The Challenges of Magnetic Recording on Tape for Data Storage (The One Terabyte Cartridge and Beyond)

Richard H. Dee
Storage Technology Corporation, One StorageTek Drive, Louisville CO 80028-4274
Tel: +1-303-673-3976, FAX : +1-303-673-8406, richard_dee@storagetek.com

Abstract
Operating points to achieve Terabyte tape cartridge capacities and beyond drive both linear and track densities to values not perceived possible a few short years ago. The primary contributors to the issues related to these high capacities are the physical and magnetic properties of the tape media itself. The total magnetic moment of the recorded bit, driven by the magnetic coating thickness, dominates the recording process and determines the linear recording density possible. Moving a thin tape at high speeds and the mechanical stability in the cross track direction provide engineering challenges for increasing track densities in combination with many parallel channels for high data rates. These issues and trade offs are the main focus of this paper.

1. Introduction

Storing and retrieving data on magnetic tape is driven by (a) capacity (Gbytes/cartridge) primarily because of the cost of storage ($/Gbyte), (b) data rate (Mbytes/second) as people don't want to wait forever and (c) reliability (the data has to be there!). This paper complements the presentations given by Ted Schwarz in past years [1-2] with a little more technical depth. The capacity of a tape cartridge is simply the areal density of the data multiplied by the area of the media used but is often preferably computed in tape by using the relation

$$C = \frac{NbL\varepsilon}{8} \quad \ldots (1)$$

in bytes, where ***N*** is the number of tracks across the tape, ***b*** is the linear recording density in bits per inch, ***L*** is the length of the tape (in inches) and ***ε*** is a formatting/ECC overhead efficiency factor (typically about 0.7). The 8 assumes 8 bit bytes. The date rate is given by

$$D = \frac{nbV\varepsilon}{8} \quad \ldots (2)$$

in bytes/second, where ***n*** is the number of parallel channels used and ***V*** is the speed of the tape (in inches/second). These two relations capture the main parameters in increasing capacities to terabyte levels and data rates to 100's of Mbytes/sec. The linear density (***b***) appears in both calculations and thus is a strong contributor to the problem. The number of tracks (***N***) in the capacity and number of channels (***n***) in the data rate are parameters that may be in conflict when radically increased as will be discussed later.

Capacity (TB)	0.5	0.5	1	1	5	5	10	10
Data Rate (MB/sec)	**60**	**120**	**110**	**220**	**150**	**300**	**280**	**559**
No. of Pll Data Channels, *n*	16	32	16	32	16	32	16	32
No. of Data Tracks, *N*	768	768	1344	1344	4750	4750	4140	4140
Trk. Pitch (µm)	14.0	14.0	8.0	8.0	2.3	2.3	2.6	2.6
Channel Pitch, c_p (µm)	109	55	109	55	109	55	109	55
Rd. Track Width (µm)	7.0	7.0	4.0	4.0	1.1	1.1	1.3	1.3
Tape Speed, ***V*** (m/s)	4.8	4.8	8.0	8.0	9.0	9.0	10.0	10.0
Bit Density (kbpi)	224	224	248	248	298	298	500	500
Track Density (tpi)	1812	1812	3172	3172	11211	11211	9771	9771
Areal Density (Gb/in^2)	0.41	0.41	0.79	0.79	3.35	3.35	4.89	4.89
Bit Cell (nm)	114	114	103	103	85	85	51	51
Bit Cell (ns)	23.7	23.7	12.9	12.9	9.5	9.5	5.1	5.1
Write Eq. Pulse (nS)	9.5	9.5	5.2	5.2	3.8	3.8	2.0	2.0
Tape Length (m)	865	865	865	865	1000	1000	1400	1400
Write Time per Cart. (min)	144	72	152	76	550	275	604	302

Table 1. Terabyte operating points

Table 1 shows scenarios for a 0.5, 1, 5 and 10 Terabyte capacities for various data rates for a normal IBM3480/STK9840/LTO/DLT ½ inch wide tape cartridge form factor. Some tradeoffs between the parameters given in equations 1 and 2 have been included for illustrative purposes and one can easily see where a different set of trade offs could yield the same result depending on which aspect of the tape system you wished to stress more. The stress points are boxed for the cases shown and it is these challenges that are discussed below in relation to the media aspects, the heads and the channel in order to accomplish these operating points.

