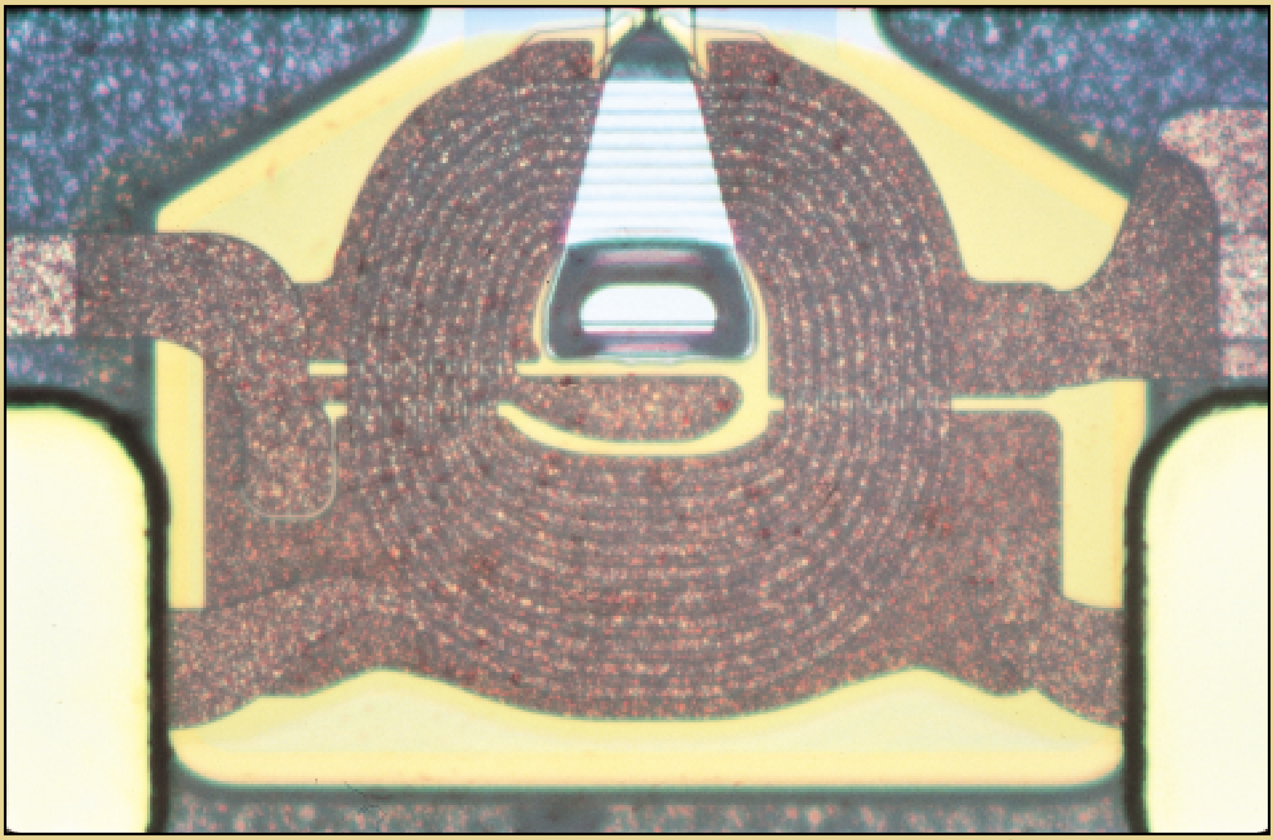


Optical micrograph of an IBM giant magnetoresistive head for writing and reading data on a computer's magnetic hard disk drive

MEMS-Based Integrated-Circuit Mass-Storage Systems

Abandoning the rotating disk paradigm, simple miniature microelectromechanical systems position probe tips over the storage media, potentially creating a new generation of nonvolatile rewritable mass storage devices, as well as support for multitude of "intelligent" gadgets.

DISK DRIVES' 100+%-PER-YEAR CAPACITY GROWTH RATE CONTINUES TO OUTPACE THE 60%-PER-YEAR GROWTH RATE OF SEMICONDUCTOR MEMORIES, ACCORDING TO THE SEMICONDUCTOR INDUSTRIES ASSOCIATION, ALLOWING DISK DRIVES TO MAINTAIN AND REINFORCE THEIR STATUS AS THE MOST COST-EFFECTIVE NONVOLATILE STORAGE SOLUTION AVAILABLE FOR COMPUTERS AND FOR FUTURE CONSUMER ELECTRONICS PRODUCTS. UNFORTUNATELY, DECREASES IN disk drive access times have been minimal, creating a significant performance problem for common applications accessing many small pieces of data, such as those in transaction-processing workloads. However, though the minimum entry cost of disk drives has declined in recent years due to decreases in the number of disks and heads per drive, their cost is still much too high for many consumer applications.

