

VLBA Sensitivity Upgrade MEMO 22

The Evolution of Hard Drive Prices Over 2000 Days

Walter Brisken, Jon Romney (National Radio Astronomy Observatory)

2008 July 16

Abstract

The price evolution of hard disks over about 2000 days (5.5 years) is presented and interpreted. Some predictions of disk prices in the near future are made.

1 Introduction

Since the early Mark5 days I have been periodically (roughly once per week) collecting the prices of PATA and SATA disks from the Pricewatch (<http://www.pricewatch.com/>) web page. This web page automatically mines the web sites of many computer vendors and reports the lowest price for any particular part. For hard disks, sizes are kept distinct, but all brands of drives are considered together, as are models within a particular brand, so the brand and model associated with each size will vary with time — there is no attempt to track this here. The prices reported often have strings attached, such as “one drive per purchase”, “credit card purchase only”, and “no written quote available”, and the price of shipping is typically not included, so the prices are not perfectly relevant to the purchasing needs of a VLBI operation. In some cases clearly inaccurate prices were censored as there were some discrepancies at the order of magnitude level, but for the most part all data were retained. That said, the long term history of prices is thought to track the trends within the market, both over time and disk size, fairly accurately.

2 The data

All collected data are taken from http://www.pricewatch.com/hard_removable_drives/. This web page appears to have no archival data, so prices from before the beginning of this exercise are not available; a request for archival prices was sent to the maintainer of the web page without any response. Data for PATA disks were collected starting Jan 14, 2003. The 250 GB disk was brand new at this time and it was the largest available drive until August 8, 2003 when the 300 GB disk was introduced. Collection of SATA disks started immediately after their availability, on April 20, 2003.

3 Price evolution of a particular disk

The price history for 250 and 500 GB disks is shown in figures 1 and 2 respectively. The exponential law like evolution that characterized the price to size ratio (which will be demonstrated in 4) does not apply well to individual disk sizes over the entire period of availability, but appears to apply during the product mid-life, when the disks are most interesting from an economic point of view. At the introduction of a new disk size the “first kid on the block” phenomenon drives prices upwards. In the case of 250 GB disks it took about 2.5 years for the price to size ratio to be competitive with smaller (typically 120 GB) disks. As a disk size becomes old the demand becomes minimal so the price drops to a very slowly changing price which is likely dominated by the production cost.

4 Best value evolution through time

A useful metric for long term planning is *best value*, the smallest ratio of price to size of any available disk (i.e., U.S. Dollar per GB). At any given time a particular disk size will maximize this metric, though it is often the case that two disk sizes have nearly the same value and the choice of optimal size to purchase may then depend on other factors; VLBI applications will typically favor the larger drives in these situations to minimize complete module and shipping costs. The plot of hard disk best value is