2. Magnetic Recording

Figure 1 shows a block diagram of a tape recording system from data in from a host computer channel, onto and off the tape and back to the host upon a data read [3]. This figure summarizes the main components and systems needed for the tape system to function. All the subsystems (write method, read equalization and detection, servo, head and tape handling) serve to deal with the unique properties of the tape media itself. This is from both a magnetic and mechanical perspective. The media dictates how the rest of the system is designed in order to achieve high-density data recording and thus is the main contributor to limitations thereto.

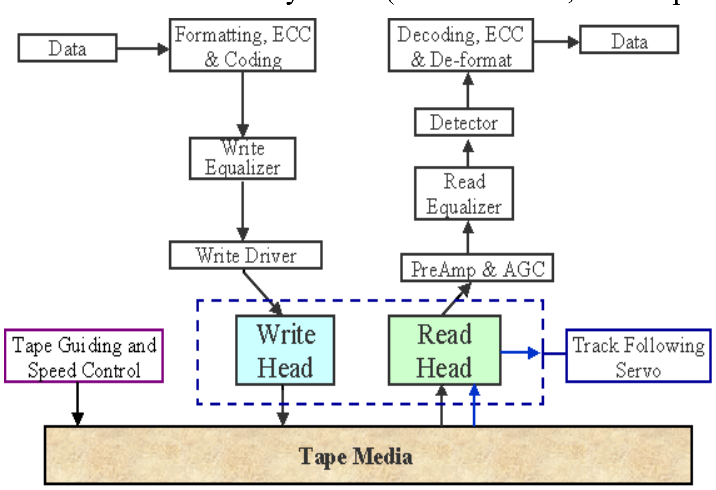

Figure 1. Block diagram of a tape recording system

Fundamental to recording digital data on magnetic tape is the analog magnetic recording that takes place between the

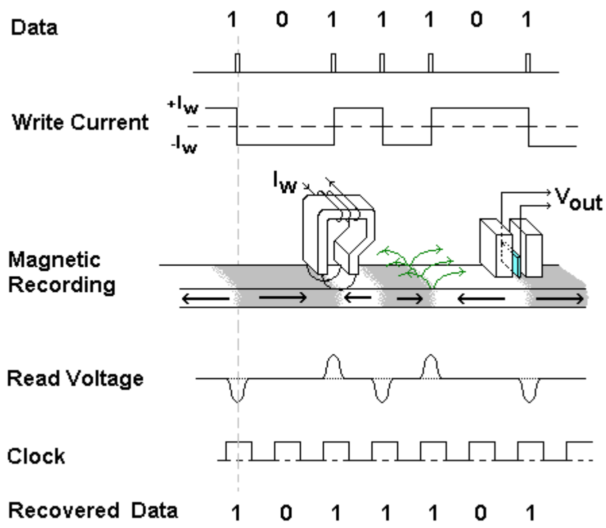

Figure 2. Magnetic Recording

head and the media. These two magnetic components in combination can make or break a reliable data recording system. Figures 2 illustrates magnetic recording on tape and its digital interpretation. The digital interpretation is that a transition between a region on the tape magnetized in one direction to the opposite direction is interpreted as a logical '1' and the absence of the transition a '0' when referenced to a data clock. This interpretation depends on the logic used by the detection system and coding design. For instance a PRML channel (Partial Response Maximum Likelihood) interprets the recorded transitions in a different way by partial amplitude sampling in order to increase the bit density using somewhat lower magnetic transition densities than in straight peak detect channels as illustrated. Such channels increase the logical bit density up to twice that of the recorded magnetic transition density.

3. Recording Technology Challenges

Fundamentally, an increase in linear recording density requires the transitions to be closer and closer together on the media and the ability to resolve them. Table 1 indicates the length of a logical bit (bit cell (nm)) for the various scenarios given for reference (~50 – 100nm). Tape media to date has had the magnetic coating somewhat thick (0.5µm or more) compared to these dimensions which gives broad written transitions due to the generation of transitions curving into the depth of the magnetic coating and the demagnetizing effect of sizeable opposing magnetic poles. These effects are summarized in the equation for the transition length parameter (the 'a' parameter) thus:

$$a = 2\left[\left(\frac{2}{\sqrt{3}}\right)\frac{M_r\delta}{H_c}\left(d+\frac{\delta}{2}\right)\right]^{\frac{1}{2}} \quad \ldots (3)$$