In order for mass-storage devices to pierce both the access-time and entry-cost barriers, researchers have turned to hybrid approaches leveraging the best of both semiconductor memories and disk drives. From semiconductor memories, the hybrid approaches have adopted the parallel wafer-fabrication process that keeps unit costs low. From disk drives, the hybrid approaches have adopted recording heads using mechanical positioning to address data stored on thin-film material instead of the lithographic definition required by today's semiconductor memories. But for compatibility with

L. Richard Carley, Gregory R. Ganger, and David F. Nagle

IBM ALMADEN RESEARCH CENTER, SAN JOSE, CA silicon-wafer-fabrication processes, these hybrid approaches typically abandon the rotating disk paradigm in favor of simple microelectromechanical systems (MEMS) to position probe tips over the storage media.

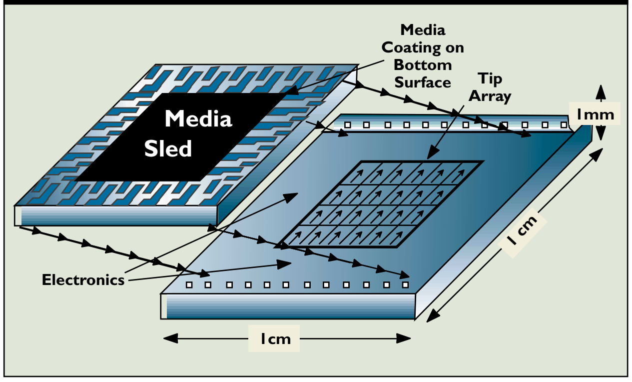
MEMS-based storage systems are potentially a whole new storage technology capable of a dramatic decrease in entry cost, access time, volume, mass, power dissipation, failure rate, and shock sensitivity. More important, these devices can integrate computation with storage, creating complete system-on-a-chip solutions—including mass storage. Integration enables many new applications that exploit the low unit cost and extremely small size of these new hybrid devices, including intelligent appliances, sophisticated teaching toys, biomedical monitoring devices, civil infrastructure monitoring devices, micro- and nanosatellites, highly integrated archival storage systems, and highly secure systems. The technologies needed to build these hybrid mass-storage devices are emerging today, making it likely that a broad market for nonvolatile rewritable mass-storage devices will develop within the next five years.

Here, we examine MEMS-based data storage technology and how it can be applied to computer systems from the perspective of a prototype MEMS-based data storage system being developed at Carnegie Mellon University. Where possible, we compare and contrast this system with others being developed at such major industrial research laboratories as Hewlett-Packard, IBM, and Kionix.

MEMS-based Probe Storage Architecture

Figure 1 shows the Carnegie Mellon prototype MEMS-based system [3, 4]. Like disk drives, the device has recording heads and a recording media surface that moves. However, the heads are actually MEMS probe tips fabricated in a parallel wafer-level manufacturing process. This prototype employs magnetic storage media much like that used by disk drives. But the media surface does not rotate; instead, it moves linearly in the X and Y directions to seek the appropriate data. Data access is accomplished by moving the media at a constant velocity in the Y direction while data is read or written by the stationary probe tips. This design avoids problems with "stiction," or the force required to begin moving one body whose surface is in contact with a fixed surface, that occurs in rotating bearings at very small geometries. Stiction problems can prevent precise nanometer position control, as elements tend to move by alternatively sticking and slipping. The design in Figure 1 also avoids the potential wear

Figure 1. Conceptual view of MEMS-actuated data storage devices with X-Y motion of the media and Z motion of a 4x5 array of probe tips.

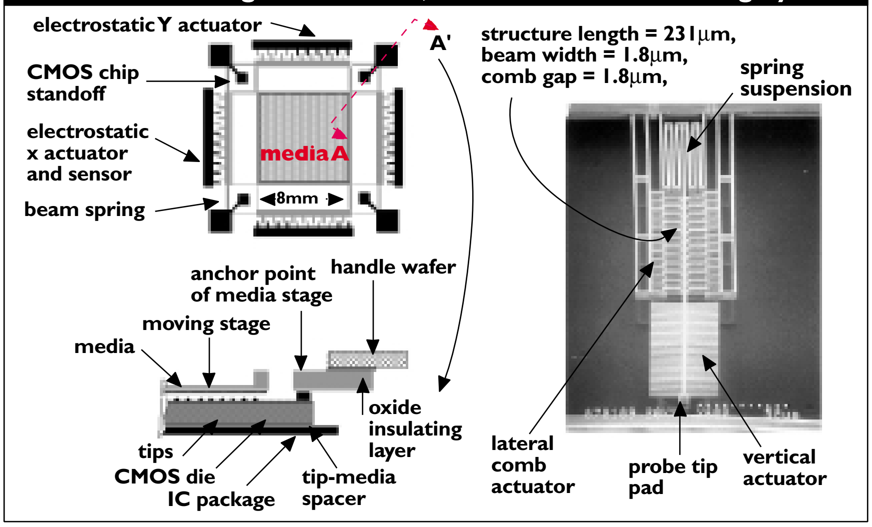


(MEMS bearings tend to have short life spans) arising when micromechanical surfaces come into contact. The media for the prototype is deposited on a large (8mm×8mm×500µm) square plate (called the "media sled") and is held above the probe-tip array by a network of springs. (For more on the project, please see www.chips.ece.cmu.edu.)

