EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

MASS DATA STORAGE FOR ON-LINE EXPERIMENTS

B.G. Taylor

(to be presented at the All Union Seminar on Physics
Data Processing, Erevan, USSR, 9-16 September, 1977)

EP Division

Geneva,
August 1977

MASS DATA STORAGE FOR ON-LINE EXPERIMENTS

B.G. Taylor

Indexing terms: Mass data storage, Video-tape recording, Optical holography

Abstract

The paper discusses present and future recording technologies applicable to the mass storage of digital data at on-line elementary-particle physics experiments.

First, a set of target system (TS) performance specifications is established for a realistic device which would be well adapted to the requirements of this specific data recording application. The parameters considered are on-line data capacity, data-rate, access time, media cost, storage compactness, data retrieval capability, and system availability.

The performance of a currently-implemented system using digital video-tape (VT) recording is then compared with the TS specification, and the operating experience and device shortcomings are reviewed.

Finally the expected capabilities of an alternative technology using optical holographic (OH) data recording on film media are outlined, based on the published developments and the tentative specifications for a storage device which will be commercially available in the next 12 months.

The performance capability comparison allows the relative match of VT and OH characteristics to TS specifications to be evaluated with a view to future requirements for mass data storage at on-line experiments.

1 Introduction

The recording of experimental data for subsequent analysis and study is a fundamental facet of scientific endeavour. The recording process serves to interface the time-scale of the observations, which is detector and physical process - related, to the time-scale of their interpretation; which is paced by the computational tools of the scientist and his conceptual inspiration.

When Rutherford and his contemporaries were studying large-angle single alpha-particle scattering, digital magnetic recording was already an infant technology, Poulsen having demonstrated a magnetic wire telegraph signal recorder(1) as early as 1900. Half a century later, scintillation counter and magnetic recording techniques were merged when the first computer magnetic tape units were introduced and used on particle physics experiments on accelerator beams.

Since the mid-1950s, magnetic surface recording has dominated the mass memory (>10^7 bits) and mass storage (>10^{10} bits) sectors in spite of intensive research efforts in alternative technologies. Over 10^5 half-inch magnetic tape drives are in use today(2), and current state-of-the-art allows automated tape files exceeding 10^{12} bits per facility - a data recording capacity equivalent to 10 million pages of the laboratory notebooks of the founders of modern physics.

The reason for the continued widespread utilization of magnetic storage is not lack of progress in the development of the materials and components required for other methods, but rather the continuing technical advances in magnetic recording technology which have served to maintain its economically competitive leadership. For many years, it has been recognized that optical data storage technology(3) holds promise as a high-performance contender for certain areas of the memory hierarchy. High-density optical data recording has been researched in the engineering laboratories for over a decade(4). Two large bit-by-bit archival systems(5,6) and a number of small products reached the market during this period, but it is only now that the first read-write optical holographic memory system is approaching the commercial product stage.

On the other hand, the projections of researchers suggest that substantial performance improvements are still plausible in magnetic surface recording[7]. If signal-to-noise ratio and head dynamics problems can be overcome, the use of small magnetic particles could in theory permit recording densities of up to 250,000 bpi, an order of magnitude higher than what is achieved in the laboratories, and two orders of magnitude above current commercial practice. Track density projections of 2000 tracks/in have also been made, and the resultant areal density of 5×10^8 bits/in^2 is comparable with the upper bound imposed in optical memories by the diffraction limit of light, which for visible wavelengths and using f/1 optics is approximately 0.5 μm. At this density the entire content of a conventional magnetic disk pack having 20 recording surfaces could be stored on 1 in^2 of medium.

This paper discusses the factors which determine a realistic set of performance characteristics for a storage device for bulk data recording at on-line particle physics experiments. We then review briefly how these requirements are matched by the relative merits and shortcomings of the current technology, high-density magnetic tape recording, and a possible future technology, using optical holographic techniques. The nascent non-mechanical storage technologies, such as charge-coupled devices (CCDs), magnetic bubbles and domain tip (DOT) memory, and electron-beam-accessed MOS memory (BEAMOS), are not considered as it is not expected that they are likely to challenge the upper end of the storage hierarchy during the next decade.

2 Target system (TS) parameters for on-line data recording

In general, a high-performance mass memory recorder should be characterized by large on-line data capacity, high recording and replay data-rates, low access-time, low media cost-per-bit, compact data storage, and high reliability in terms both of data retrieval and system availability. The requirements for an on-line data recorder for physics experiments differ in several respects from those of a general-purpose computer mass memory and may be summarised as follows.

2.1 TS on-line data capcity

The data storage capacity which is required to be on-line, ie. which can be accessed automatically by the system without operator intervention for activities such as reel-changing, may be estimated separately for the data acquisition and data analysis phases. In our experience, on-line capacity for 1 hour of uninterrupted recording should be considered a minimum. Less than this creates undesirable demands for operator intervention by a small team who are busily engaged on other aspects of experiment control, while capacity for many hours of data-taking is not essential because over such a period it is frequently the case that experimental apparatus or beam conditions are changed in a way requiring some operator action to signal the new conditions of the run. At data rates up to 8×10^6 bits/sec (see Section 2.2), this implies a minimum on-line capacity of about 3×10^{10} bits.

During the analysis phase, a time between operator interventions of greater than 10 hours is desirable so that the system can be left to operate overnight without any operator present. The time expansion factor from acquisition to analysis is a function of many parameters such as the selectivity of the initial trigger, the multiplicity of the events, the formatting of the read-out data, the uniformity of the magnetic fields in which track reconstruction is required, the power of the computer and whether or not it is aided by a special hardware processor. Taking a factor of 50 as a realistic estimate, the replay requirement dictates an on-line capacity of 6×10^9 bits. The recording requirement of 3×10^{10} bits is thus the dominant one.