where $\boldsymbol{M_r}$ is the remanent magnetic moment of the medium, $\boldsymbol{\delta}$ the magnetic thickness, $\boldsymbol{H_c}$ the magnetic coercivity of the medium, and $\boldsymbol{d}$ the head to tape spacing. This relation comes from assuming that the transition follows and arctangent function shape [4]. In order to reduce this transition length parameter the ratio $\boldsymbol{M_r\delta/H_c}$ must be reduced. This can be done either by increasing the coercivity, $\boldsymbol{H_c}$, which physically means it is harder to push the magnetized regions apart or by reducing the medium thickness, $\boldsymbol{\delta}$, which lowers the total magnetic moment and hence the force which is pushing the regions apart.

Reducing ***M_r*** is a little more difficult using iron particles (as currently used in MP tape), as this would mean reducing the number of particles in the magnetic coating, which would have the side effect of reducing the signal-to-noise ratio (SNR). An acceptable reduction in ***M_r*** could only come from a different particle; for example barium ferrite (BaFe) or a different media construct (such as thin film media). The coercivity of tapes is in fact on the upswing with prototype MP media pushing 2500 Oe compared with today's 1650 Oe 9840 media and 1850 Oe DLT/LTO media. Figure 3 shows how linear density has indeed gated tape products in the past according to media coercivity together with a projection for future systems based on published roadmaps. (The data here are taken from existing IBM, STK, Quantum DLT and LTO tape products). Excessive increases in coercivity would however begin to challenge the available magnetic pole materials used in the write head where the saturation flux density is limited. This would eventually degrade recording performance if the coercivity increases much beyond 3500-4000 Oe.

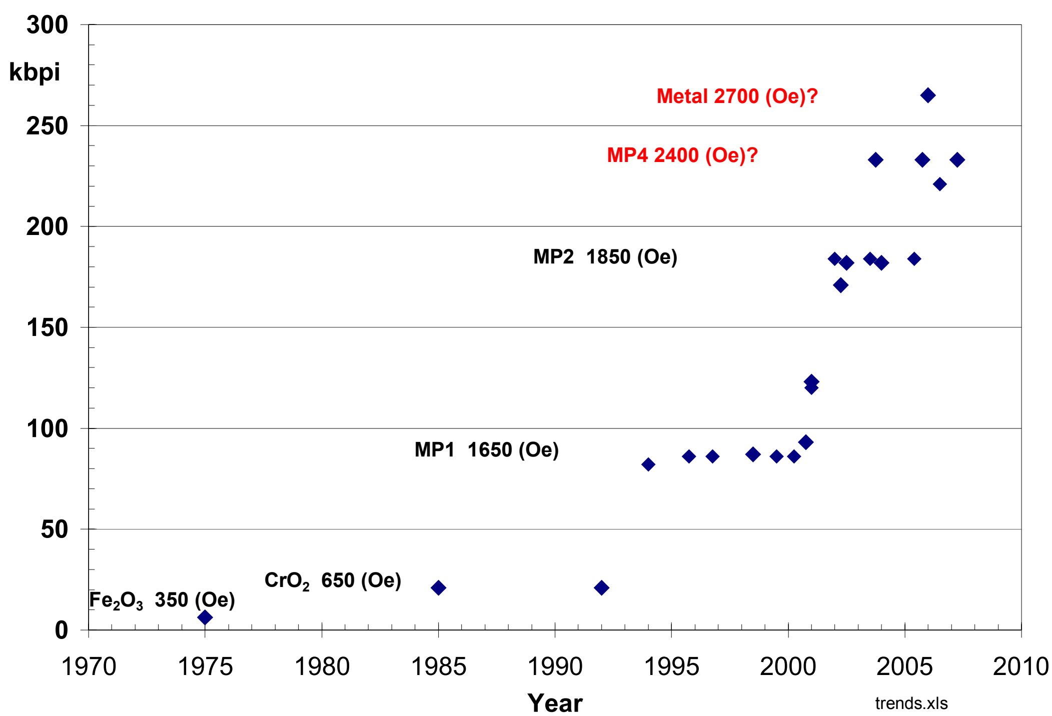

Figure 3. Linear density versus year for linear tape systems

Reducing the thickness is the primary direction to pursue and recently this has been achieved in particulate media by using a dual coating process. Here the magnetic portion of the tape coating is spread thinly over a simultaneously coated non-magnetic under layer. This effectively provides a thick physical coating for smoothing purposes coupled with a reduced thickness magnetic layer as illustrated in figure 4. This has enabled coatings to be produced as low as 100nm and progress is being made to reduce this further [5]. This technique, however, will eventually run out of steam for the particle in binder tape medium concept. One quickly gets to very few 20-30nm thick particles stacked on top of one another in a <100nm coating with the resultant SNR reduction. For areal densities greater than a few Gb/in^2 , the move to thin film media will have to be