A force is applied to the sled using electrostatic actuators, though in principle, electomagnetic or thermal actuators could be used. Unfortunately, such reciprocating motion is usually limited to a small fraction of the size of the structure. With typical motions being 10% or less of the suspension/actuator length, a single probe tip "sweeps" only on the order of 1% of the media sled. However, by using a large array of probe tips, all of the media area can be addressed, as long as the media sled moves in X and Y, by the pitch of the probe-tip array. A large array of probe tips also provides a significant increase in the data rate and reliability for the overall system.

Probe-tip positioning. Because the media surface is not perfectly flat, and individual probe-tip heights can vary across the probe-tip array due to both manufacturing variations and the curvature of the probe-tip wafer, nearly all approaches to MEMS-based storage incorporate some form of tip-height control. One exception is the approach described by Hewlett-Packard researchers in which the probe-tip-to-media separation is larger than the variations just described [10]. Though this larger separation may limit the choice of media type and reading and writing mechanisms, it dramatically reduces system complexity. Nearly all other approaches employ either passive or active independent positioning of probe tips in Z. For example, the IBM Millipede probe-based mass-storage system uses a 32×32 array of probe tips, each placed on the end of a cantilever [9]. When the tips are brought into contact with the media, the cantilevers each bend a different amount to make up for varia-

Figure 2. Examples MEMS implementations of (a) the media actuator and (b) the probe head-positioning actuator. The media actuator has a 100μm-stroke and is fabricated using deep Si reactive ion etching. The footprint is 14mm x 14mm. Anchored regions are black; the movable structure is gray.



tions in initial Z separation. (Positioning on X and Y is taken care of by a large media sled suspended by springs and moved using electromagnetic actuators.)

Although the complexity and level of integration with the required electronics is much greater for non-contact approaches, mechanical contact inhibits precision position control of the tip and may degrade the system's reliability and longevity. The Carnegie Mellon prototype design provides for independent active control of the Z motion at every probe tip. Individual probe tips are placed on cantilevers that are electrostatically actuated to a fixed distance from the media surface using a local Z-positioning feedback loop [8]. In an active Z-compensation approach, the probe-tip suspension must have enough range of actuation to "servo-out" all variations in Z.

See Figure 2(b) for a prototype Z control system designed, fabricated, and tested for use in the Carnegie Mellon system. The probe is mounted on a pallet at the bottom of the structure. Above the probe tip is an array of "fingers," or short metal stubs, that provide parallel plate electrostatic Z actuation upward (toward) the media surface. This same plate is used as a capacitive displacement sensor to determine the probe-tip-to-media separation [4, 8]. The serpentine flexure structure (top of Figure 2b) reduces the spring constant, softening the spring without making it longer. On either side of the spring, a set of comb fingers provide for a small lateral actuation (~0.5μm). Fingers attached to anchors are "interdigitated" with ones attached to the movable structure. This fine lateral actuator enables fine track-following and can cancel out temperature differences between the silicon tip array and the silicon media sled in large designs. Experimental measurements on the design in Figure 2(b) indicate a Z-mode resonant frequency for the spring-mass system of about 10kHz. Movement of the pallet by 0.5μm in the Z direction requires only 8V on the Z actuator fingers. A 0.25μm static lateral motion requires only about 5V on the lateral actuator fingers (near the middle of Figure 2b).

Wiring this MEMS-based storage system's 6,400 probe tips to servo and channel electronics requires integration of the electronics directly into the same die as the probe tips. This integration greatly improves the bandwidth and sensitivity of the capacitive sensors integrated into the probe tips to determine their Z positions relative to the media. To achieve a highly integrated CMOS+MEMS process, the system's designers have developed a series of post-processing steps following a standard CMOS fabrication process that turns conventional interconnect of a CMOS integrated circuit into movable mechanical structures [5, 8].

Extensions to this integrated CMOS+MEMS process are being developed to fabricate the read/write probe heads. A major challenge is fabricating the necessary read/write probe tips in a way that is compatible with the underlying CMOS circuitry—including thermal compatibility with CMOS, geometrical compatibility with CMOS+MEMS, compatibility with current constraints due to electromigration, and chemical compatibility with the release processes.

Media positioning. The most effective and efficient use of the media requires the media sled move by the probe-tip actuator pitch in X and Y. The system's current targets are a probe-tip array with 100μm centers in X and Y; hence, the media actuator must move at least ±50μm. (Figure 2(a) shows a conceptual layout of the media actuator design and a partial cross section of the entire assembly.) This design is based on a decoupled-mode X-Y microstage, originally conceived in 1996 for use as a vibratory-rate gyroscope. A box-spring suspension decouples the two lateral directions of actuation, so comb fingers can be used for X and Y actuation without mechanical interference.