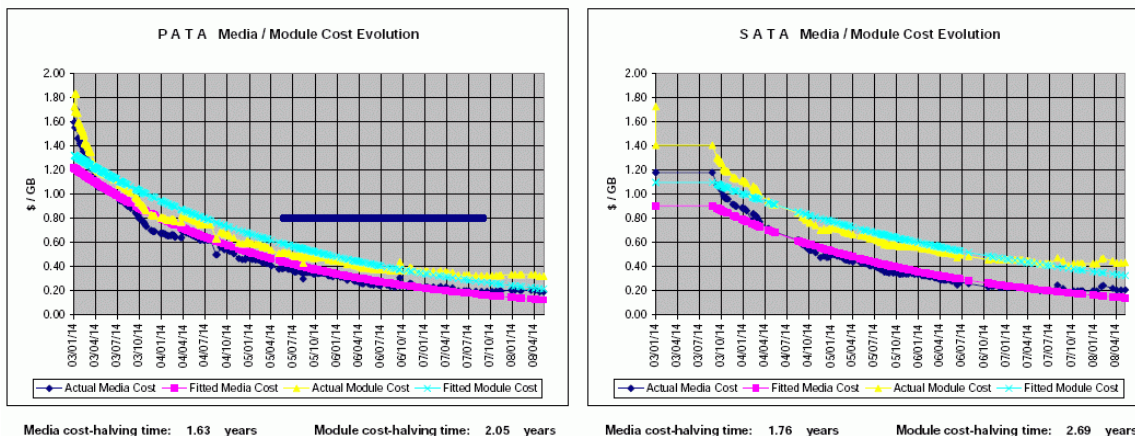


Figure 1: The price history of 250 GB hard disks. PATA and SATA disk prices are plotted separately in the left and right panels respectively. Four curves are plotted on each panel: blue triangles are the actual media cost in U.S. Dollars per GB of storage; the yellow triangles track the price per GB when including the Mark5 module price; magenta and cyan are best fit exponentials to the blue and yellow respectively. The thick horizontal bar on the PATA plot signifies the period of time when this disk size was optimal in the price per size sense.

shown in figures 3 & 4. An attempt to fit an exponentially decreasing price per size causes one to suspect a break in the exponent around year 2004.2, where the price halving timescale changed from 1.33 years to 2.28 years. A model predicting the best value purchase as a function of time is given by

$$v = \begin{cases} 1.0 \exp(-d/500) & \text{if } d < 438 \text{ (before 2004.2)} \\ 0.77 \exp(-d/1200) = 0.53 \exp(-(d - 438)/1200) & \text{otherwise} \end{cases} \quad (1)$$

where d is days since 2003 Jan 01 and v is best value in \$ per GB.

Note that for the first 3 or so years SATA disks commanded a premium, but now the cost of the two types is nearly identical. The market has clearly shifted towards favoring SATA. The two recent disk size grades (750 GB and 1 TB) were available in SATA form well before PATA form. SATA offers greater bandwidth and improved connectorization and is clearly the choice product for most applications.

5 Comparison with purchase prices

Figure 3 shows in addition to the `pricewatch.com` prices for hard drives the prices actually paid for discs ordered for the assembly of Mark5 modules; the prices are also tabulated in table 1. The quantities of disks ordered ranged from 40 to 600 disks; there is no apparent dependency on the order size on the price per unit. With the exception of a purchase of 500 GB disks around year 2006.56, a time when their novelty still commanded a premium, all purchases were between 39 and 48% more expensive than the model of the `pricewatch.com` metric.

6 Future predictions

The goal for the VLBA is to reach 4 Gbps recording rate in 2011. For lack of a better date I will use Sep 30, 2011 (day 3194), the end of that fiscal year, as the fiducial day for prediction of disk prices and sizes for that year. The `pricewatch.com` measurements predict that media will cost \$0.054 per GB. If our realized purchase price continues to be about 44% greater than this, we will be paying \$0.078 per GB.

