

European Commission

IST programme

Future and Emerging Technologies

Microelectronics Advanced Research Initiative

MELARI NANO

**Technology
Roadmap for
Nanoelectronics**

Editors:

R. Compañó
L. Molenkamp
D.J. Paul

Contents

1. Introduction	4
1.1 Roadmaps	4
1.2 CMOS Technology	4
1.3 Present Semiconductor Technology.....	7
1.4 Emerging Technologies	7
1.5 Markets and Applications	9
2 Devices	11
2.1 Reference Point: MOSFETs.....	11
2.1.1 Theoretical limits.....	11
2.1.2 Technology limits.....	11
2.1.3 Economical limits.....	12
2.1.4 Major Challenges and Difficulties	13
2.2 Emerging Devices.....	14
2.2.1 Single Electron Tunnelling (SET) Devices	14
2.2.1.1 SET	14
2.2.1.2 Nano-flash device.....	16
2.2.1.3 Yano memory.....	16
2.2.1.4 Summary of Device Parameters.....	16
2.2.1.5 Major Challenges and Difficulties for SET.....	18
2.2.2 Resonant Tunnelling Diodes (RTDs).....	19
2.2.2.1 RTD	19
2.2.2.2 RTD Performance	20
2.2.2.3 Major Challenges and Difficulties for RTDs	21
2.2.3 Rapid Single Flux Quantum Logic	22
2.2.3.1 Principle of Operation	22
2.2.3.2 Technology, Critical Dimensions and Performance	22
2.2.3.3 Major Challenges and Difficulties for RSFQ	23
2.2.4 Intramolecular Nanoelectronics	24
2.2.4.1 Definition	24
2.2.4.2 Electric-field controlled molecular electronic switching devices	24
2.2.4.3 Alternative Approaches to Molecular Switching	25
2.2.4.4 Molecular Wires	26
2.2.4.5 Major Challenges and Difficulties for Intramolecular Electronics	27
2.2.5 Spin Devices	28
2.2.5.1 Operation	28
2.2.5.2 Performance.....	28
2.2.5.3 Challenges and Difficulties	29
3 Nanofabrication	31
3.1 Lithography for CMOS technology	33
3.1.1 Optical lithography	33
3.1.2 Extreme Ultraviolet lithography	33
3.1.3 X-ray proximity lithography.....	34
3.1.4 E-beam projection lithography and the SCALPEL system	34

3.1.5 Ion beam projection	36
3.2. <i>Emerging nanofabrication methods</i>	37
3.2.1. <i>Top Down Approaches</i>	37
3.2.1.1 <i>Electron-Beam Nanolithography</i>	37
3.2.1.2 <i>Scanning Probe Methods (SPM)</i>	39
3.2.1.3. <i>Nanoimprint</i>	39
3.2.2. <i>Bottom Up Approaches</i>	40
3.3 <i>Comparison of Fabrication Techniques</i>	41
4. Circuits and Systems	43
4.1 <i>Design Strategies: Interconnect Problems and Design Complexity</i>	43
4.1.1 <i>Impact of Increasing Clocking Frequencies on Nanoscale Circuits</i>	45
4.1.2 <i>Local Architectures and Synchronisation</i>	45
4.2 <i>Novel Logic Circuitry</i>	46
4.2.1 <i>Resonant Tunnelling Device Circuits</i>	46
4.2.2. <i>QCA Circuits</i>	49
4.3 <i>Current Trends in Novel System Architectures</i>	52
4.3.1 <i>Starting Point: Systems on chip and Innovations in Microprocessor Designs</i>	52
4.3.2 <i>Parallel Processing</i>	53
4.3.2.1 <i>Propagated Instruction Processor</i>	53
4.3.2.2 <i>Reconfigurable Hardware</i>	54
4.3.2.3 <i>Defect and Fault Tolerant Architectures</i>	54
4.3.2.4 <i>DNA Computing</i>	55
4.3.3 <i>Artificial Neural Networks</i>	56
4.3.4 <i>Quantum Information Processing (QIP)</i>	57
4.4 <i>Comparison of the Various System Architectures</i>	57
5. Conclusions	60
5.1 <i>Introduction</i>	60
5.2 <i>Devices</i>	60
5.3 <i>Fabrication</i>	61
5.4 <i>Architectures</i>	62
5.5 <i>Outlook</i>	62
References	63
Annex I:- Comparison Between Technologies	69
I.I <i>Comparison of memory devices in 1999</i>	70
I.II <i>Predictions of memory devices in 2006</i>	71
I. III <i>Predictions of memory devices in 2012</i>	72
I. IV <i>Comparison of logic and high frequency devices in 1998 (in production)</i>	73
I.V <i>Comparison of logic and high frequency devices in 1998 (research)</i>	74
I. IV <i>Predictions of logic and high frequency devices in 2006 (research)</i>	75
I.VII <i>Predictions of logic and high frequency devices in 2012 (research)</i>	76
I.VIII <i>Comparison of circuit performance in 1998</i>	77
I.IX <i>Predictions of circuit performance in 2006</i>	78
I.X <i>Predictions of circuit performance in 2012</i>	79
Annex II:- MEL-ARI Projects and Contributors	80

1. Introduction

1.1 Roadmaps

A “Roadmap” is an extended look at the future of a chosen field of inquiry composed from the collective knowledge of researchers in that field. The composition of a roadmap can encompass trends in the area, links and comparisons between different fields, identification of discontinuities or knowledge voids, highlight potential major show-stoppers along with giving many other compactors. History has shown, however, that the major breakthroughs in many fields have seldom been predicted beforehand. The roadmap should therefore be considered as a helping guide to thought, breakthrough and comparison in research rather than a strict document for the research field.

This document has been compiled with the input of the MEL-ARI projects, which consist in total of more than 50 distinct European research groups working in different areas of nanoelectronics. The document has been edited by Laurens Molenkamp, Douglas Paul and Ramón Compañó, who have compiled the different information sources and extracted the information for the roadmap. In addition, Huguette Launois has reviewed the chapter on nanofabrication and Karl Goser and Christian Pacha have compiled the chapter on architectures. While the writing of the state of the art is a common exercise, the analysis of its future is not and the perspectives are even more unpredictable in this non-mature field. Due to the nature of this document, this roadmap cannot reflect all views, but represents an attempt to find a consensus between most of them.

Since the beginning of the seventies, the microelectronics industry has followed Moore’s law, doubling processing power every 18 months. This performance increase has been obtained mainly by decreasing the size of circuit features obtained by optimization and improvement of existing technology. In this light the semiconductor industry through the Semiconductor Industry Association (SIA) bases a Microelectronics Roadmap on the future of CMOS technology. The minimum feature size is fast approaching 100 nm in the next decade with switching charges containing 1000 or less electrons. Physical limits (quantum effects and non-deterministic behaviour of small currents) and technological limits (such as power dissipation, design complexity and tunnelling currents) may hinder the further progress of microelectronics on the basis of conventional circuit scaling. Technological problems, together with strongly increasing investment costs for conventional CMOS technology, may reduce the entry barriers for alternative device concepts.

It is the purpose of this roadmap to investigate non-CMOS technology for nanoelectronics. It is helpful, however, as a reference point and guide to provide a brief review of the SIA Roadmap.

1.2 CMOS Technology

The 1997 SIA Roadmap provides data predominantly on the scaling of dynamic random access memory (DRAM), microprocessing units (MPU) and applications specific integrated circuits (ASICs) manufacture. The major research and development inside the main semiconductor manufacturers is predominantly in line with the SIA roadmap predictions. The SIA roadmap predicts a steady reduction in the feature size (figure 1.1) combined with a steady rise in density (figure 1.2) until the year 2012 (table 1.1). There are still a number of technological challenges required to be solved if CMOS is to reach the 35 nm and 10^8 transistors per cm^2 in 2012. It should be noted that important technological challenges have been identified in the SIA roadmap while no solutions have been presented or outlined. This lack of ideas for solutions may offer a window of opportunity for alternative devices.

Assuming that semiconductors will hold track with the SIA predictions, for a new technology to compete with CMOS, it must be able to have by 2012:

Year of first DRAM shipment	1997	1999	2001	2003	2006	2009	2012
Bits / chip - DRAM	256M	1G	-	4G	16G	64G	256G
Transistors / chip - MPU	11M	21M	40M	76M	200M	520M	1.4G
Minimum feature size DRAM (nm)	250	180	150	130	100	70	50
Minimum feature size MPU (nm)	200	140	120	100	70	50	35
Local clock (GHz)	0.75	1.25	1.5	2.1	3.5	6.0	10.0
Across chip clock (GHz)	0.75	1.2	1.4	1.6	2.0	2.5	3.0
Wafer size (mm)	200	300	300	300	300	450	450
No. levels of interconnect	6	6-7	7	7	7-8	8-9	9
Chip size - DRAM (mm²)	280	400	445	560	790	1120	1580
Chip size - MPU (mm²)	300	340	385	430	520	620	750
Power dissipation (W)	70	90	110	130	160	170	175
Power dissipation - hand-held / portable (W)	1.2	1.4	1.7	2	2.4	2.8	3.2
Cost / bit packaged DRAM (µcent)	120	60	30	15	5.3	1.9	0.66
"Affordable" cost/transistor MPU- packaged (µcent)	3000	1735	1000	580	255	110	50
Cost of fab	1.5B\$				>5B\$		

Table 1.1:- Selected data from the 1997 SIA Roadmap

- Feature sizes of order of 30 nm or smaller
- over 10^8 transistors per cm² for logic
- over 10^{10} bits per cm² for memory
- a cost of less than 50 µcent per transistor for logic
- a cost of less than 660 ncent per bit for memory
- operate at 10 GHz
- 1.4 billion transistors must use less than 170 W power in a MPU

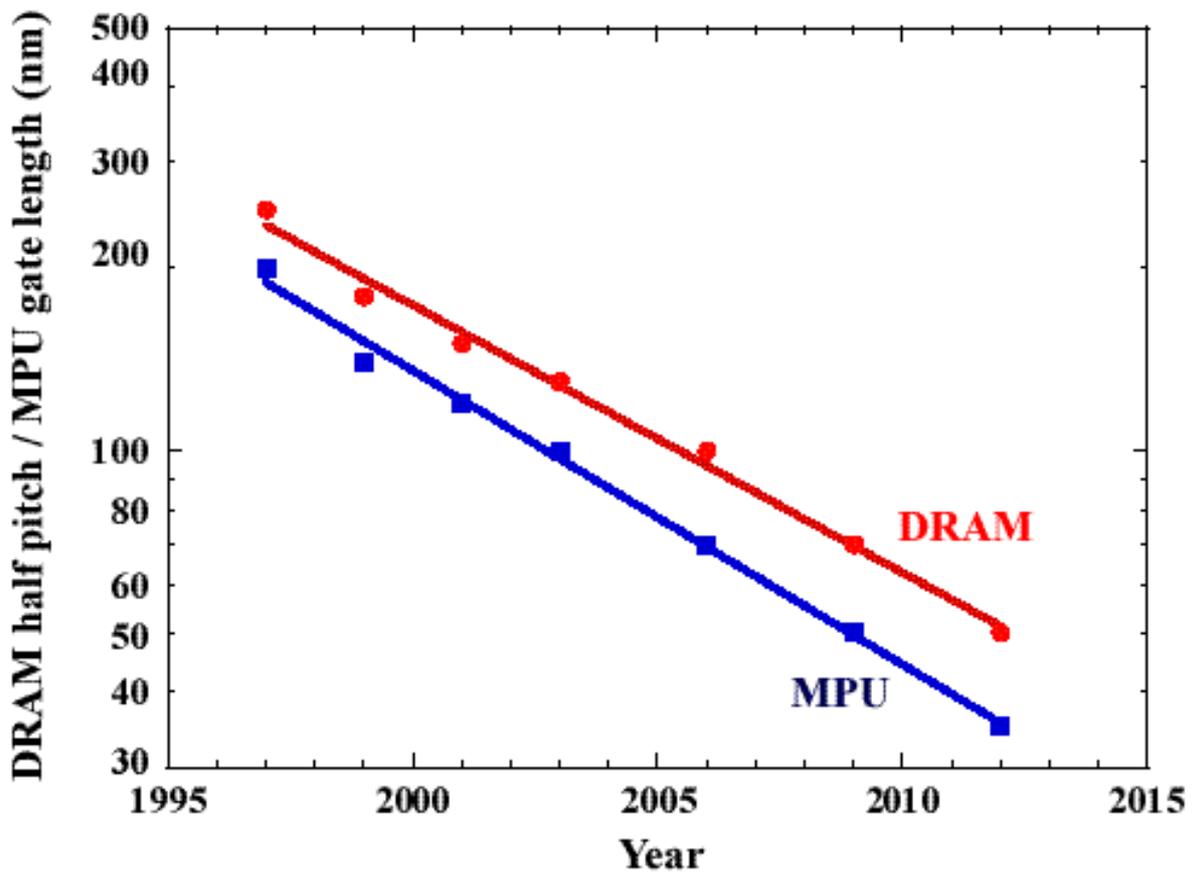


Figure 1.1:- The minimum feature size plotted against year from the 1997 SIA Roadmap.

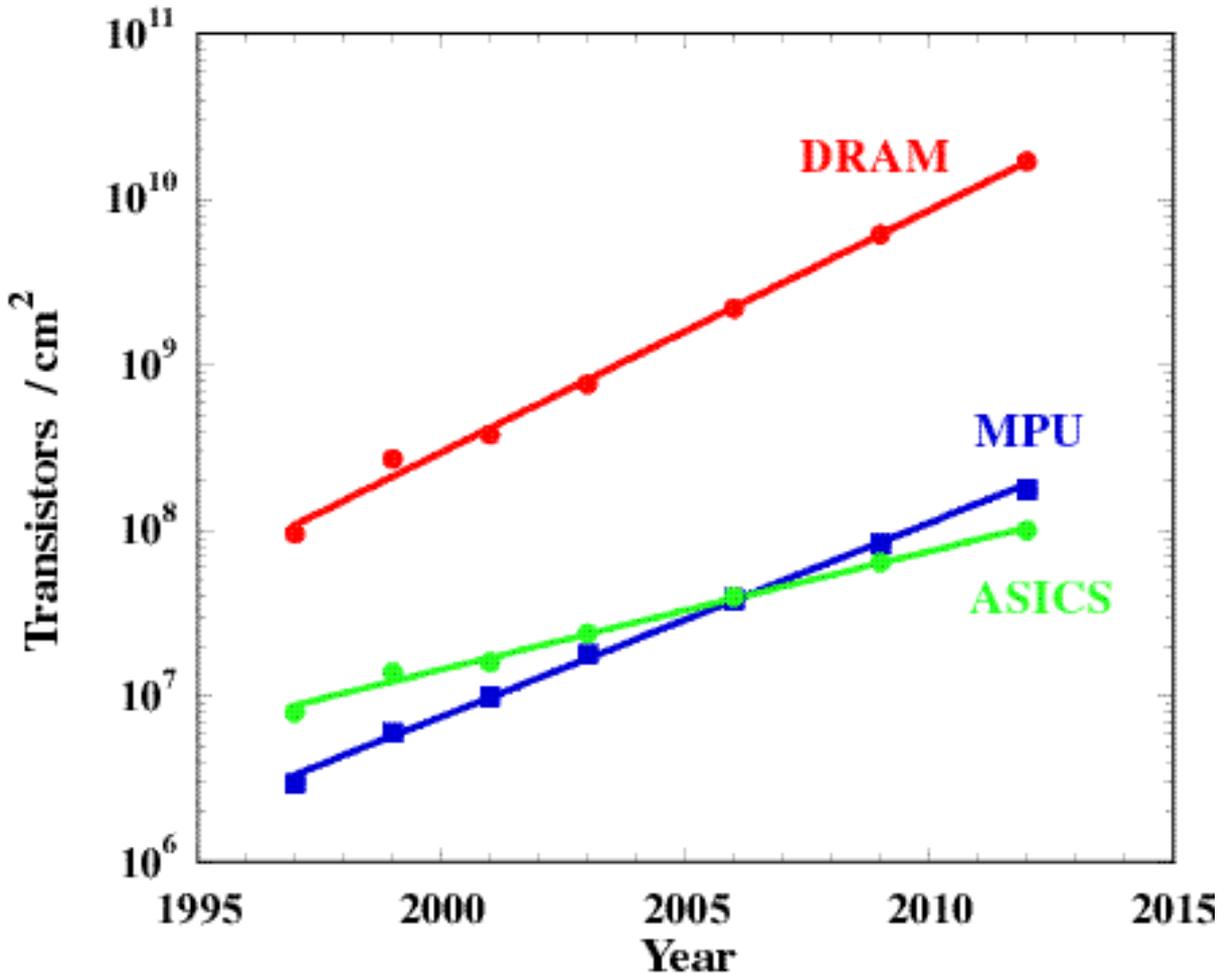


Figure 1.2:- The number of transistors per cm² plotted against the year of first manufacture.

1.3 Present Semiconductor Technology

From the above target figures it seems likely that CMOS will have a dominant market position even after 2012. The SIA roadmap, however, only deals with DRAM, MPU and ASICs and does not cover other aspects, such as III-V materials, telecommunications, analogue microelectronics, low-power portables, optoelectronics or other technological areas. As will be discussed later in this document, the requirements as suggested from the SIA roadmap may be relaxed if new architectures and functions may be integrated which allow similar functionality at reduced device numbers. For example, III-V materials are dominant in optoelectronics and at present in radio frequency, where CMOS cannot compete. This roadmap may therefore be perceived as looking at alternative routes to mainstream CMOS, that may become economically viable in some areas *of microelectronics in the longer term*.