2.2 TS data rate

In present experiments at CERN, CAMAC data acquisition rates rarely exceed 8×10^6 bits/sec (500 Kw/sec). The variety of modules incorporated in a typical experiment currently requires appreciable software activity to read-out each event, and many minicomputers in common use have synchronous 1 μs cycle-time memories and significant DMA latency times. Higher average data rates are to be anticipated in the future with the introduction of faster asynchronous computer memories and intelligent CAMAC controllers[(8)] capable of autonomous operations.

It is very desirable that the recording rate of an on-line data storage device should match that of CAMAC data acquisition. In experiments on a DC machine such as the intersecting storage rings (ISR), this results in the data recording activity making no contribution to the experiment read-out dead time; while on a pulsed machine like the SPS it reduces the otherwise large computer buffer memory requirement. The pulsed machine situation can be effectively transformed to the DC one by the use of multiple autonomous controllers operating in parallel into a multi-port buffer memory, which is accessed globally by the mass data-storage device.

During the replay phase, a high data-rate is unnecessary because with only a double event buffer the reading of further data can proceed in parallel with the much slower data analysis activity.

2.3 TS access time

Most mass storage systems are designed for applications requiring random access to specifically addressed data volumes rather than sequential data recording and replay, and much system complexity in the form of parallel sensors and data paths and high-performance media displacement mechanisms is directed towards this end. For on-line experiment data, not only is a random access facility superfluous, but (for a given run or set of experimental conditions) it is not even necessary to retrieve data volumes corresponding to physical events in any particular sequence, as long as they can be flagged such that they are not inadvertently analysed more than once. It is only necessary that the retrievable contiguous data volume be sufficient to accommodate the complete event. In present experiments this is typically less than 10^4 bits, but in high-multiplicity

$p - \bar{p}$ experiments currently under discussion could be as high as 10^5 bits per event.

During data analysis, reasonably rapid access is required to the data corresponding to the set of runs which may be specified in the job schedule. An acceptable performance criterion is that the total time spent in such accesses be negligible relative to the data processing time - say less than 500 seconds per 10 hour session. In general, the job schedule will not specify a set of runs which are randomly distributed throughout the medium, but which are in adjacent or near-adjacent areas. The number of runs which can be analysed per night session is limited by the time expansion factor, but it is not unusual to require the analysis of an incomplete selection of the data from several runs. Taking an upper limit of 20 accesses per analysis session, each spanning an average of 10% of the on-line data volume, a random access time of 125 seconds is sufficient to meet the above criterion.

2.4 TS media cost

While it is not essential that a mass memory recorder should have lower media cost-per-bit than conventional systems if it offers important performance features in other areas, it is natural to expect a high density technology to be associated with media cost savings. If total media usage is large, such savings may help to offset the higher capital cost of the device. An actual crossover into overall system economy is, however, not to be anticipated at the data volume of a single experiment, unless factors representing storage and handling costs, and effective lost beam-time for the lower speed device, are included in the balance.

A media cost target could be set at 3×10^{-9} SF/bit, representing one order of magnitude lower than conventional magnetic-tape recording at 6250 bpi.

2.5 TS data storage compactness

The volume of the present magnetic tape archive at CERN is such that it is highly desirable that substantial savings of media storage space should result from the use of on-line mass memory recorders. A satisfactory situation is attained when the typical archive of an experimental group (say 10^{12} bits) can be accommodated in a desk cabinet (10^4 in^3), corresponding to a packing density of 10^8 bits/in^3. Assuming a laminar medium of minimum thickness 10^{-3} in for strength and stability reasons, and allowing 50% of the cabinet space for container mechanics, this implies an areal density of at least 2×10^5 bits/in^2. While 3-dimensional storage media have been described(9), the technique is not readily expandable to large capacity systems.

Experience suggests that significant reliability problems arise when an integrated mechanical system must manipulate an active recording area exceeding 10^5 in^2 (equivalent to 7 magnetic tape drives or 80 disk spindles). To provide an on-line capacity of 3×10^{10} bits, this criterion indicates a minimum recording density of 3×10^5 bits/in^2, comparable with that indicated by the storage volume considerations.

2.6 TS data retrieval reliability

In general data-processing applications, it is desirable when using bit-by-bit optical or magnetic recording techniques to sacrifice potential storage capacity and speed in order to add the redundancy necessary to achieve a very high probability of successful individual data retrieval through error correction. This is because typical bulk storage media exhibit 'raw' error rates of the order of 10^{-8} bits, while system error rates of 10^{-11} bits or better are desirable, particularly in commercial applications. In 6250 bpi magnetic-tape group code recording(10), for example, over 30% of the storage capacity is used for checkbits to achieve this performance.

In the mass storage of particle physics experiment data, on the other hand, the individual data retrieval capability is not a critical performance criterion, because the data are composed of stochastic events contributing independently to the statistics measured. Instead, optimum system performance is attained by maximising the overall error-free event data throughput and capacity. For typical media error statistics and event sizes, this is achieved by the use of error detection without error correction, only sufficient redundant checkbits being added to render the probability of undetected error insignificant. Unlike error-correcting codes which, to be effective, must be matched to the error statistics[11], the undetected error rate for error-detecting codes generally converges rapidly within a small range of redundancy (< 10%) for all random, burst and synchronisation errors.

In holographic as opposed to bit-by-bit recording, the fact that the Fourier transform of the spatially modulated light beam is recorded, rather than the page composer image itself, results in a natural distribution of the contribution of each bit-cell over the entire hologram surface. This spatial distribution is analogous to the enlarged temporal distribution of the influence of each bit which is created by adding redundant checkbits, and it results in the same improved immunity to data loss due to localised media defects. This inherent redundancy, an attractive feature of the holographic technique, may be expected to result in essentially error-free data retrieval for defects up to a critical size, above which the entire data page is lost as the signal-to-noise ratio of the reconstructed image is degraded below a threshold.

