IBM Systems Storage

Storage Futures
Technologieausblick

Dr. Axel Koester
Technologist, STG Consultant
axel.koester@de.ibm.com

Thinking Beyond Today

Was macht eigentlich... "Millipede" ?

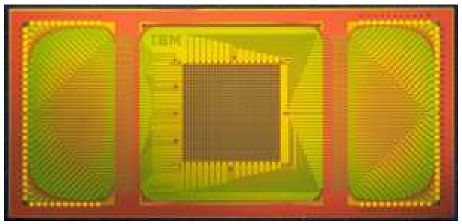
Siegeszug des Solid-state Speichers ?

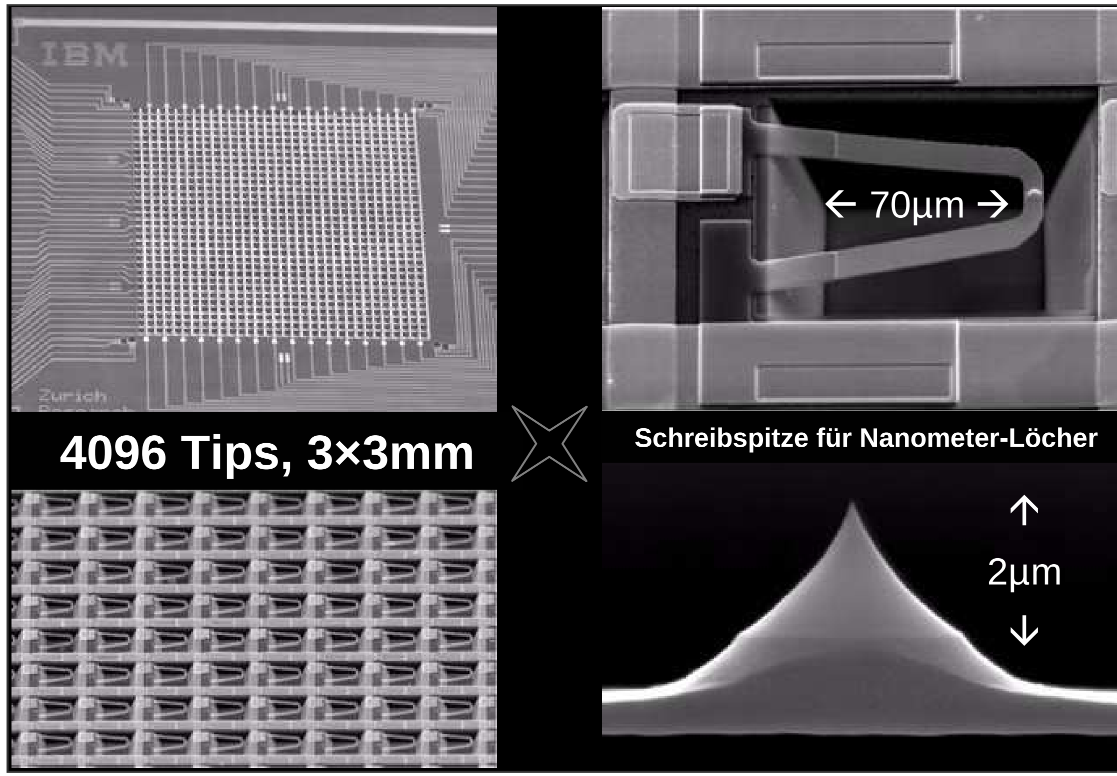
Innovationen in der Chiptechnologie

Nichtflüchtiges RAM

Was macht eigentlich... "Millipede"?

Nanomechanischer Speicher : "Lochkarte v2.0"





1 Spitze liest ~2000 Löcher/s
128x128 Prototyp: 120 Mbit/s
Verbrauch: <100 milliwatt
Schreibzyklen: >100.000

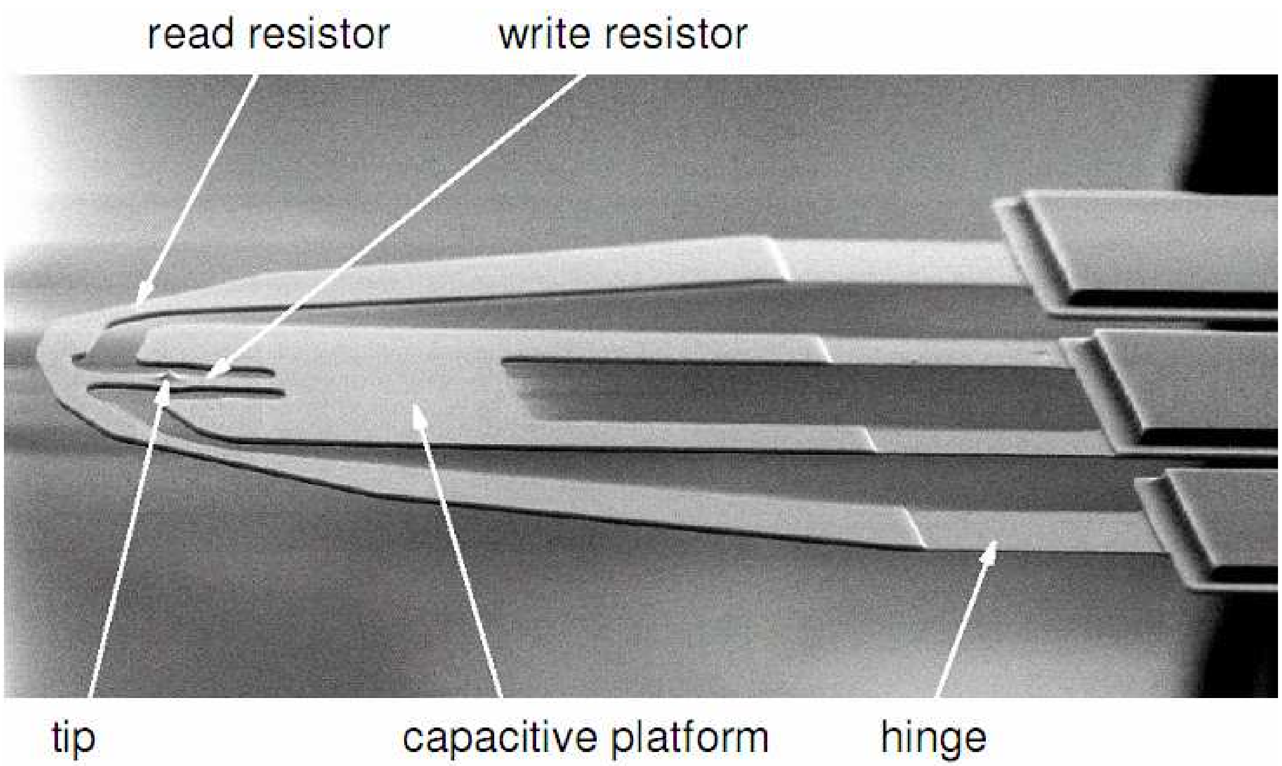
Gerd Binnig/IBM gewann 1986 den Nobel Preis für seine Erfindung "*scanning tunneling microscope*", Basis des heutigen "Millipede"

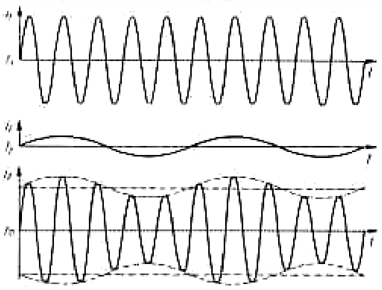
Neuer *Millipede* Schreib-Lesesensor

Write resistor – erhitzt lokal die Schreibspitze

Read resistor – mißt die von der Schwebehöhe abhängige Wärme-ableitung (±1nm)

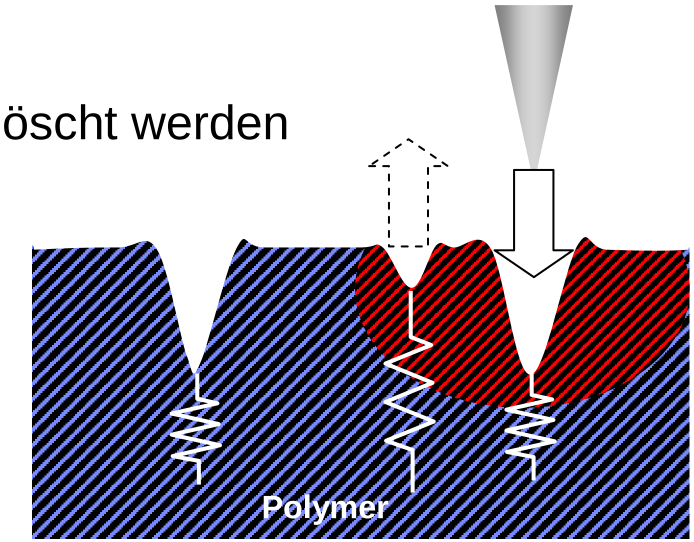
Capacitor – zieht den Cantilever zum Substrat durch statische Anziehung