The Moore's law decrease in size and increase in density may slow as 2012 is approached, but changes in system design would allow the systems performance to continue to increase. Perhaps the most important driver for new technologies is the ability to reduce the cost per function on a chip. There is therefore great scope for technologies which may be integrated on a CMOS chip and enhance the functionality of the CMOS chip (Figure 1.3). Already Si bipolar and Si_{1-x}G_x heterojunction bipolar transistors have been integrated with CMOS (BiCMOS). Numerous demonstrators of radio frequency circuits integrated with CMOS processing using the BiCMOS processes have been published [Cressler 1998]. The ability to integrate optoelectronic components may also be of significant benefit. As yet, Si optoelectronics is at a very immature stage compared to III-V materials [Soref 1997] and the poor Si emitter efficiencies may open some opportunities for III-V materials. Optical interconnects for both intra- and inter-chip communication are good examples and are discussed elsewhere [Esprit MEL-ARI Optoelectronics Roadmap].

There is a great preference in the semiconductor industry for the system-on-a-chip with as many different functional Si-based (and perhaps other) devices on the one Si chip (figure 1.3). Nanotechnologies which may be integrated onto a CMOS chip would be the preferred route for companies. Multi-chip modules (MCMs) may also play a large role because it may be cheaper to have 2 or 3 chip solutions than 1 chip [Wieder 1998]. Intel has already demonstrated the problems of manufacturing both logic and memory in one process with yields of Pentium processors dropping because the optimised production process of one is not fully compatible with the other and vice versa. Another problem is crosstalk specially when digital and analogue components are placed on the same chip. Moreover, it is also easier to test smaller units, hence 2 or 3 chip solutions in MCMs may be preferable to complete systems-on-a-chip. The application will determine the final chip count. This opens up possibilities of some nano-technologies being produced on chips to be integrated in MCMs with CMOS and other technologies. Again, if an efficient Si emitter cannot be found then a III-V chip for optoelectronics may have to be used in the MCM for optical communications.

1.4 Emerging Technologies

Theoretical modelling has suggested that conventional single-gate MOSFETs may operate down to channel lengths of 30 nm [Taur 1997]. Using double-gate MOSFETs, channel lengths below 10 nm may be achieved. For Terabit/chip technology radical alternatives to CMOS are needed and should be available by 2012 to avoid a dislocation in progress. In order to have maximum impact it is important that devices can operate at room temperature, in all but research environments. Single electron tunnelling (SET) devices are perceived by some to be the natural successor to the MOSFET although the research is also perceived by some companies as a method to investigate the ultimate limits of the MOSFET as single electron effects will limit normal transistor action at the smallest length scales. SETs are reviewed in section 2.2.1. Functional Si devices are still at a

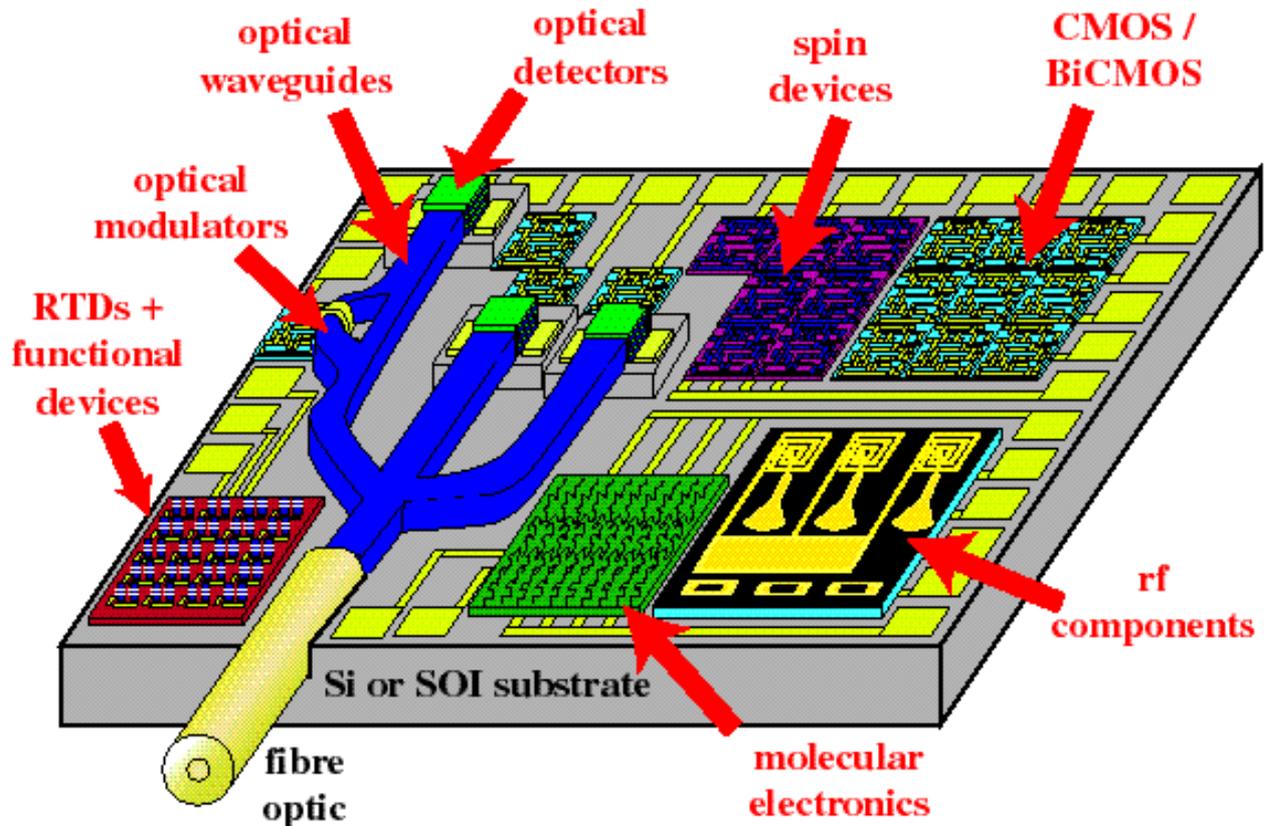


Figure 1.3:- The system-on-a-chip of the future? The ability to integrate new technologies with CMOS or BiCMOS is important but may be too expensive for some technologies or suffer technological problems such as crosstalk like rf solutions. MCMs may therefore be more cost effective for some technologies.

relatively immature stage but structures such as the charge injection transistor (CHINT) allow complete integration with the present SiGe BiCMOS layer structure and processing [Kasper 1984] while reducing transistor counts.

Section 2.2.2 reviews one functional device family, resonant tunnelling diodes (RTDs) where substantial optimism about devices has arisen. They allow reduced transistor counts and either high speed or low power, and are presently being demonstrated close to production levels by Raytheon (formerly TI). If such functional devices can be integrated into a CMOS process and allow a reduction in cost per function then they will be used on the same chip as CMOS.

Section 2.2.3 reviews rapid single flux quanta (RSFQ) devices which offer extremely high circuit speeds but require cryogenic temperatures for operation and therefore may be limited to the highest performance computing. Intramolecular electronics are reviewed in section 2.2.4. These devices may be fabricated by chemical or biological reactions. If appropriate molecules and reactions can be found then the ultimate goal would be to self-assemble whole circuits using reactions. Such an approach is potentially very inexpensive and circumvents a number of the lithographic and fabrication problems perceived below 100 nm in CMOS. The technology is presently at a very immature stage. Finally, section 2.2.5 reviews spin devices. Numerous types of spin devices have been proposed mainly for nonvolatile memory applications.

Consumer	Military	Space	Attribute	SRAM	DRAM	RTD	SET	MRAM
✓	✓	✓	Low power		✓	✓	✓	✓
✓	✓	✓	Non-volatile					✓
✓	✓	✓	Fast write	✓	✓	✓		
	✓	✓	Radiation hard	✓				✓
		✓	Unlimited R/W	✓	✓	✓	✓	✓
✓	✓	✓	High density	✓	✓	✓	✓	✓
✓	✓	✓	Highly reliable	✓	✓	✓		✓
✓	✓	✓	Random access	✓	✓	✓	✓	✓
		✓	Fast access	✓	✓	✓	✓	✓
✓	✓	✓	Fast read	✓	✓			✓
✓	✓	✓	Low write power	✓	✓			✓
		✓	Long term storage	✓	✓			✓
✓	✓	✓	Noise					
✓	✓	✓	Testing	✓	✓			✓
	✓		High power					

Table 1.2:- Comparison of memory applications and technologies

Consumer	Military	Space	Attribute	CMOS	SET	RTD	RSFQ
✓	✓	✓	Low power	✓	✓	✓	✓
	✓	✓	High power	✓			
✓	✓	✓	Radiation hard			✓	
	✓	✓	High density	✓	✓		✓
		✓	Reliability	✓			✓
✓	✓	✓	High speed	✓		✓	✓
✓	✓	✓	Noise	✓		✓	✓
✓	✓	✓	Testing	✓	✓	✓	

Table 1.3:- Comparison of logic applications and technologies.

1.5 Markets and Applications

In the past, the main driving forces for microelectronics research have been consumer electronics and military applications. The first is by far the larger market, while the latter has a strategic role and has often been the bridge to civil introduction. In addition to these two applications, others such as electronics for research purposes do play an inspirational role. Among these niche applications, space applications may become strategically important in the future especially as markets such as satellite communications increases.

Much of the progress in CMOS memory and logic has been driven by the need for ever faster personal and high performance computers as the complexity, power and volume of software and networking has increased. It is not clear how much longer this trend can continue although appearing technologies such as voice recognition may give new impulses. Rapidly developing consumer markets, which may take over include telecommunications, home entertainment (real-time video processing) and portables. Much of telecommunications requires high speed switching although many solutions are mixed with optics due to the fibre optic cable and hence III-V solutions may dominate. Portable products are driving to low power technology where battery performance is almost static and higher performance must be achieved in the microelectronics.

At present, mobile phones are one significant driver for Si (and SiGe, GaAs) technology. System-on-a-chip approaches would allow cheaper phones with longer battery lifetimes and higher data transfer rates. As both radio frequency and infrared technologies become cheaper, market share is climbing for both technologies. Such technology would also allow wireless computer networking where convenience over the clutter of cables around the office or home, especially with portable computers, suggests another market.

State of the art in data processors and memory, and particularly in multilevel chips and complex packaging, is limited by technology and driven by the computer industry rather than by telecommunications. But, as happened in optical technology, it may be that the telecommunications industry might take the driving seat to meet the need for high speed signal and data processing. Possible applications could be real-time video processing at home or fast routers. Ever increasing demands for high performance computing and possible new telecommunication applications give confidence that there will be a market for nano-scale integrated circuits. Other markets are also predicted to increase substantially such as portables, where low power solutions and nonvolatile RAM are of prime importance.

In a rough manner, tables 1.2 and 1.3 summarise the requirements for different applications that technologies must address if they are to be successful in a number of different markets. Without advancing the conclusions of this document, the alternative device concepts may find the following applications:

Single electron tunnelling (SET) devices are predominantly aimed at high-density, low-power memory markets especially since a number of the designs are a miniature version of flash memory technology. There are a number of designs for low power SET logic, but as yet none have been demonstrated at room temperature.

Resonant tunnelling diodes (RTDs) have demonstrated numerous applications and potential markets including digital to analogue converters (DACs), clock quantisers, shift registers and ultralow power SRAM. The RTDs may be designed for much higher speeds than CMOS for DACs, etc. typically in the speed range 10 to 100 GHz or for much lower power than CMOS such as the SRAM technology.

Rapid Single Flux Quantum (RSFQ) is a digital circuit technology, which offers high speed in the GHz regime, while producing low dissipation. As the superconducting effects upon which the principle is based works already with feature sizes in the micrometer regime, RSFQ has the potentiality to enter the market of applications, where Si-CMOS cannot achieve the same frequencies. One major application is in high-speed analogue to digital and digital to analogue conversion. Unfortunately, RSFQ as systems based on superconducting materials needs cooling, augmenting the overall costs of the whole system.

Numerous magnetic memories have been demonstrated which are aimed at the nonvolatile memory market such as that required in portable electronics. Because magnetism is an inherent property of the material in the memory, the devices have long retention times without requiring power to retain the memory states. At present, access speeds are slow compared to DRAM but predictions show that similar speeds to DRAM are possible and the technology may ultimately be scaled to smaller dimensions and higher densities than DRAM.

2 Devices

2.1 Reference Point: MOSFETs

2.1.1 Theoretical Limits

In order to understand the theoretical limits of information storage, it is useful to place CMOS in the context of a general purpose computation system. There are three important limits, which determine the ultimate performance of such systems. These limits are the thermal limit, the quantum limit, and the power dissipation limit (see figure 2.1). The energy necessary to write a bit determines the thermal limit. This energy must be bigger than the average energy of the thermal fluctuations, kT , otherwise bit errors will occur. For CMOS, the energy necessary to write a bit is about 10^{-13} J (10^6 eV) which corresponds to a temperature of 10^{10} K. The trend in CMOS is to decrease this energy and thereby decrease the power dissipated. The optimum value for the energy to write a bit for room temperature operation is about 4×10^{-19} J (2 eV), which is one hundred times greater than kT .

The energy needed to read or write a bit and the frequency of the circuits are limited by the uncertainty principle, $\Delta E \Delta t \geq h$. To prevent bit errors, the circuit cannot operate too close to the minimum uncertainty product. The quantum limit is then approximately given by $E/f = 100 h$ where E is the energy needed to write a bit and f is the clock frequency. CMOS circuits operate far above the quantum limit but as scaling continues below 100 nm, the limit will eventually be approached as E decreases and f increases (See Table 1).

The next important limit is the power dissipation limit, determined by the power density $p = EfnP$, where n is the device density, and P the probability that the device switches in a clock cycle, typically $P \sim 0.1$. The maximum tolerable power density is about 100 W/cm^2 . This means that the energy, the frequency, and the packing density are limited by $Efn \sim 100 \text{ W/cm}^2$. At high frequency, a high device density is desirable, and therefore a low energy per bit is desirable. CMOS circuits operate near the power dissipation limit. These three limits can be summarised in an energy - delay diagram as shown in figure 2.1.

2.1.2 Technology limits

Present results indicate that there is still room for MOS devices to follow the exponential scaling trends that have governed the microelectronics industry over the past decades. Already 100 nm groundrule technology and even 70 nm groundrule technology have both been demonstrated on research lines. While a few problems still exist in implementing the technology, essentially these results demonstrate that CMOS can and will be produced at such length scales on production lines. It is speculated that changes in architecture may gain up to three generations of devices. Another possibility would be to go ahead with fully three-dimensional integrated circuits. The major limitation, however, is likely to be power dissipation where 3D designs are not efficient at dissipating heat. It is obvious that, should industrial lithography, at the beginning of the next century, remain at a design rule of just below $0.1 \mu\text{m}$, this has immediate consequences for those quantum devices that directly depend on nm-scale lithography, most notably the SET devices.

Currently the most advanced MOSFET which has been realised is a double gate transistor and has an ultra short poly length (down to 30 nm) using a special electron-beam resist (calixarene film) and RIE with CF_4 gas. The extension regions consist of ultra shallow inversion layers created by a second gate. Very low off-state currents below $20 \text{ pA}/\mu\text{m}$ are realised at room temperature. The on-state currents however are limited by source/drain resistance effects. For the smallest poly

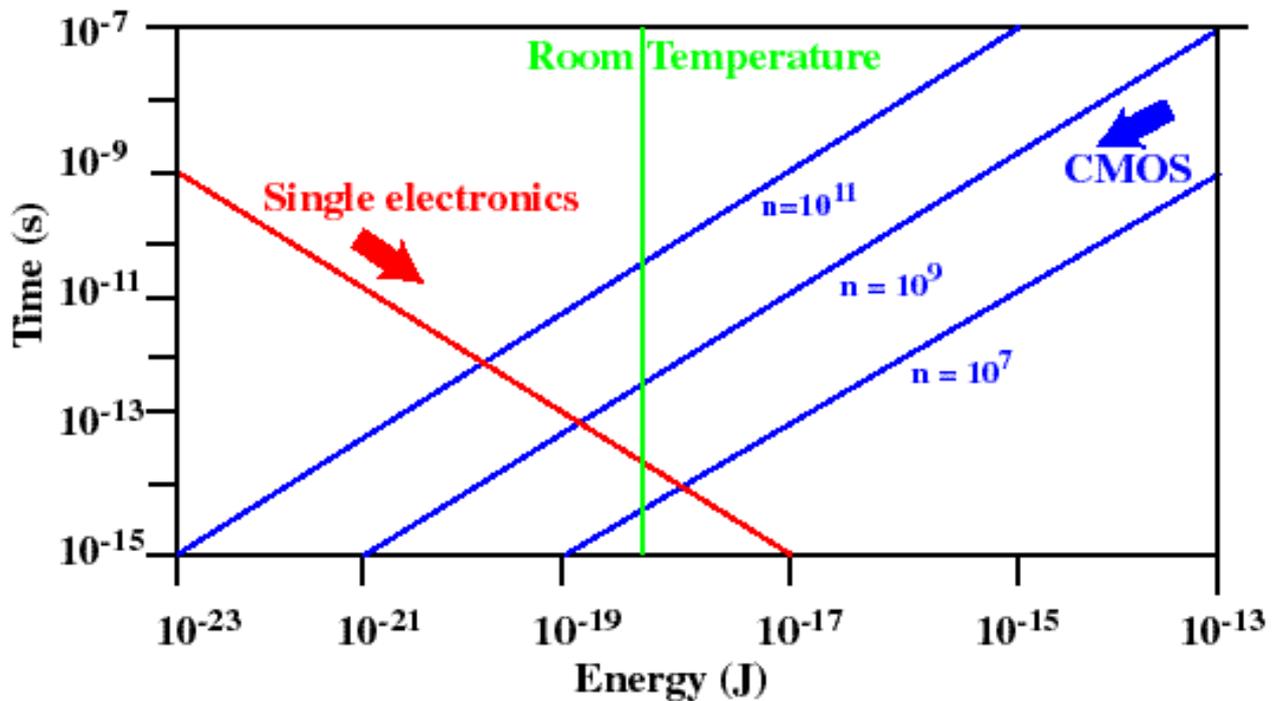


Figure 2.1:- An energy - delay diagram for single electronics. Room temperature operation can only be achieved to the right of the line labelled room temperature. The lower left-hand corner of the diagram is inaccessible due to quantum fluctuations and the lower right-hand corner of the diagram is inaccessible due to dissipation. The dissipation limit is represented by three lines, each corresponding to a different device density, n . The current trends in CMOS and single electronics are indicated in the diagram [Hadley 1996].

lengths the devices show tunnelling effects at 5 K [Kawaura 1998].

