figure 6. The area above and to the left of the AB line segment of constant part failure rate (1×10^{-10} per hour) represents simplex computer design while the area below and to the right is infeasible for simplex design assuming a part failure rate limitation of 1×10^{-10} per hour for the 1975 period.

The above two realistic examples demonstrate a continued need for redundancy in selected military and space computers. The QUAD technique will continue to be applied for those applications requiring the highest feasible reliability and for long unmanned missions with a maintenance at failure policy. TMR will be applied for relatively short mission requiring reliability unattainable by simplex computers. Because of the high potential for ease of trouble shooting, TMR will be preferred for manned missions and those applications requiring considerable usage prior to the first mission. Multiple mission requirements will also call for TMR.

CONCLUSIONS

The design problems represented by TMR and QUAD redundancy have been identified and overcome for the Saturn Computer and the OAO Primary Processor and Data Storage. Reliability predictions, aided by computer programs have been made and field verification has been accomplished at an engineering judgment level. TMR and QUAD redundancy will continue to find selective application in future military and space systems.

BIBLIOGRAPHY

1. J. vonNeumann, "Probabilistic Logics and the Synthesis of Reliable Organisms From Unreliable Components" Automata Studies, Princeton University Press 1956.

2. E. F. Moore and C. E. Shannon, "Reliable Circuits Using Less Reliable Relays" Journal of the Franklin Institute, vol. 262, part I (September 1956) p.p. 191-208 and part II (October 1956) p.p. 281-297.

3. W. E. Dickinson and R. M. Walker, "Reliability Improvements by Use of Multiple Element Circuits", IBM Journal 2, 142-147 (1958).

4. R. E. Lyons and W. Vanderkulk, "The Use of Triple Modular Redundancy to Improve Computer Reliability" IBM Journal 6, 200-209 (1962).

5. H. L. Ergott and D. P. Rozenberg, "On the Analysis of Reliability Improvement Through Redundancy" IBM Corp., FSD Space Guidance Center Report 62-533-001, May 1962.

6. M. Cohn, "Redundancy in Complex Computers", Proceedings of the National Conference on Aeronautical Electronics, Dayton, Ohio, May 1956, pp 231-235.

7. M. M. Dickinson, J. B. Jackson and G. C. Randa, "Saturn V Launch Vehicle Digital Computer and Data Adapter", Fall Computer Conference, San Francisco, October 1964.

8. C. V. McNeil and G. C. Randa, "Self-Correcting Memory - The Basis of a Reliable Computer", Electronic Design, August 30, 1965.

9. J. E. Anderson and F. J. Macri, "Multiple Redundancy Applications in a Computer", Proceedings 1967 Annual Symposium on Reliability, Washington, D. C. , 553-563 (January 10-12, 1967).

10. R. B. Coffelt, "Automated System Reliability Prediction", Proceedings 1967 Annual Symposium on Reliability, Washington, D.C., 302-305 (January 10-12, 1967).

11. M. Sobel and J. A. Tischendorf, "Acceptance Sampling With New Life Test Objective", Proceedings Fifth National Symposium on Reliability and Quality Control, Philadelphia, Pa. 108-119 (January 12-14, 1959).

12. C. J. Creveling, "Increasing the Reliability of Electronic Equipment by the Use of Redundant Circuits", Proceedings IRE, Vol. 44, 509-515 (April 1956).

13. A. A. Sorenson, "Digital-Circuit Reliability Through Redundancy", Electro-Technology, Vol. 68, 135-140 (July 1961).

14. R. M. Fasano and A. G. Lemack, "A QUAD Configuration-Reliability and Design Aspects", Proceedings Eighth National Symposium on Reliability and Quality Control, Washington, D. C. 394-407 (January 9-11, 1962).