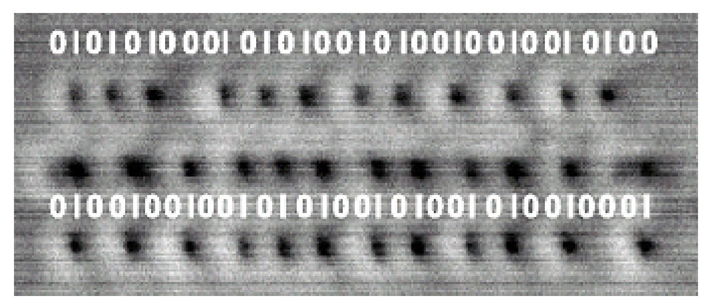


Neues Bit-lösch-Verfahren für "Millipede"

- Bits konnten bislang nicht individuell gelöscht werden
- Das Schreiben sehr eng aufeinander folgender Bits (<20 nm) **löscht** jedoch **die unmittelbaren Nachbarn**

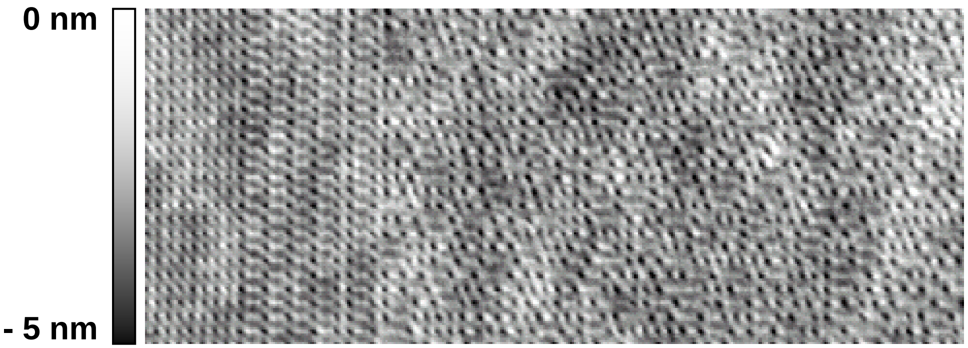


- Ein neues Codierverfahren erlaubt das Schreiben ohne vorheriges Löschen (vorhandene 1 wird durch benachbarte 1 wieder gelöscht)



Codierung ohne 11 Bitfolgen!

Herausforderung: Ultraglattes Trägermaterial

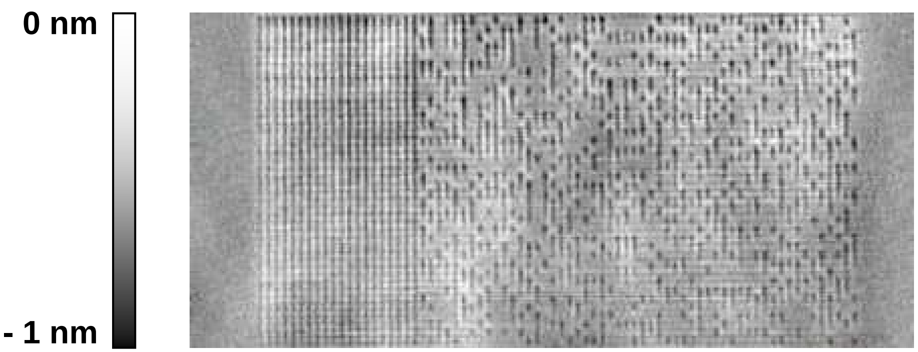


- Altes Polymersubstrat

  Bitabstand 13 nm, Tracks 26 nm

  1,2 Tbit/inch²

  Bit-Tiefe 5 nm, S/R ~9dB



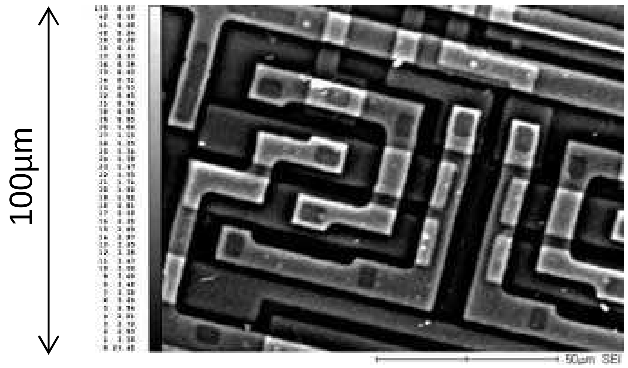
- Neues ultra-glattes Polymer

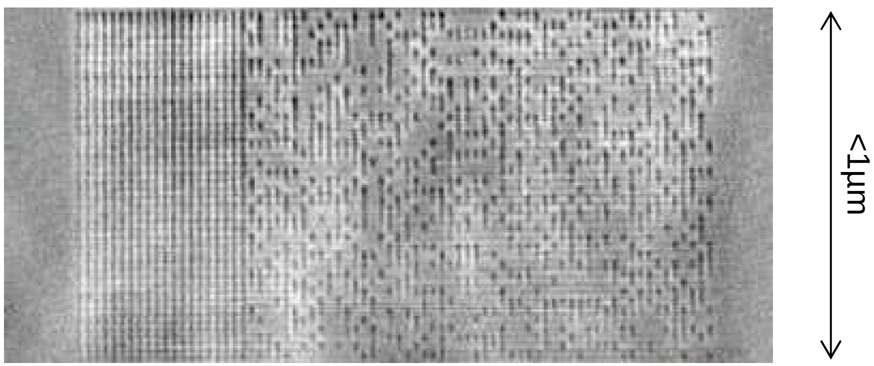
  Bitabstand 9 nm, Tracks 18 nm

  2,7 Tbit/inch²

  Bit-Tiefe 1 nm, S/R ~9dB

Patterned Media oder **Unpatterned** Media?





- Lithografisch geätzter Chip

  Minimale Strukturgröße = **45 nm**

  Natürliche Grenze erreicht ?