Mass storage data retrieval reliability can be considered acceptable if the total loss of event data in the device is negligible compared with the loss due to other inefficiencies in the experiment, provided the loss is not correlated with the data values. In practice, all devices tend to exhibit some data pattern sensitivity[12], so to safeguard against bias we shall consider the loss of 1 event in 10^3 as the minimum satisfactory performance. For event lengths of 10^4 to 10^5 bits this corresponds to a random error rate of 10^{-7} to 10^{-8} bits.

According to the system vulnerability, the overall data retrieval reliability may be degraded by rare catastrophic media damage caused by operator mishandling or mechanical malfunction.

2.7 TS system availability

In the present state of the art, it is not feasible to construct data storage systems having the desired minimum on-line capacity of 3×10^{10} bits without electromechanical media-displacement within critical tolerances. As a result, scheduled maintenance is required at relatively frequent intervals. In addition, such devices necessarily comprise a number of relatively complex subsystems in series, and successful system operation depends on all the subsystems functioning correctly. System failure rate will therefore generally be high enough to have a significant impact on operations.

The effect of failures is more severe during the data acquisition than the analysis phase, in that any interruption of recording represents an irrecoverable loss of physics beam-time, while an interruption to replay can generally be made up later. On the other hand, back-up operation at reduced rate is normally possible during data-taking by switching to conventional magnetic-tape recording, while a breakdown during analysis halts the operation completely when the data are only available on the mass storage system.

CERN ISR operating schedules can result in periods of up to 65 hours continuous data-taking. To ensure a 90% probability of successful system operation for such a period, we require a failure rate per hour of $-(1/65)\log_e 0.9 = 1.62 \times 10^{-3}$, corresponding to a mean time between failures (MTBF) of 617 hours. This is probably an attainable figure where considerable attention is paid to simple and efficient design, low component stress ratios, and quality assurance controls such as component preconditioning and screening. For a system reliability of 99% per 65 hour period, an MTBF of 6470 hours would be required - a performance which is unattainable without special measures involving the duplication of critical system elements and the automatic re-assignment of serviceable resources.

To set a realistic acceptable MTBF, we can target a figure of

1000 hours, although this means that breakdowns during physics data acquisition will occur about 1 run in 16. The loss of system availability caused by these failures is influenced by the response time of engineering support and the mean-time-to-repair (MTTR); which is the total of the time to set up test equipment, the time to locate the fault, and the time to replace the defective part and carry out any alignment and verification necessary before resuming normal operation. An MTTR of 1 hour would not be considered too serious an interruption to data acquisition in most cases. To attain this figure, good diagnostic facilities must be prepared in advance, and replacements for every part of the system must be available on-site. The replacements for critical mechanical parts in particular must be made at the sub-assembly level, because component-level replacement may involve re-building and alignment work of several hours' duration.

2.8 TS specification summary

The set of target system performance characteristics which have been derived in Sections 2.1 - 2.7 is listed in Table 1. It is emphasized that these do not represent 'ideal' specifications in the sense that higher capacity and speed, lower error and failure rates etc would be welcome advantages. Rather they represent a realistic set of parameters which have been chosen to match the most important requirements of the application envisaged. The major factor of system cost has not been included. Evidently a storage system meeting all the specifications given would gain general acceptance only if it represented an investment which experimental groups could readily afford.

3 Digital video-tape (VT) recording

Having discussed the requirements for an on-line data recorder for physics experiments, we now examine to what extent these can be met by practical systems. From the small range of current technologies (various forms of magnetic surface recording), and the wide range of potential future technologies, only two are considered. The first is

Table 1 Summary target specifications for an on-line mass storage device

On-line data capacity:	3×10^{10} bits
Data rate:	8×10^{6} bits/sec
Access time:	125 sec
Media cost:	3×10^{-9} SF/bit
Storage density:	3×10^{5} bits/in^2
Data error rate:	10^{-8} bits
MTBF:	1000 hours
MTTR:	1 hour

high-density magnetic-tape recording using video-tape techniques, which we have implemented at CERN and operated at an on-line experiment, and the second is optical holographic recording on film, which is a technique expected to enter commercial use during the next 12 months.

The video-tape data storage device is in regular use as part of an integrated data handling system[(13)] (see Fig. 1) employing an HP2100A minicomputer for control and CAMAC data acquisition, and an IBM 360/44 for data analysis. The video-tape recorder (VTR) is a helical-scan machine employing 1-inch magnetic tape, on which data are recorded by Miller coding[(14)] at a linear density of 11,000 bpi. The scanline width and separation results in an areal density of 1.2×10^6 bits/in^2. The machine was conceived in 1970 and follows VTR practice current at that time. It carries additional ferrite scanning heads for erase and write-verify functions, and fixed heads for the longitudinal recording of address and status information as well as synchronising signals. While tape motion is normally capstan-controlled at 6.91 ips, a reel-drive mode allows computer-controlled tape searches at velocities up to 400 ips.

The system has run for 8000 hours and, in spite of its shortcomings, has established an enthusiastic acceptance by the experimental group for which it was developed. The performance differs from the target specification of Table 1 in a number of respects which are discussed in the following.

3.1 VT on-line data capacity

The on-line data capacity for one 7000 ft reel of tape is 7.2×10^{10} bits, over twice that apparently necessary. While this suggests that smaller tape reels could be used, with consequently improved reel-servo performance, there are practical reasons why the large reels are more convenient. For example, they reduce the overhead task of preparing a video tape for first use, when a transparent mylar leader must be spliced to it and a set of 7 coded reflective markers accurately applied.