The system's designers recently fabricated a first-generation prototype actuator using deep-Si reactive-ion-etch (DRIE) technology. The high-aspect-ratio silicon structures in the springs provide media stability in the Z direction. Electrostatic actuators with greater than 200 interdigitated fingers in three ranks with 500μm thickness and 16μm gaps are designed to achieve the ±50μm displacement in each direction at an actuation voltage of 120V. This voltage is far above the rated voltage for the CMOS process; but since there are only four drive voltages, two for X and two for Y, these high-voltage drivers can be implemented with off-chip high-voltage transistors.

Electrostatic force generated by the comb-drive actuators is balanced by a combination of spring force and inertial forces. If the actuator potential is increased, more force is applied. At first, the additional force accelerates the media sled. Then, as it approaches its desired location, an opposite force is applied to decelerate the sled to stop at the desired spot. A servo feedback system senses the lateral position by measuring changes in the capacitance of the same lateral comb structure used for actuation. The capacitance sensor has a gain of around 700fF/μm, making possible closed-loop media X-Y position control with less than 10nm of positioning error. The operating shock-limit target in the prototype system is 10g, corresponding to 10% of the maximum actuating force available for media-sled movement. In future generations, shock tolerance could be increased by using smaller media sleds or by selectively hollowing-out portions of the media sled.

The system's engineers have carried out detailed mechanical finite element analyses on this actuator structure. The lowest mechanical resonant frequency in X and Y is 739Hz, and the lowest resonant frequency in Z is 2,940Hz. These resonant frequencies render the positioning system relatively shock-tolerant, making it possible to achieve access times of less than a millisecond.

Storing, reading, and writing bits. These engineers, as well as engineers in other laboratories, have explored an extremely wide variety of probe-tip/media combinations for use in writing and reading bits in MEMS-based storage systems. IBM's Millipede is one of the best-developed probe-based mass data storage approaches [9]. Writing in this system is achieved by melting pits in a polymer; a readout signal is elegantly obtained by measuring the thermal conductance of the tip to the medium. IBM has demonstrated the integration of a large number of probes (32$\times$32) on a 3mm$\times$3mm surface, only a little bigger than the diameter of a pencil lead.

There are, however, several potential disadvantages to melting pits in polymers using tip heating. The probe tips are constantly in contact with the media, raising concerns about the wear and durability of both the tips and the media. In principle, erasing media can be achieved by heating to remelt the surface and remove the pits. It should also be possible to erase either the complete media or parts of a thermally segmented media. However, the practicality of this form of rewriting and the upper maximum number of rewrite cycles that would be allowed is currently unknown.

Table 1. Parameters for system-level characterization of the media sled.

Parameter	Value
Sled mobility in X and Y	100um
Bit cell width (area)	40nm^2
Number of tips	6,400
Simultaneously active tips	1,280
Tip sector length	80 bits
Servo overhead	10 bits
Deice capacitiy	4.0GB
Sled acceleration	82g
Per-tip data rate	700Kb/s
Settling time constants	1
Spring factor	75%

Another promising approach is the use of media that can be converted from an amorphous to a crystalline state and back to an amorphous state (hence the name "phase-change media"). This media has been used in rewritable CDs, whereby laser beams heat the material (see McDaniel's "Magneto-Optical Data Storage" and van Houten and Leibbrandt's "Phase Change Recording" in this section). There is a considerable body of knowledge concerning these materials. For example, Hewlett-Packard researchers recently described a MEMS-positioned probe-based storage device employing phase-change materials in which field emission currents can be used to heat the media locally [10].

Other less-mature candidates for MEMS-based mass-storage systems include ferroelectric memories, charge-trapping devices, atomic motion in alloys (such as Pt atom migration), and charge motion within molecular media layers. However, none of them has yet demonstrated the combination of long-term stability of the memory element, a high degree of rewritability, and a practical probe-based method for reading and writing. In order to start with a storage mechanism that has already demonstrated all such capabilities, the prototype Carnegie Mellon system is designed to store bits using vertically oriented magnetic media, similar to the media in magneto-optical disk drives.

Small permanent magnet probe tips have been used to image domains in magnetic thin films using a technique called "magnetic force microscopy" (MFM), which detects magnetic details with diame-

Table 2. Performance of the MEMS-based storage system prototype.