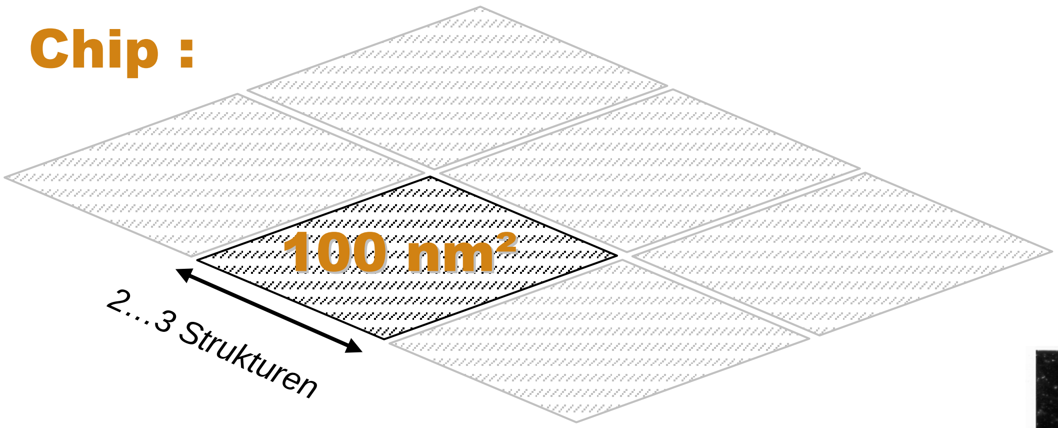
- Polymer-Datenaufzeichnung

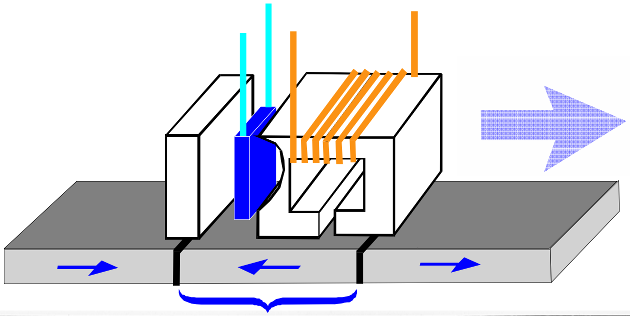
  Analoge Auflösung, derzeit **9 nm**

  2,7 Tbit/inch²

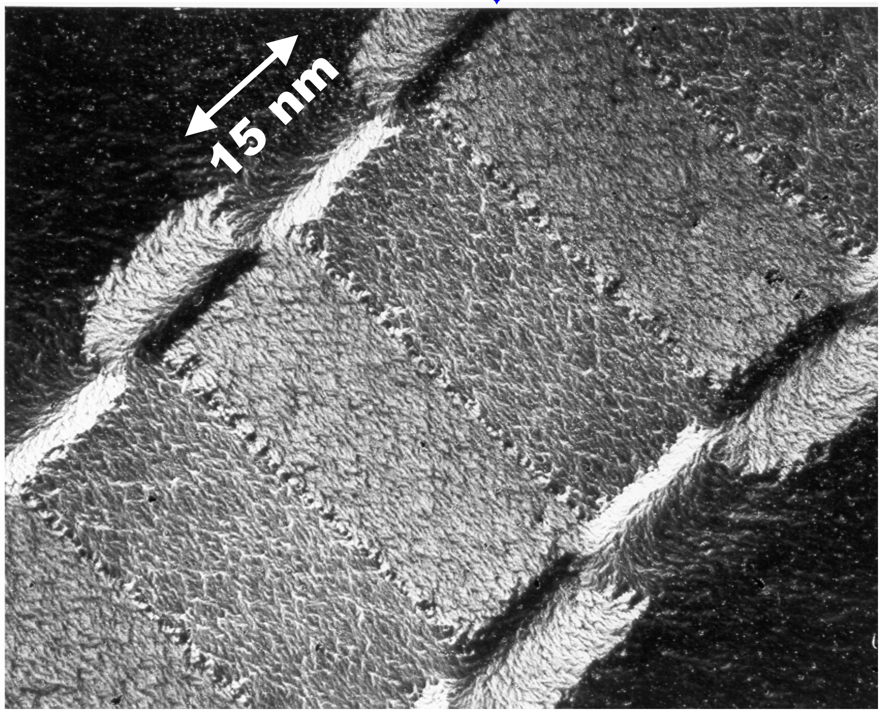
Chip-Storage oder magnetischer Speicher ?

Chip :



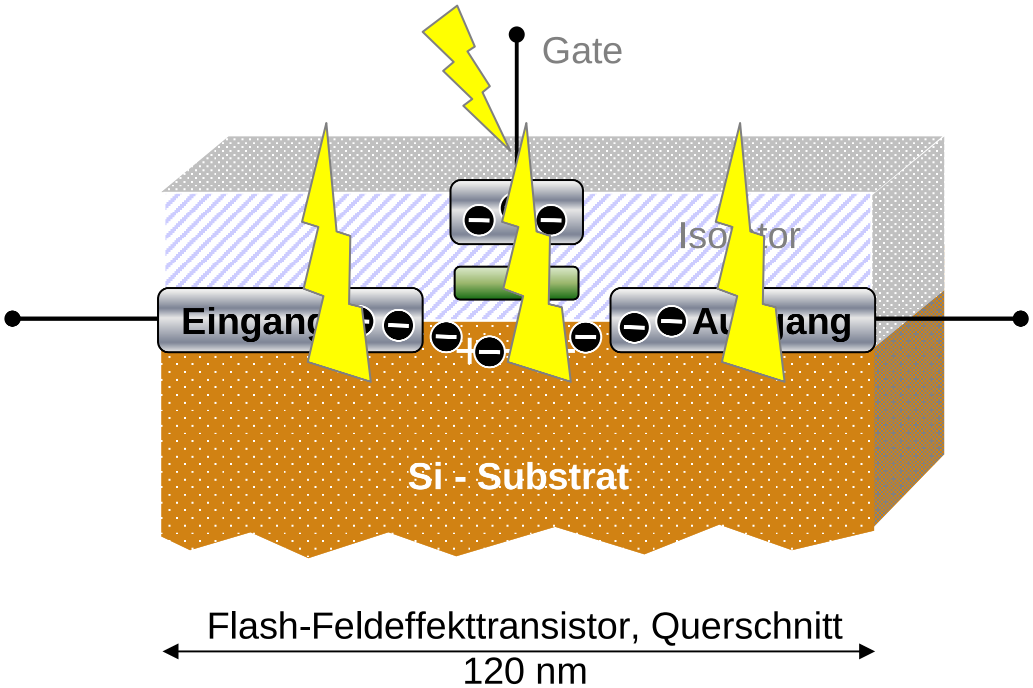


Disk :

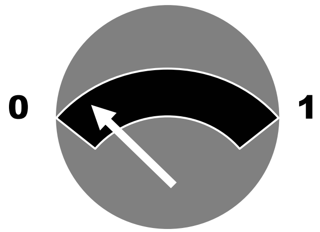


Disk versus(?) Flash

Funktionsweise von Flash-Speicher *(EEPROM)*



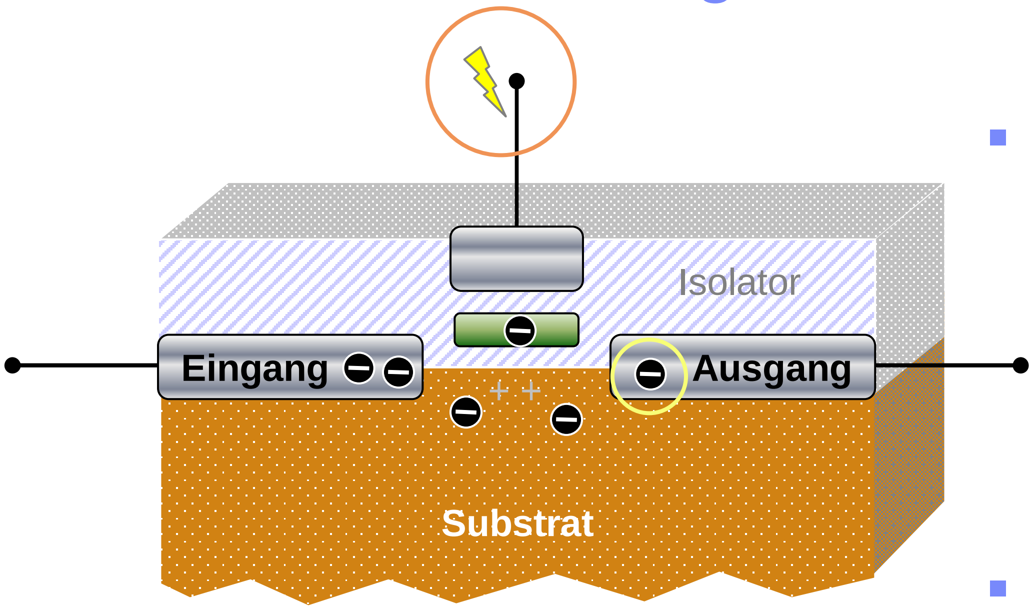
- **Floating Gate** (isoliert)



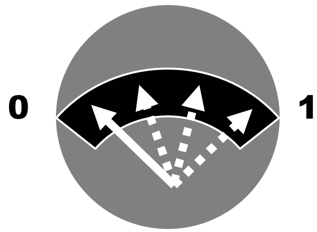
- Es können nur EINSEN* geschrieben werden
- Individuelles Löschen nicht möglich, nur Blocklöschung
- Löschen = Alterung

* Elektrisch korrekt = es können nur NULLEN geschrieben werden

Multilevel Flash Storage MLC

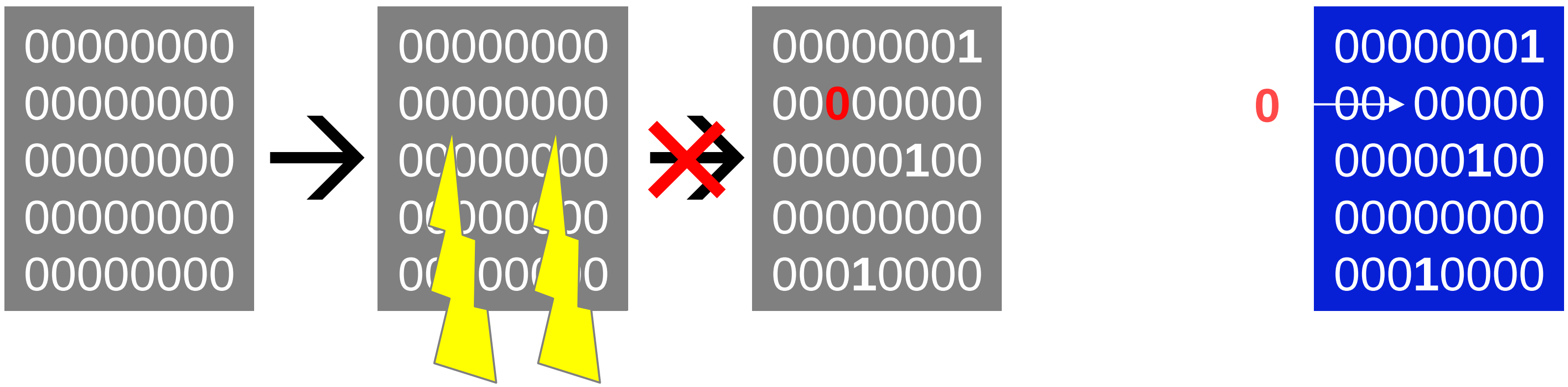


- Reduzierter Schreibimplus



- 2 Bits je Zelle durch vier unterscheidbare Zustände
- 3…4 Bit je Zelle in Planung
- Sparsam in Wafer cm², aber langsamer & unzuverlässiger ("USB-Stick Technologie")