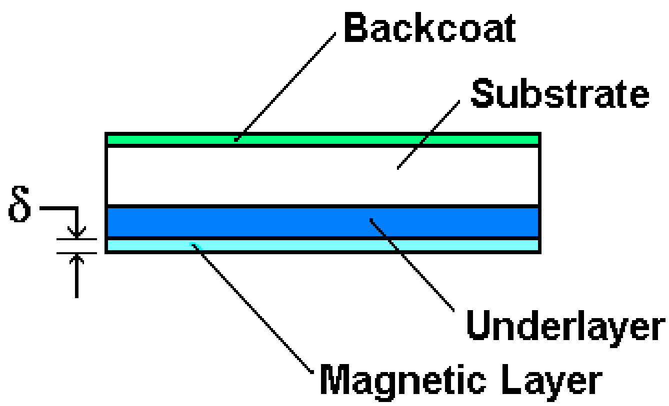

Figure 4. Diagram of a cross section of dual coat tape recording media

made as it was for magnetic disk. (Tape is indeed fortunate that magnetic disk has already demonstrated solutions to high areal density magnetic recording.)

The other parameter that figures into the areal density is track density. Again the number of particles contained within the bit becomes squeezed as the track narrows. As the SNR is related to the total number of particles contained in the bit volume [4] an estimate for the areal density limit, $\boldsymbol{A_{lim}}$, for metal particle tape can be made from an SNR standpoint and input from media producers on what might be the maximum particle density (smallest thermally stable dispersible particle). Following Mallinson [4] it can be shown that

$$A_{\lim} = t^{\frac{1}{2}} \left(\frac{2pSNR}{3} \right)^{\frac{1}{2}} \qquad \ldots (4)$$

where $\boldsymbol{t}$ is the track density, $\boldsymbol{p}$ the magnetic particle density in the media and $\boldsymbol{SNR}$ is the signal-to-noise ratio requirement. Using for example 3000 tracks/cm (7620tpi), 10^{17} particles/cm and 20dB we get an areal density of approximately 10Gb/in^2. This assumes that the whole written track is read, no spacing loss and one logical bit per transition. Using a write wide read narrow scenario, as linear tape currently does, and invoking a PRML channel you come out with a very similar number or maybe slightly higher depending on the $\boldsymbol{SNR}$ and desired raw bit error rate. (PRML channels operate at lower effective SNR values.) The areal densities in the cases shown in Table 1 approach 5Gb/sq.in. and the question arises as to how close to the computed limit can you engineer particulate media for this, or is thin film media prompted as it was in disk.

The other main parameter in equation 3 is $\boldsymbol{d}$**,** the head to medium spacing. This also figures heavily into the wavelength response upon read back. Loss of resolution of the shortest wavelengths is severe ($\boldsymbol{e^{-kd}}$, where $\boldsymbol{k}$ is the wave number) and the resultant signal loss is normally given in dB form by the relation [6]

$$Loss = -54.6\frac{d}{\lambda} \quad \mathbf{dB} \qquad \ldots (5)$$

In combination with spacing on write, the multiplier in equation 5 (-54.6) is closer to -100! Although we run the tape in physical contact with the head, the 'magnetic' spacing seen is due to media roughness, recession of the magnetic elements in the head and any adherent (or temporary) debris or stains on the head. Current systems appear to have up to 70nm of magnetic spacing while in apparent physical contact and this will have to come down if we want to resolve high density terabyte recordings and not suffer the resulting loss in signal amplitude and resolution.