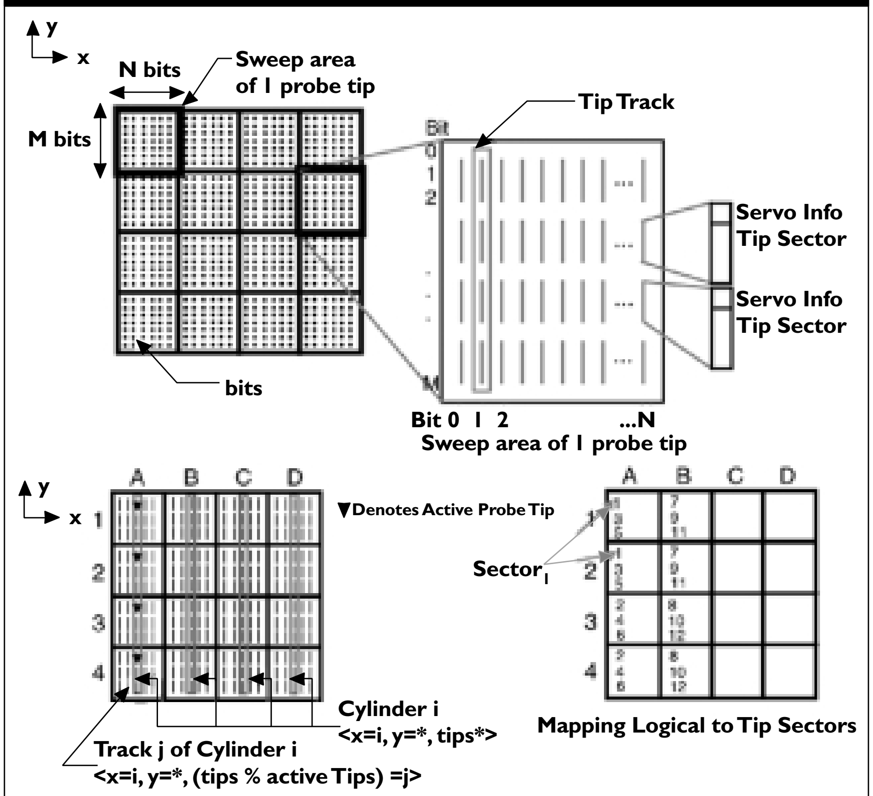
Performance	Time (ms)
Average service time	0.67ms
Maximum service time	1.92ms
Average seek time	0.53ms
Maximum seek time	0.73ms
Average X seek time	0.52ms
Maximum X seek time	0.73ms
Average Y seek time	0.35ms
Maximum Y seek time	0.73ms
Settling time	0.22ms
Average per-request turnaround time	0.06ms
Maximum per-request turnaround time	0.50ms

ters down to 30nm. Writing can be achieved in the MFM-mode simply by tailoring the material properties in such a way that a domain is nucleated when the tip is nearly in contact with the media. This writing scheme can be modified by applying additional external fields or by locally heating the sample. However, it should also be possible to write bits without physical contact between the tip and the medium (for the sake of system reliability and longevity).

Although very promising with respect to achievable data density, the MFM-mode of reading and writing suffers from two major drawbacks: a separate, oppositely magnetized probe tip must be used to erase the media before it can be rewritten; and the data rate is limited by the rate at which the cantilever can move up and down in Z. In principle, the second limitation should not have a serious effect in massively parallel data storage systems. Even assuming a data-rate of 50Kb/s per probe tip, an 80×80 probe-tip array should be able to achieve an aggregate data rate of over 320Mb/s (approximately 32MB/s). However, the separate erase probe-tip requirement is a barrier to implementation. A more desirable approach—being developed for the prototype Carnegie Mellon system—is a miniaturized magnetic recording head similar in function to the ones used in disk drives. That means an inductive write coil with a soft magnetic probe-tip flux guide for writing and a magneto-resistive element for reading. Large arrays of such read/write heads can be created on an array of probe tips [1–3].

The actual limit on the magnetic storage density for a MEMS-positioned storage device is set by the super-paramagnetic limit, which is a lower bound on the size of the individual magnetic switching units. If they are too small, background thermal energy has too high a probability of switching the stored bit. In advanced magnetic media being developed today by researchers at places like IBM and Seagate, this limit on unit size is on the order of 20nm. However, the size of the written regions that indicate the data bits must be much larger to ensure an acceptable media signal-to-noise ratio (SNR). For example, using today's most advanced magnetic technology, the mark size limit is on the order of 80nm×80nm. As the disk drive industry develops more stable media, we expect that by 2003 the minimum mark size will decrease to 40nm×40nm.

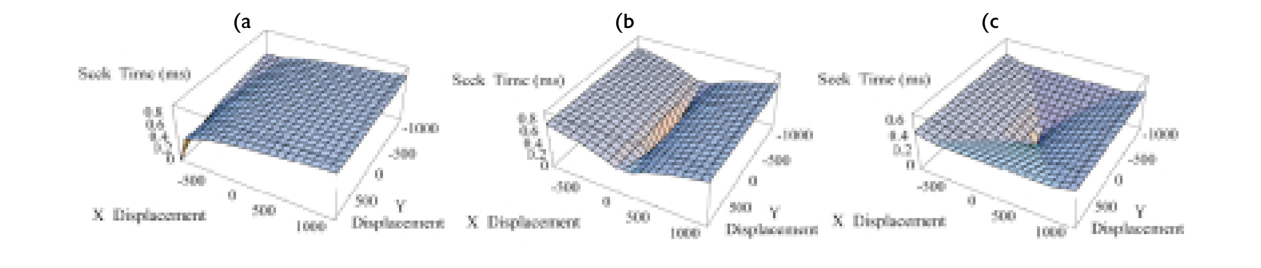
Figure 3. Data layout on the media sled. Each distinct row/column pair corresponds to a different probe tip.