500 GB

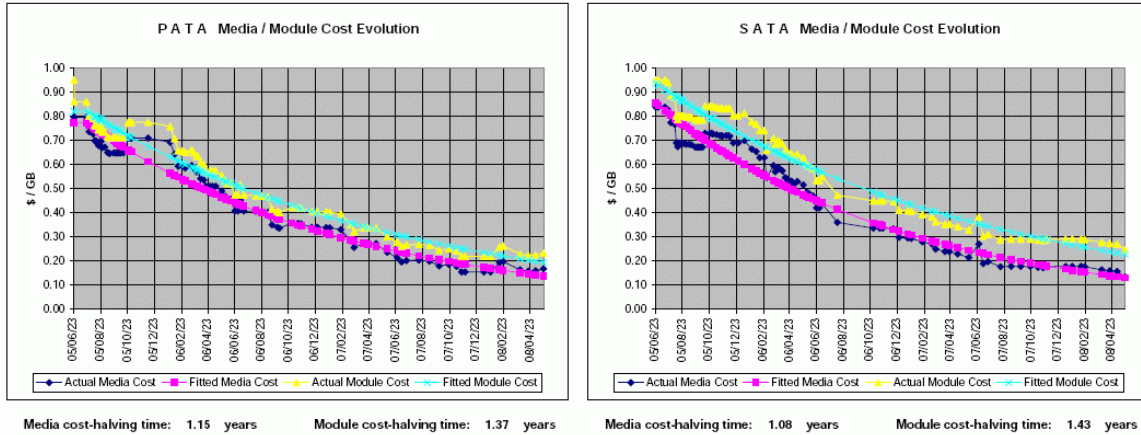


Figure 2: The price history of 500 GB hard disks. PATA and SATA disk prices are plotted separately. See figure 1 for explanation.

At this time the largest individual disks will be between 3 and 4 TB, but the larger of these will still be excessively expensive, so it is reasonable to assume that we would be purchasing roughly 2.5 TB disks, each of which will cost \$194, but the exact details will depend on which disk sizes the industry chooses to standardize. If one assumes a SATA module shell still costs \$435, then a 20 TB SATA module will cost \$1987. To calculate the total media that would be required for the VLBA, I assume a 30 day disk turn-around period (i.e., a complete record and correlate cycle), 75% observing duty cycle, 10% spares on 10 antennas, resulting in a total of 11000 TB, or 550 disk modules, for a total cost of \$1.1M. A substantial fraction of the disk purchases would occur much earlier than 2011.75, so the total investment in disks will certainly be substantially greater.

Note that two 20 TB modules will be able to record at 4 Gbps for 78 ksec, or just under 1 day. Continuous uninterrupted recording over 64 hour unattended weekends would require two 60 TB modules (probably not available until late 2013 at the earliest) or four 30 TB modules (and 2 Mark5 units per station) which will be available sometime in 2011, but likely at a slight cost premium.

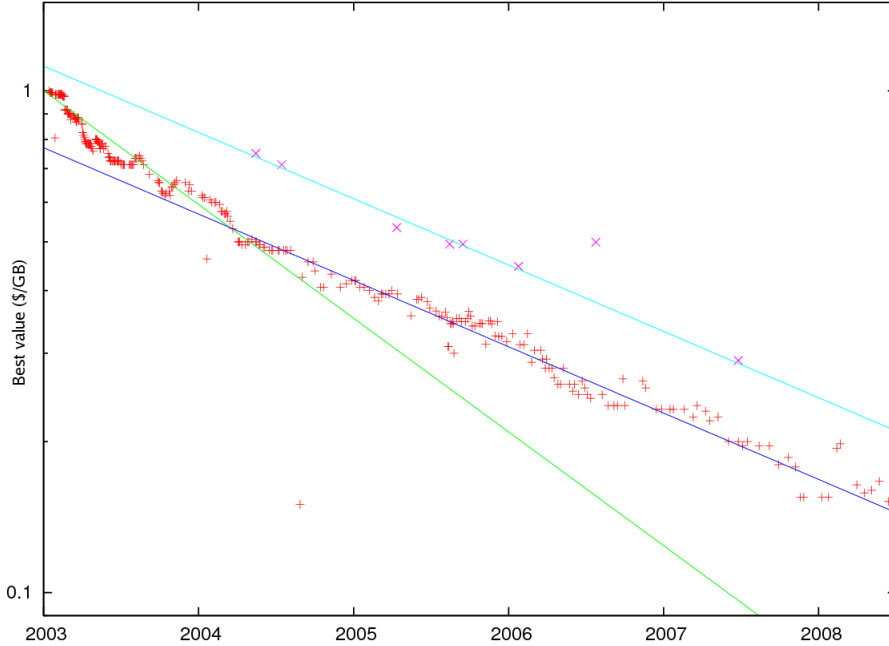


Figure 3: The evolution of hard disk through time. Red marks (+) show the lowest price per size of any disk available on the particular day. This metric is well fit by a broken exponential law shown with the green and blue lines; equation 1 parameterizes the two exponential curves. In magenta are the actual price per size paid for disks. All purchases were made after the 2004.2 exponential law break, and all (except one – see Sec 5) are well fit by the cyan exponential law with the same exponent that characterizes the `pricewatch.com` disk prices, but with a magnitude about 44% higher.