Head technology appears to have enough precedents and product introductions (again as seen in disk magnetic recording) that tape head offerings should be able to readily respond to new media types as they are developed. A classic example would be the shift to all thin film write heads and thin film shielded read heads as well as merged pole/shared shield structures commonly used in disk and now being seen in tape applications. Examples are shown in figure 5. The main issues facing the tape head

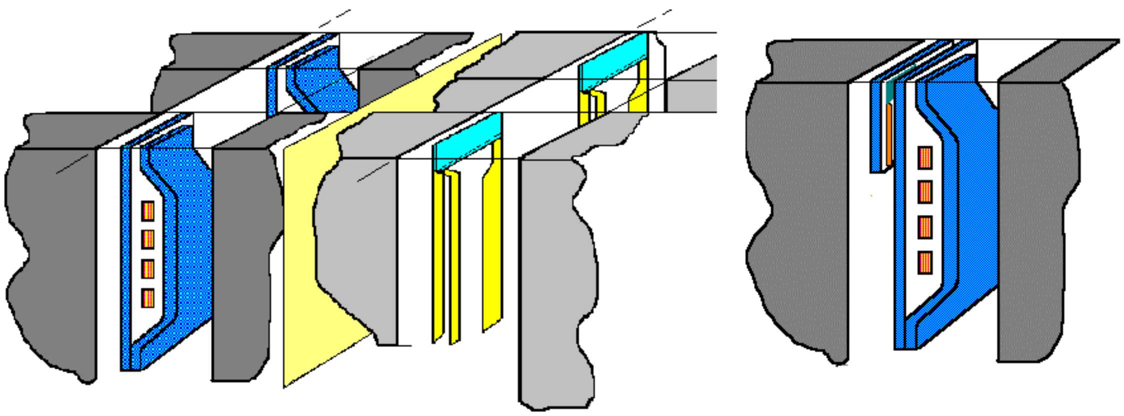

Figure 5. Diagrams of thin film tape head types showing thin film write, MR read and combination shared shield devices

concern the consequences of using multiple channels simultaneously in read-while-write mode. I.e., direct write to read feed through and read element off-track due to tape static and dynamic azimuth. A future example of disk like technology for tape would be the introduction of the GMR spin valve read sensor now prevalent in desktop systems in disk drives. This would be predicated by the availability of a suitable media that would be compatible with high-density recordings and these very sensitive devices, as well as environmental issues seen in tape usage. Alternatively, new designs of spin valve sensors customized for tape could be used with the still somewhat higher $M_r\delta$ values that may persist. The switch to spin valves will be driven by the need for raw signal amplitude to overcome the unique noise sources in the multi-channel read-while-write tape environment (such as write-to-read feedthrough) as the read element width and hence signal amplitude is reduced.

Another issue raised in Table 1 is the time scale of the recording. For high bpi and fast tape speeds, the bit cell time is reduced to <10nS. If write equalization persists as a favorable recording method (which it will if the $M_r\delta$ is not reduced significantly) then the recording system (write current, write head magnetics and media magnetization) will have to respond on the 1nS time scale. For a 3nS write equalization pulse the media has to see the field at least 2nS of that time to stand a chance of responding. Figure 6 shows how magnetic media (in this case MP1 media) changes its effective coercivity for fields

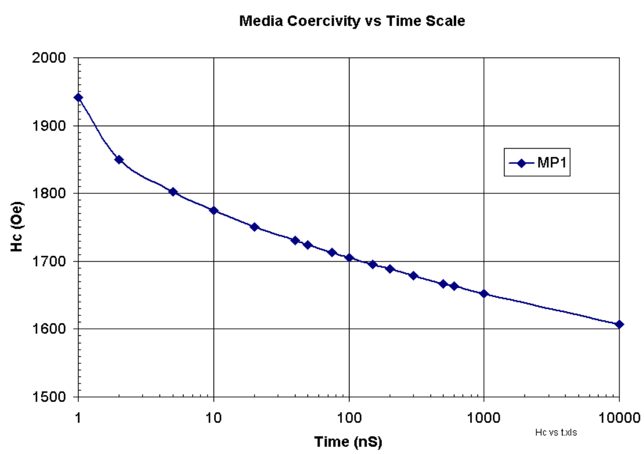

Figure 6. Coercivity of MP tape versus time scale of the applied magnetic field.

applied at very short times. The rise in coercivity means that we would have to overdrive the system to affect the recording in the required way presuming that the head magnetic core can provide the specified field in response to the drive current. The issue of getting the drive current into an inductive load like a write head exacerbates the problems in a multi-element tape head where stray capacitance paths can shunt the coil current. Core materials for the head appear to be available to provide this response and current production head types such as the StorageTek T9840B write head have demonstrated good performance down to 10nS. Data for this is shown in figure 7 for such a write head, which uses cobalt based amorphous alloy poles.