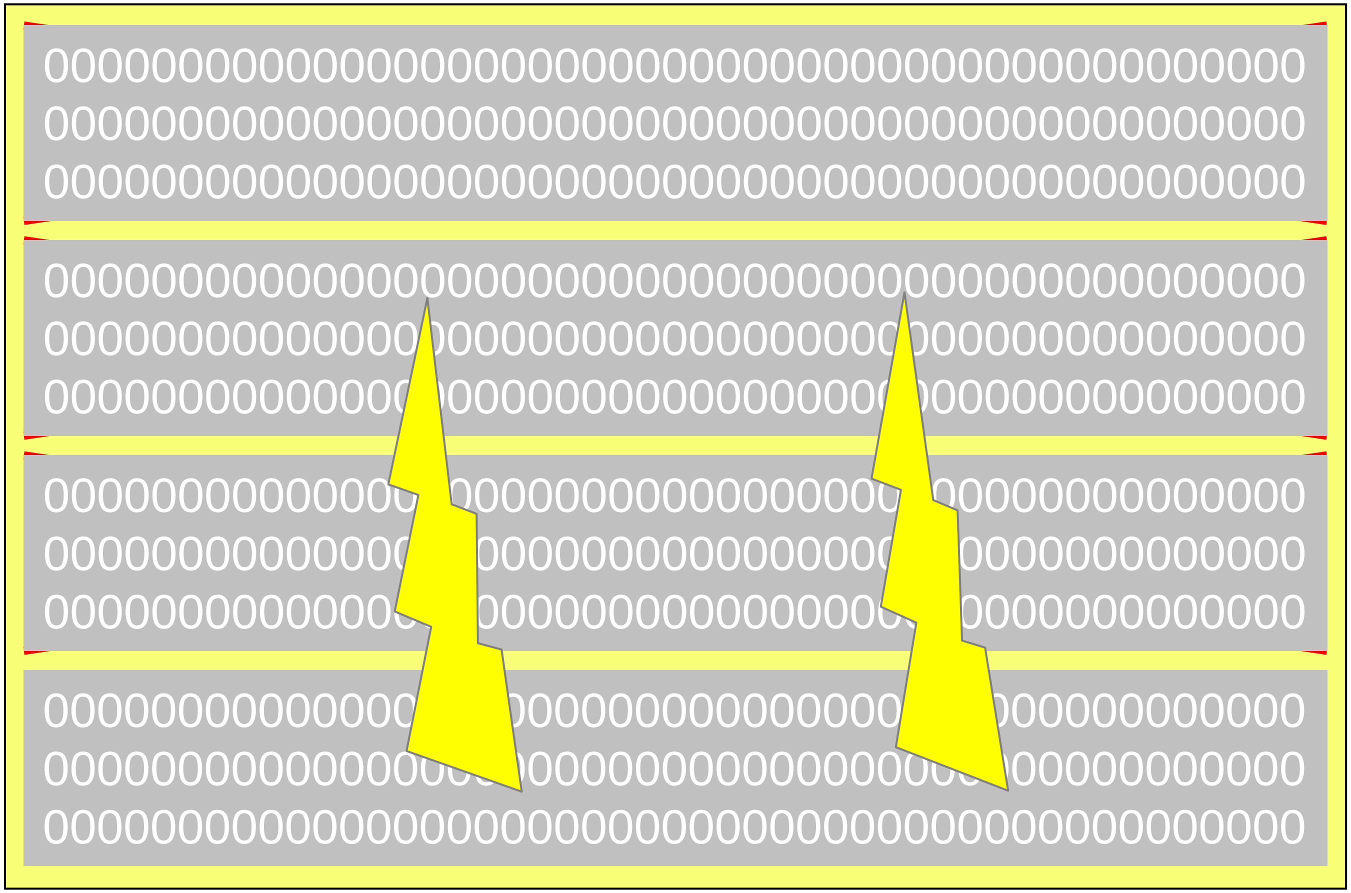
Beschreiben & Löschen von NAND-Flash



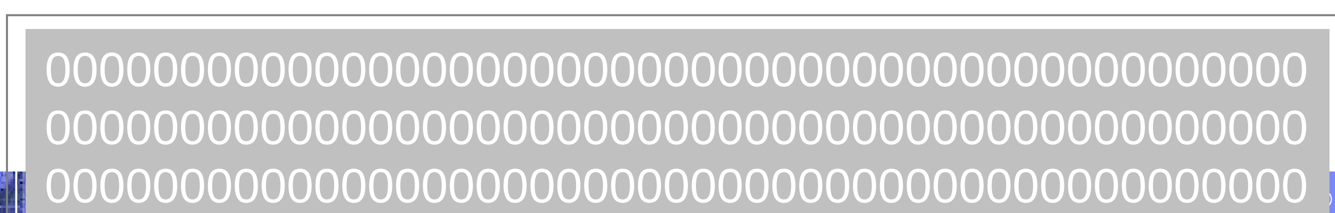
- Für **jede Löschoperation** ist eine Blockkopie notwendig
- **Random Write** ist nicht der optimale Workload für NAND Flash

- Das Verlagern oft genutzter Blöcke verringert vorschnelle Alterung

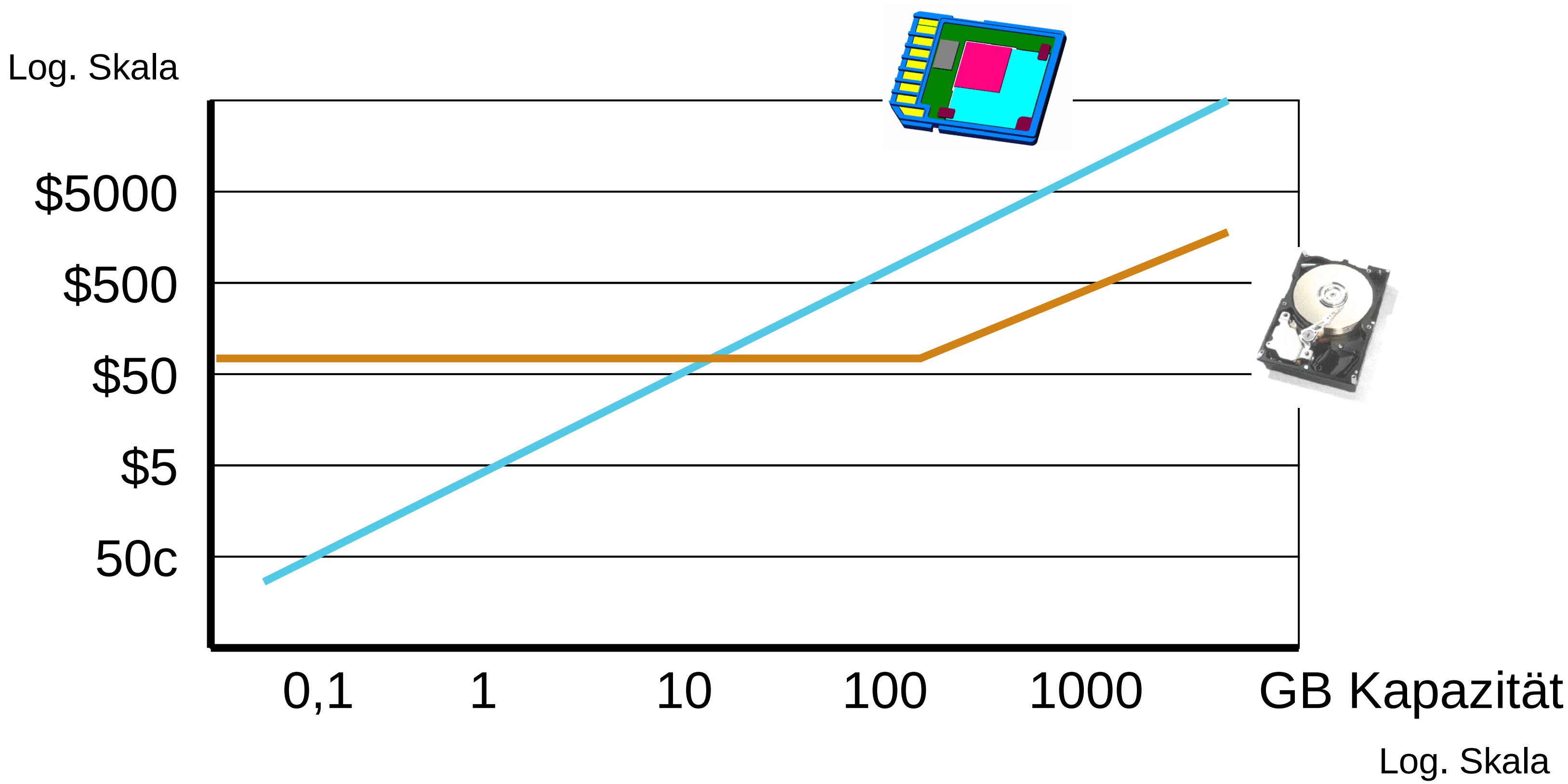
Block-Overprovisioning in Solid State Drives (SSDs)