Integrating MEMS into Computer Systems

Ultimate market acceptance of MEMS-based data storage devices depends on their price and performance gains over disk drives in terms of data rate, power, and access time. Exploring how the Carnegie Mellon system might be incorporated into various otherwise-conventional computing systems, we

Figure 4. MEMS-based actuator performance.



describe the system's basic performance characteristics and potential for improving overall system performance (see the sidebar and Figure 3 for how data is organized in MEMS devices).

Performance of MEMS-based storage. Using the basic physics equations governing a MEMS storage device's behavior [7], we built a performance model into the DiskSim storage simulator (a flexible storage system simulation infrastructure [6]) and evaluated its performance, given the device parameters in Table 1. The results (see Table 2) show that the system's average service time achieves better than an order-of-magnitude improvement over disk drives (0.52ms versus 10.1ms), with X seek time being the dominant factor. Figure 4 illustrates the importance of X seek time, showing the seek time variations from the sled's corner (4a) and center (4b). Most seeks have little dependence on Y-dimension movements, except for the shortest X-dimension seeks; 4c shows the same seeks as 4b but without any settling time.

The springs supporting the media sled also impart a position-dependent force on the sled that influences the seek time. The importance of the spring force depends on the ratio of the spring force at maximum sled displacement to the maximum actuator force available. Designers of MEMS-based mass-storage systems need sufficient actuator force to accelerate the media sled, overcome spring force, and overcome external accelerations.

Finally, device bandwidth depends on the number of simultaneously active tips together with the per-tip data rate. Like conventional disks, a MEMS-based storage system has to switch tracks (or cylinders) when media transfers cross track boundaries. Unlike conventional disks, whose rotational speed is independent of seek-arm positioning, a MEMS-based storage device's track switching time depends directly on access velocity. Specifically, the sled must turn around each time a media transfer crosses a track boundary. Reversing direction requires decelerating, changing direction, then re-accelerating to the access velocity. As the access velocity increases, turnaround time increases.

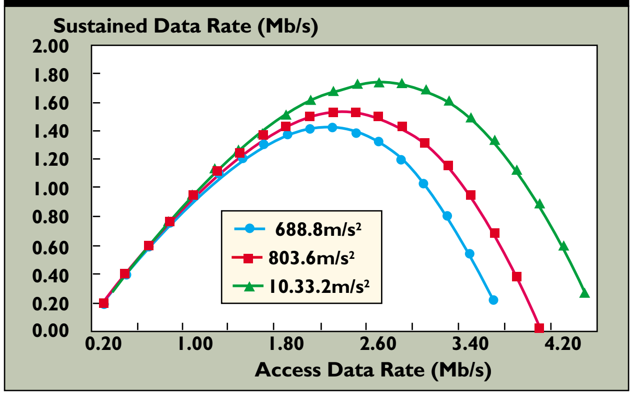
Figure 5 shows the sustained bandwidth of a single tip, given increasing per-tip data rates and three different values of maximum actuator acceleration. For each actuator acceleration, there is a maximum data rate, after which turnaround times dominate transfer rates. This upper bound is an important result, because it indicates that the recording head and channel need not handle ever-higher data rates, making them simpler to manufacture and improving the SNR caused by electronics noise sources.

Application performance. One notable benefit of MEMS-based storage is its great potential improvement in end-to-end performance. Table 3 shows the application performance improvements for the Carnegie Mellon system over conventional disks. This data was generated with the SimOS complete-machine simulator combined with the MEMS-based

Table 3. CPU time on real applications for disks and MEMS-based storage.

Benchmark	Storage Device	Compute Time	I/O Time	Total Tine
POSTMARK	Atlas 10k	10	730	740
	Superdisk	10	400	410
	MEMS Store	10	215	225
TPC-D (query 4)	Atlas 10k	3	27	30
	Superdisk	3	11	14
	MEMS Store	3	7	10

Figure 5. Maximum long-term average data rates as a function of channel-reading data rate for several actuator maximum accelerations.



storage-device model [12]. PostMark is a file-system benchmark corresponding to file activity in Internet servers (for email servers, newsgroup access and storage, and Web-based e-commerce applications). TPC-D is a large-scale database benchmark that exercises complex, long-running decision-support queries against large complex data structures.