Date	Size (GB)	price (\$)	\$/GB	model \$/GB	ratio
2004.36	250	187.50	0.750	0.508	1.48
2004.54	250	178.13	0.713	0.482	1.48
2005.28	250	133.50	0.534	0.385	1.39
2005.62	250	123.75	0.495	0.347	1.43
2005.71	250	123.75	0.495	0.338	1.46
2006.06	300	134.00	0.447	0.303	1.48
2006.56*	500	249.75	0.500	0.260	1.92
2007.48†	500	145.00	0.290	0.200	1.47

Table 1: The actual purchase prices for disks compared to the model (see equation 1) for hard drive best value. With the exception of the purchase of “professional” model disks (marked with a *), the purchase price was consistently about 45% higher. Taking this correction factor into consideration allows one to reliably predict the purchase price of disks, and hence Mark5 modules, probably several years into the future. The † indicates a purchase of SATA disks.

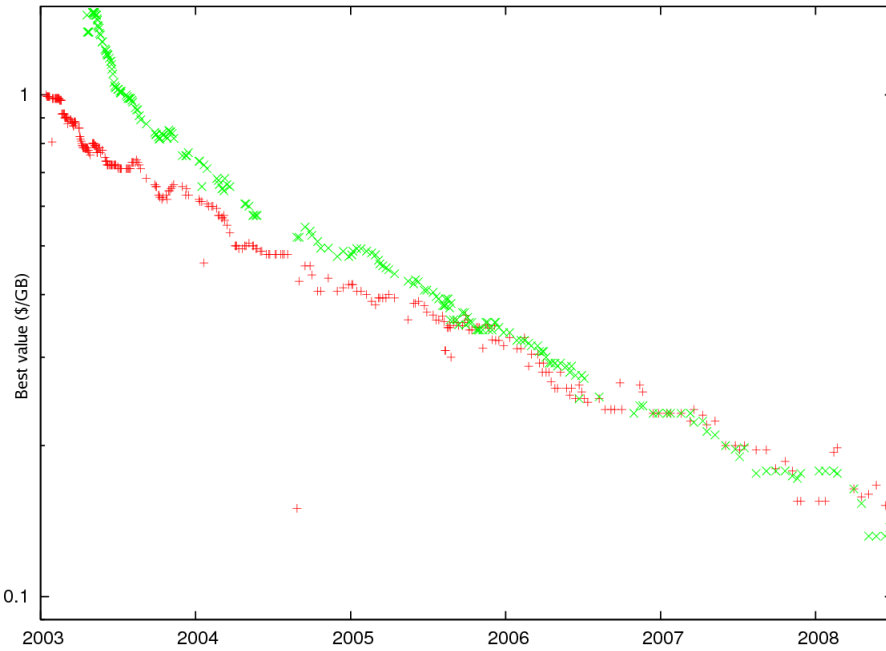


Figure 4: The evolution of SATA disk prices in comparison to PATA prices. Each PATA disk price measurement is shown with red + and each sata measurement with a green x. Until mid 2005 SATA disks commanded a premium, but since that point SATA disks have been comparably priced, with a possible recent trend toward being cheaper.