Hier: 1x löschen für 4x Random Write – Mehr Leistung auf Kosten der Kapazität



Preisentwicklung Flash versus Disk nach Kapazität



Nicht maßstabsgetreu; 2008 Schätzwerte

Kosteneffektive Random IO Leistung

1. **Flash PCI-Memory** selber verwalten
2. Random IOs **serialisieren**, selten löschen
3. IO Muster auf Systemebene optimieren, nicht erst auf Drive-Ebene

Quicksilver

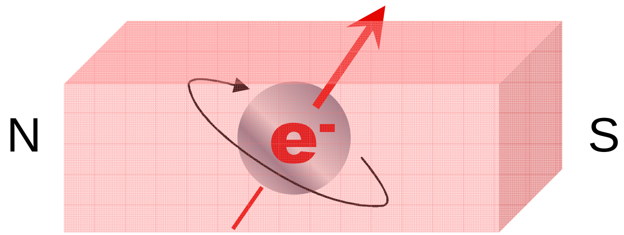
ADTRON® SSD 2,5" 160GB
$80-$115 *je Gigabyte*

PCI-based Flash Memory
$30 *je Gigabyte*

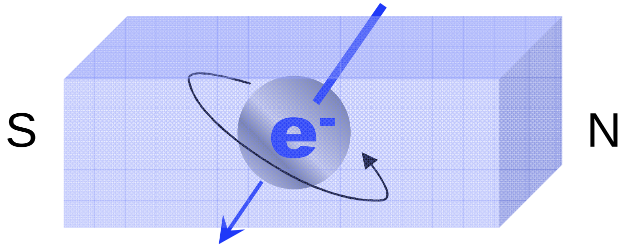
Chip-Technologie im Labor

Image © DOCB

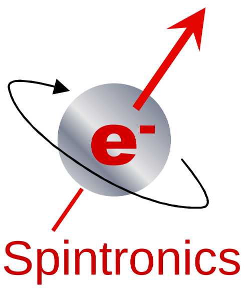
Eine neue Klasse von Festkörperspeichern

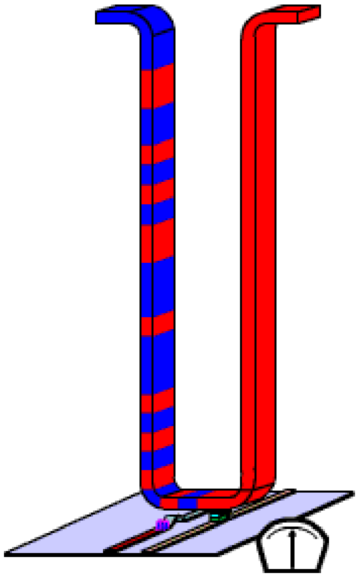


Spintronics

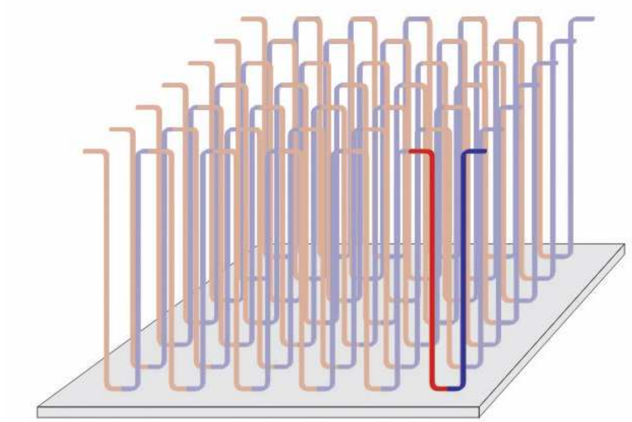


Speichern in Elektronenspins: "RaceTrack"

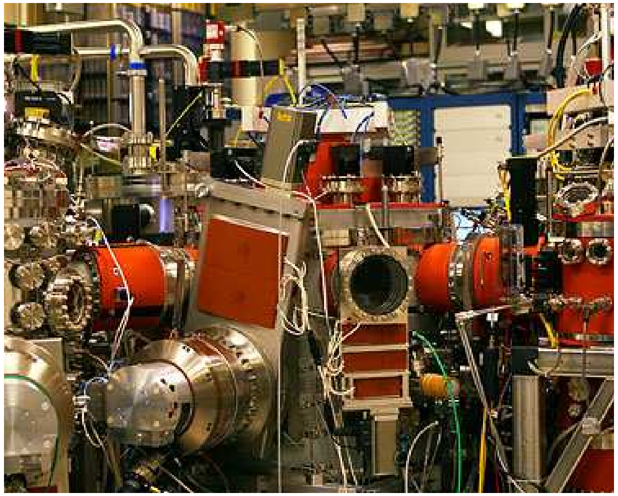




"Großer" Schreib/Lese-Kopf, "kleine" Spin-Datenfelder auf ferroelektrischem Nanodraht

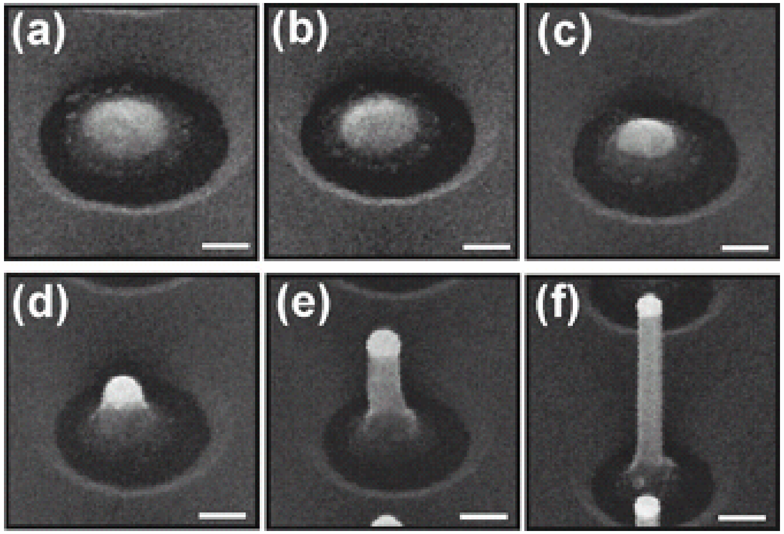


RaceTrack Storage Array: Hohe Datendichte in 3D



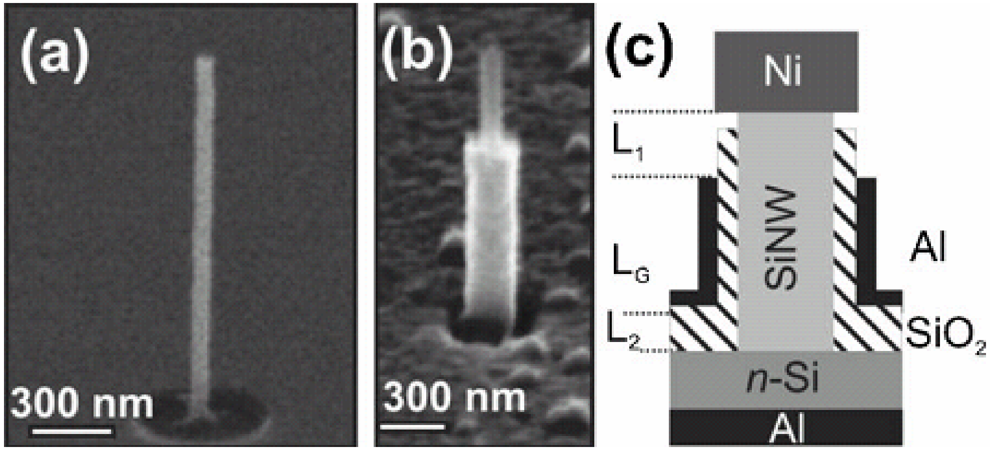
IBM Fellow Stuart Parkin, Erfinder der GMR Leseköpfe, erforscht "Racetrack Memory"