The eventual limitations for conventional single gate MOSFETs are expected to be for minimum feature sizes of about 30 nm on SOI substrates [Taur 1997], before degradation in device performance can no longer be compensated. The limit is dictated simultaneously by Zener breakdown of source / substrate junctions as well as by leakage across the gate oxide, due to the need to compress vertical dimensions in order to maintain good electrostatic control of the channel current. For MOSFETs below 30 nm, one must change the basic design of the MOSFET, either by using a back-gate (the dual gate MOSFET [Taur 1997]) or by using a second gate to create shallow (2DEG) ohmic contacts to reduce short channel effects [Kawaura 1998]. The dual gate may be scaled down to about 5 or 6 nm, while the gated-contact MOSFETs are predicted to achieve approximately 15 nm [Iwai 1998].

2.1.3 Economical Limits

CPU and ASIC manufacturers already state that CMOS production continuing on down Moore's law past 100 nm groundrules will be inevitable [Siemens, IBM, Lucent] but DRAM producers are more concerned that economics may limit future DRAM generations even before 100 nm is reached [Toshiba, Samsung]. The major companies are predicting that Moore's law will slow towards 2010, but the reduction in cost per function on the chip will continue at the same rate due to changes in systems design including self-test and error tolerant architectures along with increased integration levels leading to the system-on-a-chip. While cross-talk and the lack of a Si light emitter are present limitations to increased functionality and reduced cost per function on chips, eventually, radio frequency and optical functions (optical inter-chip / inter-system and on-chip interconnects) will all be integrated onto CMOS chips to reduce systems costs (Figure 1.3).

2.1.4 Major Challenges and Difficulties

The SIA roadmap produces comprehensive lists of the major challenges and difficulties for CMOS in the future. The following six points seem to be important in respect of the context of this document:

- Power management impacts at all levels.
- New architectures are required to overcome fatal bottlenecks in interconnects.
- New materials required for gate dielectric.
- Increased channel doping at short gatelengths will ultimately limit the drive currents and require fully depleted SOI or dual gate structures around 30 nm.
- Sub 100 nm lithography with throughput at costs which can sustain increased productivity.
- Semiconductor factory cost with escalating factory capitalisation and operational costs.

2.2 Emerging Devices

2.2.1 Single Electron Tunnelling (SET) Devices

2.2.1.1 SET

In order to avoid confusion over the concept of “single electron devices” we define the terminology used in this document:

- a SET device is a three terminal device based on the Coulomb blockade where the number of electrons on an island or dot is an integer number controlled by a gate. The dot may have up to thousands of electrons depending on the size and material.
- a Yano-type memory is a two terminal device where information is stored in deep traps in poly-Si.
- a Nano-flash memory is a three terminal device without a tunnel barrier between source and drain but with the addition of a floating gate in between the driven gate and the transistor channel. When fabricated at nanoscale dimensions, the increase of charge by one electron causes an abrupt shift in the turn off voltage.

In the following sections, mainly SET will be discussed. Nano-flash and Yano-type devices are essentially scaled down versions of conventional flash memories and therefore do not fully comply with the concept of nanoscale devices solely exploiting quantum effects and will therefore be discussed only superficially.

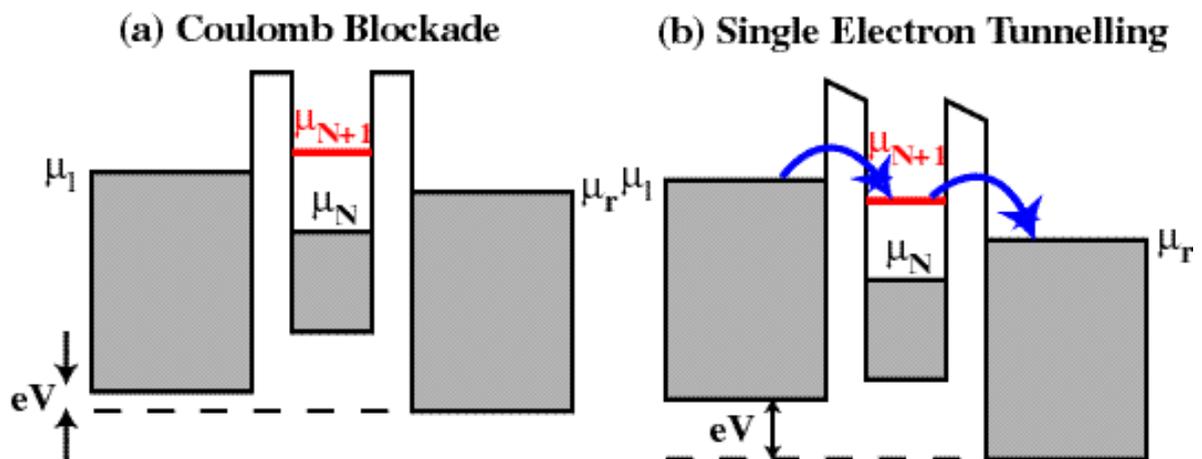


Figure 2.2:- For an island of total capacitance C with N electrons, being μ_N the chemical potential of the highest filled electron state, μ_{N+1} the chemical potential of the first available empty state for an electron and μ_l and μ_r the chemical potentials of the left and right electrodes respectively, it may be shown that the energy to add an electron to the island is $\mu_{N+1} - \mu = e^2/C$. Therefore provided $e^2/C \gg k_B T$ (the thermal energy - i.e. C is small) and the tunnelling resistance, $R_T \gg R_K \approx 25.8 k\Omega$ (i.e. the electron wavefunction may be localised on the island) for a voltage V applied across the electrodes, no electrons may flow if $\mu_{N+1} > \mu_l$ and μ_r - the state known as Coulomb blockade (a). If a larger bias is applied across the electrodes such that $\mu_l > \mu_{N+1} > \mu_r$, then empty states may be populated in the island and single electrons may tunnel through the island (b). A gate may be used to change the Fermi level of the island and therefore switch the single electron current on or off.

SETs may be produced in a number of different ways, the most common are metallic islands or semiconducting quantum dots. The basic operation (figure 2.2) requires an island of electrons with a capacitance C which is small enough that a charging energy for the island, (e^2/C) is much larger than the thermal fluctuations in the system, $(k_B T)$. Electrons may only flow through the circuit by tunnelling onto the first unoccupied energy level, μ_{N+1} . Therefore, electrons will only flow one by one if the bias voltage, V is increased such that $\mu_1 > \mu_{N+1} > \mu_r$ or a gate is used to change the electrostatics of the island to produce the same tunnelling conditions.

Numerous SET based memory circuits have been suggested and a number demonstrated at helium temperature [Tsukagoshi 1998] and at 40 K [Takahashi 1998], which are aimed at, low power, rather than high-speed applications. The first uses binary decision diagrams to produce an AND gate while the second uses oscillatory characteristics of a multiple gate SET transistor to produce a XOR gate. SET devices are at present believed to be useful predominantly for memory, electrometer and metrology applications.

To make a SET device operational at room temperature, it is estimated that the charging energy of the island, (e^2/C) should exceed the thermal energy $k_B T$ by at least a factor of 10. This suggests that the island of the SET device must be of the order of 10 nm. For reliable circuit operation, however, e^2/C should exceed the thermal energy $k_B T$ by a much larger factor and hence the feature sizes must be smaller than 10 nm. Simulations of complete SET circuits using a conventional type of architecture and incorporating perturbation by background charge fluctuations suggest that it will be necessary to go to dimensions of the order of 2 nm, and that liquid nitrogen cooling may be necessary [Korotkov 1995]. The feature dimension will therefore depend on the control of the background charge fluctuations. Recent calculations suggest that quantum dot array structures (such as multiple tunnel junctions) are less susceptible to disorder and background charge effects [Müller 1998].

The energy needed to read or write a bit and the frequency of the circuits are limited by the uncertainty principle, $\Delta E \Delta t \geq h$. To prevent bit errors, the circuit cannot operate too close to the minimum uncertainty product. The quantum limit is then approximately given by $E/f = 100 h$. Here E is the energy needed to write a bit and f is the clock frequency. SET based circuits operate at the quantum limit. The trend in SET devices is to increase E and f simultaneously so that as the speed and operating temperature increase, the ratio E/f remains constant and the circuits stay at the quantum limit (figure 2.1). CMOS circuits operate far above the quantum limit but are approaching it as E decreases and f increases (figure 2.1). As discussed in section 2.1.1 there are three important limits, which determine the ultimate performance of such systems, in particular the thermal limit, the quantum limit, and the power dissipation limit (figure 2.1). The energy necessary to write a bit determines the thermal limit. This energy must be bigger than the average energy of the thermal fluctuations, $k_B T$, otherwise bit errors will occur. For the present SET circuits, this energy is 10-22 J (10^{-3} eV) which corresponds to a temperature of 10 K. The trend in SET circuits is to increase this energy and thereby to increase the operating temperature of the circuits. The optimum value for the energy to write a bit for room temperature operation is about 4×10^{-19} J (2 eV), which is a factor 100 greater than $k_B T$.

In principle, the speed of SETs is limited by the RC time constant that, for capacitances of 1 aF, corresponds to a switching speed of 0.1 ps. To take advantage of these speeds, however, the logic architecture would have to be local so that the SETs would not have to drive a high capacitance line across the chip. Logic circuits would possibly be based on local architectures, such as binary decision diagram (BDD) logic or cellular automata, are theoretically and experimentally under investigation (Chapter 4). In practice, when a SET device has to drive an external load, such as a word or bit line in a memory cell, there are RC delays that limit the operating frequency and it is likely that even using graded tunnel barriers in a Yano type device, only sub-ns access times may be achieved. Due to the high impedance required for Coulomb blockade, SET devices are easier to implement into memory structures, than logic circuits [Lutwyche 1994].

2.2.1.2 Nano-flash device

Nano-flash devices are basically three terminal devices where a floating gate is charged and the charge produces a large change in the threshold voltage of the transistor channel. The design allows an intermediate between DRAM and Coulomb blockade potentially allowing higher density than DRAM at lower power and higher operating temperatures.

In addition, nonvolatile Dram-like memories based on the Coulomb blockade effect are intensively investigated. Both Hitachi's PLED [Nakazato 1997] and Likharev's NOVORAM being [Likharev 1998] are prominent examples. The key issue is the creation of extremely flexible tunnel barriers, for instance by multiple barriers or sandwiched barriers.

2.2.1.3 Yano memory

The Yano type memory is a 2 terminal device where information is stored in deep traps in poly-Si. The devices are created on a 3 nm thick Si film using 0.25 μm technology where one or more dots are formed naturally in the vicinity of a FET in which trapped charge modulates the threshold voltage of the FET [Yano 1998]. The device can be operated at room temperature and has been integrated in very large-scale memories (128 Mb in 8k x 8k x 2 units of which half was operational) although it is not certain if Coulomb blockade is of any relevance for device operation. One of the major problems of this type of memory is relying on the natural formation of dots and the resulting poor control of device characteristics. This may be a major hurdle to manufacturability [Yano 1998]. The advantage is a small cell size of $2F^2$, one quarter of a folded-data line DRAM cell size.

Since 1994 at least three major companies have introduced technologies for room temperature Coulomb blockade memory cells [Yano 1994] [Hanafi 1996] [Tiwari 1996] [Nakajima 1997]. They are compatible with CMOS process and integration on 250 nm technology level was demonstrated in one case [Yano 1998]. Their properties place these memories between today's DRAM and flash EEPROM.

Hitachi presented the first single electron based integrated circuit by making an 8 x 8 memory cell with read / write operation. The operation voltage is 15 V and the device is based on ultra thin polysilicon wires (3 nm x 100 nm) in which the memory node consists of an isolated poly grain representing a potential well. The presence of charge in this well modulates the conductance of naturally formed current paths between the grains. Because of the compatibility with "classical" silicon processing these results offer a real breakthrough. The device operating principle, however, relies upon the statistics within the poly wire. Recently, Hitachi extended this technology to demonstrate a 128 Mb SET memory using 0.25 μm CMOS processing [Yano 1998].

2.2.1.4 Summary of Device Parameters

Silicon DRAM has complexity of 3×10^8 bits/chip at present and is projected to reach 3×10^{11} bits/chip by 2012. The drivers are complexity and access time; power dissipation is not a major issue for DRAM. CMOS DRAM speed, however, is close to saturation, rising from 100 MHz at present to 150 MHz by 2012, and this will become a major limitation. There is clearly scope for other memory technologies with better access times that could reach this level of complexity. SRAM is faster and scales better (0.2 - 0.6 GHz now, 1 GHz in 2012), but dissipates power in much the same way as logic. Verification and error correction are important issues in present DRAMs as well as flash memories. Use of these techniques [Yano 1998] contribute significantly to the memory reliability.

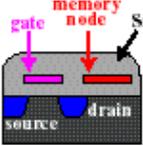
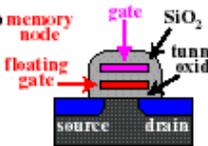
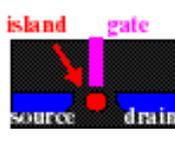
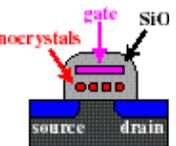
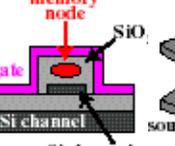
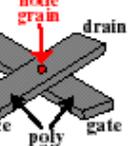
	Conventional Memory		Quantum Dot Memory			
	DRAM	Flash	SET	Nano-flash		Yano-type
				Multidot	Single dot	
device structure						
read time	~10 ns	~10 ns	1 ns	~10 ns	~10ns	~20 μs
write time	~10 ns	~1 ms	1 ns	~100 ns	<1 μs	~10 μs
erase time	< 1nm	~1ms	< 1nm	~1 ms	<1 ms	~10 μs
retention time	~1 s	~10 years	~ 1s	~1 week	~5 s	~1 day
endurance cycles	infinite	10 ⁶	infinite	10 ⁹	10 ⁹	10 ⁷
operating voltage	3 V	15 V	1 V	5 V	10 V	15 V
voltage for state inversion	0.2 V	~5 V	< 0.1 V	0.65 V	0.1 V	0.5 V
electron number to write bit	10 ⁵	10 ³	1 (excluding no to change gate potential)	10 ³	1 (excluding no to change gate potential)	2 (excluding no to change gate potential)
cell size	~12 F ² /bit	~9F ² /bit	9-12 F ² /bit	9F ² /bit	9F ² /bit	2F ² /bit

Table 2.1:- Performance comparison among conventional and proposed memories (modified after data from Hitachi Cambridge Laboratory). For complete overview see table I.I in Annex I

Table 2.1 is an extract of the tables in Annex I and compares conventional DRAM and flash memory with known data from a number of SET based quantum dot memory devices from the literature. The schematics demonstrate that most of the SET devices are really small scale examples of conventional DRAM or flash memory so that the capacitance of the memory node is small enough to produce single electron effects. Most of the proposed and realised SET memories store information in gain cells rather than the conventional DRAM 1T cells, which on one hand is an intrinsic advantage but on the other complicates direct comparison between these memory technologies.

To compete in this market, alternative technologies will have to reach a high level of maturity.

2.2.1.5 Major Challenges and Difficulties for SET

- ❑ Background charge fluctuations remain the biggest technological bottleneck. In order to reduce the perturbation of these effects on SET circuits the critical dimension must be on the order of 2 nm. Unless significant progress can be made in controlling the background charges, it seems unlikely that Coulomb blockade circuits can be integrated on a large scale.
- ❑ The required uniformity of devices is extremely demanding, raising doubts if they can be manufactured with the required tolerances at a reasonable price.
- ❑ Even assuming that large scale integration is possible, solutions must be found on how to overcome the electrostatic interactions between devices.
- ❑ Error tolerance for Coulomb blockade devices has not been investigated in great detail but it seems likely that in order to have adequate tolerances the device must operate either at lower temperature or higher voltage (and hence power).

Due to these difficulties, SET appears to be, for the present, far from being an alternative to CMOS. Only time will demonstrate if the technological bottlenecks can be overcome and perhaps nano-flash devices may bridge the gap between MOSFETs and SET.

2.2.2 Resonant Tunnelling Diodes (RTDs)

2.2.2.1 RTD

Resonant tunnelling is a device concept whose inherent multistability allows for very compact circuit design (not hindered by a need to go much below a 200 nm design rule), and its inherent speed paves the way into the GHz regime. The basic concept of the RTD is illustrated in Figure 2.3. Negative differential resistance (NDR) is produced in a double barrier structure with a resonance peak at some voltage, V which corresponds to resonant tunnelling of electrons through a subband energy, E_0 in the quantum well between the barriers. Multiple NDR devices may also be produced by designing systems with multiple subbands or extra wells.