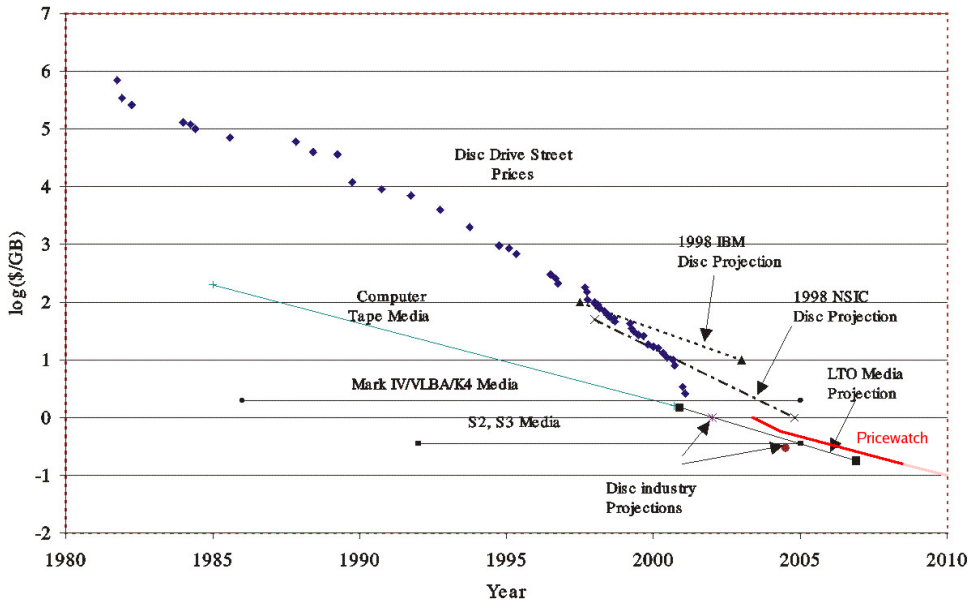


Figure 5: Thirty years of magnetic media price evolution. The data presented here are shown in red in the lower right. Extrapolation of the new data continues in pink to the year 2010. Original figure credit: Alan Whitney.

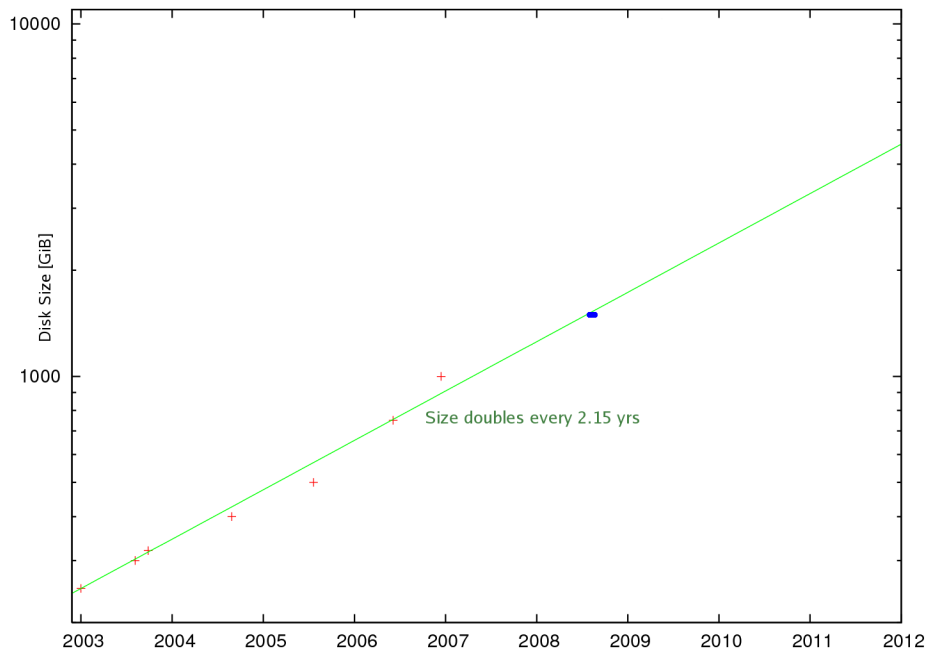


Figure 6: The size of the largest available hard disk as a function of time over the last 5.5 years. The slope is consistent with a size doubling every 2.15 years. Note that as of the submission of this memo, the next largest disk sizes is just being announced: 1.5 TB SATA drives from Seagate will be available in August 2008. The blue oval on this plot marks this expected event.