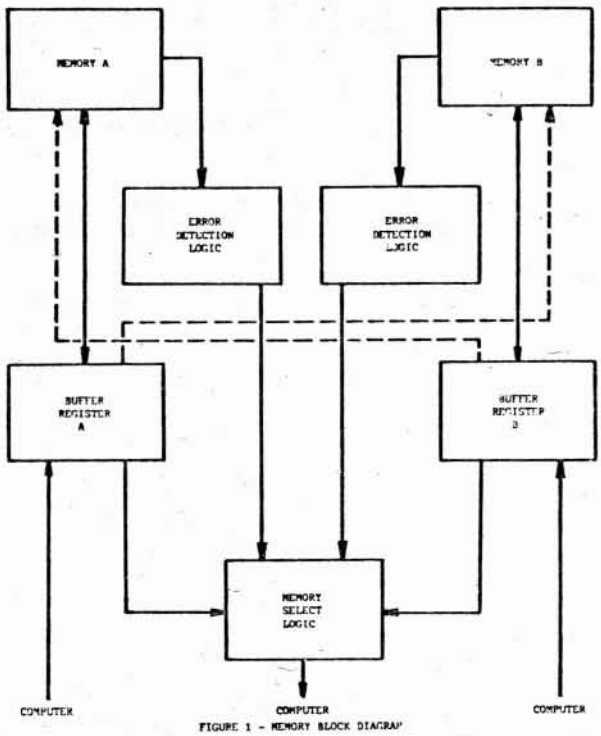


FIGURE 1 - MEMORY BLOCK DIAGRAM

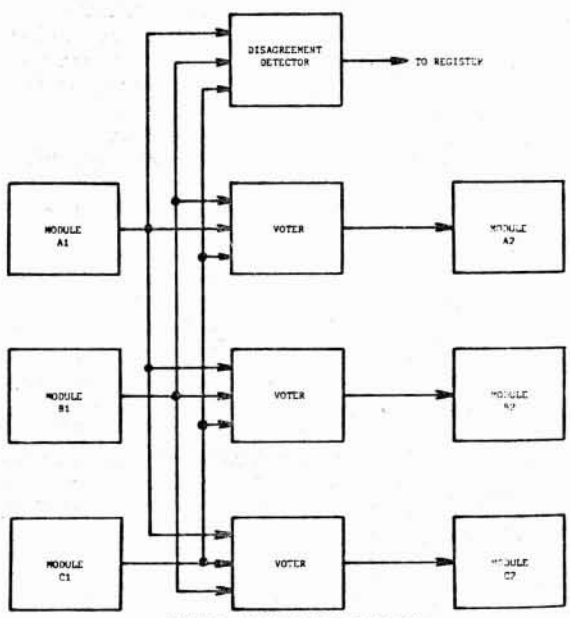


FIGURE 2 - TMR WITH DISAGREEMENT DETECTOR

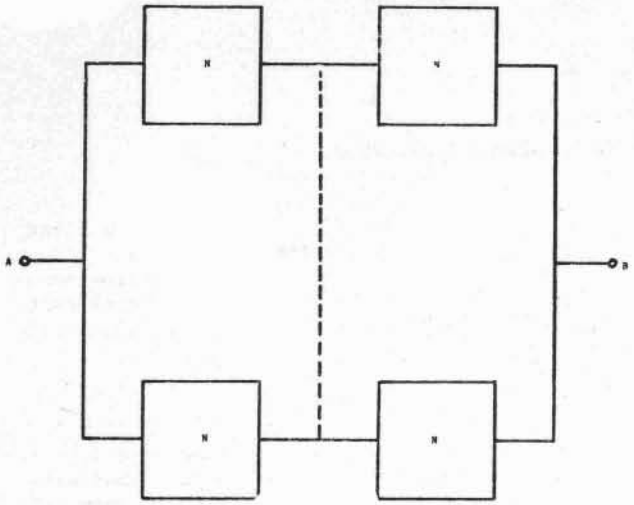


FIGURE 3 - QUAD COMPONENT CONFIGURATION

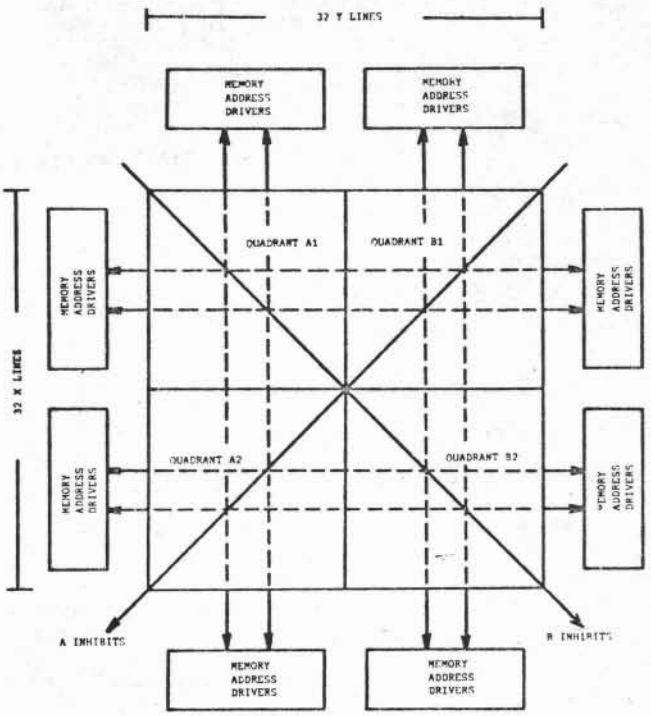
