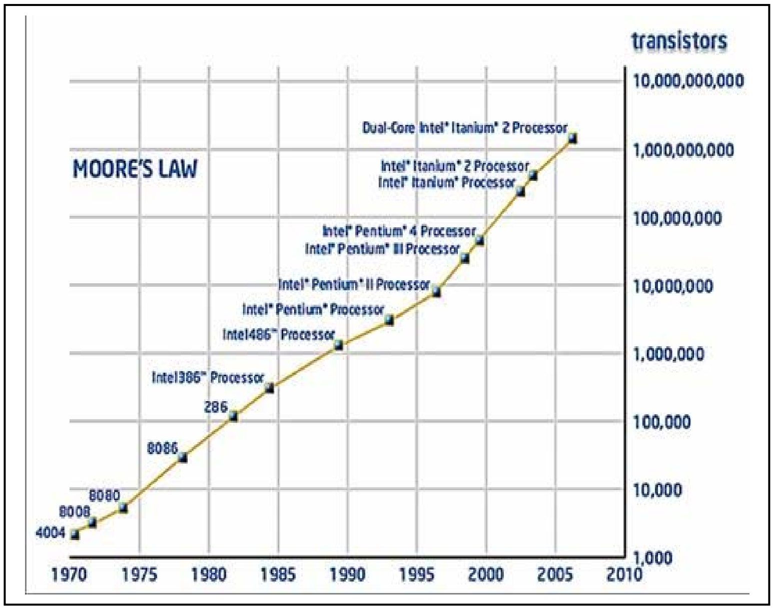
Scientific Data - The gift that keeps on giving!



NASA is scheduled to launch its Solar Dynamics Observatory mission in late 2009. The three instruments on board the satellite will take lots and lots of pictures of the Sun. So it will require lots and lots of disc space to store the data.

Imagine you are the engineer responsible for purchasing the computer systems to handle that data. You must figure out how many computer discs and tapes you'll need, and what they will cost, so you can ask NASA for the right amount of money in your mission budget.

8 bits	= 1 byte	
1 kilobyte	= 1 thousand bytes	10^3
1 megabyte	= 1 million bytes	10^6
1 gigabyte	= 1 billion bytes	10^9
1 terabyte	= 1 trillion bytes	10^{12}
1 petabyte	= 1 quadrillion bytes	10^{15}

Data rate:	130 megabits / second
Hours of data/day:	24 hrs
Size of disc drive:	1 terabyte
Cost per disc drive:	$300
Size of backup tape:	800 gigabytes
Cost per backup tape:	$60

Problem 1 - How many bytes will your discs need to hold each day? (Express your answer in terabytes)

Problem 2 - NASA would like you to keep about 60 days of data online (i.e. on disc). Data older than 60 days will be archived and copied to tapes. How many bytes of drive space will you need to hold 60 days of data?

Problem 3 - How many disc drives will you need to purchase, and how much will they cost?

Problem 4 - How many terabytes of tape storage will you need to archive a year's worth of data?

Problem 5 - How many tapes will you need, and what will they cost you?

Problem 6 - There is a rule of thumb (called Moore's Law) that says the costs for electronics halve every 18 months. If you can wait 18 months before you purchase your discs and tapes, estimate how much money you might save.

Answer Key

Problem 1 - Answer: The data rate is 130 megabits per second. We first convert this to a daily rate. There are 60 sec/min and 60 min/hr and 24 hr/day so there are 60 x 60 x 24 = 86,400 sec/day. Since the data rate is 130 x 10^6 bits/sec or 1.3 x 10^8 bits/sec, in one day there will be 1.3 x 10^8 bits/sec x 86,400 sec/day = 1.12 x 10^{13} bits/day. Since there are 8 bits in 1 byte, we have 1.12 x 10^{13} bits/day x (1 bytes/8 bits) = 1.4 x 10^{12} bytes/day. Since 1 terabyte = 10^{12} bytes, we have 1.4 x 10^{12} bytes/day x (1 terabyte/10^{12} bytes) = 1.4 terabytes/day.

Problem 2 - Answer: For 60 days at 1.4 terabytes/day there will be 60 days x 1.4 terabytes/day = 84 terabytes of data.

Problem 3 - Answer: 1 disk drive stores 1 terabyte of data so you will need 84 disk drives. Since the price is $300 per disk drive, the total cost will be 84 drives x $300/drive = $25,200.

Problem 4 - Answer: At a rate of 1.4 terabytes/day for 365 days you will have 1.4 terabytes/day x 365 days = 511 terabytes of data to archive.

Problem 5 - One tape holds 800 gigabytes of data. We first convert this to terabytes. Since 1,000 gigabytes = 1 terabyte, we have 800 gigabytes x (1 terabyte/1000 gugabytes) = 0.8 terabytes per tape. Then for 511 terabytes of data in one year, we need 511 terabytes/year x (1 tape/0.8 terabytes) = 638.75 tapes, but since we cannot buy a fraction of a tape ,we round up to 639 tapes. Each tape costs $60, so the total annual cost for backup tapes will be 639 tapes x ($60/1 tape) = $38,340

Problem 6 - Answer: The cost for drives will fall from $25,200 to one-half this cost or $12,600, and the cost for the backup tapes will fall from $38,340 to only $19,170. For the combined purchases, you would save $12,600 + $19,170 = $31,770.