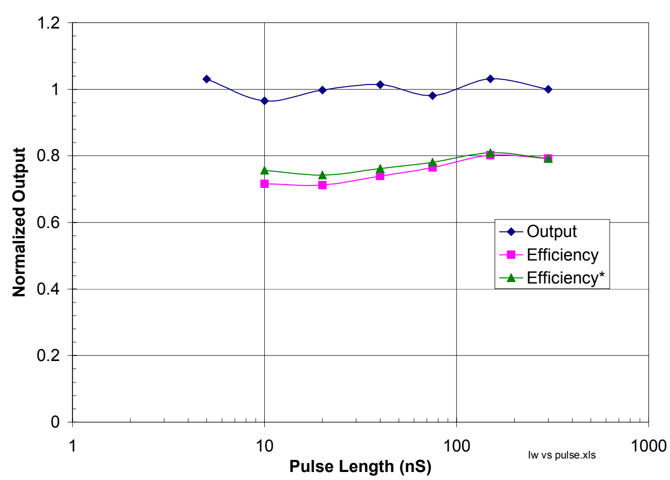

Figure 7. Head efficiency and readback output vs. write pulse length

This data shows that the read back amplitude remains the same when the pulse length is reduced and that the head efficiency does not roll off significantly. The two efficiency curves represent the directly measured head efficiency and the head efficiency corrected for the media coercivity shift according to figure 4 (this curve is indicated with an *). These time scale issues are not at any fundamental limits imposed by the laws of physics but provide the engineer with interesting challenges. The particulate and metal based tape media respond at 1nS and I think operating near 1nS will be avoided in any case with the eventual elimination of write equalization.

4. Mechanical Issues

The tape speeds used in Table 1, for the high data rates, provide the tape path and motion control with some challenges. This is especially so considering the tape lengths needed for the capacities which in turn means the tape thickness needed for the cartridge size approaches 5µm (or even a little thinner). This thickness (or rather thinness) means a relatively low tape tension with which to achieve these speeds with adequate lateral guiding and tape pack management. On top of this, the bandwidth of a track following servo system would have to increase together with its capability to achieve the track pitch targets. Again, no real fundamental limits here, just a solid engineering problem. Unfortunately, these factors get much less attention than the more intuitive limits imposed on track density by the dimensional stability of the media itself. Very narrow tracks, coupled with multi-channel heads that span a significant portion of the width of the tape, result in track mis-registration (TMR) numbers that imply roadblocks beyond mere electronics. There is an interesting trade off between data rate and capacity that can be made as outlined in the 1998 NSIC tape roadmap [7]. Given a fixed tape length and achievable linear recording density, capacity can only be increased by increasing the track density (narrower tracks). This means the allowable off-track capability (OTC) is reduced. For higher data rates the only adjustable parameter, once the tape speed is set, is the number of parallel channels in the head. The more channels side by side the wider the span across the tape and more likely the end tracks will exceed the OTC as the tape dimensions change with time, tension, temperature and humidity. The results of the calculation of this trade off is formulated as

$$D = \frac{2(OT)LWVb^2\varepsilon^2}{64Cc_p m_c} \qquad \dots (6)$$

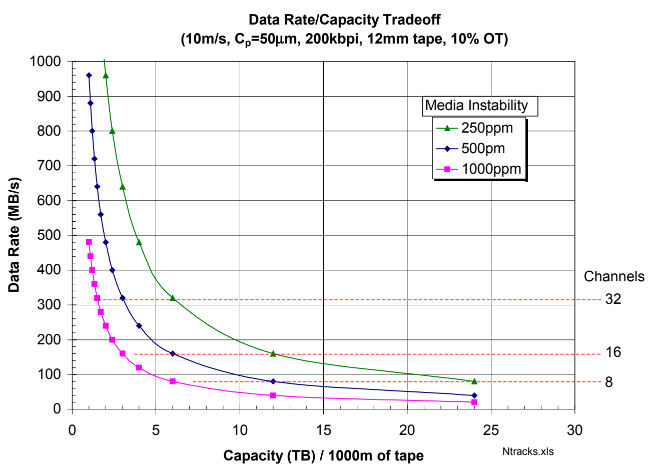

Figure 8. Data Rate/Capacity trade off for a linear tape system.

