

# Reliability Analysis of a Highly-Integrated Multiprocessor System

Venkat K. R. Chiluvuri

Israel Koren

Electrical & Computer Engg Dept  
University of Massachusetts  
Amherst, MA 01003

Electrical & Computer Engg Dept  
University of Massachusetts  
Amherst, MA 01003

## Abstract

*The relationship between the reliability of a multiprocessor system and the integration level of its components is analyzed. Failure rates are calculated and the relationship between the die size, power dissipation, terminal count and the thermal characteristics of the package is discussed. An optimal level of integration is proposed for a multiple-bus multiprocessor system. IC yields are estimated at various integration levels to compare the relative impact of integration level on system reliability and yield.*

## 1 Introduction

With the advancements made in VLSI technology during the last decade, very high level of integration is achieved in ICs. Chips with a million transistors are already in commercial production [18]. The trend in VLSI technology indicates that multiple processors, memory modules and other peripheral logic can be realized on a single chip. In other words, almost a complete multiprocessor computer system can be realized on a single chip. An experimental 250-MHz on-chip multiprocessor system consisting of four 32-bit processors and 32-Kbyte cache was designed by Hitachi [11].

Implementing more and more logic on a single chip offers many advantages in terms of performance, compactness, ease of maintenance etc. Ideally, one would like to have the entire system on a single chip. This approach however, tends to make systems more expensive due to yield related problems, loss of general purposefulness etc. Other important aspects are the problems associated with the thermal management, pin out requirements and reliability. In order to achieve higher integration levels these problems must be dealt with.

For many highly computing intensive applications nowadays, multiprocessor systems are preferred to super computers because of their better cost-performance ratio and expandability of the system. A typical multiprocessor system consists of several processors and memory modules connected by an interconnection network. This system can be realized using individual processor and memory chips or by integrating several processor and memory modules on a single chip. An example of a multiple-bus multiprocessor system is presented in Section 2 and the impact of integration on failure rates of the chips is analysed with reference to this system.

The failure rate of an IC is a function of chip parameters such as technology, integration level, i.e., the number of transistors/gates/memory cells, pin count, the size, packaging, heat dissipation and design/process maturity. Most of these parameters are related to one another. Heat dissipation is the major contributing factor to failures in chips and its relationship with failure rates is well established. Heat dissipation, in turn, is a function of technology, clock speed, packaging, die size and pin count. Failure rates of chips with different integration levels are calculated in Section 6 using the Military Handbook MIL-HDBK-217E [23]. The failure rate model and its parameters are discussed in Section 3.

During the last decade, the IC feature size has diminished from  $5\mu\text{m}$  to submicron associated with multifold increase in circuit integration and power dissipation. As the state of art VLSI technology allows integration of millions of transistors on a single chip, heat fluxes have already reached the range of  $10\text{-}30\text{W}/\text{cm}^2$ . It is expected to reach  $70\text{-}100\text{W}/\text{cm}^2$  in the near future [6]. These heat fluxes pose serious problems of thermal management and packaging techniques. The power dissipation problem is further aggravated by very high operating speeds of the high performance logic in the range of  $50\text{-}100\text{MHz}$ . For reliable operation, heat must be removed from the chips so that the junction temperatures can be maintained below a specified limit beyond which the functionality of the chip is not guaranteed. The failure rate of the chip increases by a factor of about two for every  $10^\circ\text{C}$  increase in junction temperatures. Therefore, suitable packaging and cooling techniques must be employed to improve the reliability of the chips and systems. The thermal characteristics of the VLSI chips and their impact on reliability are presented in Section 4. Since power dissipation is the major contributing factor for chip failures, accuracy of the reliability analysis depends on the accuracy of the power dissipation calculations. In Section 4, procedures for power dissipation calculations and associated problems are discussed.

As the level of integration goes up terminal count also increases for a chip. The shared multiple-bus architecture, proposed in Section 2, requires only a modest increase in pin count. As pin count increases, power dissipation and die size also increase. The effect of die size on thermal resistance, yield and packaging

styles of the chip are presented in Section 5.

In Section 6, various system configurations of the multiprocessor system, proposed in Section 2, consisting of chips with different integration levels are analysed. Die size, pin count and power dissipation are estimated for each system configuration. The impact of these parameters on the system reliability, chip yield and packaging are analysed and an optimal level of integration with respect to reliability is proposed. Conclusions are presented in Section 7.