On the PostMark benchmark, the system with MEMS-based storage completes three times faster than the disk drive, primarily due to much faster positioning times. Using a Quantum Atlas 10K disk, PostMark's average access time is 5.49ms; using the MEMS-based storage system, it is only 0.67ms. PostMark is largely characterized by many small accesses, mostly to file system metadata. The MEMS-based storage system's shorter seek times are especially beneficial for this type of access pattern. Moreover, repeated writes to the same blocks, common in metadata updates, perform poorly on disks with large rotational penalties, whereas a MEMS-based storage system could reverse at any time. TPC-D's similar speedups are especially impressive considering that the MEMS-storage-device model lacks an on-board prefetching cache; the on-board cache hit rate is almost 84% for the Quantum Atlas 10K.

Future Opportunities

There are many opportunities for MEMS devices in the storage hierarchy (see Hesselink's introduction to this section). Besides replacing disks, MEMS-based storage devices could serve as a nonvolatile disk cache, absorbing write traffic at much greater speed than conventional disk drives. The cache could also be explicitly exposed to and managed by software, allowing the software to make customized allocation decisions based on the performance needs

Device Characteristics and Data Layout

The MEMS media sled is organized into rectangular regions (see Figure 3), each storing N×M bits (2000×2000 bits in the default model) accessible by a single probe tip. The smallest accessible data unit is a "tip sector" consisting of servo information (10b) and encoded data+error correction (80b = 8B encoded data). Multiple tip sectors are grouped into logical sectors, similar to a SCSI disk's logical blocks. Unlike conventional disks, however, multiple probe tips access the media in parallel; many tip sectors can be read or written simultaneously.

To organize the low-level media contents, each bit is identified by the triple <x,y,tip> where <x,y> represents bit coordinates within the region addressable by <tip>. Each active tip reads or writes data within a column of bits called a "tip track," as in Figure 3. Drawing an analogy with disk drives, each bit with identical values for <x> is called a "cylinder." Therefore, a cylinder consists of all bits accessible by any tip without moving the sled along the X axis. Because power dissipation limits the number of probe tips that can be active simultaneously, cylinders are divided into tracks (all bits within a cylinder that can be read or written by concurrently active tips). In Figure 3, tips A1, A2, A3, and A4 are active, and the corresponding track is indicated.

Multiple simultaneously active tips allow logical sectors to be "striped," or spread across tip sectors (in multiple tip tracks) to reduce access time. Figure 3 shows how each logical sector is striped across two tip sectors. In order to read logical sectors 1 and 2, tips A1—A4 are activated, while the sled seeks to the top of cylinder 2 and moves down (in $-y$) across the first tip sector. Tip A1 reads half of logical sector 1; tip A2 reads the other half; and tips A3 and A4 read logical sector 2. In the default model, logical sectors of 512B are striped across 64 tip sectors of 8B each.

Positioning the sled for read or write involves several mechanical and electrical actions. To seek to a desired sector, the appropriate probe tips must be activated, the sled must be positioned so the tips are under the first bit of the pre-sector servo information, and the sled must be moving in the correct direction with the correct velocity ($v_x=0$ $v_y=+/-v_{access}$). Managing this activity can be tricky; whenever the sled moves in X, the sled's rapid acceleration and deceleration causes the spring-sled system to oscillate in X, imposing a settle time before $v_x \sim 0$. In addition, the spring-restoring force (possibly as large as 75% of the sled-actuating force) makes sled acceleration a function of instantaneous sled position.

Media access requires constant velocity in the y dimension, as determined by the desired per-tip read and write rates, the bit width, and the sled actuator force. In order to switch tracks during large data transfers, the sled performs a turnaround (reversing direction such that $\langle x,y\rangle_{final} = \langle x,y\rangle_{initial}$, and $v_{final} = -v_{initial}$). Turnarounds are also needed for seeks whenever the initial or final direction of motion is not the one that's needed. The spring-restoring force makes turnaround time a function of both instantaneous sled position and the direction of motion. C and access patterns of various data objects, including metadata, small files, and files with real-time constraints (such as video).

For many "portable" applications, including notebook PCs, PDAs, and video camcorders, MEMS-based storage provides a more reliable and lower-power solution than conventional disk drives. Unlike rotating storage, which cannot cope with device rotation (as in a rapidly turning PDA) and is relatively sensitive to shock (as in dropping a device), MEMS-based storage can be engineered to be immune to gyroscopic effects and to absorb much greater external forces. Moreover, MEMS-based storage creates a new low-cost entry point for applications in the 1–10GB capacity range. This low cost is a result of the MEMS-based storage device being manufactured using integrated-circuit-style wafer-level parallel fabrication.

With new applications creating massive amounts of data, MEMS researchers are also exploring how MEMS-based data storage devices can help solve data archival problems, including capacity, time to access data, and long-term data retrieval. For example, medical imaging generates gigabytes of data per patient. For cost reasons, this data is usually stored on tape. With "areal" densities, or bit count per unit area, 10X greater than high-capacity tape, it should be cost-effective to build storage "bricks" consisting of thousands of MEMS-based storage devices. Each brick would hold petabytes of data accessible in less than a second. By incorporating computing logic into the MEMS-based storage device, it would also be possible to process data directly within the storage brick. Large numbers of storage bricks would offer massive computational parallelism, creating the ultimate "disk" [11].