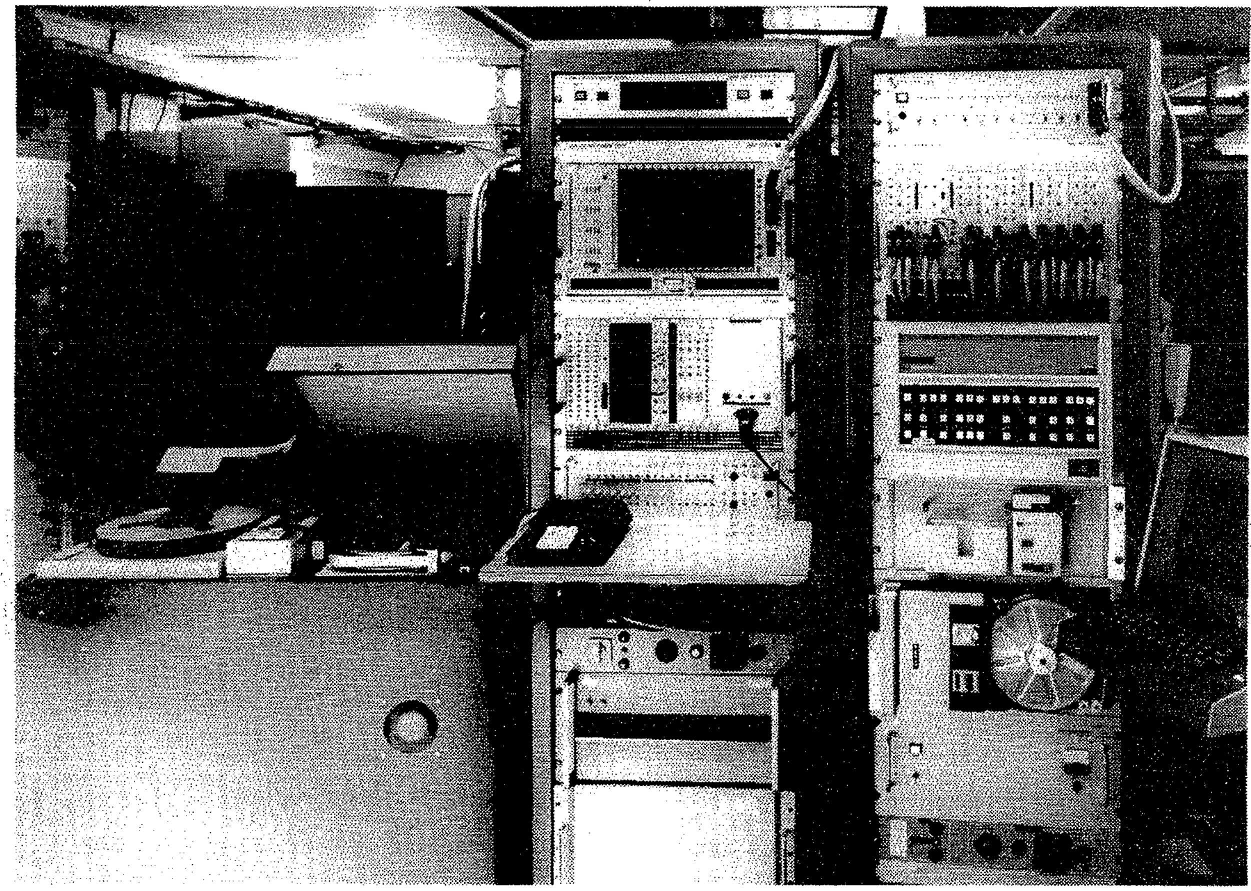

Fig. 1

Digital video-tape data handling system

3.2 VT data rate

While the recording rate at the VTR scanner is 8.1×10^6 bits/sec, the system does not attain our target specification data-rate. Deductions must be made for the checkdigits added to the data by the controller for error-detection purposes, and for the time durations of the preamble (16.8% of each scanline) and postamble (6.3%), when only synchronising information is written to tape. The resulting maximum continuous data throughput is 5.8×10^6 bits/sec. This has not proved a limitation in the present experiment, but it could become so in future applications when the developments referred to in Section 2.2 are implemented.

3.3 VT access time

The system access time is 105 seconds, compared with our target specification of 125 seconds for 42% of the data volume. There would therefore appear to be a case for slowing the machine, which would reduce the head and tape wear associated with high search velocities. The maximum speed is a convenience for rewind/unload operations when changing reels.

3.4 VT media cost

The VTR was designed for use with a tensilized polyester-base tape with an intrinsic coercivity of 500 oersteds and retentivity of 1350 Gauss. This tape is no longer manufactured. As a replacement, we have employed a video tape with similar magnetic coating properties but with a non-tensilized backing. Although the yield strength of this tape is only 13 lb/in, compared with 23.5 lb/in for the original, we have in practice experienced no serious scanline tracking problems attributed to tape geometry.

Media cost is somewhat sensitive to purchase volume, which is low with only one system in service. 20 video-tapes have been written in the current experiment, equivalent to about 5000 reels of standard 1600 bpi computer tape. At this quantity the cost is 5.8×10^{-9} SF/bit, which is twice as great as our target, but still cheaper than standard 1600 bpi tape by a factor of 18. The actual cost saving for the above data volume is SF 143,900; which implies that, if development manpower costs are included, the video-tape system cannot yet be said to have paid for itself by media cost savings alone.

3.5 VT data storage compactness

With a recording density of 1.2×10^6 bits/in^2, and the overheads mentioned above, the density of useful data storage corresponds to 8.6×10^5 bits/in^2, or almost three times the target specification. This suggests that, from the point of view of compactness of data storage, the current system represents a degree of overkill. On the other hand, while a reduction of the system recording density could result in an improvement in error rate, it would have an adverse effect on the media cost savings quoted above.

3.6 VT data retrieval reliability

The average signal-to-noise ratio of the replay video signal is such that data errors are almost entirely due to localised media defects or contamination. This is demonstrated by the fact that essentially zero-error performance can be repeated consistently on defect-free lengths of tape. Individual tapes generally exhibit a relatively uniform distribution of defects along their length, while there are substantial variations in quality from tape to tape. In one case, where a set of tapes was known to have been slit from the same web, the tapes were uniform in quality, suggesting that the variations occur from one manufacturing batch to another. Although the tapes are made for television use and are not certified for digital recording, no tapes have been received which were too poor to be usable, and the average raw data error rate is between 10^{-7} and 10^{-8} bits.