Die weltweit schnellsten & sparsamsten Transistoren



Wachstum von Nanosäulen

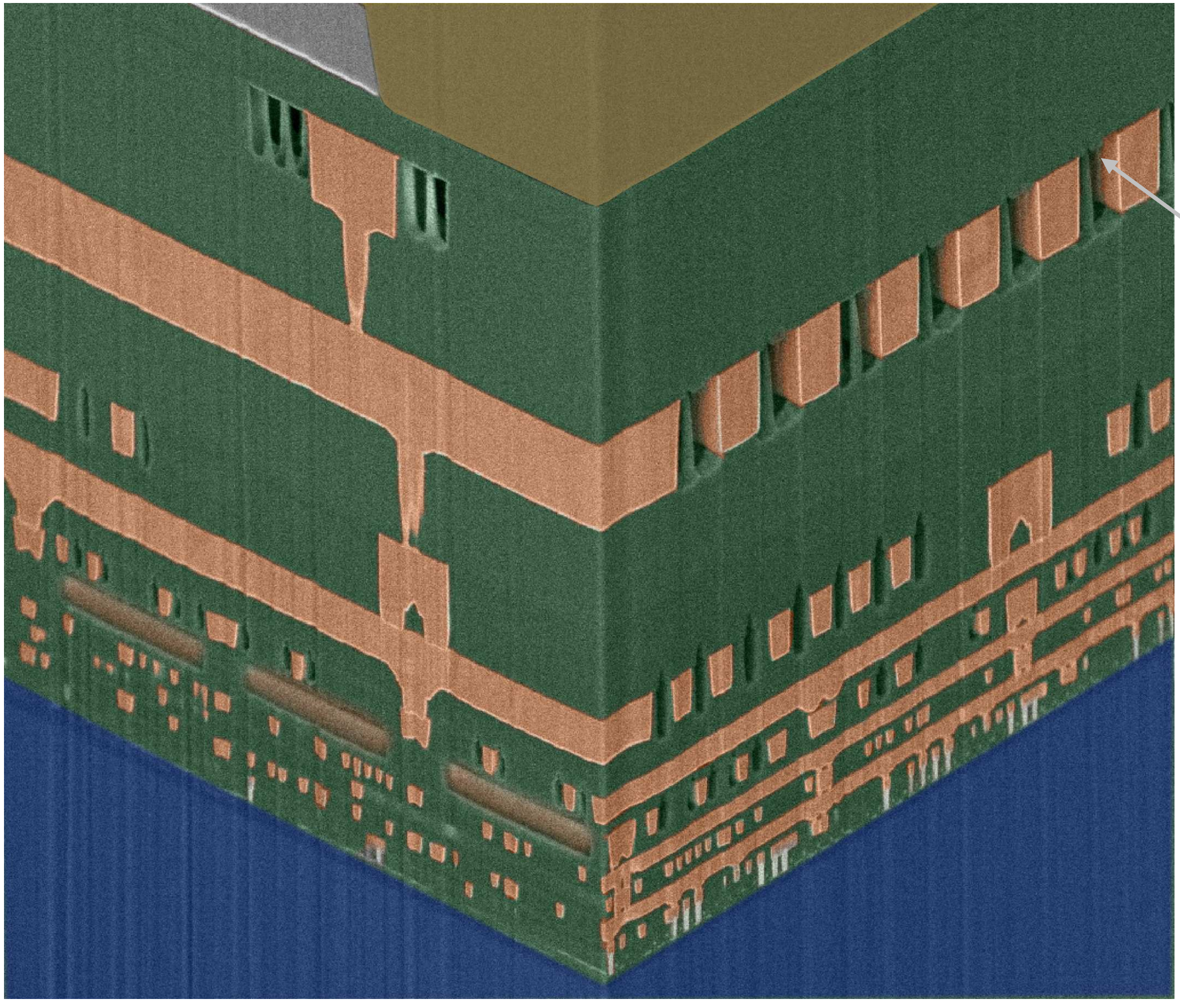
Schnell sperrender Transistor



Mit Mantel-Gate versehene Nanosäule

- Minimaler Verbrauch
- Höchste Schaltgeschwindigkeit
- TeraHertz nicht ausgeschlossen

Kompakter, schneller, störungsfreier : Air-Gap Chip

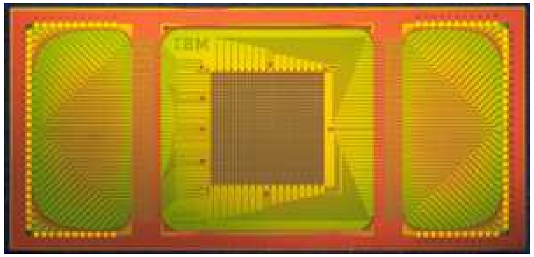


Trick !

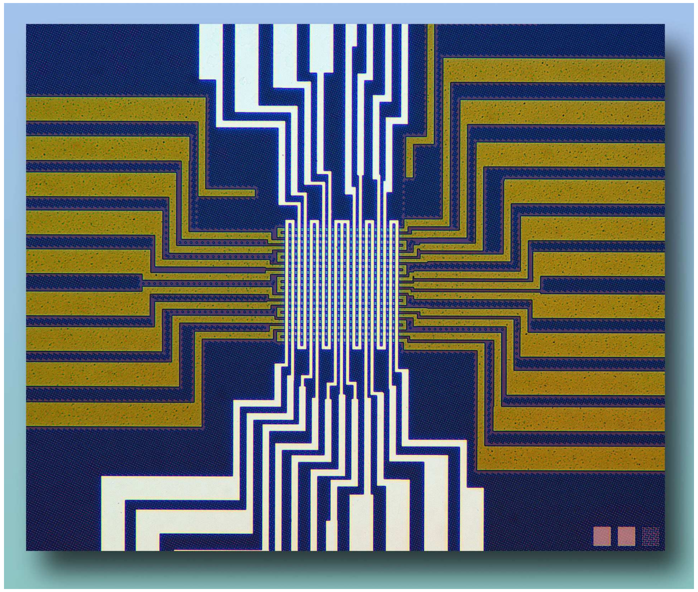
Wie trägt man Schichten über Löchern auf ?