The RTD is normally a vertical tunnelling transistor with the vertical dimensions produced by growth (i.e. MBE or CVD) and the lateral dimensions produced by lithography, although some lateral tunnelling transistors have also been demonstrated. The dimensions in the tunnelling direction are therefore typically a few atomic layers thick while the lateral dimensions are restricted by lithographic sizes. The width of the devices determines the current and hence the power dissipation. Therefore smaller widths typically produce lower power devices.

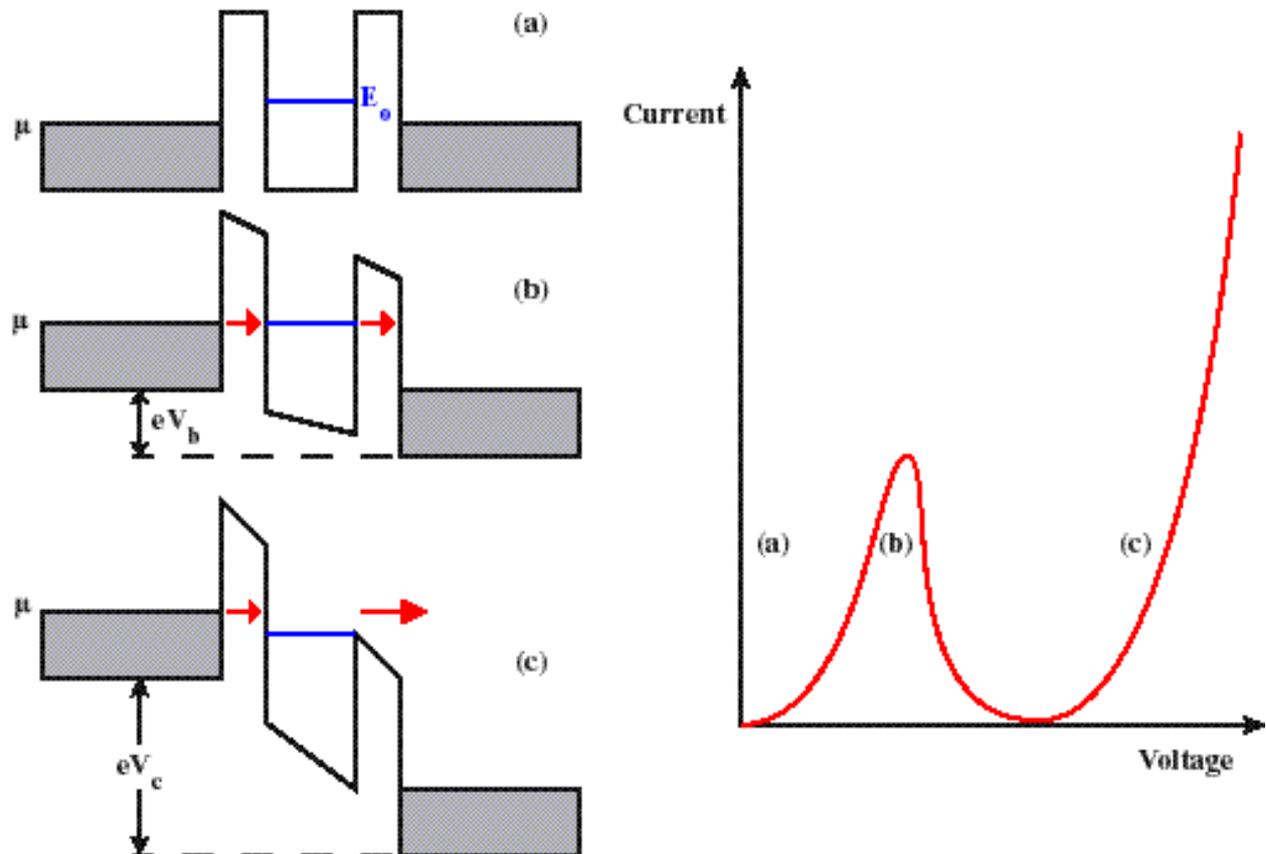


Figure 2.3:- Basic concept of the RTD. The subband energy E_0 is approximately inversely proportional to the square of the well thickness. The peak in the I-V curve occurs when the incident electrons match the energy of the subband and the electrons resonantly tunnel from the source to the drain.

2.2.2.2 RTD Performance

RTD device technology is now quite mature with many demonstrations of circuits. These are the first quantum transport devices to have made it into (pilot) production. The design of the RTD depends on the specific application in question. To reduce the power in memory and logic, one of the important issues is to produce the minimum valley current density possible at a low voltage. The peak to valley ratio (PVR) must be large enough to allow an appropriate memory or logic function with a reasonable noise margin; but a too large PVR would result in a slower circuit as it takes too long to reach the peak current density. The switching time of the RTD depends on the RC time constant of the device and the charging time of the RTD is determined by its peak current density. Hence there is a trade-off between high speed and low power in the design of the circuits. For high speed logic applications, higher peak current densities are required, typically $>10 \text{ kA/cm}^2$ allowing a maximum clocking frequency of 6.25 GHz compared to the low power TSRAM where the peak current densities are $<0.16 \text{ A/cm}^2$ which results in a maximum clocking frequency of 592 kHz [Pacha 1998]. Table 2.2 demonstrates the trade-offs between some of the important parameters in RTD designs for logic and memory applications.

Resonant tunnelling has proven fruitful in III-V devices and has shown the following results:
for Devices:

- ❑ 712 GHz InAs/InGaAs oscillator with 0.3 μW output power [Brown 1991]

for Circuits:

- ❑ 50 nW TSRAM cell with a $150 \mu\text{m}^2$ footprint (200x lower power than GaAs SRAM) [Van der Wagt 1996]
- ❑ generic logic circuits operating at 12 GHz with 20 μm minimum feature size [Williamson 1997]
- ❑ memory circuits - see chapter 4 [Watanabe 1992]

Parameter	High Speeds RTD logic	Predicted high speed logic	Low power RTD memory	Predicted low power memory	Nanometer scaled RTDs	SiGe RTD	SiGe interband tunnel diode
Peak to Valley ratio	4	3	2	3	3	1.2	2.1
Peak current density (A/cm^2)	40k	10k	0.16	0.1	10k	0.4k	22k
Minimum feature size	2000 nm	200 nm	500 nm	200 nm	50 nm	100 μm	18 μm
Peak Voltage	0.35	0.16	0.20	0.20	0.2	0.4	1.0
Maximum clocking frequency	12.5 GHz	6.25 GHz	592 kHz	56.8 MHz	6.25 GHz		
RTD time constant	0.02 ns	0.04 ns	422 ns	4.4 ns	0.04 ns		
Ref.	Pacha 1998	Pacha 1998	Raytheon 1998	Pacha 1998	Pacha 1998	Ismail 1991	Rommel 1998

Table 2.2:- Comparison of different RTD parameters (expanded from Pacha [1998])

-
- ❑ multivalued logic circuits have been demonstrated [Mizuta 1995]
 - ❑ monolithic 4-bit 2 Gsps analogue to digital (ADC) converters [Broekaert 1998]
 - ❑ 3 GHz clocked quantisers with 40 dB spur free dynamic range
 - ❑ 40 GHz static binary frequency dividers [Umeda 1995]
 - ❑ clock circuits
 - ❑ shift registers
 - ❑ 2 GB/s photodetectors with low switching energies of 30 fJ

2.2.2.3 Major Challenges and Difficulties for RTDs

- ❑ The major problem is the extreme sensitivity of device characteristics to the thickness of the tunnelling well as the tunnelling current depends exponentially on the thickness of the tunnel barrier. At present, the thickness fluctuations across a typical wafer are too large to allow for ULSI production [Evers 1996]. They are also one of the few devices which have demonstrated switching at speeds greater than high speed III-V FETs of 350 GHz, although the power output is still too low to replace electron tubes in some applications. Unless these problems are solved, RTDs will remain a niche product for high speed switching, ADC, DAC, etc. and low power applications.
- ❑ operation can only be realised in III-V semiconductors at present although recent demonstrations of interband tunnelling in Si/SiGe devices have been demonstrated with peak-to-valley ratios (PVR) of 2.05 and a peak current density of 22 kA/cm² [Rommel 1998]. RTDs in SiGe have also been fabricated but with PVRs of 1.2 and current density of 400 A/cm². The preferred system would be Si/SiO₂ RTDs [Klimeck 1997] which would allow CMOS compatible processing and integration of RTDs with CMOS circuits but there are problems with growing single crystal Si between SiO₂ barriers.
- ❑ For THz oscillator applications, a high output power from the RTD device is important. At present the output powers are quite low (μW) and require improvements (> mW).

2.2.3 Rapid Single Flux Quantum Logic

2.2.3.1 Principle of Operation

Rapid Single Flux Quantum (RSFQ) Logic is based upon a superconducting quantum effect, where a flux quantum is used as a bit. The basic switching elements are Josephson junctions (JJ). Being a quantum device, RSFQ does not suffer from the 200 nm optical lithography barrier. RSFQ is the only known technology in which circuit speeds above 100 GHz and power dissipation down to 1 μ W/gate can be envisaged.

There are two basically different technologies to produce RSFQ circuits, depending on whether low temperature superconductors (LTS) or high temperature superconductors (HTS) are employed. Due to its superconducting principle, a RSFQ device needs cooling and it should be remembered that the operation temperature of the device is lower than the critical temperature of the bulk superconductor material. The availability of adequate cooling systems, which comply with needed specifications (temperature, size, weight, dimensions, etc.) in the limits of reasonable prices, is one of the most important drawbacks for the market introduction of this technology.

2.2.3.2 Technology, Critical Dimensions and Performance

A) Technology employing low temperature superconductors (LTS)

The LTS technology uses superconductor-insulator-superconductor (SIS) tunnel junctions. The maturest technology is a three-layer process using Nb as superconductor and AlO_x as insulator. Nb-based devices must be cooled down to liquid helium temperature. Circuits with several thousand junctions, produced with a 3.5 μ m design rule, and frequencies up to 40 GHz can be obtained commercially. As the critical dimensions are limited by the magnetic penetration depth, Nb miniaturisation should be possible down to 100 nm.

An alternative is to use NbN as superconductor, where circuits can operate at 10 K. Current circuits have up to 1000 junctions, i.e. one order of magnitude less than the conventional Nb technology. This technology is less mature mainly due to the lack of self limiting tunnel barriers.

For most current LTS circuits 10 mW (@ 4K) of cooling power is sufficient. In terms of input power this corresponds to about 1 kW of electrical power, which makes LTS suited only to large systems where the energy consumption and size of cooling engines is of minor importance or to applications exploring quantum effects of superconductors, such as in metrological standards. Very large computers such as the Teraflop / Petaflop (10^{12} and 10^{15} floating point operations per second, respectively) computer, are examples of the first case. They would consume enormous amounts of power if realized in silicon, while the use of LTS reduces the cooling problem and offers a more compact layout.

B) Technology employing high temperature superconductors (HTS)

HTS technology uses mainly YBCO (or other 123 cuprates) as superconductor and mostly PBCO as insulator. Presently, only simple circuits, such as shift registers, with less than 30 junctions have been demonstrated to operate in the order of 40 GHz and at temperatures below 50 K. This is two orders of magnitude smaller complexity than current LTS technology. Due to the fact that YBCO is an anisotropic material, the fabrication is far more complex than with LTS. There are diverse techniques to produce HTS Josephson junctions, such as using the interface of bicrystals, c-axis microbridge and ramp type junctions. Which fabrication method will be chosen is not clear as all of them have pros and cons. For example, ramp type junctions have the best reproducibility, while microbridge systems are more easily scalable, a factor that may become important in the future.

The advantage of HTS in respect to LTS is that it is supposed to operate at higher frequencies (up to 500 GHz¹) and higher temperatures (50K), when the circuits are produced with a 500 nm groundrule. The feature size 500 nm is limited by the magnetic penetration depth.

The higher temperature allows for less extensive cooling efforts, while the higher frequency may open a window for new applications. As HTS circuits will be operated at higher temperatures than LTS it is expected that the risk of bit error rates increases. Measurements of these values have not been performed.

2.2.3.3 Major Challenges and Difficulties for RSFQ

- ❑ In principle, RSFQ circuits can operate at frequencies up to several hundreds of GHz. From a client's point of view, the important issue is not only the frequency of the RSFQ circuits, but also the input / output operation frequency to the outside world. In this respect, a challenge is to find appropriate interfaces to Si-CMOS or other devices that can benefit from the potential of very high-speed circuits.
- ❑ For LTS, the design and implementation of VLSI circuits with more than 10^5 gates per chip is a major challenge. For acceptable operating margins, very narrow spreads of junction parameters will be needed (probably in the region $1\sigma < 3-5\%$). Apart from the technological problems, economical considerations may produce some drawbacks. In view of the fact that the market for applications requiring VLSI, such as very high performance computing, is small, it is not evident if the large financial investments needed to upgrade facilities to produce VLSI circuits would pay off.
- ❑ HTS technology seems to be about 10 years behind LTS. In the short term, i.e. in the next three years, the main challenge is to fabricate and test circuits in the order of 100 junctions. For this to be achieved, major efforts must investigate the development of junction technology with a small parameter spread (desirable $1\sigma < 8\%$), the development of small inductances, reliable resistors and the reproducible connection of these circuit elements. In the medium term (5-8 years), the feasibility of small circuits operating at frequencies of around 100 GHz should be definitely demonstrated and integrated into a demonstrator / application.
- ❑ The main obstacle for a broad implementation of superconducting electronics is the need for cooling. Most potential clients are reluctant to accept liquid cryogenic cooling, because it is expensive, mostly not reliable enough and requires bulky compressors. Therefore market acceptance will be strongly linked to the availability of low cost, highly reliable and compact cooling systems. Technical progress has been made in producing improved cryocoolers, such as Joule-Thomson, Stirling or pulsed tube refrigerators, but at present there is no real market and hence no serious investment in developing cryocooler systems.
- ❑ To be economically competitive, HTS must be operational with a single stage Stirling and Joule Thomson cryocoolers. In terms of performance, this requires that the RSFQ circuit must be operational below 40 K and requiring at a maximum 100 mW power dissipation.
- ❑ Nb-based circuits require cooling at 4 K. These temperatures can be achieved employing liquid helium cooling or using a three stage cryocooler, such as a two stage Gifford-McMahon plus a Joule Thomson stage. All options are expensive and are bulky. Financially, a major progress would be obtained if a two stage system could be used. NbN-based or other LTS superconductors technology may fulfil the requirements to be operational at 10 K, but the technology is not very advanced.

¹ This does not imply that a clock is used at this speed, as at these speeds circuits are likely to block asynchronous and probably completely asynchronous [Hou 1995].

2.2.4 Intramolecular Nanoelectronics

2.2.4.1 Definition

The term intermolecular nanoelectronics is used to distinguish switching at the single molecule level compared to the molecular electronics at the bulk, large molecular number level as used in displays. Aviram and Ratner proposed in 1974 that suitable molecules could be used as functional electronic components, specifically for rectification [Aviram 1974]. Ashwell [1990] demonstrated such rectification behaviour using zwitterionic molecules and a magnesium evaporated electrode. While only a few major results have appeared in the field, the potential for mass self-assembly of circuits using molecular sized switching elements by cheap chemical or biological reactions is one of the few technologies which has the potential to match the ultimate densities of CMOS. In particular, intramolecular nanoelectronics is attempting to reduce logic and memory elements of circuits to the sizes of molecules and seeking:

- ultimate bit densities of 10^{12} bits/cm² [Hopfield 1992]
- decreased switching cycle times to 10 ps [Hameroff 1989]
- reduced energy per bit cycle to 1 eV [Hopfield 1992].

Intramolecular nanoelectronics is still in its infancy, but holds promise in that it potentially provides a bottom-up way to produce sub-nm sized functional devices without the inherent tolerance problems of a lithographic (top-down) approach [Goldhaber-Gordon 1997]. First manufacturable solutions are predicted to be bolt on technologies on top of a CMOS chip [Wieder 1998] with the CMOS chip providing input / output and communication to the rest of the world.

2.2.4.2 Electric-field controlled molecular electronic switching devices

Among molecular switches, electric-field controlled molecular electronic switching devices are closest to conventional semiconductor devices and therefore the most likely candidates for applications. The basic idea is to use the central part of molecules for functions (switching, RTD, transistor, rectification, etc....) while the endgroups are used to self-assemble. Diblock copolymers, for example, may be self-assembled into chains and purified to very high yields. Thiol endgroups can be used to anchor the copolymers to metal electrodes.

The mechanism for the functional behaviour can be very diverse. As already indicated, attaching electrophile and electrophobic groups to either end of the molecule induces rectifying behaviour. Aviram [Aviram 1988] proposed a number of molecules that switch between conducting and insulating states by oxidation / reduction. One example is tetrathiafulvalium (+) with a partially filled highest occupied molecular orbital to allow free states for conduction and tetrathisfulvalene which is an insulator because of its filled highest molecular orbital. The transport properties are very similar to Coulomb blockade (Figure 2.4). Current voltage data from single designer molecules have been demonstrated such as the rectification properties of a benzene ring attached to two gold electrodes using S atoms (Benzene-1, 4-dithiolate) [Reed 1997]. More complicated structures have also been demonstrated where the rotation of a single bond of the π orbitals allows the conductivity to be changed by over 6 orders of magnitude, providing a switching mechanism in a polymer chain [Reed 1998].

The first “molecular transistor” actions were shown by R.P. Andes [Andes 1996] using self-assembled gold particles on top of a,a'-xyldithiol (XYL) molecules on an Au(111) surface. The second electrode was a STM tip and a Coulomb staircase function was demonstrated at room temperature. A second demonstration was using carbon nanotubes placed across two electrodes with a controlling back-gate [Tans 1998]. Clear Coulomb blockade and also single electron transistor properties were demonstrated at room temperature. In semiconducting nanotubes, “classical” field effect transistor action could be obtained [Tans 1998, IBM]

These experiments demonstrate the basic transistor principles but neither technique is suitable for mass manufacture. A number of proposals have appeared in the literature for molecular devices but few have been demonstrated. No real molecular three terminal transistor device defined in this roadmap has yet been demonstrated. Significant progress, however, which has been demonstrated in the literature is discussed below.