In the present application, event sizes do not exceed 10^4 bits, so that the resulting event loss rate is acceptable even for the higher error-rate extreme. If event sizes do increase to 10^5 bits (ie. 1 event per scanline), either the target specification can be enforced by tape selection or surface treatment, or the device pattern-sensitivity can be studied quantitatively to determine whether or not the specification is unnecessarily conservative. If pattern-sensitivity is not serious, the poorest quality tapes are acceptable because even a 1% loss of events is without significance if there is no bias. If pattern-sensitivity is appreciable, a number of technical remedies are feasible such as a change from Miller coding to recently discovered codes[(15)] having much reduced low-frequency components, or accepting the capacity and throughput penalty of applying error correction.

3.7 VT system availability

A reliability analysis[(16)] of an early design configuration of the VTR indicated an MTBF of 694 hours. This figure has not been achieved in practice, although reliability continues to improve steadily. The disparity between predicted and actual performance appears to be due to the fact that the analysis based on component statistics does not take into account the occurrence of breakdowns associated with improper construction or weaknesses of design. During the early phase in particular, failures of this kind were predominant, and it is only after further design development that they are being eliminated.

The least reliable part of the machine is the tape transport subsystem (calculated MTBF 2100 hours), with a high electromechanical content, followed by the image electronics system (2400 hours) with its delicate scanner. Then there is the reel servo subsystem (4200 hours) comprising motors and power semiconductors. Other major subsystems having a high integrated circuit logic content, such as address electronics, capstan and scanner servos and control logic, have proved substantially more reliable, as also the VTR controller and computer interfaces.

An MTTR of 30 minutes is theoretically feasible for the VTR, based on the on-site availability of 44 field-replaceable subassemblies. Some items such as printed circuit cards can be exchanged in a few minutes. Long repair times apply to specific assemblies which are either difficult of access (eg. the crowbar unit) or require lengthy set-up procedures (eg. the scanner assembly). Of all the target specifications, the desirable MTTR is that from which the video-tape system performance departs most widely. This is because of spares supply problems, and the fact that it would be uneconomic to maintain a full range of replacement parts on-site to support a single facility. A down-time of several days can therefore occur when some essential component must be obtained from the manufacturer, and this can be even more prolonged in the event of shipment errors or other difficulties.

In spite of the numerous engineering difficulties involved in implementing an advanced performance on-line mass storage system, the major shortcoming of the present facility is not essentially technical, but rather the failure to support a short MTTR. It will be noted that this is a factor unrelated to the technology involved, except in so far as a low-cost technology would reduce the investment necessary to carry comprehensive on-site spare parts.

4 Optical holographic (OH) data storage

Compared with magnetic-tape recording, optical data storage techniques feature very high potential recording densities without transducer-medium contact problems; and a replay signal-level determined by readout laser intensity rather than the stored energy of the medium. We restrict our consideration to the interferometric Fourier transform holography technique[17], because this offers the same features as bit-by-bit optical recording, plus the inherent tolerance to localised medium defects and dust particles, which have the effect of reducing the aperture rather than corrupting individual bits. In addition, the technique has some immunity to optical misalignments because the exact