## 2 Multiprocessor system

A multiple-bus multiprocessor system described below is used as an example to illustrate the impact of integration on reliability and yield. Similar analysis can be done for any multiprocessor system. A multiple-bus multiprocessor system contains  $N$  processors and  $M$  memory modules connected by multiple buses. In the conventional organization, individual processor and memory chips are assembled on printed wiring boards which are then assembled into a system. With the state-of-the-art VLSI technology, chips with two or more processors and memory modules can be fabricated. The performance improvements of this approach are discussed in [8]. In this approach, the multiprocessor system is partitioned into several chips, each containing a smaller number of processors, memories and buses. These chips are then interconnected through global buses. Thus, the buses in this multiprocessor system are partitioned into local and global buses. The local buses interconnect only the on-chip components providing a high-speed local communication. Since these buses are local to components within a chip, they need not be brought out of the chip. Therefore, the pin out requirement is reduced in this bus architecture. This system is formally specified as follows.

A multiple-bus multiprocessor system consists of  $K$  chips, each containing  $n$  processors,  $m$  memory modules and  $B_l$  local buses.  $B_g$  global buses connect these chips for global references [8]. The overall multiprocessor system contains  $N$  processors,  $M$  memory modules and  $K$  chips, where  $N = K.n$  and  $M = K.m$ . This system is referred to as a  $(N \times M \times (B_g + K.B_l))$  system. The organization of the system is shown in Figure 1.

## 3 Reliability prediction

Reliability prediction during design/feasibility evaluation of systems has several advantages. It helps in comparing several competing designs/approaches and in identifying the potential reliability problems during the initial stages of the system design. To predict the overall system reliability, failure rates of individual components are calculated first. Then, the failure rates of the subsystems (board level) are calculated. Any redundancy provided in the system should be accounted for while calculating failure rates. One of the best data bases of the failure rates of various electronic components is the reliability handbook [23]. The reliability prediction procedure using this handbook is briefly outlined below.

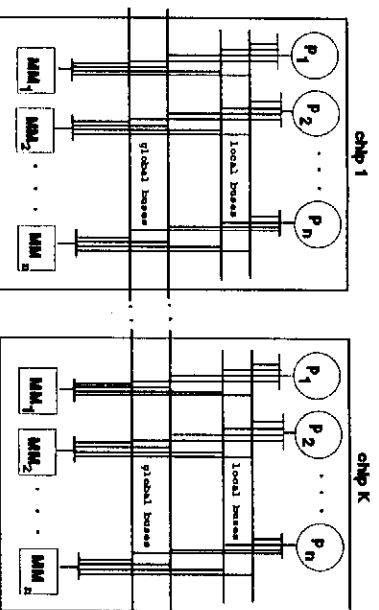


Figure 1: A modular and highly-integrated multiple-bus with  $B_l = B_g = 2$

### 3.1 Reliability prediction using MIL-HDBK-217

This handbook provides a consistent, uniform database for making reliability predictions during the design phase of the electronic equipment. It provides two methods for calculating component level failure rates, the "parts stress method," and "parts count method". The parts count method requires only limited information where as the parts stress method requires significant information in order to calculate the failure rates of the components. In this analysis, failure rates are calculated using the parts stress method.

Since failure rate parameters are obtained from field failure data and life test data, typically it takes about 4 to 5 years to include the failure rate data of new components/technologies in the handbook. In updating this document, priority is given to component types and styles most commonly used by the military. For example, though surface mount technology components (SMT) have been in the market for the past 7 - 8 years, their failure data will be included for the first time in the next version - 217F which is expected to be released by 1992 [25].

Failure rate models of microprocessors and dynamic RAMs are taken from this handbook. Since this document does not contain information on failure rate parameters of recent technologies/components, manufacturers' data is used to the extent possible to calculate various parameters of the model that is described below. For the component types not included in this handbook, i.e., a single chip with processors and memory modules, parameters are calculated from the closest equivalent models given in the handbook.

#### 3.2 Failure rate model

The first mathematical model for IC failure rate derived from the Arrhenius relationship was given in MIL-HDBK-217A. This basic model was modified several times based on field data in subsequent releases of MIL-HDBK-217. The failure rate model which includes several multiplicative adjustment factors reflecting device technology, packaging, screen-

ing/testing level, use environment, operating voltage, process/design maturity etc., is given below.

$$\lambda_P = \Pi_Q (C_1 \Pi_T \Pi_V + C_2 \Pi_E) \Pi_L \quad \text{Failures}/10^6 \text{ hours.} \quad (1)$$

where  $\lambda_P$  is the device failure rate in failures/ $10^6$  hours,  $\Pi_Q$  is the screening/testing level factor,  $\Pi_T$  is the temperature acceleration factor, based on technology,  $\Pi_V$  is the voltage stress derating factor,  $\Pi_E$  is the application environment factor,  $C_1$  is the circuit complexity factor based on bit count,  $C_2$  is the package complexity factor and  $\Pi_L$  is the device learning (process/design maturity) factor.