Air-Gap Chip Herstellung

Solid-state in RAM Geschwindigkeit



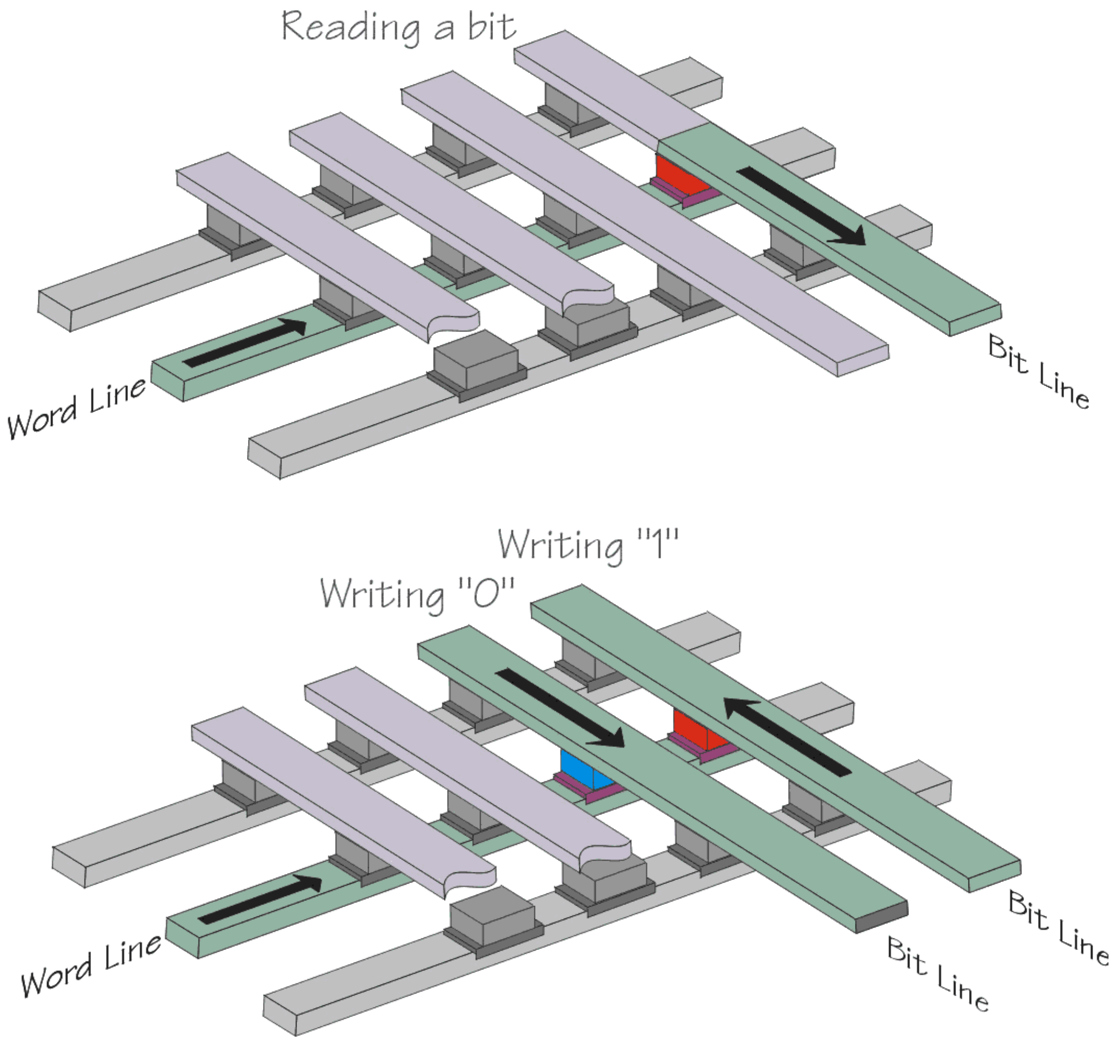
Magnetisches RAM



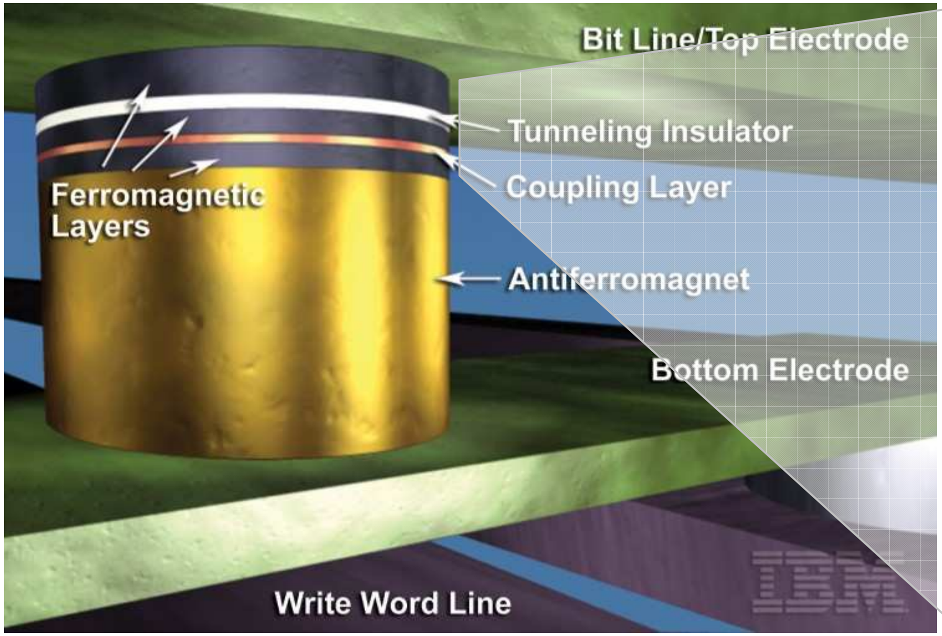
IBM Prototyp 199..

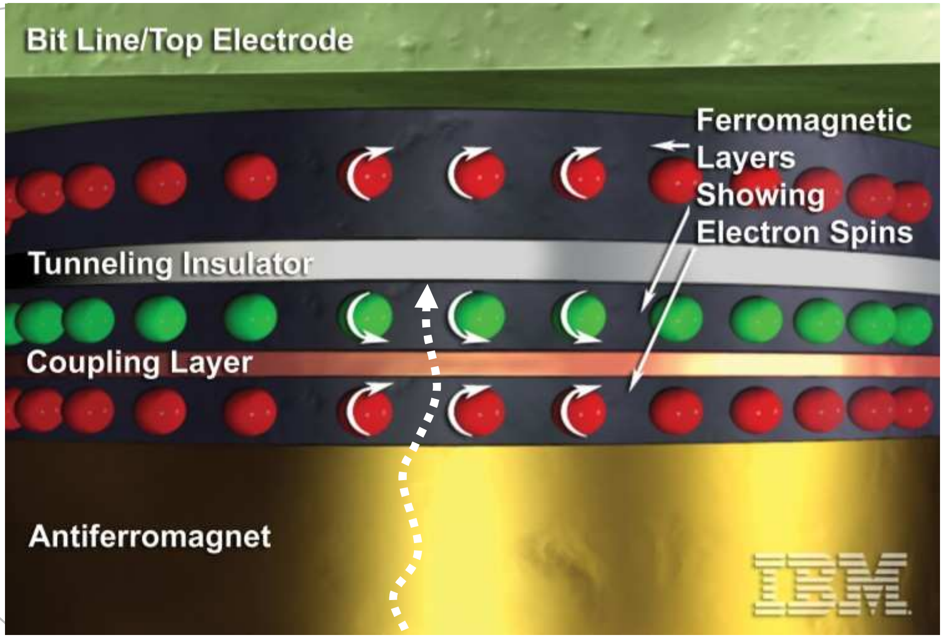
Hersteller: Freescale Inc.
"MR2A16A" 4Mb nichtflüchtig
@ 35 nsec Zugriffszeit.

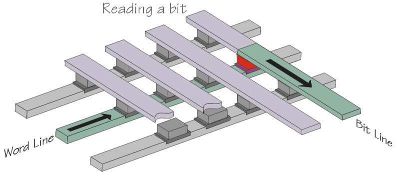
Neuer IBM Demonstrator:
2 nsec Zugriffszeit



Magnetisches RAM



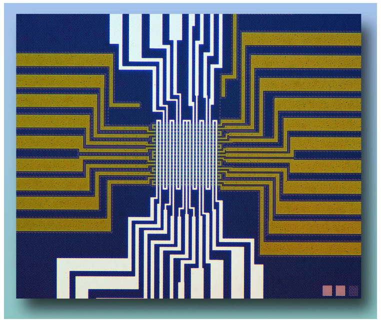


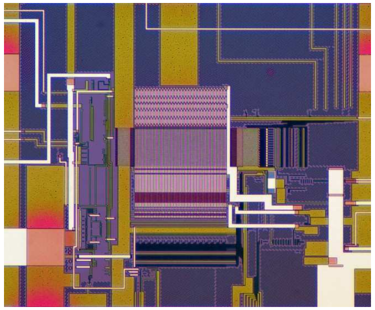


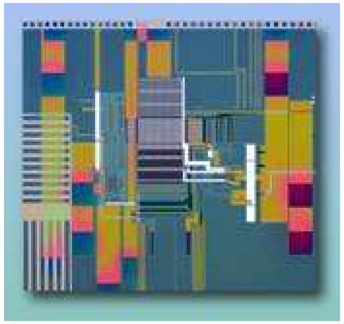
Neuer IBM Demonstrator:
2 nsec Zugriffszeit

(10 mal schneller als heutiges DRAM)

Nicht-flüchtiges RAM = IT Revolution !

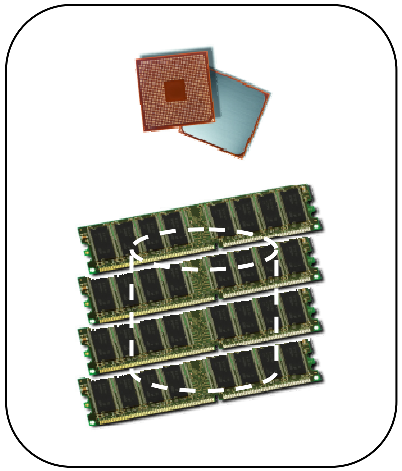


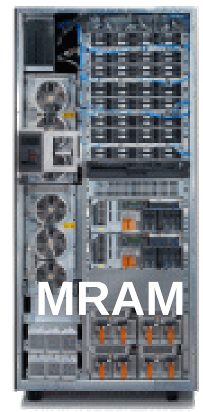




- Magnetic RAM / Ferroelectric RAM / Phase-Change RAM
- Flash Memory = Übergangstechnologie ("marktbereitend")

Wie verbaut man ***schnellen*** Solid-State Speicher?