where ***OT*** is the allowable off-track expressed as a fraction of the track width, ***W*** the width of the tape, ***C*** the capacity of the cartridge, c_p the channel pitch in the head and m_c the media instability coefficient. Figure 8 shows the situation for various media stability numbers (from ref. 7) and is considered somewhat optimistic as it considers only writing the tracks in the correct location and not any read-while-write or realistic read back scenarios. Also current feedback from media suppliers is that the stability numbers will probably not improve as significantly as suggested here anytime soon. The implications of this chart are simple to interpret. If you want very high capacity, i.e., very narrow tracks, the number of parallel channels laid side by side will have to be reduced, lowering the possible data rate.

The only way around this is to change the way we parallel up head stacks to avoid the excessive head span or change in some other way we lay data on tape. Super DLT and helical systems for instance use dual azimuth recording on adjacent tracks allowing a larger OTC. This is one reason the areal density demonstrated by helical scan systems (e.g. SONY) already exceeds that projected for linear systems. Helical technology uses a single channel or few channels approach, high head-tape interface speeds, dual azimuth and short length tracks, which circumvent these media related problems. Unfortunately helical technology has suffered head and media wear problems and there is a perception of poorer reliability compared to linear systems, the basis for which is somewhat clouded. The challenge for the multi-channel linear head here is the reduced channel pitch. 50µm as indicated in Table 1 is certainly achievable, but beyond that, a new approach over

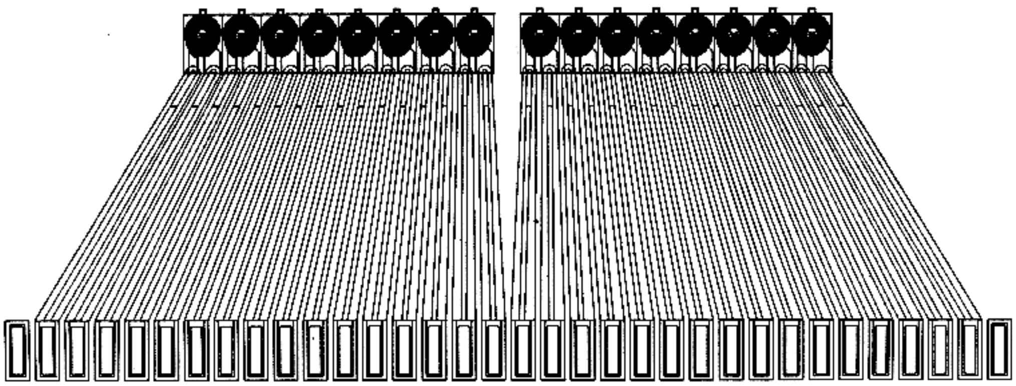

Figure 9. Example of a multi-channel, side-by-side architecture, thin film tape write head as used in today's linear tape heads [3]

today's norm (figure 9) is expected.

Finally, figure 10 summarizes the areal density progress and trend extension for linear tape based on past and present systems and published roadmaps. As mentioned before, heads and media in combination are the primary drivers for this parameter. The coercivity rise from oxide tapes to MP tapes and in the future thinly coated particulate or metal film tapes have been responded to by heads moving from ferrites to thin films and high moment thin films to write these tapes. This is in conjunction with MR and eventually GMR read heads to deliver appropriate signal quality. Also shown is the SONY helical 6.5 Gb/in^2 demonstration on metal evaporated (ME) tape and subsequent 16.4 Gb/in^2 point using spin valve heads [8], and the estimated MP limit using today's assumptions.

5. Conclusions

It is clear that the medium has a significant if not the primary impact on the density growth in magnetic tape recording. As demonstrated by disk magnetic recording the $\boldsymbol{M_r\delta}$ has to be reduced in order to increase the linear density. Significant reduction in this parameter would allow closely spaced magnetic transitions and enable the use of more sensitive read head sensors such as spin valves to boost the sagging raw signal amplitude as both the bpi and tpi increase. Calculating a limit for MP tape throws down the gauntlet for media, head and channel developers to counter this, as was seen recently in magnetic disk. There the areal density limit was calculated to be 36Gb/in^2 in 1997 [9], which is now exceeded in today's normal production disk drives!