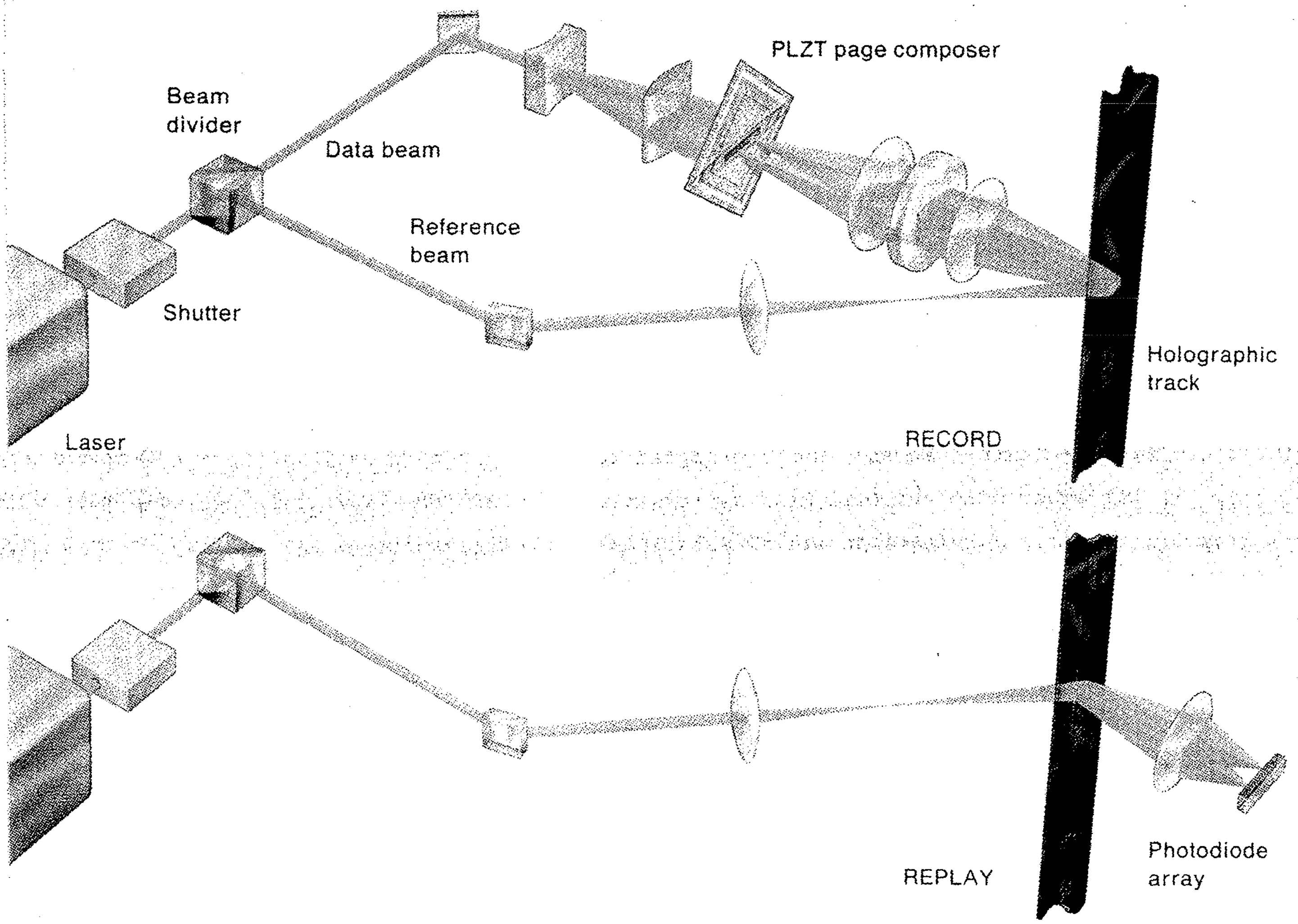

Fig. 2

<u>Principal components of a tape-organized optical holographic data recorder</u>

position of the readout beam across the hologram is not critical and the possible depth-of-field relatively large[18].

4.1 OH on-line data capacity

Much of the research in optical holographic storage has been directed towards the development of completely non-mechanical devices with few-microsecond access-times for disk replacement. Realisable solid-state detector dimensions limit the capacity of such memories to 10^8 bits[19], two orders of magnitude less than our target requirement. Of the many techniques[20] which have been proposed to attain higher on-line data capacities, the moving-tape medium configuration of magnetic-tape recording is probably the most practical for a data recorder. The optical configuration of one such device under development is indicated schematically in Fig. 2, and a prototype model is shown in Fig. 3. With such a tape (roll-film) approach, the page composer (which spatially modulates the laser object beam with the data pattern) need only be one-dimensional. This is a significant advantage because PLZT (lanthanum-doped lead zirconate-titanate ceramics) and other electro-optic materials tend to suffer from non-uniformity and instability.

Current recording densities allow the storage of the target capacity on less than 200 ft of film-tape, so that a conventional type of tape drive could readily handle one order of magnitude more on-line data.

4.2 OH data rate

The principal factors limiting the attainable recording rate in optical holographic memories are the laser peak-power and medium sensitivity, and the frequency response of the beam modulator (page composer).

Thermally induced recording techniques, such as thermomagnetic recording with magneto-optic readout, require the highest write energy

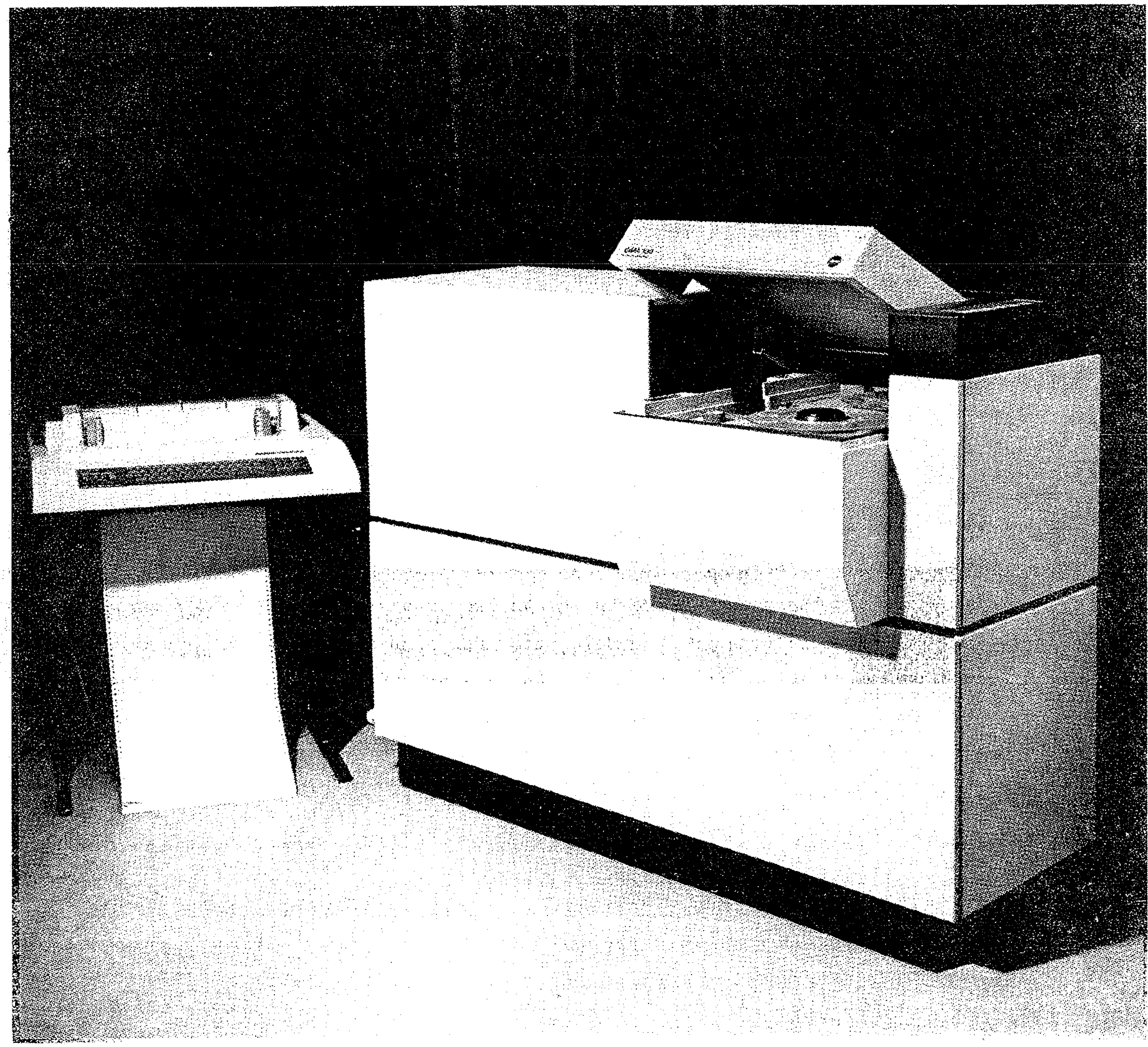

Fig. 3

Prototype model commercial optical holographic data recorder densities - typically 1 nJ/μm^2. Recording on such media at high data rates would thus demand very high peak powers, which are obtainable only with solid-dielectric lasers such as the ruby laser. These devices operate only at low repetition rates, so that unless a reliable page composer with a very large number of elements becomes available, thermally induced recording should be considered inappropriate for high data-rates.