Another application domain for MEMS-based storage is bulk nonvolatile storage for embedded computers. Single-chip "throwaway" devices storing very large datasets can be built for such applications as civil infrastructure monitoring (on, say, bridges, walls, and roadways), weather and seismic tracking, and medical applications. One futuristic application is temporary storage for microsatellites in very low earth orbit. Given that satellites in a very low orbit move quickly, communication is possible only in very short bursts. A small, high-capacity, nonvolatile storage device is needed to buffer data.

MEMS-based storage devices could also add huge databases to single-chip continuous speech-recognition systems and be integrated into low-cost consumer or mobile devices. These chips could be completely self-contained, with hundreds of megabytes of speech data, custom-recognition hardware, and only minimal connections for power and input/output. **C**

References

1. Abelmann, L., Khizroev, S., Litvinov, D., Zhu, J.-G., Bain, J., Kryder, M., Ramstock, K., and Lotter, C. Micromagnetic simulation of an ultrasmall single-pole perpendicular write head. *J. Appl. Phys. 87,* 9 (May 1, 2000), 6636–38.
2. Abelmann, L., Zhu, J.-G., Bain, J., Ramstock, K., and Lotter, C. Micromagnetic simulation of a flux guide for a read head with sub-100nm resolution. *J. Appl. Phys. 87,* 9 (May 1, 2000), 5538–40.
3. Carley, L., Bain, J., Fedder, G., Greve, D., Guillou, D., Lu, M., Mukherjee, T., Santhanam, S., Abelmann, L., and Min, S. Single-chip computers with microelectromechanical systems-based magnetic memory. *J. Appl. Phys. 87,* 9 (May 1, 2000), 6680–85.
4. Carley, L., Guillou, D., Ganger, D., and Nagle, D. System design considerations for MEMS-actuated magnetic probe-based mass storage. *IEEE Transact. Magnet.* (2000).
5. Fedder, G., Santhanam, S., Reed, M., Eagle, S., Guillou, D., Lu, M.S.-C., and Carley, L. Laminated high-aspect-ratio microstructures in a conventional CMOS process. *Sens. Actuat. A,* 57 (1996) 103-110.
6. Ganger, G. et. al. *The DiskSim Simulation Environment Version 1.0 Reference Manual.* Technical Report CSR-TR-358-98, The University of Michigan, Ann Arbor, Feb. 1998; see www.ece.cmu.edu/~ganger/disksim/.
7. Griffin, J., et. al. Modeling and performance of MEMS-based storage devices. In *Proceedings of ACM SIGMETRICS2000* (Santa Clara, Calif., June). ACM Press, New York, 2000, 56–65.
8. Guillou, D. *Design of a MEMS-Based Data Storage System*, Ph.D. thesis, Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, Pa., 2000.
9. Lutwych, M., Cross, G., Despont, M., Drechsler, U., Durig, U., Haberle, W., Rothuizen, H., Shutz, R., Widmer, R., Binnig, G., Vettiger, P. Millipede: A highly parallel dense scanning-probe-based data-storage system. In *Digest of Technical Papers, IEEE International Solid-State Circuits Conference* (San Francisco, Feb. 7–9). IEEE, Piscataway, N.J., 2000, 126–127.
10. Morehouse, C. A decade away: Atomic resolution storage. Sidebar in Toigo, J. Avoiding a data crunch. *Sci. Am.* (May 2000).
11. Riedel, E. *Active Disks: Remote Execution for Network-Attached Storage*, Ph.D. thesis, CMU-CS-99-177, Carnegie Mellon University, Pittsburgh, Pa., 1999.
12. Schlosser, S., Griffin, J., Nagle, D., Ganger, G. Designing computer systems with MEMS-based storage. In *Proceedings of the 9th International Conference on Architectural Support for Programming Languages and Operating Systems* (Boston, Mass., Nov. 13–15, 2000).

L. Richard Carley (carley@ece.cmu.edu) is a professor of electrical and computer engineering at Carnegie Mellon University in Pittsburgh, PA.

Gregory G. Ganger (ganger@ece.cum.edu) is an assistant professor of electrical and computer engineering at Carnegie Mellon University in Pittsburgh, PA.

David F. Nagle (bassoon@ece.cmu.edu) is a research scientist in the School of Computer Science at Carnegie Mellon University in Pittsburgh, PA.

This research is sponsored in part by the National Science Foundation under grant no. ECD-8907068, by DARPA under the AFRL, Air Force Materiel Command, USAF, under agreement F30602-97-2-0323, and by an NSF Graduate Student Fellowship. The U.S. Government has certain rights to this material. IBM and Intel donated equipment that helped support this research.
