A Case for Redundant Arrays of Inexpensive Disks (RAID)

David A Patterson, Garth Gibson, and Randy H Katz

Computer Science Division
Department of Electrical Engineering and Computer Sciences
571 Evans Hall
University of California
Berkeley, CA 94720
(pattrsn@ginger berkeley edu)

Abstract *Increasing performance of CPUs and memories will be squandered if not matched by a similar performance increase in I/O While the capacity of Single Large Expensive Disks (SLED) has grown rapidly, the performance improvement of SLED has been modest Redundant Arrays of Inexpensive Disks (RAID), based on the magnetic disk technology developed for personal computers, offers an attractive alternative to SLED, promising improvements of an order of magnitude in performance, reliability, power consumption, and scalability This paper introduces five levels of RAIDs, giving their relative cost/performance, and compares RAID to an IBM 3380 and a Fujitsu Super Eagle*

1 Background: Rising CPU and Memory Performance

The users of computers are currently enjoying unprecedented growth in the speed of computers Gordon Bell said that between 1974 and 1984, single chip computers improved in performance by 40% per year, about twice the rate of minicomputers [Bell 84] In the following year Bill Joy predicted an even faster growth [Joy 85]

$$MIPS = 2^{Year-1984}$$

Mainframe and supercomputer manufacturers, having difficulty keeping pace with the rapid growth predicted by "Joy's Law," cope by offering multiprocessors as their top-of-the-line product.

But a fast CPU does not a fast system make Gene Amdahl related CPU speed to main memory size using this rule [Siewiorek 82]

Each CPU instruction per second requires one byte of main memory,

If computer system costs are not to be dominated by the cost of memory, then Amdahl's constant suggests that memory chip capacity should grow at the same rate Gordon Moore predicted that growth rate over 20 years ago

$$transistors/chip = 2^{Year-1964}$$

As predicted by Moore's Law, RAMs have quadrupled in capacity every two [Moore 75] to three years [Myers 86]

Recently the ratio of megabytes of main memory to MIPS has been defined as *alpha* [Garcia 84], with Amdahl's constant meaning alpha = 1 In part because of the rapid drop of memory prices, main memory sizes have grown faster than CPU speeds and many machines are shipped today with alphas of 3 or higher

To maintain the balance of costs in computer systems, secondary storage must match the advances in other parts of the system A key measure of magnetic disk technology is the growth in the maximum number of bits that can be stored per square inch, or the bits per inch in a track times the number of tracks per inch Called M A D , for maximal areal density, the "First Law in Disk Density" predicts [Frank87]

$$MAD = 10^{(Year-1971)/10}$$

Magnetic disk technology has doubled capacity and halved price every three years, in line with the growth rate of semiconductor memory, and in practice between 1967 and 1979 the disk capacity of the average IBM data processing system more than kept up with its main memory [Stevens81]

Capacity is not the only memory characteristic that must grow rapidly to maintain system balance, since the speed with which instructions and data are delivered to a CPU also determines its ultimate performance The speed of main memory has kept pace for two reasons

(1) the invention of caches, showing that a small buffer can be managed automatically to contain a substantial fraction of memory references,

(2) and the SRAM technology, used to build caches, whose speed has improved at the rate of 40% to 100% per year

In contrast to primary memory technologies, the performance of single large expensive magnetic disks (SLED) has improved at a modest rate These *mechanical* devices are dominated by the seek and the rotation delays from 1971 to 1981, the raw seek time for a high-end IBM disk improved by only a factor of two while the rotation time did not change[Harker81] Greater density means a higher transfer rate when the information is found, and extra heads can reduce the average seek time, but the raw seek time only improved at a rate of 7% per year There is no reason to expect a faster rate in the near future

To maintain balance, computer systems have been using even larger main memories or solid state disks to buffer some of the I/O activity This may be a fine solution for applications whose I/O activity has locality of reference and for which volatility is not an issue, but applications dominated by a high rate of random requests for small pieces of data (such as transaction-processing) or by a low number of requests for massive amounts of data (such as large simulations running on supercomputers) are facing a serious performance limitation

2. The Pending I/O Crisis

What is the impact of improving the performance of some pieces of a problem while leaving others the same? Amdahl's answer is now known as Amdahl's Law [Amdahl67]

$$S = \frac{1}{(1-f) + f/k}$$

where

S = the effective speedup,
f = fraction of work in faster mode, and
k = speedup while in faster mode

Suppose that some current applications spend 10% of their time in I/O Then when computers are 10X faster--according to Bill Joy in just over three years--then Amdahl's Law predicts effective speedup will be only 5X When we have computers 100X faster--via evolution of uniprocessors or by multiprocessors--this application will be less than 10X faster, wasting 90% of the potential speedup