### 3.3 Example failure rate calculations for a microprocessor

Failure rate calculations of Intel-80486, a 32 bit microprocessor are shown in this section to illustrate the various parameters of the model. Thermal specifications of this processor shown in Table 1 are taken from the Intel technical manual [13].

Table 1: Thermal specifications of Intel-80486 microprocessor

Parameter	Specification
power dissipation at 25 MHz	5 Watts (absolute maximum)
Thermal resistance from junction to case ( $\theta_{JC}$ ) - case to ambient ( $\theta_{CA}$ )	2.5 Watts (typical)
	1.5 $^{\circ}$ C/Watt 3.4 $^{\circ}$ C/Watt (with heat sink and air flow of 1000 ft/min)

Based on the chip specifications and end-use environmental conditions, appropriate values are selected from the handbook for different parameters of the model given in equation 1. Individual parameter values and calculations are shown in Table 2. By substituting these values in equation 1, the failure rate ( $\lambda_P$ ) of the 80486 microprocessor is estimated as 0.5178 failures/million hours.

## 4 Thermal aspects

Very high speed technologies coupled with high integration levels pose serious problems of thermal management and packaging techniques. Most of the high performance general purpose microprocessors consume power in the order of 2 to 8 Watts and most of it is converted to heat. Chips with 25 Watts maximum power rating are fabricated for air cooled environments [5]. Since failure rates of the chips increase by a factor of about two for every 10 $^{\circ}$ C increase in junction temperature, suitable packaging and cooling techniques must be employed to improve the reliability of the devices/systems.

Table 2: Failure rate calculations of Intel-80486 Microprocessor

Parameter	Value	Specifications
$\Pi_Q$	10.0	no additional screening
$C_1$	0.12	32 bit microprocessor
$\Pi_V$	1.0	Vdd is 5 Volts
$\Pi_E$	0.38	airconditioned environment
$\Pi_L$	1.0	mature CMOS-IV process
$C_2$ calculations:		
$C_2 = 2.8 \times 10^{-4} (N_p)^{1.08}$ (leadless chip carrier)		
$= 0.063$ (Active pins( $N_p$ ) = 151)		
$\Pi_T$ calculations:		
$\Pi_T = 0.1 \exp(X)$		
$X = -A \left( \frac{1}{T_j + 273} - \frac{1}{298} \right)$		
$A$ is 6373 (hermetic CMOS) and $T_j$ is the junction temperature)		
$\Pi_T = 0.2327$		

### 4.1 Thermal resistance

The size of the die, where heat is generated, is a very small fraction of the IC package from where heat is dissipated to the external environment. Due to variations in thermal characteristics of the package material and the die, thermal resistance exists between these two materials. Similarly, there is a thermal resistance between the case and external coolant, i.e., air. The total thermal resistance of a package is normally expressed as the sum of internal and external thermal resistance [1]. The internal thermal resistance from junction to case ( $\theta_{JC}$ ) is the characteristic of the package material and package type. The external thermal resistance from case to ambient ( $\theta_{CA}$ ) is dependent on the package style, cooling method employed on the chip (heat sinks) and in the system [27].

Chips perform satisfactorily upto a specified maximum temperature depending on the technology used. For example, TTL plastic devices perform well upto a maximum junction temperature of 125 $^{\circ}$ C. Therefore, the case temperature of a chip must be maintained such that junction temperatures do not exceed the maximum limit. Ambient temperature ( $T_A$ ) limits can be calculated from the thermal resistance characteristics ( $\theta_{JA}$ ) using the equations given below.

$$\begin{aligned} \theta_{JA} &= \theta_{JC} + \theta_{CA} \\ T_C &= T_{jmax} - \theta_{JC} * P_d \\ T_A &= T_C - \theta_{CA} * P_d \\ T_A &= T_{jmax} - \theta_{JA} * P_d \end{aligned} \quad (2)$$

where  $T_C$  is the case temperature,  $T_{jmax}$  is the maximum junction temperature of the devices in the chip, and  $P_d$  is the power dissipation of the chip. Chip manufacturers specify case temperature limits for their devices so that junction temperatures do not exceed 70 $^{\circ}$ -85 $^{\circ}$ C.

Since  $\theta_{CA}$  is dependent on cooling provisions made in the system, reliability of the devices can be improved by providing heat sinks and forced air cooling.

The effect of cooling on  $\theta_{CA}$  for two different packages of the same chip [14] is shown in Figure 2.