Recognition of conformation and adaptation of individual molecules has been demonstrated where the STM is used to understand bond rotations in molecules and to determine how they adapt or can switch their state. Such conformational processes are used in natural systems and form a basis for molecular switches [Jung 1997]. It has been demonstrated that molecular systems such as fullerene based systems can be incorporated in two and three terminal devices and that interesting electronic characteristics can be obtained, i.e. macroscopic quantum state in a bucky tube [Tans 1997].

The area of self-assembly has for some time been subject to designer molecular concepts

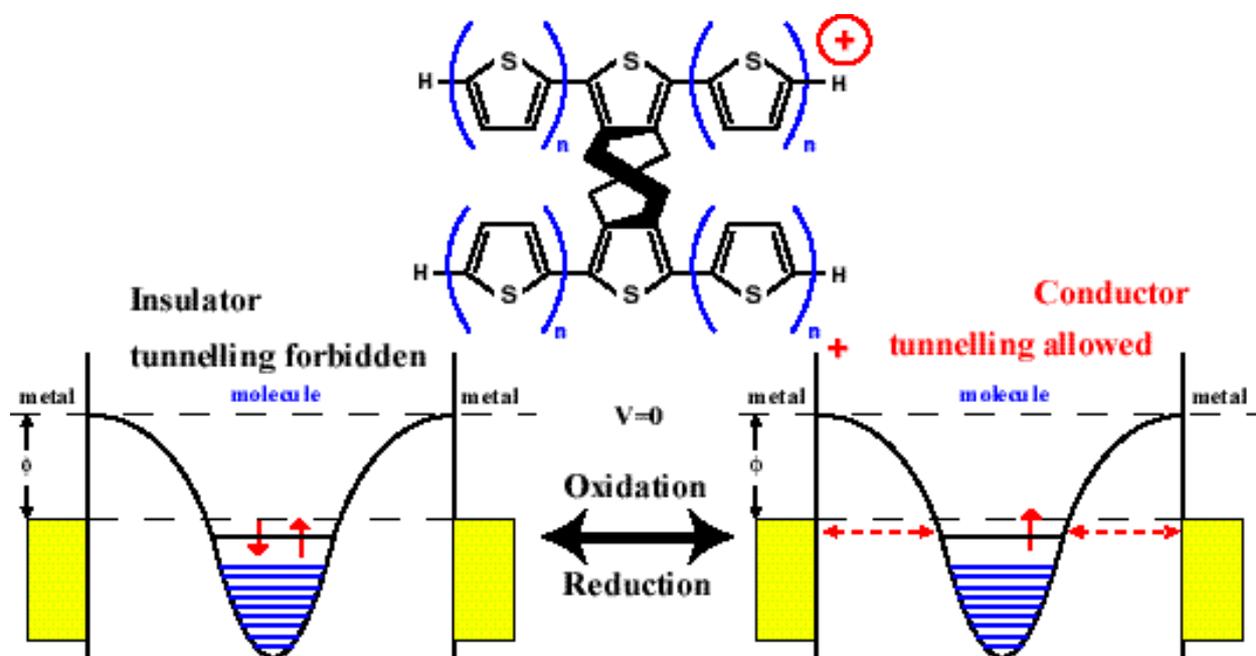


Figure 2.4:- The transport mechanism responsible for the conduction and insulation transformation through the reduction / oxidation of a molecule. The mechanism is similar to Coulomb blockade.

through optimisation of molecular chain length and omega functionalisation as well as techniques such as contact printing to define molecular patterns on the scale of tens of nanometers. Lehn's [Lehn 1995] supramolecular concepts have also been explored in the design and self-organisation of grid-like metal co-ordination arrays using molecular systems [Schubert 1997].

The assembly of devices and arrays may be expected to be radically different from current approaches. For instance, at the macroscopic scale chemical interactions can be used for the three dimensional self-assembly of millimetre-scale components [Terfort 1997]. A different form of molecular recognition is to use DNA to create a binding interaction between nanocrystals of colloidal metals in solution (or potentially at a surface) [Mirkin 1996].

2.2.4.3 Alternative Approaches to Molecular Switching

Electromechanical Switching

Several illustrations of STM based molecular manipulation have in recent years been proposed as a means for performing electronic computation. Although the need of UHV and the intrinsic

sically low speed of the STM precludes any applications, we briefly mention certain examples of these thought-provoking experiments.

The main consideration of electromechanical switching at the molecular level is that the properties are controlled by deforming or reorientating a molecule rather than moving the electrons on the molecule. An example is the C_{60} based single molecule electromechanical amplifier, which has been demonstrated by deforming the molecule using a STM tip. The device relies on the controlled vertical deformation of the C_{60} cage resulting in mechanical modification of resonance tunnelling bands. This demonstrates a fundamentally new approach to switching and amplification using molecular mechanics and quantum processes. The speed of such a device would be limited by the vibrational frequency of C_{60} at over 10 THz (10^{13} Hz) although much lower frequencies (10 Hz) have been demonstrated [Joachim 1997].

Another example is the Molecular Abacus device using the fullerene C_{60} as beads, a 0.25 nm high monatomic step as the rod and STM tip as the finger to both reposition in one dimension and count by imaging. The abacus itself has an active area defined as around 1 nm x 13 nm and is a molecular mechanical device. [Cuberes 1996]. By moving a single atom out or into an atom wire researchers at Hitachi have simulated a two state electronic device of atomic dimensions [Wada 1993]. The device is called an atom relay or molecular relay. Two-state devices based on the change of a molecular conformation can also make a relay such as a rotamer [Takeda 1982].

Photoactive / photochromatic switching

A number of molecules, specifically a number of proteins, may have their electron distribution changed by the absorption of photons to produce switching effects. The biological photochrome bacteriorhodopsin has been suggested as one possible type of molecule for holography, spatial light modulators, neural network optical computing, nonlinear optical devices, and optical memories [Birge 1995]. The significance of bacteriorhodopsin stems from its biological function as a photosynthetic proton pump in the bacterium *Halobacterium halobium*. A combination of serendipity and natural selection has yielded a native protein ideal for optoelectronic applications. Other examples of optoelectronic biological molecules include visual rhodopsin [Birge 1995], chloroplasts [Greenbaum 1992] and photosynthetic reaction centres [Boxer 1992].

2.2.4.4 Molecular Wires

A further important problem, which has escaped much attention, is that most polymers are very poor conductors and hence molecules must be found with good conducting properties if circuits are to be built up. Note that utilising the delocalised electrons on aromatic and acetylene groups is not sufficient to obtain good conducting properties as required by the microelectronics industry in circuits. The delocalised π electronics in such compounds occupy the “homo” (highest occupied molecular orbital) fully, so that for charge transport an extra electron must be added to the “lumo” (lowest unoccupied molecular orbital), which is typically a few eV away. Thus, the compound acts as a semiconductor and good conducting properties can only be obtained by doping the systems. The only metallic wire like molecules that have been studied in more detail are carbon nanowires, where the bandstructure allows metallic behaviour. These compounds seem quite promising for some applications, but ways must be found to connect them, through conducting junctions at specified positions, to the active switching element. The RNA synthesis pursued by the Technion group could be a useful approach in this direction.

2.2.4.5 Major Challenges and Difficulties for Intramolecular Electronics

Intramolecular electronics are far behind other alternative concepts as none of the following issues have yet been demonstrated:

- self-assembly of single devices at reasonable (or small) integration densities let alone integration with conducting molecular wires;
- appropriate yield of devices from chemical or biological reactions for fabricating or manufacturing of circuits;
- high data bandwidths for interconnects;
- the potential for low cost per bit has not been demonstrated.

Currently only a few circuit components at the single molecular stage have been demonstrated, but as some architectures (chapter 4) are appropriate for molecular electronic systems there is hope for integrating the components into elementary circuits.

2.2.5 Spin Devices

2.2.5.1 Operation

Magnetic RAMs (MRAMs) are nonvolatile memory (NVM) devices whose operating principle rest upon the differences in conductance of multilayered stacks of nanoscale magnetic materials. Several concepts have been proposed in the past including giant magnetoresistance (GMR) devices (functioning on the basis of ferromagnetic / antiferromagnetic coupling) or device concepts based on the magnetic behaviour of the electron. The magnetic state does not require external power to be stabilised and hence applications are in NVM or static RAM. Since spin is already defined on a single electron, the devices should be able to downscale to very small dimensions.

MRAM circuits are built on existing CMOS and use the GMR effect in magnetic multilayers. The effects exploited in MRAM are based on the differences in transport properties and density of states for spin-up and spin-down electrons in a magnetic metal. This difference in properties of minority and majority spins in a material with a high number of available carriers for extremely small structure sizes (<10 nm) triggers the quest for new device concepts.

In principle, MRAMs are an array of ultra-small magnetic sensors with hysteresis constituting the bits. The bits are “set” by a magnetic field (order 10 Oe) generated by 2 crossing conductors carrying a current in the order of 10 mA. The bits consist of magnetic thin multilayers exhibiting magnetoresistance, which enables one to read their memory state through measurement of their resistance (read-action). The signal magnitudes (order 5 mV) are read by Si sensing circuits using auto-zero schemes to level out the transistor imbalances and enable correct reading of the small signals.

2.2.5.2 Performance

MRAM based on magnetoresistive effects in metallic structures presents some advantages compared to existing NVM approaches. Scaling rules of MRAM are dictated by Si-technology. New materials development (e.g. tunnel magnetoresistance) will improve density, speed and power performance. Access times of the order of 300 ns have been reported so far, but it should be possible to reach 10 ns (at room temperature) within a few years.

VS Nonvolatile Electronics Inc. demonstrated MRAM with 200 nm cells producing 8% output in resistance change with a clear memory effect. For a future implementation of the memory circuit some additional requirements have to be met: careful control of the interlayer thickness, thin insulation, a magnetic keeper structure and conductor materials with higher allowed current densities ($>10^7$ A/cm²). Analysis shows that sense line width less than 100 nm and array densities of 5×10^8 bits/cm² are possible with GMR.

Honeywell Inc. has been pursuing the AMR-MRAM route since 1984 (based on the Anisotropic Magneto-Resistance effect in permalloy NiFe) and foresees the transition to GMR elements to result in 4 Gbit/cm² NVM technology. Their objective is full CMOS compatibility to ensure lowest power, highest reliability and rapid insertion into existing markets.

Lower cost can be achieved in MRAM compared to the equivalent semiconductor memory due to the need for only 2 to 3 metal interconnect levels. Also, because of the compact architecture process feature sizes that are 1 to 2 generations behind semiconductor memory / DRAM can be applied to achieve competitive densities.

A new development in the MRAM area is the use of memory nodes based on Tunnel Magnetoresistance (TMR). In brief, these components are a tri-layer structure consisting of two ferromagnetic metals, spaced by a thin tunnel oxide (mostly Al₂O₃, cf. Josephson Junctions). Similar to spin-valve structures, these tri-layers show a difference in resistance dependent on the alignment of the moments in the ferromagnetic layers. The main difference between TMR structures and the

presently used all-metal GMR structures is their much higher resistance and higher (relative) resistance change in the low-field regime. Further, TMR devices are compatible with CMOS. Most of the successful TMR devices have been fabricated using shadow masking technology. To enter the sub-micron regime and achieve IC-compatibility standard processing techniques can be applied, $100 \times 100 \text{ nm}^2$ in industrial laboratories.

2.2.5.3 Challenges and Difficulties

Down to 180 nm, the presently known MRAM concepts will be applicable because magnetic properties can be treated as bulk-properties and programmable logic functions will be demonstrated using these materials. Main challenges will be to control the magnetic switching behaviour and micromagnetism in sub-micron structures, as in this regime, magnetic materials show quantum effects, that may hinder proper operation. Also important for the integration of devices is the control of the magnetic switching properties of small-sized magnetic contacts as well as the control and/or use of stray-fields from nanoscale magnets. When the size of a magnetic particle is decreased below some critical value (10 nm for Co), the superparamagnetic effect sets in. Below this size, the thermal activation is strong enough to overcome the ferromagnetic ordering in the structure and ferromagnetic behaviour is lost.

Besides the above metallic devices, the integration of magnetic nano-structures and semiconductors in novel device structures and the exploitation of minority-majority spin effects offers new perspectives on magneto-electronics. These include the control of spin-polarised current in components using magnetic field, electric field, or light modulation.

Magneto-electronic devices have no inherent electronic gain. Either one reads the resistance of an impedance matched device (e.g. TMR, magnetically modulated devices) or a transimpedance concept must be developed (e.g. spin-valve transistor and spin-injection devices) to enable signal read-out.

Multiple stable magnetic states are possible in a multilayer of magnetic material, opening the possibility of fabrication of multi-value MRAM storage nodes. This route is at present not being addressed. It should be clear that the success of present day MRAM lies in the simplicity and density of the design: crossing metallic conductors constitute the memory array. There is no need to include a single transistor in a memory node. This allows for back-end processing, buried electronics and high density, limited by metal-distance and line width rules. When a transistor is part of the magneto-electronic memory-cell, areal density is sacrificed in favour of larger fan-out and functionality.

Cell Width (μm)	Cell Area (μm^2)	Capacity (Mbit)	Die Size (cm^2) 40ns cycle
0.4 μm	1.90	16	0.9
0.2 μm	0.55	64	1.0
0.1 μm	0.19	256	1.4

Table 2.3:-Estimates of GMR memory density and performance: (based on 6 % GMR and 15 W / square resistivity)

The difficulties related to spin devices can be summarised as follows:

- ❑ For future applications metallic MRAM will have to be linked to CMOS circuits. For high speed MRAMs the tunnel junction should be as thin as possible (in the range of 50 nm) for good communication with fast CMOS circuitry
- ❑ Compatibility of ferromagnetic materials in a CMOS line.
- ❑ Future high density TJ-MRAM would require thin tunnel junctions of the order of a monolayer in order to have optimum resistances of about 50 Ω . The spread of the junctions must be below the monolayer regime in order to guarantee the uniformity needed for high density. This requires a high precision in manufacturing which is not achievable at the moment.
- ❑ Spin injection, which may lead to novel magnetic devices, has not been demonstrated.

3 Nanofabrication

As already discussed in section 2.1 on MOSFETs, optical lithography has been a key driver in the increase of CMOS integration. Year on year the predicted physical limits, related to diffraction, have been pushed down using smaller wavelengths in addition to top surface imaging, silylation of resists and sophisticated mask technology including phase-shift masks and optical proximity corrections. It is currently believed that optical lithography can be used for groundrules down to 150 nm and might even be used for the 100 nm generation and below. This would imply an increasing process and mask complexity, increasing the cost of ownership. Alternative methods are actively being explored. The so-called “Next Generation Lithography” methods include extreme ultraviolet (EUV), X-ray, projection electron-beam and projection ion-beam lithographies.

There are a number of alternative routes to create smaller structures. Some of these employ the basic ingredients of the traditional technology: replica generation using some kind of mask. Recently, a number of technologies have been proposed which leave out the mask as an intermediate step. The resolution limits of some of these alternative lithography methods under practical and production considerations are summarised in table 3.1 and 3.2.

Industrially applicable nanofabrication techniques have the following minimal requirements: they must be able to produce millions or billions of these small structures in a quick, reliable and cost-effective way. These techniques must be able to connect these structures in a predefined manner. For this reason, parallel techniques, such as conventional circuit lithography techniques using mask alignment and pattern transfer seem the only currently realistic way to achieve highly integrated circuits. Serial techniques, such as scanning probe techniques or electron beam lithography, may be useful for mask making, or single component fabrication, but do not supply adequate throughput for large-scale integrated circuits.

In the following we will briefly review the optical and next generation lithography methods, developed for their potential application to CMOS technology. Then a summary of current trends in nanofabrication will be presented, one of which may yield crucial technology for future device generations. We make a subdivision between top-down approaches, where through extensive technology small structures are fabricated by deposition followed by lithographic patterning and etching, and bottom-up methods where fabrication starts at the basic molecular or single device level and assembles circuits by pulling devices together.

Year of Introduction (Production)	1999	2001	2003	2006	2009
Minimum Feature Size (nm) Dense Lines	180	150	120	90	65
Overlay (nm)	65	55	45	35	25
Optical 248nm					
Optical 193nm					
Optical 157nm					
Extreme Ultraviolet (EUV)					
X-rays					
Electron beam projection					
Ion beam projection					

Table 3.1:- Comparison of different lithographic systems indicating expected applicable resolutions.

Table 3.2:- Practical and ultimate resolution limits for lithography methods

3.1 Lithography for CMOS technology

The potential lithography solutions to be used in the next ten years are listed in Table 3.1, on the basis of the 1998 update of the SIA roadmap [Canning 1997 and 1998].

3.1.1 Optical lithography

At present optical lithography is the dominant technology. Using projection optics, the resolution varies as $k_1\lambda/NA$ where k_1 is an empirical process parameter, λ the source wavelength and NA the numerical aperture of the optical system. The depth of focus varies as $k_2\lambda/(NA)^2$. The decrease of the critical dimension is based on decreasing the wavelength, successively using mercury G-line (436 nm), mercury I-line (365 nm), followed by excimer lasers:

KrF (248 nm) presently in use for the 180 nm generation, ArF (197 nm), F₂ (157 nm).