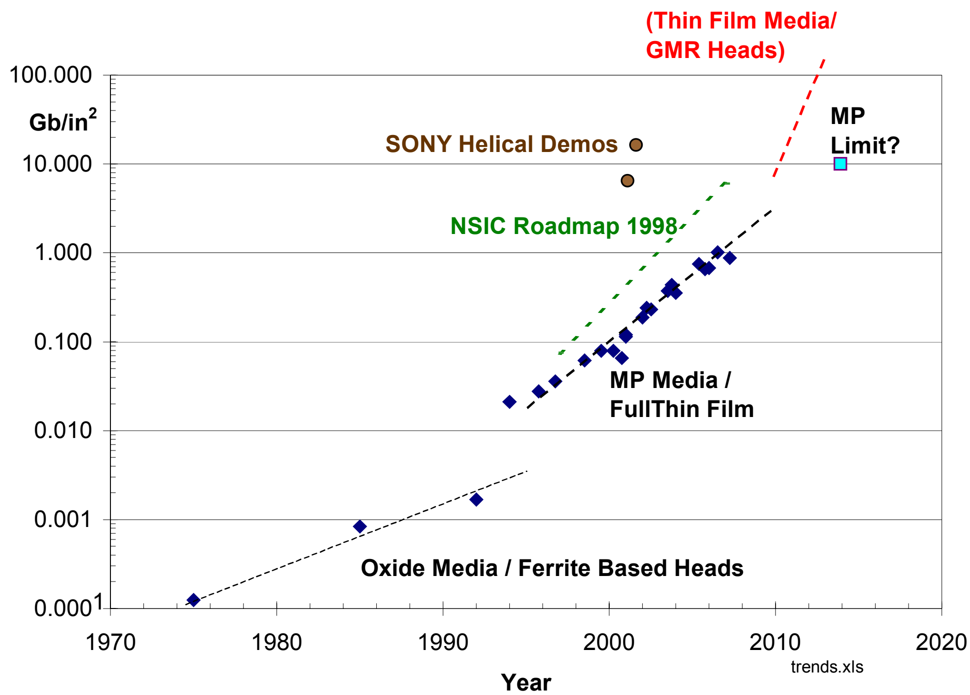

Figure 10. Areal density trends in linear tape systems

Increasing the data rate by increasing the number of parallel channels involves trade offs with tpi (i.e. capacity) if we remain with the side-by-side head stack architecture due to the increasing span of the active channels across flexible media, which is accepted as having somewhat poor dimensional stability. Head technology appears to be available to meet the challenge of the multi-terabyte capacity cartridge but this target is gated by media type and availability, and overcoming the engineering challenges of handling the magnetic and physical properties of the media. Tape is not nearing any fundamental scientific limits as seen in magnetic disk. Given the rather moderate areal densities currently seen in tape systems and optimism with regard to the development of tapes with thinner magnetic coatings, data storage systems using tape are poised to make some rapid advances in capacity *and* data rate.

6. References

[1] "Magnetic tape as the mass storage medium". T, Schwarz, Mass Storage Conference (2000).

[2] "The future of magnetic tape", T.Schwarz, Mass Storage Conference (2001).

[3] "Magnetic Tape Recording Technology and Devices", R. H. Dee, Proc. Int'l Non-Volatile Mem. Tech. Conf. pp55-64 (1998). (IEEE Cat. No. 98EX141).

[4] "The Foundations of Magnetic Recording", J. C. Mallinson, (Second Ed.), Academic Press (1993).

[5] "Investigation of particulate media with an ultra-thin magnetic layer suitable for MR heads on a rotating drum" K. Ejiri et al, IEEE Trans. Magn. Vol.37, No. 4, pp1605-1607 (2001) and Fuji Film press release on 'Nano Cubic Technology' www.fujifilm.com (2001).
[6] "The reproduction of magnetically recorded signals", R. L. Wallace Jr., Bell Syst. Tech J. vol.30, 1145 (1951). See also Ref. 2 p.88.
[7] "Tape Roadmap", National Storage Industry Consortium (June, 1998).
[8] "Beyond 6.5Gbit/inch^2 recording using spin valve heads in tape systems", T. Ozue et al (SONY), Proc. TMRC Conf., Minneapolis, MN (Aug, 2001).
[9] "Thermal stability of recorded information at high densities", S. Charap, P.Lu and Y.Lee, IEEE Trans. Magn. Vol. 33, No.1, pp. 978-983 (1997).