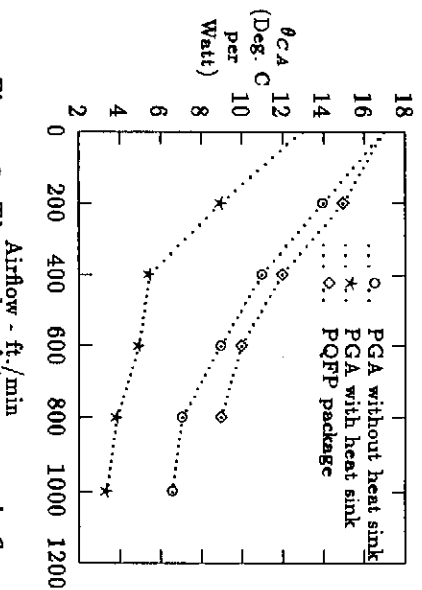


Figure 2: Thermal resistance vs. air flow.

$\theta_{CA}$  depends on the material and the style of the package. It is reduced from 17°C/Watt to 13°C/Watt for pin grid array(PGA) package by providing heat sink and is reduced from 17°C/Watt to 7.1°C/Watt by providing 800 ft./min airflow. For plastic quad flat pack(PQFP)  $\theta_{CA}$  is reduced from 17°C/Watt to 9°C/Watt by providing the same amount of air flow. With the above changes in  $\theta_{CA}$ , reliability can be improved by about three times. Further improvements in thermal resistance are possible by applying better cooling methods, i.e., liquid coolants [17].

The relationship between the thermal resistance and failure rate of a chip is given in equations 1 and 2. Failure rates of a 32 bit microprocessor (Intel 80486) are calculated with different values of  $\theta_{JA}$ . The results are shown in Figure 3. It is evident that the device reliability can be improved with innovative packaging and cooling techniques.

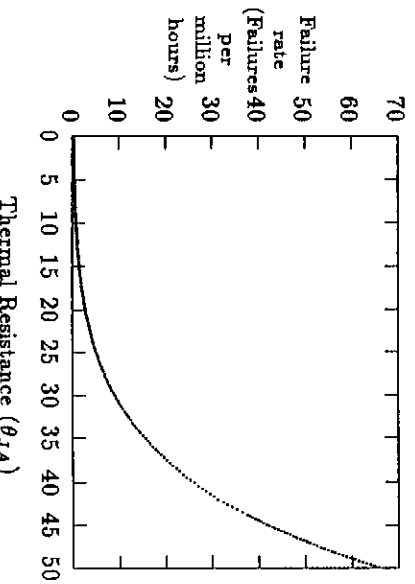


Figure 3: Failure rate vs. Thermal resistance.

#### 4.2 Power dissipation

Power dissipation of VLSI devices vary over a wide range. It varies from few microwatts to milliwatts per gate for different technologies. Estimating the actual power dissipation in VLSI circuits is somewhat complicated due to the fact that actual power dissipation

is application dependent. However, maximum power dissipation can be estimated from layout details. For example, power requirements of a CMOS inverter can be estimated using equation 3 given below.

$$P_S = V_{dd}^2 f C_g (W_p L_p + W_n L_n) \quad (3)$$

where  $P_S$  is the power dissipation of the inverter,  $V_{dd}$  is the supply voltage,  $f$  is the frequency of switching,  $C_g$  is the average gate capacitance per unit area of the channel,  $W_p$  and  $L_p$  are the width and the length of the p-channel, respectively, and  $W_n$  and  $L_n$  are the width and the length of the n-channel, respectively.

It is evident that the power dissipation of a chip is a function of operating voltage, operating/switching frequency, capacitive load and feature size. From this basic inverter power dissipation, total power dissipation can be estimated. The disadvantage of this method is that there can be many types of circuits, having structures different from basic inverter. Also, in this method power requirement is overestimated by many-fold. To have realistic power estimates, simulation of typical application environments are suggested [9]. Device manufacturers calculate the maximum power dissipation by measuring supply current ( $I_{cc}$ ) under worst case conditions.

Power dissipation for different functional blocks vary within the chip of a particular technology. Power dissipation depends on

- Number of transistors used per gate/cell.
- Capacitive loads to be driven.
- Regularity of the structure.

For example, I/O buffers dissipate power in the order of milliwatts where as memory cells consume very little power, i.e., in the order of a microwatt, thereby allowing very dense memory chips [26]. Power dissipation of random logic is in between these two extremes. Power dissipation values of these functional blocks (derived from [10, 12, 22]) are given for three technologies in Table 3.