Large efforts are being made in research on optical materials in order to obtain aberration free lenses with large values of NA in addition to resolution enhancement techniques to yield smaller values of k_1 and k_2 . They include off-axis illumination, attenuated and alternating phase-shifting masks, optical proximity corrections, and top surface imaging techniques. The introduction of optical proximity correction (OPC) and phase-shifting in mask technology involves structures with lateral dimensions smaller than four times the critical dimension. This results in the loss of the benefit of reduction techniques for mask definition and they also have to be controlled in thickness. The decrease in depth of focus by using these methods has to remain within an acceptable manufacturing process window.

If optical lithography reaches the sub 100 nm region, which seems physically feasible, many issues such as mask fabrication and repair, sensitive high resolution resists, overlay... remain to be demonstrated. Research and development have to be quick enough to meet the predicted introduction into production, and the technology has to be cost-effective compared to the alternative technologies presented below.

3.1.2 Extreme Ultraviolet lithography

Extreme Ultra Violet (EUV) lithography or soft-X-ray lithography is a natural continuation of conventional optical lithography in its course towards smaller wavelengths [Gwyn 1997]. Major changes are the use of reflection optics and masks at wavelengths in the range 10 to 15 nm, using multilayer mirrors. EUV lithography is supposed to be introduced, if ready, for production at 90 nm and may be extended (although this has not yet been demonstrated) down to 30 nm.

The EUV source must achieve high power. Corresponding solutions are synchrotron radiation and plasma laser sources. Synchrotron radiation is a mature technology, providing the necessary technical solutions, but suffers from a lack of popularity in the semiconductor world. Laser plasma sources providing high power without debris are under development. The EUV system also involves multilayer collectors and mirrors providing the reduction factor. Surface and interface roughness, with an accuracy better than 0.25 nm, zero-stress control, and robustness to thermal effects are today major challenges for multilayer EUV optics. The very low value of the numerical aperture of the system allows a better depth of field than in conventional optical lithography. Mini-steppers have been realised for studies of the capability of the method.

The reflective mask is achieved by patterning an absorbing metal layer deposited on a multilayer (typically Mo/Si). Repairing a multilayer is not feasible, and control of defects during growth is essential. Two orders of magnitude have to be gained on the best performance of today.

Demonstrations of 100 nm line-and-space and 70 nm isolated lines have already been shown. The method is attractive, but has a long way to go before being introduced into production.

3.1.3 X-ray proximity lithography

X-ray proximity lithography [Silverman 1998], using a typical wavelength of 1 nm, represents the last step in the decrease of photon wavelength for nanolithography purposes. It has been intensively developed by academia and industry during the past 18 years, and is nearly ready for use. In the absence of appropriate X-ray optics, it is based on a direct 1:1 wafer imaging behind a mask. In order to give sufficient throughput, the X-ray source is either synchrotron radiation or plasma laser source. The former is commercially available and has demonstrated a very low downtime per year, while the latter is under development. Commercial steppers are also available, even if they need further improvements for future generations. The mask technology has made good progress: an X-ray mask consists of relatively thick (a few hundred nm) high atomic number material structure, on a large area 1000 nm thick SiC membrane. Stress control of these masks, essential for pattern placement accuracy, is not a problem. Zero defect masks have been demonstrated. A number of integrated circuits or device demonstrations have been produced with dimensions not yet reached by competing methods.

Some technical issues remain open, such as the availability of an electron-beam pattern generator with a high throughput for writing the 1:1 mask, the stability of the masks after the high number of irradiations required to make the method cost-competitive, and a full demonstration of extendibility to 70 nm ground rules.

Additional advantages of X-ray lithography include the easy single layer resist process and high reproducibility. But a clear, enormous disadvantage, are the huge costs.

3.1.4 E-beam projection lithography and the SCALPEL system

Electron beam lithography (e-beam) makes use of the fact that electrons can be deflected and modulated by electrostatic or magnetic fields to produce an image. The fundamental resolution limit is given by the Heisenberg uncertainty ($\Delta x \Delta p \geq \Delta h/2\pi$) offering the possibility to achieve patterning below 10 nm (for electrons in the eV to keV region the resolution can be expressed by the following expression: $\Delta x \propto 1/E^{1/2}$, being E in eV and x in nm). There are two ways to perform e-beam lithography, either by scanning the beam to generate patterns or to perform electron imaging through masks.

Current e-beam lithography systems are mainly serial technologies which are limited by the scanning speed of the e-beam in the pattern generation. The throughput may be increased by a matrix of parallel e-beams. This increases the complexity of the system and puts high requirements for the individual beams in terms of brightness, energy spread and beam opening, as well as on the optics of the lenses. For example, new sources must be developed that allow array emission: Spindt type field effect arrays would offer the possibility to in the low voltage regime (less than 250 eV at currents smaller than 500 A), but the current fluctuations from the tips cause constantly changing field emissions that can not be tolerated. On the other side, micro-fabricated of Si tips offer good control of the tip location and would provide the possibility of an active matrix, but the surface state of the emission area is difficult to control having effects on the field emission uniformity. An additional approach is to use carbon surfaces (diamond like carbon or nanotubes) for electron emission. Ideally, any appropriate source must be robust and able to operate at hard environment conditions, i.e. insensitive to the degassing from non-clean wafers and intrinsic to the fabrication process (resin deposition, chemical bath treatments), a good brightness, a low energy spread and operate at a low energy.

The last issue dealing with the electron energy is related to the fact that high energy electrons have the effect to damage samples and to destroy the gate structures by the dissipation of the build up charges. One of way to circumvent this is to use low energy electrons that have a low penetration depth. In particular e-beams with an energy below 300 eV (mean free path < 1nm) are considered to a solution to overcome sample charging and surface contamination.

To achieve arrayed focused e-beams two main ways are currently pursued: microcolumns and microguns. A microcolumn consists of a non planar technology construction of about 4 mm in height including a Schottky tip as e-source, a filter and a set of microfabricated lenses and detectors. On the contrary, in the microgun concept the electrodes, the deflectors and the detectors are micro-fabricated planar on a Si wafer. For the microcolumn ($V = 1\text{ kV}$) approach, Chang (Etec 1999) has estimated the required performance of the system for a 100 nm lithography with a pixel spacing of 50 nm in the following way: if a throughput of 10 wafers (300 mm) per hour are envisaged, the system would require 50 columns operating at 30 nA per column (or 200 columns at 0.8 nA/col). If the throughput should be increased to 25 wafers / hour the all numbers would scale with the factor 2.5 and so on. In view that the limiting factor is to offer a high exposure per area, the tendency in both techniques, microcolumns and microguns, is to increase the number of array elements while offering a high current per beam. In particular, there are two major bottlenecks: First to produce a stable and reliable sources which work at low energy, high coherence, etc that are reliable to work at in an array structure under fabrication environment and second to develop a microfabrication technology to produce lenses, detectors in a precise and economically effective way.

Current research efforts are aimed to optimise array e-beam projection systems that would be applicable for 100 nm lithography. Very early results with a low energy ($< 300\text{ eV}$) single electron beam indicate that resolutions of 30 nm are achievable with a single spot. In the long term, a resolution in the order of 40-50 nm could be the practical limit of an array system with an industrially adequate throughput.

The scattering with angular limitation projection electron-beam lithography (SCALPEL) approach combines the high resolution and wide process latitude inherent in electron beam lithography with the throughput of a parallel projection system [Harriot 1997]. In the SCALPEL system, a mask consisting of a low atomic number membrane and a high atomic number pattern layer is uniformly illuminated with high energy electrons (100 keV). The entire mask structure is essentially transparent to the electron beam, so very little of the beam energy is deposited in it. The portions of the beam, which pass through the high atomic number pattern layer, are scattered through angles of a few milliradians. An aperture in the back focal plane of the electron projection imaging lenses stops the scattered electrons and produces a high contrast image in the plane of the semiconductor wafer.

In contrast to other approaches to electron or ion beam projection where stencil masks are used, the mask for SCALPEL consists of a continuous layer. This way even doughnut-like structures which are unavailable in stencil masks can be realised. The thin membrane (Si_3N_4) is supported by a grid of struts, that stabilise the membrane and at the same time serve as heat sinks, minimising thermal stress during exposure. The lithography is a 4:1 projection. So the mask structures for a 100 nm process are in the 400 nm range, which allows a fabrication by standard tools for optical lithography masks. Down to 100 nm feature size no proximity correction is necessary, so the mask design is an easy 4:1 copy process from the features to be produced.

As a 'die at once' exposure requires strutless masks and large field electron projection, with extensive and expensive correction systems, the SCALPEL approach uses a step and scan strategy. Between the struts, of the mask the pattern is present in stripes. The electron beam shape is a square of $1 \times 1\text{ mm}^2$. The pattern stripes are mechanically scanned through the beam and the wafer stage moves synchronously. Errors are detected by an interferometer and fed into a stitching deflector which can correct them at nanometer size. The trade off between high throughput and slow mechanics leads to an additional electronic scanning of the beam on mask, allowing for an effective size of 3 mm^2 for the projected field on the wafer ($0.25 \times 0.25\text{ mm}^2$ without scanning). Several series

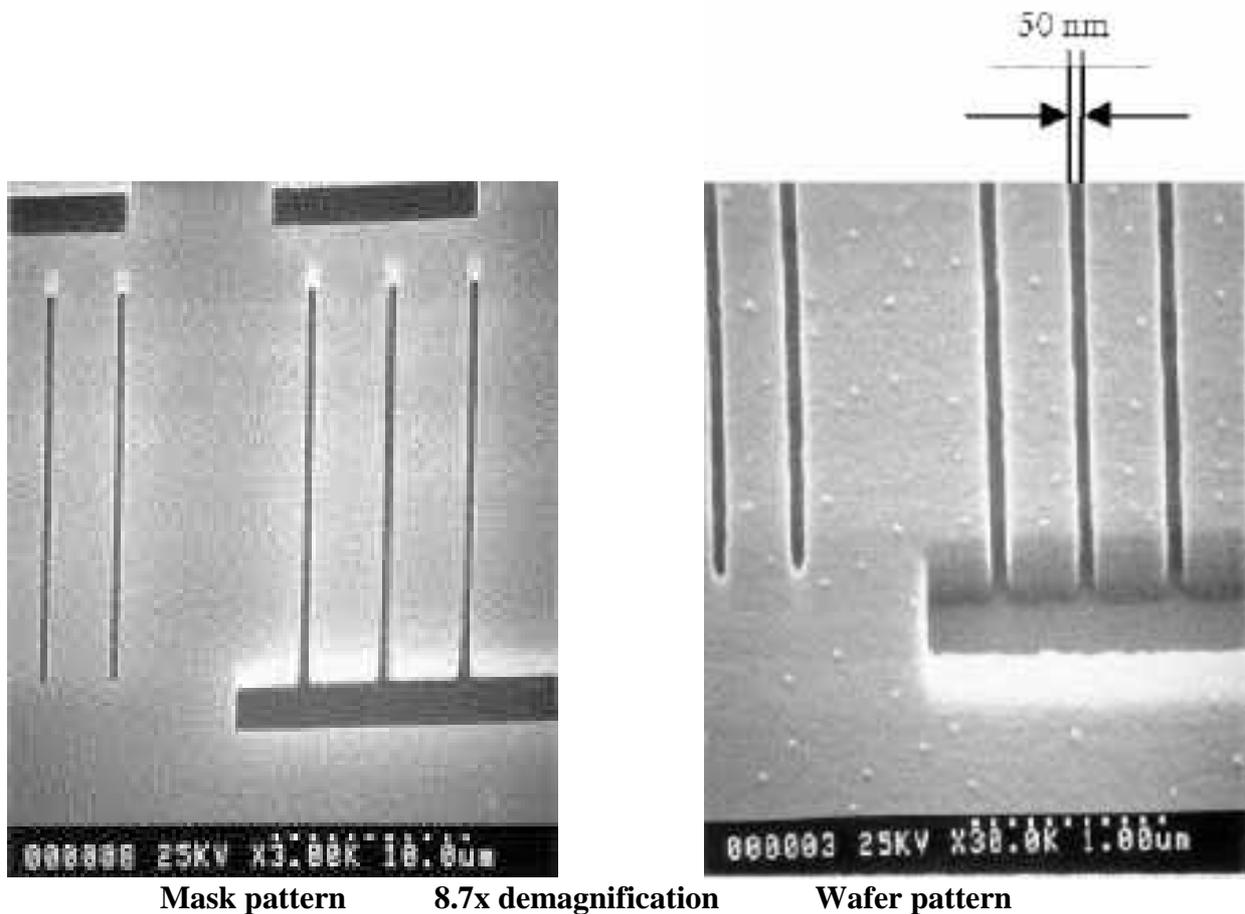


Figure 3.1: Ion protection lithography with 75 keV He⁺ ions of a pattern on an open stencil mask (Si, 2 μm thick) onto a wafer coated with 300 nm of Shipley UV II HS resist. Exposure dose: 0.3 μC/cm² with research type ion projector operated with multi cusp ion source (2 eV energy spread) [Brünger 1999].

of marks are used for wafer alignment, die alignment and stitching alignment, allowing for a high overlay accuracy.

One limitation is the throughput, which strongly depends on beam current and resist sensitivity. A sensitivity of 10 μC/cm² allows for a throughput of 1.5 x 8"-wafers per μA beam current. The interaction of the electrons inside the beam lead to an intrinsic blur, proportional to I^{2/3} and limits the resolution to about 35 nm at 10 μA. Improvements can be made by increasing the resist sensitivity. The latter includes an improvement of the stage performance, as the movement of mask and wafer may then be a limiting factor. The resolution is mainly affected by the blur in the beam and projection errors. The sensitivity to dose variation and the depth of focus (mask deformation) is much lower than optical lithography. At present SCALPEL can be expected to deliver accurate lithography results down to less than 100 nm with good linewidth at a reasonable throughput if the necessary beam current can be managed and the resist sensitivity can be further reduced.

3.1.5 Ion beam projection

Focused electron beams and focused ion beams have been used to write dimensions into resist well below 20 nm, but because of the serial writing process these techniques are far too slow for high volume production purposes.

Projection systems [Melngailis 1998], employing a mask and performing parallel printing of whole images, combine both high resolution and high throughput. Ions are particularly well suited for this because they suffer little or no scattering in a resist (small proximity effect) and the resist sensitivity is high because the range of light ions (H⁺; He⁺⁺) in the resist in the 50 – 100 keV energy

domain is typically in the same order of magnitude as the resist thickness. For chemically amplified standard deep-UV resists which are also useful for ion exposure the resist sensitivity is 10^{12} to 10^{13} ions/cm² ($0.16 - 1.6 \mu\text{C}/\text{cm}^2$), which allows for exposure times of below 1 second in a $10 - 100 \text{ mm}^2$ field.

So far with an ion projection lithography (IPL) system 50 nm wide lines have been produced in relatively thick (300 nm) resist at a reduction rate of 8.7 from mask to wafer (figure 3.1) [Brünger 1999]. With thinner resists resolution below 50 nm is expected.

Ions can also be used for resistless processes where ions directly modify the surface of a substrate. Interesting applications are the production of magnetic nano – islands or of quantum dot arrays by damage writing. The ion projection technique will enable the necessary printing speed.

For a comparison of writing speed between e-beam and ion exposure it has to be considered that ions of adequate energy 50-100 keV deposit their energy completely in the resist layer while electrons of 100 keV penetrate deeply into the substrate. This leads to quite different resist sensitivities (= dose to clear large areas in $\mu\text{C}/\text{cm}^2$). In the following table the dose to clear large areas is multiplied by the acceleration voltages of the particles to give the energy which is transferred onto the wafer¹:

Ion Projection	=	$0.15 \mu\text{C}/\text{cm}^2 \times 75 \cdot 10^3 \text{ V}$	=	$11.3 \text{ mJ}/\text{cm}^2$
e-beam (SCALPEL)	=	$4\mu\text{C}/\text{cm}^2 \times 100 \cdot 10^3 \text{ V}$	=	$400 \text{ mJ}/\text{cm}^2$
Extreme-UV	=			$10 \text{ mJ}/\text{cm}^2$
Deep-UV	=			$15 \text{ mJ}/\text{cm}^2$

3.2. Emerging nanofabrication methods

3.2.1. Top Down Approaches

3.2.1.1 Electron-Beam Nanolithography

Electron-beam lithography is the essential basis of nanostructure fabrication at present. Gaussian beam pattern generators or scanning electron microscopes are used with high-energy electrons (100 to 200 keV) and a small electron probe size (1 to 10 nm). Using high resolution and low sensitivity PMMA resist, dimensions down to 30 nm are routinely produced in many research laboratories with high reproducibility. Resolutions down to 7 nm have been demonstrated. Smaller dimensions have been shown in inorganic resists, but no adaptable transfer method has made them usable. Progress has been made for reproducibility at the nanometer scale: NTT has presented a SET operating at room temperature whose structure has been produced by e-beam lithography. The Si Island is 10 nm thick and the Si tunnel barriers 1 nm [Kurihara 1997]. Room temperature SETs using Si point contacts fabricated on a SIMOX substrate can be fabricated with an electron-beam using an anisotropic soft stopping technique that defines the point contact width [Ishikuro 1997].

The biggest disadvantage of all this technology is the large writing time, causing a low throughput that makes electron-beam lithography prohibitively expensive for the mass production of integrated circuits.

¹Note that for real throughout numbers, besides sensitivity, other limiting factors such as Coulomb interaction in dense particle beams have to be taken into account.

Table 3.3: Resolution, Sensitivity and Application of Electron Beam Nanolithography Methods

Current research is focused on this problem, with the following approaches:

- ❑ parallel exposure, and the SCALPEL method described above, is a promising approach, getting closer to a realistic throughput, but its resolution capability is not yet known.
- ❑ very low energy e-beam writing: due to the increase of resist sensitivity, the throughput can be increased by a number of orders of magnitude. The low energy electron-beam writing has not yet demonstrated resolution below 50 nm.
- ❑ parallel writing using arrays of microfabricated microcolumns.
- ❑ more sensitive resists, such as chemically amplified resists, for which the effort concerns process window and resolution limits, and in the optimisation of the electron-beam process and the development of more sensitive resists.