For read-write recording by photon-induced mechanisms, as in photochromic or thermoplastic-photoconductive media, argon or krypton ion gas lasers can provide the required output power; while archival recording on photographic film (requiring only 10^{-4} nJ/μm^2) is possible with a lower-power, longer-life, helium-neon laser. Other forms of laser(21) such as the semiconductor p-n junction type, having inherently higher efficiency, may in due course become suitable candidates for sources; but restricted lifetime and relatively poor coherence properties render them unsuitable for holographic recording at this time. A favourable configuration for high-rate recording in the present state of the art would be the use of an atomic gas laser with photographic film for archival purposes, and a photon-induced material for read-write operation. Owing to the low conversion efficiency (10^{-3} - 10^{-4}) of such lasers, an input power of several kVA is necessary for recording at the target rate of 8×10^6 bits/sec. Even with such inputs, the target rate may not be readily attained in the case of the read-write medium.

PLZT page composers now have adequate speed for the target data rate, assuming an array of more than 100 bits. Liquid crystal page composers, having a response time of 100 ms, are quite unsuitable.

The photodetector array used for read-out requires a response time matching the desired replay data rate, so that photoconductors and phototransistors are too slow. p-n planar diffused photodiodes are suitable, and Schottky-barrier p-i-n detectors with response times of a few ns are available. Monolithic arrays(22) being developed for television use can operate at frequencies higher than the data-rate limit imposed by PLZT page composers in recording.

4.3 OH access time

Owing to the very high recording density in optical holographic storage, the target access time is readily achieved by a tape transport having a search speed over 10 ips. In practice, a search speed over 100 ips would be desirable to match the much greater potential on-line data capacity. In a tape-organized memory, a beam deflector is only required to move the beam from track to track. While many non-mechanical techniques for acousto- and electro-optic beam deflection have been developed[23], simple galvanometer-type deflectors can effect track selection with high precision in less than 10 ms. In future systems, semiconductor laser diode arrays may offer the possibility of simultaneous multi-track access.

4.4 OH media cost

Media costs for holographic recording cannot be estimated with precision since the price of proprietary material is largely determined by market conditions. While high-resolution photographic film of the required sensitivity is available commercially and second-sourced, it may require special slitting, block index-perforation and spooling to suit the transport requirements. Tentative estimates for photochromic media correspond to about 5.3×10^{-9} SF/bit, similar to that for video tape. However, if the new media were to enter large quantity production, a cost decrease by a factor of 10 could be anticipated, so that the 3×10^{-9} SF/bit target would be readily attained.

4.5 OH storage compactness

Very high potential storage density is the main feature to be gained in optical data recording, especially where the problems of mechanical tolerances are alleviated by the holographic technique. Current practical densities of 2×10^7 bits/in^2 exceed our target specification by almost two orders of magnitude, and further improvement is to be anticipated as the technology evolves.

This degree of storage compactness is of course not undesirable, but it will be realised that it is bought at the price of a technology requiring a number of expensive electro-optical components, including a laser power measured in kilowatts. There is evidence here of a degree of mismatch between what the technology currently offers, and what the on-line experiment data recording application most requires.

4.6 OH data retrieval reliability

Experimenters have reported data error rates as low as 10^{-6} bits on first power-on of a new holographic storage device, readily improved to 10^{-8} by system tuning. If this performance can be maintained in the field with volume-production media it is fully acceptable for the application. While there have been no published reports of the phenomenon, it is to be expected that some data pattern sensitivity would result from non-uniformity in the page composer contrast-ratio or the photodetector array sensitivity.

Archival recording on photographic film has been seen to have attractive features compared with optical read-write recording, since the laser power required is reduced, higher write data-rates are attainable, and the medium is more stable. Unfortunately, however, it is not possible in this case to implement a periodic or continuous read-after-write check in the manner of magnetic-tape recording, because the data cannot be read from the film until it has undergone off-line processing. The danger here is that a system fault could go undetected throughout a prolonged period of data-taking, since the fact that data had been written incorrectly due to electronic or mechanical failure, or could not be retrieved because of a defective medium, might not be discovered until the film was developed at a later date.

This shortcoming appears important enough to make the use of a photochromic or other read-write medium mandatory, in spite of the attendant disadvantages. While conventional photochromic media suffer from progressive bleaching during reading, as well as fatigue effects which limit the number of times they can be recycled, it is claimed that both these defects have been overcome by recent progress in the chemistry of the material.

4.7 OH system availability

A tape-organized optical holographic memory incorporates a medium transport essentially similar to a conventional magnetic-tape drive. When the magnetic recording heads and associated electronics are excluded from the calculation, such a unit may be expected to have an MTBF of the order of 1500 hours.

In addition, the system may incorporate a microprocessor controller, which may be expected to be highly reliable, and several electro-optic components of limited lifetime. An argon-ion laser for read-write recording may be expected to have an MTBF of 3000 hours, while a lower power helium-neon laser for archival recording may achieve 5000 hours. For comparison, a BS 161 standard filament lamp has an MTBF of 1000 hours; a 'double-life' lamp, with 10% lower initial light output, 2000 hours.