Due to these variations in power dissipation, the level of integration that can be achieved is dependent both on the technology and the functional logic. Memory devices have already reached the densities of 16 Mbits per chip [12, 22]. Relationship between power dissipation and level of integration that can be achieved for processors is shown in Figure 4. It is assumed that each processor has 400000 transistors and the thermal characteristics of the package(PQFP) are the same as those shown in Figure 2. The maximum allowed junction temperature is 125°C. To simplify the calculations only processors are considered. Similar calculations can be done for any specific architecture such as the multiple-bus architecture described in Section 2.

#### 5 Die size

Even though the VLSI manufacturing technologies have reached submicron feature size, die sizes continue to grow. As more and more logic is integrated on a

Table 3: Power dissipation of different functional blocks

Parameter	Technology	
	CMOS	ECL
Internal gate power dissipation	12 $\mu$ W/MHz	240 $\mu$ W
Output driver power dissipation	25 $\mu$ W/MHz/pF	10mW
DRAM memory cell	0.1 - 2.0 $\mu$ W	-

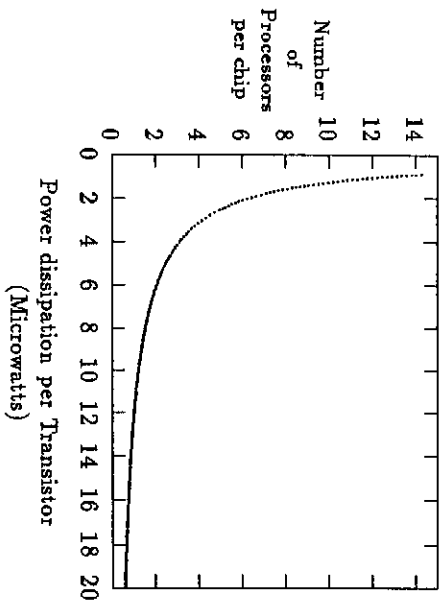


Figure 4: Level of integration vs Power dissipation

chip, the input/output terminal requirement also goes up which further increases the die size. The escalation in die size has direct impact on yield, packaging and reliability.

### 5.1 Die size and thermal resistance

The thermal resistance of a chip, as explained in Section 4, depends on die size and package style. As die size grows, the heat flux is reduced resulting in lower thermal resistance  $\theta_{JC}$ . For example,  $\theta_{JC}$  is 28°C/Watt for a dual in-line package (DIP) if the die size is smaller than 9.5 mm<sup>2</sup>. For the same package,  $\theta_{JC}$  is less than 11°C/Watt if the die size is larger than 10 mm<sup>2</sup> [24]. As die size increases, an appropriate package style is selected based on heat fluxes and terminal count. Therefore, larger dies and packages have better thermal characteristics [4, 21]. However, the die size growth precipitates many assembly and reliability problems due to enhanced stresses and strains in the package [15]. The variations in thermal expansion coefficients may create problems such as die cracking, molded plastic package cracking etc.

### 5.2 Die size and yield

The most important aspect of die size is its close relationship with the yield of the chip in the fabrication lines. For chips larger than 2 cm<sup>2</sup>, probably yield is the limiting factor for the level of integration compared to reliability. Chip sizes of the high performance general purpose microprocessors had already crossed 2 cm<sup>2</sup>.

For example, the size of the Motorola 68040 chip, a 32 bit microprocessor, is 2.25 cm<sup>2</sup> [3]. In Section 6, the chip area is estimated for different integration levels for the multiple-bus architecture. The size of the chip varies from 1.2 cm<sup>2</sup> for a single processor-memory pair chip to 8.12 cm<sup>2</sup> for a 8 processor-memory pair chip. For chips of this size, the three parameter generalized negative binomial yield model given in equation 4 was found to be more accurate than other yield models [7, 19].

$$Y = Y_0(1 + D_0A/\alpha)^{-\alpha} \quad (4)$$

where  $Y$  is the yield of the die,  $Y_0$  is the gross yield factor,  $D_0$  is the average number of defects per unit area,  $A$  is the area of the die and  $\alpha$  is the clustering parameter.