The present resolution, sensitivity, and applications of electron-beam nanolithography are summarised in table 3.3.

3.2.1.2 Scanning Probe Methods (SPM)

There has been steady progress in developing Si-based technology using scanning probe devices, although the resolution originally envisioned still remain to be demonstrated. For any application in fabrication, throughput becomes an important issue. Efforts in employing large arrays of scanning probes are certainly promising, with arrays of up to 50 SPMs having been fabricated [Minne 1997], but it remains to be seen if SPM can ever intrinsically surpass parallel lithographic methods.

3.2.1.3. Nanoimprint

Based on the Gutenberg's invention, nano-printing techniques are aimed to produce nanometer structures by physically deforming a polymer with a mold. Two main techniques do exist: imprinting (embossing) and inking (contact printing). In the first case a reusable mold is stamped at high pressure into a polymer film on a substrate and afterwards removed. The imprinted structures serve afterwards as a lift off mask. In the inking approach a flexible silicon rubber stamp is inked with a reagent solution. The stamp is put on the substrate and then removed. The parts of the stamp in contact with the substrate are covered with a self assembled thiol monolayer which offers the patterned structure.

The current state of the art for imprinting is the production of 6 nm hole and 15 nm trenches in PMMA [Chou 1997] and for inking the production of lines in an electron beam resist with 50 nm spacing using a thiol monolayer on gold and imaged with electron microscopy [Biebuyck 1997]. These results make believe that printing techniques have the potential to deliver low cost and very fast method for large throughput VLSI. The ultimate resolution is in the order of 10 nm and is limited by the radius of gyration of the polymer / ink, while the practical limit of both techniques must be about 30-50 nm mainly limited by the material properties of the polymers / ink (see also table 3.2). Printing techniques permit positive and negative patterning. Being a replica process, effort has to be spent to produce good and cheap nano-patterned masters. Other technical bottlenecks are how to overcome the alignment tolerances between layers and how to increase the stability and robustness of the polymers. These and other problems put doubts on the possible throughput as currently the cure times of the resist layers and the polymers permit only a slow manufacturing process. Moreover, it is not yet clear how much the size stamps can be increased, while offering good nano patterning, but it seems that these parameters can be optimised in the future [Hoffmann 1998]

3.2.2. Bottom Up Approaches

Generally speaking there are two ways to fabricate electronic circuits with nanometer dimensions. The most frequently used approach employs a variety of sophisticated lithographic and etch techniques in order to pattern a substrate; this is referred to as the top-down approach. The second method is the bottom-up approach that builds small structures from the atom, molecule, or single device level upward. This latter method allows in principle a very precise positioning of collections of atoms, hence functionalities. At present, however, the top-down approach has substantially better development than the bottom-up approach. It is fair to assume that for the fabrication of structures with nanometer dimensions, the bottom-up approach will play an essential role in the near future.

In order for the bottom-up approach to really become competitive, the following issues are important. Firstly, the classical ways of synthesis, either organic or inorganic, are very cumbersome if well defined, monodisperse, species of nanosize particles have to be made. A key role is foreseen for self-assembly strategies, in which information is present in a limited number of building blocks that allow the spontaneous formation of such well-defined, monodisperse species. A large number of groups now have methods exploiting non-covalent interactions in self-assembly strategies. The number of strategies, however, that allow the introduction of electronically addressable functionalities is still limited.

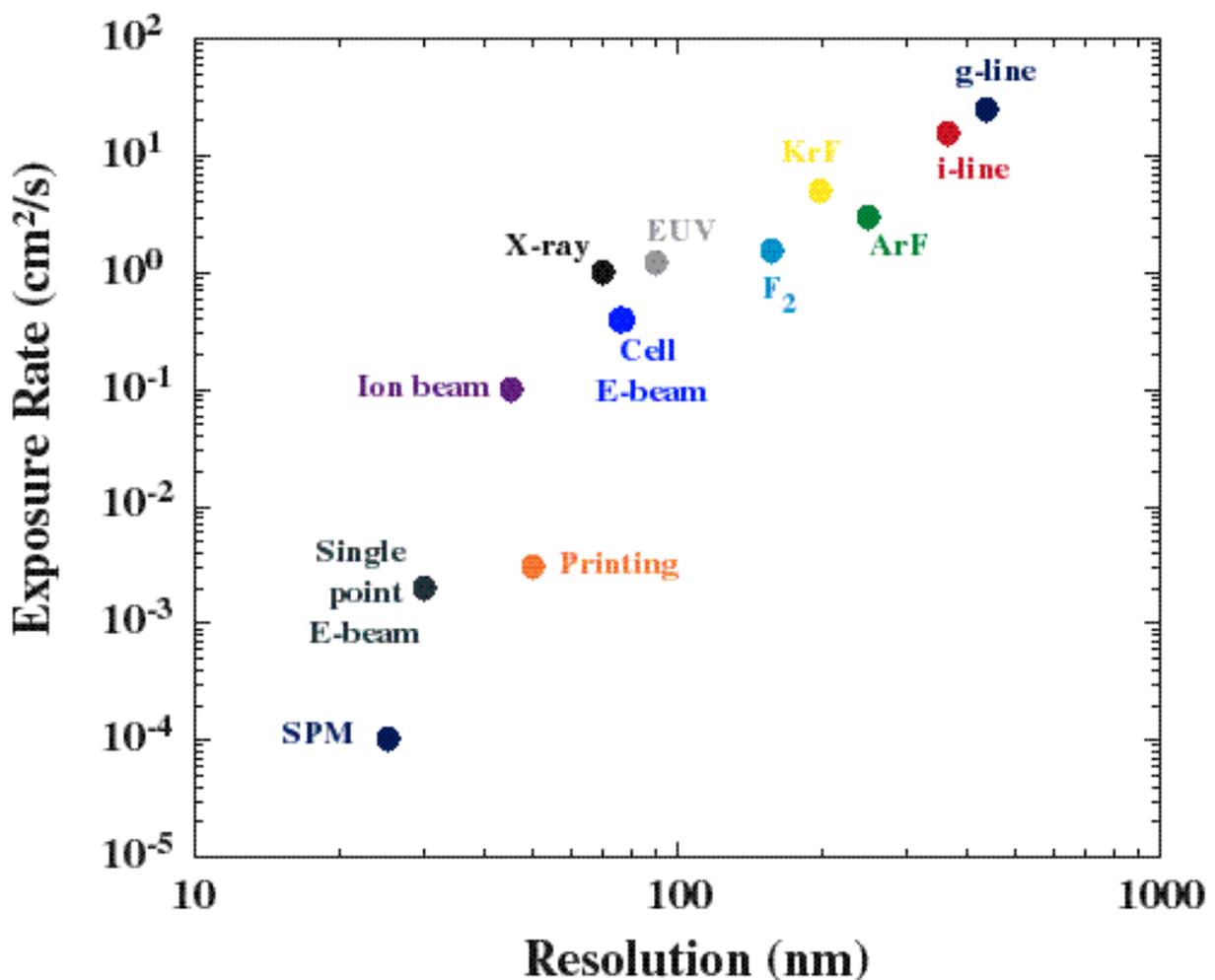


Figure 3.2: Comparison of different lithographic techniques [reworked from Asai 1997]

Secondly, the self-assembled functions have to be positioned in space. Also, here the information should be present in the assembly that allows their own confinement at the desired place. This process could be called self-organisation. The advancement in this respect is less than in self-assembly. An important key is probably played by self-assembly of monolayers on (flat) surfaces and at the same time self-assembly of the various species in this layer.

Thirdly, current integrated circuits are highly complex and therefore self-assembly and self-organisation strategies have to be developed that enable the easy formation of complex patterns. The fabrication of patterns with nanometer dimensions forms a real challenge for material scientists.

Most aspects of assembly and self-organisation, such as positioning of functional groups and their recognition properties, are better developed for organic materials than for inorganic ones, whereas the electronic properties of the latter are much better understood. It is, therefore, envisaged that especially the combination of the different material types will lead to fruitful new approaches. It may be the case that assembly will take place via recognition of organic parts followed by derivatization with inorganic materials, or that small inorganic units modified with organic anchoring points are assembled into circuits in a single process.

The ultimate dimensions of integrated circuits will be defined by the dimensions of atoms and molecules. At ambient temperature it is probably easier to confine a molecule through its functional groups than a single atom. The development of fast methods for performing chemistry on single molecules is therefore important. The various probe techniques, like AFM, STM, SNOM, etc. will play a key role in this field.

At the ultimate limits of fabrication, the bottom up molecular and atomic approach is advantageous in terms of achievable density and miniaturisation, and the following trends can be observed: at present a review of the literature reveals many designer molecular systems. Much less work is evident on the interfacing, demonstration of functionality and computation of their behaviour. It is in this crucial area of interfacing and demonstration of devices that the future progress of molecular nanoelectronics lies. Therefore self-assembly technologies have become an increasingly popular way of making nanostructures. Many examples of self-assembly can be found in the world of organic chemistry, and the ancient technique of fabricating Langmuir-Blodgett films has been successfully employed to create structures with specific transport behaviour. Moreover, the chemical synthesis of molecules can be viewed as a “self organising” way to make large numbers of completely identical nanoscale objects. Also in compound inorganic semiconductors, more well-known to the semiconductor world, self assembly is a hot research challenge, especially for the fabrication of quantum dots.

For all bottom up approaches, it should be realised that for creating useful transport devices, the self-assembled structures must be connected to the outside world in a sensible manner. This issue does deserve a lot more attention than it has enjoyed so far, as it may be a key impediment for use in commercial applications.

3.3 Comparison of Fabrication Techniques

Present day fabrication, as described above, relies almost completely on the top-down approach, using the high throughput optical lithography for dimensions down to 180 nm and electron-beam lithography down to 30 nm with a low throughput. Integrated circuit applications needing high throughput such as MPU and DRAM will very probably use optical lithography down to 100 nm. There is no doubt that optical lithography is the preferred industrial technique as it is well known and understood by industry. Figure 3.2 demonstrates throughput (by exposure rate) against resolution for comparison [reworked with new data from Asai 1997].

The situation in the range 100 to 50 nm is more open, and comparison between the extension of optics and the next generation lithography method is difficult because the development of these

methods is highly unequal. X-ray lithography and SCALPEL have both demonstrated good results in the literature but as yet still have their own problems which must be solved before it becomes clear which technology will be implemented below 100 nm feature sizes. Estimation of costs give a small advantage to X-ray compared to optical lithography at 180 nm, and the difference becomes appreciable at smaller dimension. An essential part of this interesting ownership budget is the excellent single resist process given by X-ray lithography, while the process cost of optics becomes high. But optical lithography benefits by having the advantage of being an established technology, it has an easier n:1 magnified mask fabrication, and will go on for some generations. A detailed comparison with the other next-generation lithography methods, which are still in the feasibility demonstration stage, is not yet possible. It is still worth investigating all the low cost emerging top down technologies. The introduction of lithography in fabrication might strongly depend on the time of availability of the methods. In addition, whether industry can afford the overall factory costs down to a 50 nm technology is not clear.

It is interesting to note that the dimensions supposed to be reached by these methods (30-50 nm) cover the needs of most of the device demonstrations of nanoelectronics (RTD, RSFQ, MRAM). Single electron electronics is a clear exception, needing controlled dimensions of 1-10 nm. In that case, the bottom-up approach is the only available solution, coupled to larger structures made by the top-down approach. It shall be largely the basis of the exploding field of nanotechnology.

4. Circuits and Systems

Topics of this chapter are design strategies, a review of the current state of novel logic circuitry for resonant tunnelling devices and QCAs, as well as an evaluation of novel trends in computer architectures. In this context the principal objective is to analyse in which way any of the technologies investigated in the scope of MEL-ARI, that is QCAs, RTDs, SETs or RSFQs, favours an implementation of a special kind of novel computer architecture.

The starting point of architecture evaluation is RISC machines, reconfigurable processors, artificial neural networks as well as more speculative concepts such as DNA and quantum computing which are beyond a fabrication using conventional solid-state technology. In each case the background is the current state of CMOS VLSI logic chips having 10 million transistors per die, an enormous experience of 30 years in solid-state technology, and the assumption that the integration density is expected to increase along with Moore's Law at a substantial rate. Even by the stage of 30 million transistors, the logistics of designing reticles for a 20 lithography-layer process are incredible considering the amount of data handling required.

Combining this with testing the reticles before and after manufacture provide a serious challenge to manufacturing at high integration levels. It is possible that this may be one of the limiting factors in production before lithography, dielectrics, diffusion of implants, etc. The ability to reduce the number of transistors either by new designs of logic or through multivalued logic elements would reduce transistor counts with benefits in increased performance, lower power and larger circuit yields.

A general property of many of the nanoscale devices outlined above is that they can be scaled down to the order of tenths of a nanometer. This is their main advantage over MOS devices and makes them attractive as possible building blocks of electronic circuits. Being able to downscale, however, is only one problem facing the development of nano-scale integrated circuits, the following challenges have to be taken into account:

The interconnect bandwidth problem is predicted to be a major problem. 40% of the power dissipation in a DEC Alpha CPU at present is related to distributing the clock around the chip. New architectures will be required which minimise the RC time constants of the interconnect layers. This is predicted to be required at 70 nm feature size [Iwai 1998] although the 1997 SIA Roadmap suggests 100 nm. As the number density of transistors increases, the testing of circuits becomes very difficult. Testing in some ASIC production already accounts for 60% of total costs even for 250 nm CMOS [Siemens, ITG, Hannover, 1998] and may get substantially worse as integration densities increase. Self-testing architectures are a necessity for the microelectronics industry. Defect tolerant architectures [Heath 1998] will also be required at large integration densities and certainly if technologies with poor manufacturing tolerances such as quantum dots are to be used.

4.1 Design Strategies: Interconnect Problems and Design Complexity

The economic success of the semiconductor industry during the past thirty years is primarily based on the interaction between technological progress in device fabrication and the ability to combine this technological progress with adequate logic families, circuit techniques, system architectures and software. To develop novel products from the engineering point of view, the availability of an appropriate design methodology with different levels ranging from devices to software was always of fundamental relevance and is the basis for computer aided design (CAD) tools. A serious limitation here is that the overall design productivity has been increasing by 21% per year while integration density has increased by 68% so that the economic success of nanoelectronics is not ensured even if the majority of device and circuit related problems could be solved [Hamilton 1999].

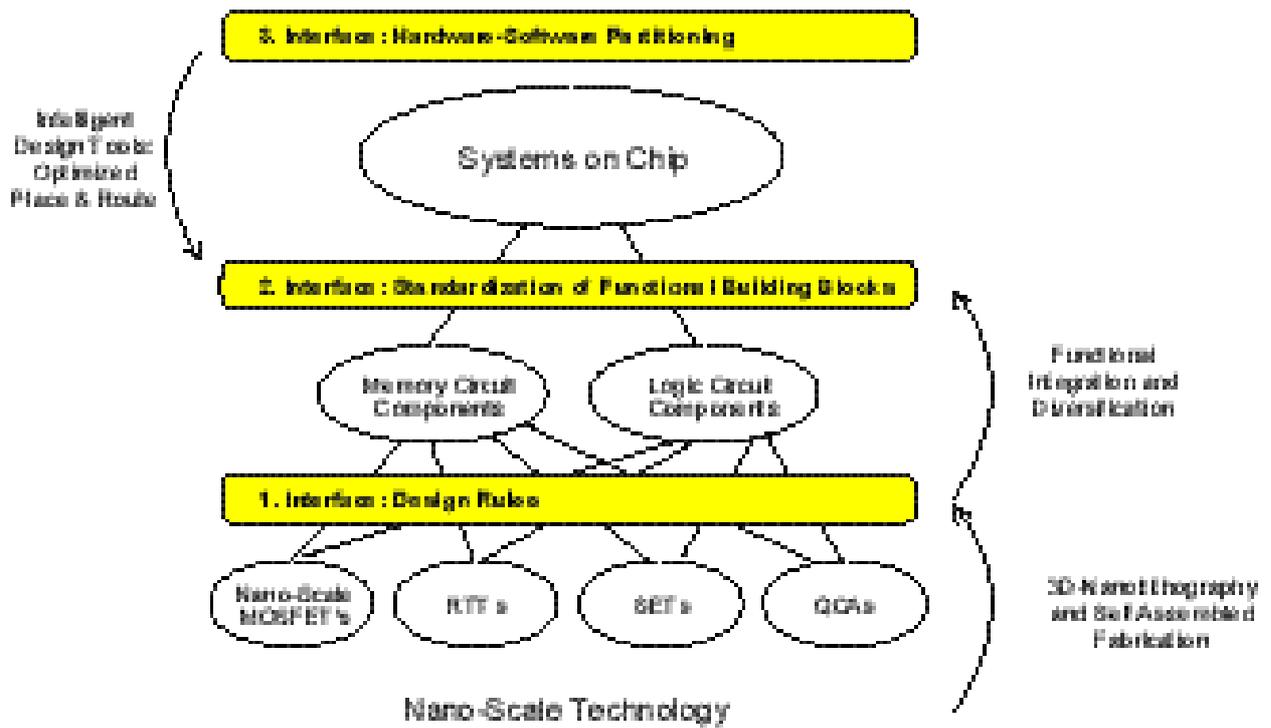


Figure 4.1:- Design hierarchies for nano-scale circuits.