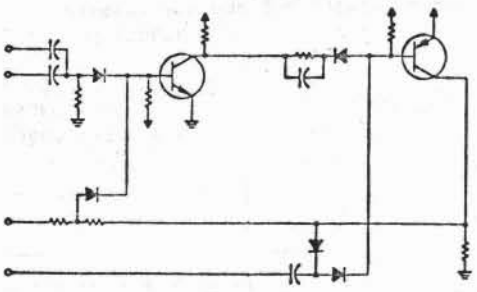


FIGURE 5 - QUAD ARRAY AND ADDRESSING



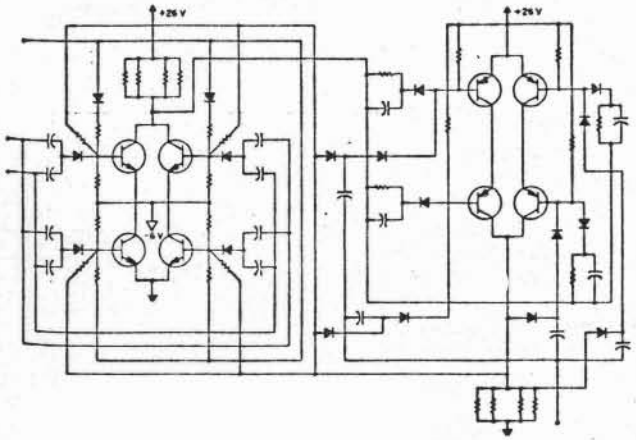


FIGURE 4 - SIMPLEX AND QUAD MEMORY ADDRESS REGISTER CIRCUITS

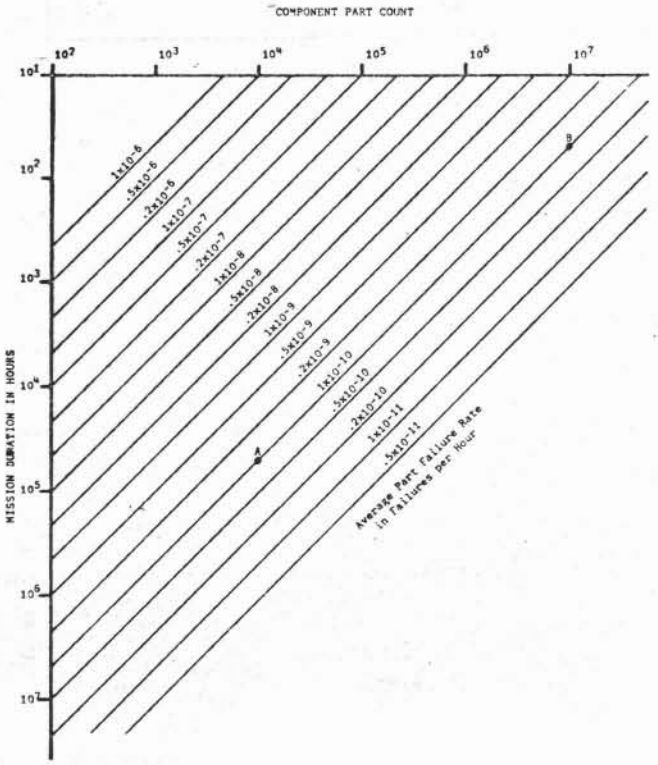


FIGURE 6 - FEASIBILITY OF 95% RELIABILITY FOR COMPUTERS