Even for a mature product and process line, the average number of defects per cm<sup>2</sup> ( $D_0$ ) is more than 1 defect/cm<sup>2</sup>. In chip fabrication lines, average defect rate of 1.5 defects/cm<sup>2</sup> is not uncommon. For a new product or process line it can be as high as 5 to 10 defects/cm<sup>2</sup>. The yield of the chips of the multiprocessor system with different integration levels is shown in Figure 5. Yield is estimated for values of  $\alpha$  between 0.3 and 3.0, which represent highly clustered and minimally clustered point defects, respectively. Two values are considered for  $D_0$  (1.0 and 1.5) and a fixed value for  $Y_0$  (0.95). For a product/process line with  $\alpha = 1.0$  and  $D_0 = 1.5$  for example, the yield is 22% for a chip with 2 processor-memory pairs (area of the chip is 2.19 cm<sup>2</sup>) and it drops to about 7% for a chip with 8 processor-memory pairs (area of the chip is 8.12 cm<sup>2</sup>). Even with the state-of-art VLSI technology, chips can not be produced at competitive prices if the yields are well below 10%. It is a long way before yield of these larger chips can be improved above 30-40%. Chip sizes are bound to increase in future and thus, new methods and techniques are to be found to improve the yield of the larger chips.

### 5.3 Terminal count and die size

The number of signal terminals (pins) to be provided on a chip depends on the type of logic, i.e., memories, random logic and gate arrays [15]. Memory chips need the least number of terminals per cell whereas gate arrays require the highest number of terminals per gate. The relationship, proposed by Rent [20], between the number of terminals and size of the

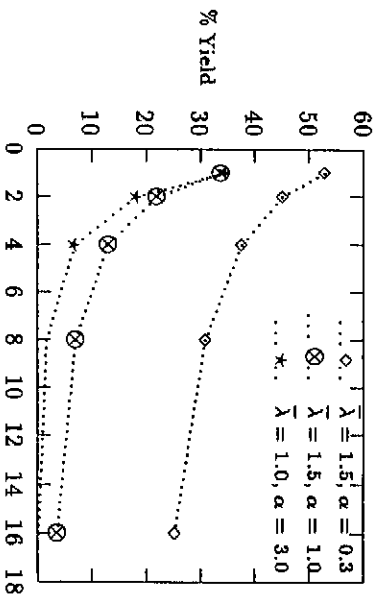


Figure 5: Number of Processor-Memory pairs vs. Yield

logic (number of logic gates) is given in equation 5.

$$N_P = K_P N_G^{K_R} \quad (5)$$

where  $N_P$  is the number of pins,  $K_P$  is a constant, the value of which depends on the ability to share signal lines,  $N_G$  is the number of gates in the logic and  $K_R$  is Rent's constant.

In multiprocessor systems, the terminal count depends on the architecture of the interconnection network. If shared local and global buses are used, the terminal count increases only marginally as the level of integration increases. In addition to signal pins multiple power and ground pins are provided in larger chips to minimize power surges. The number of power pins required on a chip depends on many parameters such as current requirements of the logic, number of output lines that might operate simultaneously, noise immunity etc. In this analysis 40% of the signal pins are assumed for power and ground pins.

## 6 Reliability analysis

The multiple-bus multiprocessor system presented in Section 2 is used as an example to analyze the impact of integration on reliability of multiprocessor systems. This system can be realized with individual processor and memory chips or with chips having several processor and memory modules. The chip count in the system depends on the level of integration of the chips used in the system. For example, the chip count for a 32 processor-memory pair system will be reduced from 64 to 32 if chips with one processor and one memory module are used in the system and it will be further reduced to 16 if chips with 2 processor-memory pairs are used. In the extreme case, the entire multiprocessor system can be realized on a single chip. However, these integration levels can not be achieved due to limitations of present day fabrication and packaging technologies. As explained in Section 3, the failure rates of the chips increase beyond acceptable limits at these integration levels. Failure rates of the system with chips of different integration levels are given below.

### 6.1 Integration level and failure rate of the chips

Chips with different configurations are considered for the reliability analysis. To estimate the failure rates of the chip, power dissipation, pin count and die size of the chip are calculated. These parameters for different configurations are shown in Table 5. The values chosen for calculating the above parameters are shown Table 4 [2, 16, 18, 28].

Table 4: Area and power dissipation specifications

Parameter	Value
Power dissipation per transistor	5 $\mu W$
Power dissipation per memory cell	0.1 $\mu W$
Power dissipation per signal pad	10 mW
Area per transistor	200 $\mu m^2$
Area per memory cell	8 $\mu m^2$
Area per signal pad	0.15 mm <sup>2</sup>

Failure rates are calculated for chips with different configurations of processor and memory modules using equation 1. Level of integration versus failure rates are shown in Figure 6. The failure rate of the overall system decreases marginally if chips, with upto 2 processor-memory pairs, are used. This decrease is with respect to the failure rate of the system when chips with individual processor and memory elements are used. It increases marginally with 4 processor-memory pairs. Beyond this level of integration, the failure rate increases drastically. The failure rate of the system increases almost 5 times from 45 to 245 failures/million hours as integration level increases 8 times, i.e., from single processor-memory pair to 8 pairs.