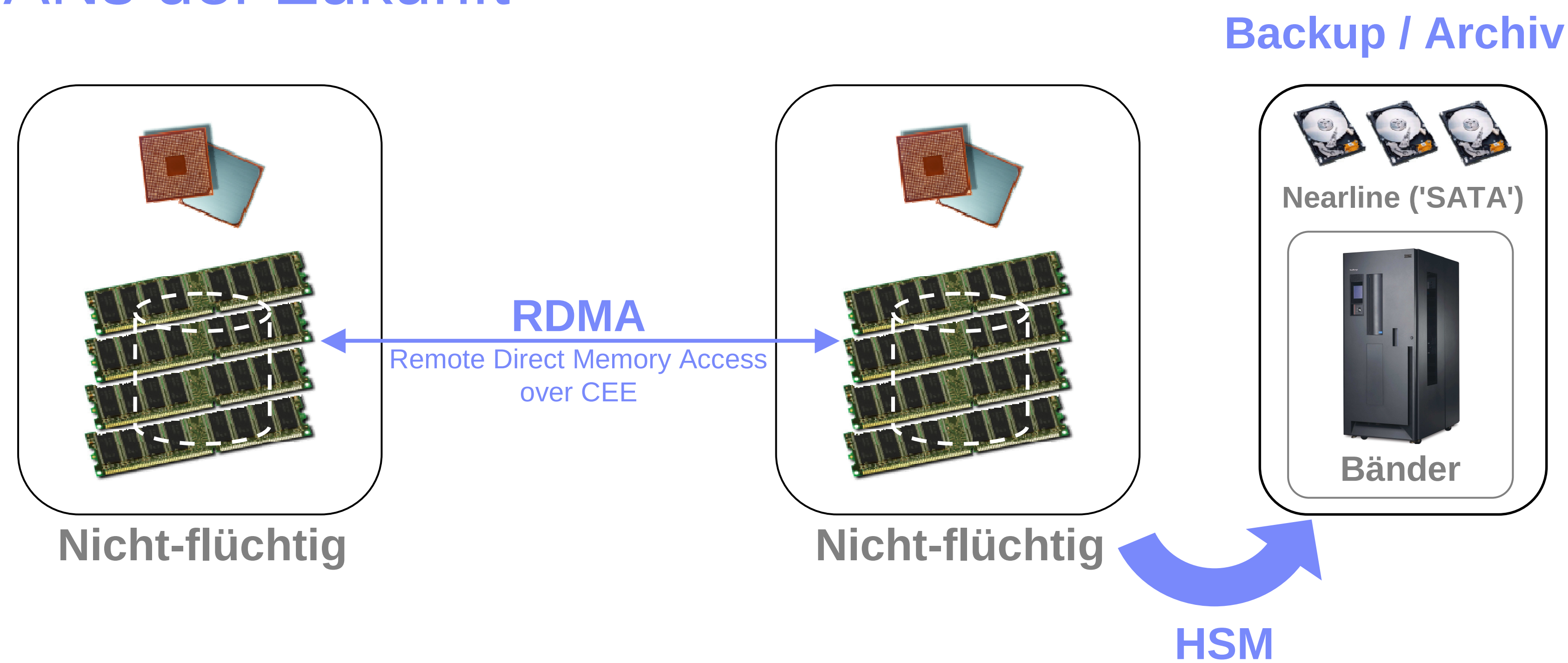




20 nsec **2 nsec**

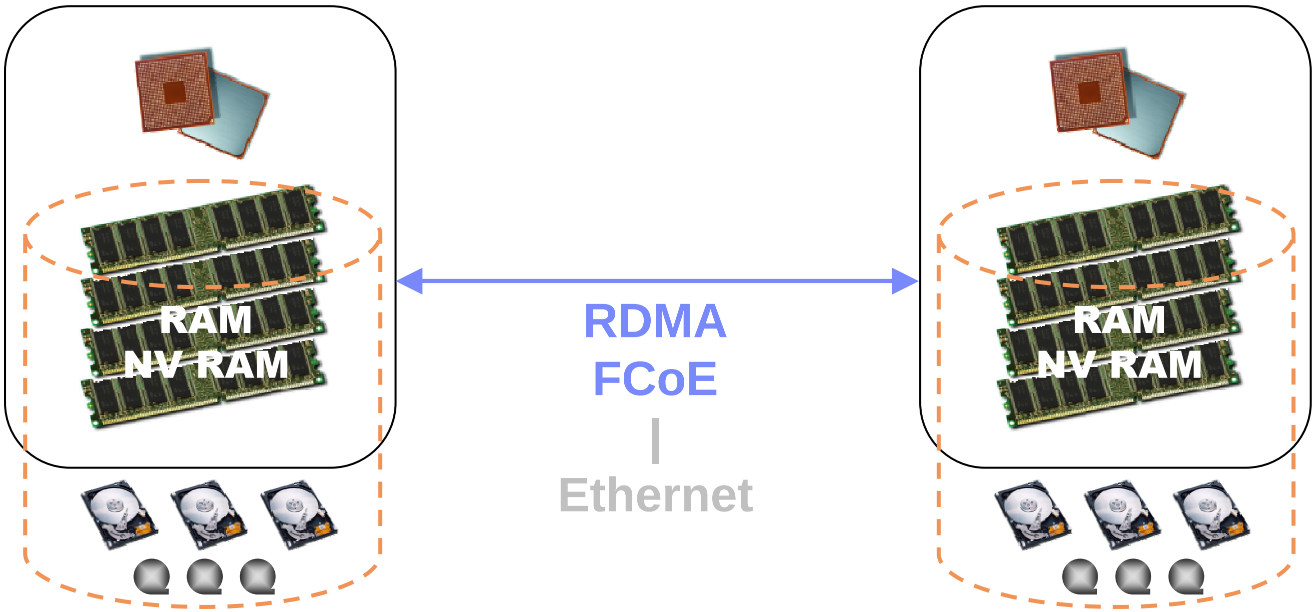
- Online Speicher zieht zum Prozessor
- Schnelle (Fibrechannel-) Platten verschwinden
- Monolithisches Design = hohe Verfügbarkeit

Die SANs der Zukunft



- Memory-to-memory SAN *(RDMA über Converged Enhanced Ethernet)*
- Platten → RAM
- Bänder → Platten
- Paging → HSM

Storage Konzepte der Zukunft

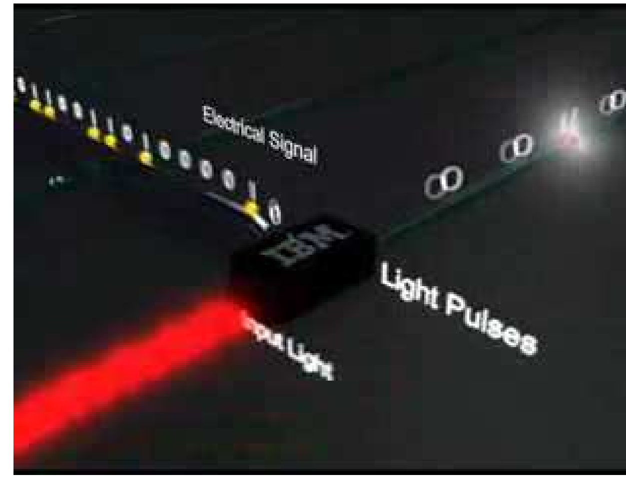


- FibreChannel over (converged enhanced) Ethernet FCoE
- NPIV and other modern concepts work transparently

Computerchips der Zukunft enthalten…

- Spintronics
- Nanostrukturen
- Möglicherweise Lichtspeicher

Kombination klassischer CMOS Technologie mit Indiumphosphid and Galliumarsenid für optische Teile. Maße: 3.25 × 5.25mm.



axel.koester@de.ibm.com

Disclaimer

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This information could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

The performance data contained herein was obtained in a controlled, isolated environment. Actual results that may be obtained in other operating environments may vary significantly. While IBM has reviewed each item for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customer experiences described herein are based upon information and opinions provided by the customer. The same results may not be obtained by every user.

Reference in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation on any non-IBM product, program or service.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR INFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (e.g. IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. IBM is not responsible for the performance or interoperability of any non-IBM products discussed herein.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.