While there is much room for improvement in the area of high-speed long-lifetime page composers, recent advances in PLZT ceramics have led to the development of devices which have survived 5000 hours of continuous operation, and have projected lifetimes up to 15,000 hours. High reliability is anticipated for the photodetector readout array, on which the incident power is only about 10 μW.

Combining these statistics suggests that it is reasonable to expect that the target 1000 hour MTBF could be attained in practice. An optical holographic memory should be able to realise an MTTR comparable with that of a magnetic-tape unit provided sufficient attention is given to this factor at the design-detail phase. A laser plasma tube may be replaced in one hour, and other parts in the optical paths can be exchanged without lengthy adjustments if they are pre-aligned and bonded to keyed supports. Fault diagnosis time may be somewhat extended by the requirement that protective shielding be in place before operating the laser, and that the transport light-cover be closed when writing film.

5 Conclusion

From the review of the capabilities of the two technologies, it is apparent that, with adequate development, both VT and OH techniques are in principle capable of meeting all of the TS specifications, although with differing features and shortcomings. Furthermore, in the case of important characteristics such as on-line data capacity, both systems actually offer a performance which exceeds the operational requirement by a substantial margin.

On the other hand, both systems achieve this performance at the expense of considerable electronic and electromechanical complexity, which implies high initial cost, relatively frequent maintenance tasks, and a limited MTBF. Media/transducer movement is fundamental to at least the reading process of magnetic surface recording, and it is necessary also in the optical case to achieve the desired capacity. Hence both techniques require a tape transport mechanism which, in spite of the advanced technology components employed in other system areas, is one of the least reliable elements in each case.

After a long period as an also-ran, the time may be approaching for OH recording to make a successful entry at the top end of the memory hierarchy. This entry will be difficult because it must displace a technology which has been dominant for two decades, and which is itself in continuous evolution. At present, the optical holographic technique appears likely to offer at least an order of magnitude improvement in storage density compared with digital video-tape recording. This is a significant progress which is most attractive in many applications, but not essential for that considered here. Should a shift in the future development emphasis result in this performance gain being traded for a substantial decrease in system cost, or an order of magnitude increase in MTBF, the inevitable inertia in changing to the new technology would be readily overcome.

6 Acknowledgments

The development of the video-tape system referred to in Section 3 of this paper was initiated by C. Rubbia. The video-tape software was written by S. Cittolin.

The author is indebted to P. Waterworth and a European manufacturer for permission to publish Figs. 2 and 3, and to quote optical holographic memory parameters given in Section 4. This company does not wish to be cited prior to a public product release, anticipated mid-1978.

7 References

1 POULSEN, V.: 'Steel tape as a recording medium', US Patent 661!619, 13 Nov 1900

2 RODRIGUEZ, J.A.: 'An analysis of tape drive technology', Proc. IEEE, 1975, 63, No 8, pp. 1153-1159

3 ESCHENFELDER, A.H.: 'Promise of magneto-optic storage systems compared to conventional magnetic technology', J. Appl. Phys., 1970, 41, pp. 1372-1376

4 ANDERSON, L.K.: 'Holographic optical memory for bulk data storage', Bell Lab Rec., 1968, 46, pp. 319-325

5 KUEHLER, J.D. and KERBY, H.R.: 'A photo-digital mass storage system', Proc. Fall Joint Comp. Conf., 1966, p. 735

6 GRAY, E.E.: 'Laser mass memory system', IEEE Trans. Mag., 1972, MAG-8, pp. 416-420

7 BATE, G.: 'The future of magnetic recording for information storage', July 1976, CERN Computer Seminar

8 COLLINS, M., GUILLAUME, C. and RAUSCH, R.: 'Real-time and autonomous microprocessing in large control and data acquisition systems', Nov 1975, CERN LAB II-CO/75-4

9 d'AURIA, L. et al: 'Experimental holographic read-write memory using 3-D storage', Appl. Opt., 1974, 13, No 4, pp. 808-818

10 'Recorded magnetic tape (9-track, 6250 cpi, GCR)', ANSI Standard X3.54-1976

11 CHIEN, R.T. et al: 'Analysis and improvement of photostore error rates', Proc. IEEE, 1968, 56, pp. 805-812

12 TAYLOR, B.G.: 'Phase-distortion compensator for high-density digital magnetic-tape recording', Electron. Lett., 1974, 10, pp.492-493

13 CITTOLIN, S. and TAYLOR, B.G.: 'Digital video-tape data handling in a CAMAC system', Proc. IERE Conf. on Video and Data Recording, Birmingham, 1976, pp. 99-110

14 MILLER, A.: US Patent 3'108'261, 22 Oct 1963

15 MALLINSON, J.C. and MILLER, J.W.: 'On optimal codes for digital magnetic recording', Proc IERE Conf. on Video and Data Recording, Birmingham, 1976, pp. 161-169

16 BRYAN, C.L.: 'Mean time between failure analysis for high information density recorder IVC-1000', Consultant's report, Saratoga CA, USA

17 GABOR, D.: 'A new microscope principle', Nature, 1948, 161, pp. 777-778

18 ZOOK, J.D. and LEE, T.C.: 'Geometrical interpretation of Gaussian beam optics', Appl. Opt., 1972, 11, No 10, pp. 2140-2145

19 VANDER LUGT, A.: 'Design relationships for holographic memories', Appl. Opt., 1973, 12, No 7, pp. 1675-1685

20 KIEMLE, H.: 'Considerations on holographic memories in the gigabyte region', Appl. Opt., 1974, 13, No 4, pp. 803-807

21 ELECCION, M.: 'The family of lasers: a survey', IEEE Spectrum, 1972, March, pp. 26-40

22 BARBE, D.F.: 'Imaging devices using the charge-coupled concept', Proc. IEEE, 1975, 63, pp. 38-67

23 TUFTE, O.N. and CHEN, D.: 'Optical memories: controlling the beam', IEEE Spectrum, 1973, March, pp. 48-53