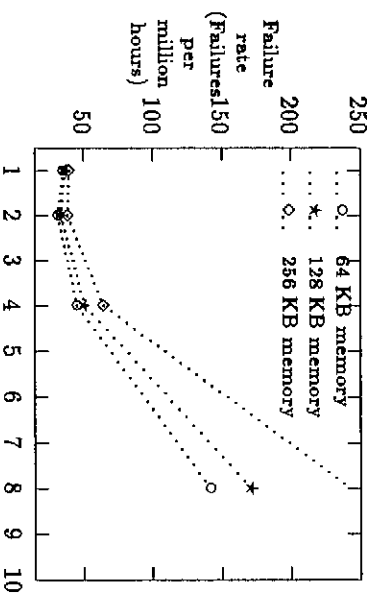


Figure 6: Failure rate vs. Level of integration.

The improvement in reliability at 2 processor-memory pairs is due to decrease in I/O pin count, i.e., pins are shared by elements. This improvement can not be sustained at higher levels of integration because of limitations of the heat dissipation capabilities of the IC packaging materials/technologies. In this

Table 5: Power dissipation and die size calculations

Level of Integration	Pin Count			Power dissipation (Watts)	Die Size ( $mm^2$ )
	Signal	Power	Total		
I 1 Processor* 1 Memory*	110	44	154	3.1	103
	24	2	28	0.45	22
II 1P + 1M	110	48	158	3.31	120
III 2P + 2M	120	51	168	5.62	219
IV 4P + 4M	140	56	196	10.24	416
V 8P + 8M	180	72	252	19.5	812

\*Processor has 400000 transistors and memory is 256 Kbytes

analysis, packages for only air cooled environments are considered. With better packaging techniques, i.e., packages with lower thermal resistance, reliability improvements can be achieved even at higher integration levels. Power dissipation is another major limiting factor in improving the reliability of the chips. Innovative IC technologies and design styles can minimize the power dissipation so that higher integration levels are possible without any significant reduction in reliability.

As the level of integration increases, overall system reliability improves due to reduction in interconnect hardware such as printed wiring boards, connectors and related input/output logic. Higher levels of integration facilitates better utilization of peripheral logic and thereby minimizes peripheral logic overhead of the system. It is also possible to minimize certain expensive logic, e.g., with on chip memory, high system performance can be achieved with no or limited cache. Similar trade-offs may be possible with different architectures.

## 7 Conclusions

The level of integration in an IC has direct impact on its performance, reliability and yield. To find an optimum level of integration we need to look at all these three parameters. Yield and reliability of the multiple-bus multiprocessor system have been analysed in Sections 5 and 6, respectively. The performance of this system is analysed in detail in [8]. In principle, a higher level of integration gives a better performance. As the level of integration increases, processors have more of the memories that they need to access, found on the same chip, i.e., they rely more on local access which is faster than a global access. Consequently, more references are satisfied faster yielding higher processing power.

As illustrated in Section 6, the reliability of the system improves upto a certain level of integration and then decreases. The optimal level of integration for a multiprocessor system depends on the complexity/size of the individual processor and memory modules, architecture of the interconnection network, circuit technology and the thermal characteristics of the package. At higher level of integration significant performance and marginal reliability improvements can be

achieved. Even higher performance is possible at a cost of marginal decrease in reliability.

Currently, the major limiting factor in achieving higher integration levels is the yield of the individual chips. As illustrated in Section 5, the yield of a chip at integration levels higher than 4 processor-memory pairs per chip can be as low as 5%. This drastic reduction in yield can not be compensated by the savings due to reduction in interconnect and peripheral logic. Therefore, systems can be prohibitively expensive at very high integration levels and performance and reliability improvements can not be justified. In order to exploit the benefits of higher integration fully, yield improvement techniques need to be addressed at the design stage.

## References

- [1] M. Aghazadeh and B. Natarajan, "Parametric Study of Heatspreader Thermal Performance in 48 Lead Plastic DIP's and 68 Lead Plastic Leaded Chip Carriers," *IEEE Trans. Components, Hybrids, Manufact. Technology*, Vol. CHMT-9, No. 4, pp. 347-352, December 1986.
- [2] R. L. Allmon et al., "CMOS Implementation of a 32b Computer," *IEEE International Solid-State Circuits Conference*, 1989, pp. 80-81.
- [3] D. Anderson et al., "The 68040 32-b Monolithic Processor," *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 5, pp. 1178-1189, October 1990.
- [4] A. Bar-cohen, "Thermal Management of Air- and Liquid-cooled Multichip Modules," *IEEE Trans. Components, Hybrids, Manufact. Technology*, Vol. CHMT-10, No. 2, pp. 159-175, June 1987.
- [5] B. D. Boschma et al., "A 30 MIPS VLSI CPU," *IEEE International Solid-State Circuits Conference*, 1989, pp. 82-83.
- [6] R. M. Burger et al., "The Impact of ICs on Computer Technology," *IEEE J. Computer*, pp. 88-95, October 1984.
- [7] J. A. Cunningham, "The Use and Evaluation of Yield Models in Integrated Circuit Manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 3, No. 2, pp. 60-71, May 1990.

- [8] S. Deshmukh and I. Koren, "A Modular and Highly-Integrated Multiple-bus Multiprocessor System," Tech. Rep. TR-90-CSE-17, ECE Dept., Univ. of Massachusetts, Amherst, 1990.
- [9] S. Devadas et al., "Estimation of Power Dissipation in CMOS Combinational Circuits," *IEEE Custom Integrated Circuits Conference*, pp. 19.7.1 - 19.7.6, 1990.
- [10] J. D. Giacomo, *VLSI handbook*, McGraw-Hill Publishing Company, NY, 1989.
- [11] M. Hanawa, T. Nishimukai, et al., "On-chip Multiple Superscalar Processors with Secondary Cache Memories," *Proc. of ICCD*, pp. 128-131, Oct. 1991.
- [12] H. L. Kater et al., "An Experimental 50-nsec 16-Mbit DRAM with 10-nsec Data Rate," *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, paper FPM 14.4 February, 1990.
- [13] Intel *i486 Microprocessor Hardware Reference Manual*.
- [14] Intel *80960CA Embedded Processor Hardware Reference Manual*.
- [15] B. C. Johnson, "High Performance Integrated Circuit Packaging," *IEEE Custom Integrated Circuits Conference*, pp. 23.1.1 - 23.1.6, 1988.
- [16] N. P. Jouppi, J. Y. F. Tang, and J. Dion, "A 20 MIPS Sustained 32b CMOS Microprocessor with 64b Data Bus," *IEEE International Solid-State Circuits Conference*, pp. 84-85, 1989.
- [17] T. Kishimoto and T. Ohsaki, "VLSI Packaging Technique using Liquid-cooled Channels," *IEEE Trans. Components, Hybrids, Manufact. Technology*, Vol. CHMT-9, No. 4, pp. 328-335, December 1986.
- [18] L. Kohn, and S. W. Fu, "A 1,000,000 Transistor Microprocessor," *IEEE International Solid-State Circuits Conference*, 1989, pp. 54-55.
- [19] I. Koren and C. H. Stapper "Yield Models for Defect Tolerant VLSI Circuits: A Review," *Defect and Fault Tolerance in VLSI Systems*, Vol. 1, I. Koren (ed.), pp. 1-21, Plenum, 1989.
- [20] B. S. Landman and R. L. Russo, "On a Pin versus Block Relationship for Partitions of Logic Graphs," *IEEE Trans. Computers*, Vol. C-20, pp. 1469-1479, 1971.
- [21] Y. C. Lee, H. T. Ghaflari, and J. M. Segelken, "Internal Thermal Resistance of a Multi-chip Packaging Design for VLSI based Systems," *IEEE Trans. Components, Hybrids, Manufact. Technology*, Vol. 12, No. 2, pp. 163-169, June 1989.
- [22] T. Mano, "Circuit Technologies for 16Mb DRAMs," *IEEE International Solid-State Circuits Conference*, 1987, pp. 22-23.
- [23] *Military Handbook "Reliability Prediction of Electronic Equipment"*, MIL-HDBK-217E, 1986.
- [24] *Notice of Change to MIL-HDBK-217E*, Notice 1, 2 January 1990.
- [25] S. F. Morris, "Use and Application of MIL-HDBK-217," *Solid State Technology*, pp. 65-69, August 1990.
- [26] C. A. Neugebauer, and R. O. Carlson, "Comparison of Wafer Scale Integration with VLSI Packaging Approaches," *IEEE Trans. Components, Hybrids, Manufact. Technology*, Vol. CHMT-10, No. 2, pp. 184-189, June 1987.
- [27] W. E. Pence and J. P. Krusius, "The Fundamental Limits for Electronic Packaging and Systems," *IEEE Trans. Components, Hybrids, Manufact. Technology*, Vol. CHMT-10, No. 2, pp. 176-183, June 1987.
- [28] T. S. Perry, "Intel's Secret is Out," *IEEE Spectrum*, pp. 22-28, April 1989.