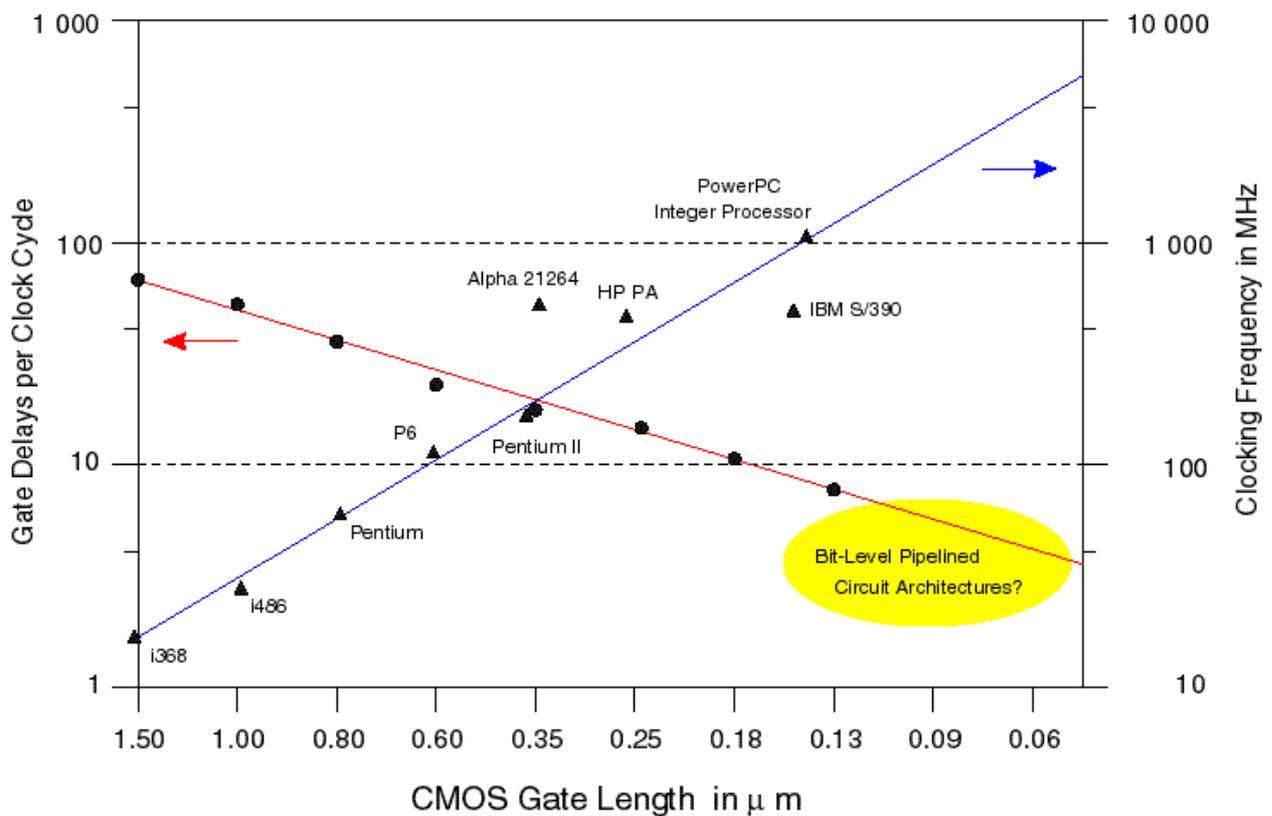


Figure 4.2:- Increasing clock frequencies and reduction of the number of logic stages (gate delays) between two synchronising latches in the data-path. Different commercial and prototype microprocessor generations are shown as reference points.

It is very likely, however, that the current approach in microelectronic system design of defining several interfaces between different design levels will also be relevant for nanoelectronics (Figure 4.1). In addition, due to the different behaviour of the devices, the interaction between the device level and the circuit level has to be strengthened to exploit for example a certain I-V characteristic (“functional circuit design”) in a circuit or to make circuits robust against device parameter variations. A prerequisite of such an interdisciplinary co-operation is the ability of the device physicist to understand the problems of the circuit designer and vice versa.

4.1.1 Impact of Increasing Clocking Frequencies on Nanoscale Circuits

Among the various problems in the field of nanoelectronic circuit design, increasing clock frequency is a good illustrative example of the way in which the complete system design is affected by novel devices, because operating frequencies in the GHz range have an impact on nearly all design levels. Due to shorter clock cycles the system designer has to change the architecture so that specific tasks, such as instruction decoding or memory access, are completed within a shorter time. Simultaneously the consideration of high frequency effects like crosstalk between several interconnecting wires is in a certain way contradictory to the high level treatment in designing a 1 billion transistor chip. Principal challenges for the circuit designer will be to choose an appropriate logic style and to ensure that a useful computation can be performed in spite of the fact that the number of logic gates between two synchronising latches rapidly decreases to 4-6 gates for 60 nm CMOS [Borkar 1998] (figure 4.2). In the case of nanometer-scaled RTDs or RSFQs, a value below 3 gate delays could be expected. A consequence for the technology is not only to optimise single transistors but also to be sure that the advantage of a high speed device is transferable to the circuit level, where interconnects and gate fan-out play an important role.

Solutions for these problems are under investigation for various circuit architectures and technologies. Examples are smart routing, systolic arrays and bit-level pipelining for SETs [Ancona 1996] and RTDs [Pacha 1999], as well as wave pipelining [Burlinson 1998]. To address the performance bottleneck of interconnects, delay tolerant circuit architectures, and the use of asynchronous sub-units while maintaining a global chip synchronisation are also discussed [Van Berkel 1999].

4.1.2 Local Architectures and Synchronisation

Locally interconnected circuit architectures allow devices with small gain to be used (such as many quantum devices). Such circuit architectures also remove interconnect bandwidth problems. Problems with synchronising clock pulses over large chips especially at high clock frequencies suggest that local architectures may have to be used with local high speed computing and lower (non-local) clock speeds for communication around chips similar to present CPU - logic board bus speed dividers. Some suggestions for realising circuit function using “local” circuit architectures for SET show great promise [Korotkov 1995, Ancona 1997]. The basic implementation of the QCA architecture, based on the spontaneous relaxation to the ground state, would not require any clock distribution, at least within each “local” functional unit; more complex implementations that have been proposed would require multi-phase clock signals, trading off simplicity for performance. There is a strong need to demonstrate an actual hardware implementation of one of these types of schemes.

Coulomb blockade phenomena (electron pumps, turnstiles, single electron transistors) are often regarded as particularly suited to these “local” circuit architectures. Accurate simulation of the electrostatics of an actual circuit layout, with all interactions taken into account, can be very difficult. The problem of random offset, or background charges remains a stumbling block for the implementation of any of these schemes.

4.2 Novel Logic Circuitry

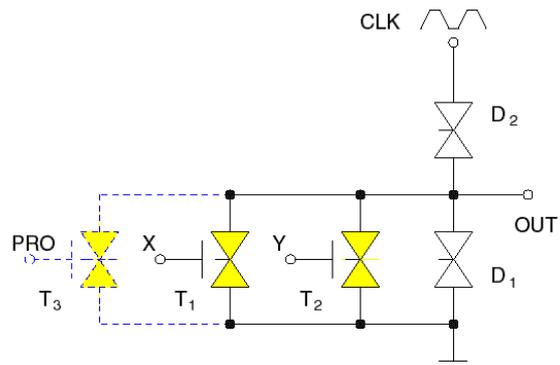
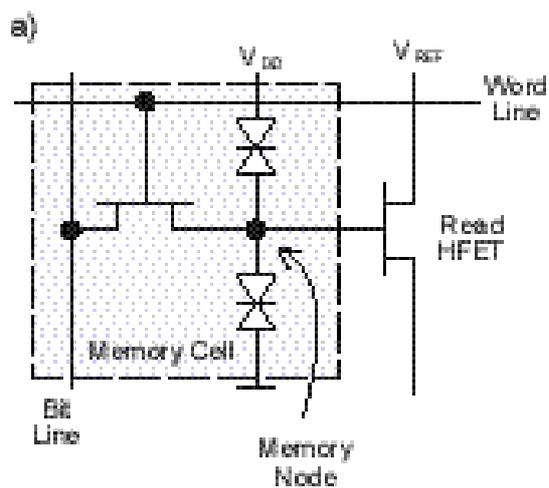
4.2.1 Resonant Tunnelling Device Circuits

The intrinsic bistability of resonant tunnelling devices was mentioned in chapter 2. In recent years the monolithic integration of resonant tunnelling diodes with heterostructure field effect and heterostructure bipolar transistors has led to a number of new circuit concepts [Mazunder 1998]. Many of these have been demonstrated and have achieved GHz clocking frequencies, a small power delay-product, and a significant reduction of circuit complexity by reducing the number of devices which are required to perform a specific computation. An example of a compact memory circuit is a low power 50 nW RTD/HFET SRAM cell with a 150 μm^2 footprint (200x lower power than a high speed GaAs SRAM) [van der Wagt 1996] (figure 4.3 a,b). A comparison of tunnelling based memory concepts implemented in III-V and Si-RTD/CMOS for high performance memory components is given in table 4.1.

Memory Technology	Speed (GHz) 1-2 μm pitch	Density (Mbit/cm ²)	Standby power (W/Mbit)
E/D HFET SRAM	1.0	0.1 - 0.4	0 - 200
H²L HBT SRAM	0.05 - 0.2	0.03	200 - 1000
1-T HFET TSRAM	0.5	1.7 - 4.4	0.01
2-T HFET TSRAM	1.0	1.0 - 2.0	0.01
Si SRAM¹⁾	0.03 - 0.1	5-10	0.01
Si DRAM²⁾	0.01 - 0.02	60-150	2 10^{-4}
1-T Si TSRAM	0.01 - 0.04	50 ³⁾	10 ⁻⁹

Table 4.1: Performance predictions for tunnelling based SRAM cells (TSRAM) and comparison with Si-memory (based on [Seabaugh 1998]). 1T(2T) refers to 1-transistor (2-transistor) cell(s).
¹⁾ 16 Mb CMOS SRAM ²⁾ 256 Mb DRAM ³⁾ based on layouts with 256 Mbit CMOS/RTD DRAMs.

For digital applications the self-latching behaviour of RTD circuits is attractive to implement high-speed dynamic logic families. The basic idea is to store information as a bistable state in contrast to conventional dynamic logic where the digital output of a gate is represented by the amount of charge on the parasitic load capacitances. By embedding a logic stage into a self-latching RTD circuit with pulsed power supply (monostable-bistable transition logic element, MOBILE) [Chen 1996] bit-level pipelined logic architectures are the preferred logic style. Their characteristic features are reduced circuit complexity and great design flexibility, as shown in figure 4.3c for a programmable NAND/NOR logic gate. Together with novel logic schemes, such as multiple-valued and threshold logic, the critical path of the circuit can be reduced. An example for the capabilities of threshold logic is the compact full adder design composed of two gates using a dynamic MOBILE based, threshold logic circuit [Pacha 1999].



Design Parameters of the RTDs

Function	RTD D ₁	RTD D ₂	RTT T ₁	RTT T ₂	RTT T ₃
NOT	1.0	1.5	1.0	-	-
NOR	1.0	1.5	1.0	1.0	-
NAND	1.0	2.5	1.0	1.0	-
NAND/NOR	1.0	2.5	1.0	1.0	1.0

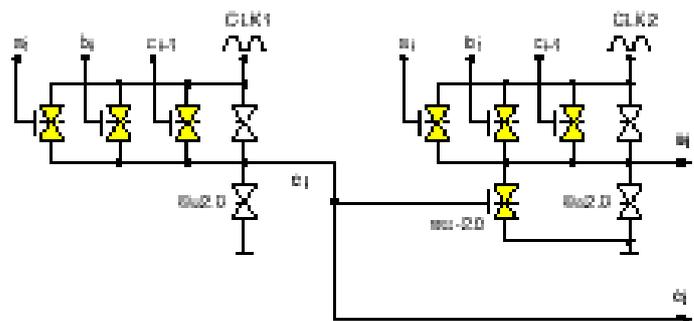
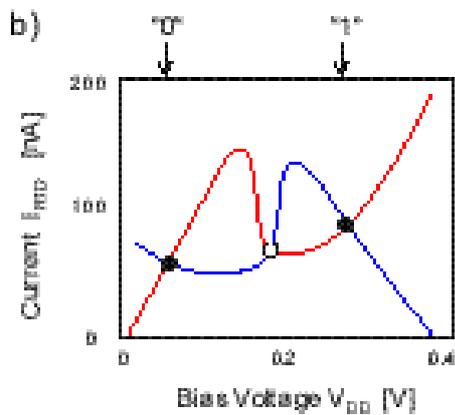


Figure 4.3: RTD/HFET circuits: SRAM cell (a), load-line diagram of a bistable RTD circuit (b), MOBILE based programmable NAND/NOR logic gates with relative RTD areas (c,) and a full adder design composed of two threshold logic gates (d).

Further demonstrated circuits are:

- ❑ generic logic circuits operating at 12 GHz with 20 μm minimum feature size [Williamson 1997]
- ❑ multivalued logic circuits [Mizuta 1995]
- ❑ monolithic 4-bit 2 Gbps analogue to digital (ADC) converters [Broekaert 1998]
- ❑ 3 GHz clocked quantisers with 40 dB spur free dynamic range
- ❑ 40 GHz static binary frequency dividers [Umeda 1995]
- ❑ clock circuits, shift registers
- ❑ 2 GB/s photodetectors with low switching energies of 30 fJ

In comparison with CMOS the potential advantages of a RTD-based approach can be estimated by an investigation of circuit designs for basic logic and arithmetic functions. For CMOS, logic families such as NORA logic and NP-Domino CMOS are the benchmarks because they are currently used for high performance logic [Rabaey 1996, Jou 1997].

Table 4.2 shows several characteristic features of different 2-bit adder designs, including the number of devices as well as an approximated equivalent area A_o . It is assumed that the devices are designed with minimum feature size A_o whenever possible. The area overhead due to interconnections has been neglected.

RTD designs can offer a reduction in component count by up to 40% when compared with the equivalent CMOS logic family. Due to the threshold logic design style, which enables an increased low-level parallelism of computing multiple input signals, the advantage of the RTD circuits increases when equivalent circuit areas are compared. The main reason is that the threshold logic design eliminates the series connections of transistors in NAND-like gate configurations. Particularly in CMOS, long series connections of p-MOSFETs result in a large area consumption. The estimates in table 4.2 are in agreement with some experimental circuits such as shift registers and multiple value to binary decoders [Mazumder 1998]. In addition to reducing the number of active devices, the logic depth of a circuit is important to estimate the data throughput. Since the 20 GHz operation of a RTD-Schottky Diode Logic family has been reported, a data throughput of about 1-10 Gbit/s should be achievable for larger circuits such as adders and multipliers for digital signal processing applications.

Circuit Family	Device Count	Area A_o	Logic depth 2-bit adder	Logic depth MSB carry	Clock Freq. (GHz)
Non-pipelined NP Domino CMOS	56	808	4	3	0.2
Non-pipelined RTD adder (fig 4.3c)	22	224	3	2	5
Pipelined NORA-CMOS	48	592	4	3	0.8
Pipelined RTD adder	40	336	4	2	5

Table 4.2:- Comparison and performance predictions of 2-bit adder designs.

At the current state of technology with minimum lateral feature sizes of about 1 μm and RTD peak currents of about 5 mA, static power dissipation limits the integration density. If RTD circuits, however, could be scaled to minimum lateral feature sizes of about 200 nm, one may expect that dynamic power dissipation due to the charging and discharging of load capacitances will become the major source (90%) of power dissipation in the GHz frequency range. In that context adiabatic clocking schemes are one method to limit the dynamic power dissipation.

At this point a critical constraint on overall system performance is the increasing gap between memory access time and processor (logic) speed. Among the various possible technological needs, scalable memory architectures having a constant memory cycle time regardless of the memory size are required. Here the complete memory array is divided into $N \times N$ memory blocks, which are operated in a pipelined (systolic) fashion so that the memory cycle is limited by the access time of a single memory block and a high data throughput is obtained [Jeong 1998]. Furthermore the systolic operation avoids long word and bit lines crossing the complete memory array. When searching for suitable architectures these scalable memories are an important innovation and should be taken into consideration for high performance SRAMs implemented with nanoscale RTDs or SETs.

4.2.2. QCA Circuits

Logic circuits based on the Quantum Cellular Automata (QCA) paradigm offer an interesting alternative to traditional architectures used for computation. The first consistent scheme the “Notre Dame architecture” for the implementation of logic functions with two-dimensional QCA arrays was proposed in 1993 [Tougaw1993].

The basic building block of the Notre Dame architecture is a cell made up of four or five quantum dots, containing two electrons, which can align along one of the two diagonals. Coulomb repulsion between the electrons in a single cell causes the charge in the cell to align along one of two directions, giving rise to two possible “polarisation states” (see Fig. 4.4), representing binary data. It has been shown that cell polarisation can propagate along a chain of cells, minimising the electrostatic energy and that properly assembled two-dimensional arrays of cells allow the implementation of logic functions. The results of any computation performed with such arrays consist in the polarisation state of some output cells.

The Cellular automata is a promising concept for possible circuit applications of quantum devices because it overcomes some of their intrinsic problems, such as the limited available fan-out and the inability of efficiently driving interconnect lines. Problems such as controlling the number of electrons in each cell or interfacing the cellular automata system with the outer world and, in particular, with conventional electronics without perturbing its operation, do exist, but it has been determined that cell operation without any performance degradation is possible using $4N+2$ excess electrons in each cell, N being any integer. Contrary to some existing hypotheses, there is a seamless transition from two-electron cells to cells made with metallic dots, which contain a very large number of electrons. An example of the logic functions that can be performed with QCA arrays is the majority voting gate (Fig. 4.5). The three input configuration is given by the polarisation states of the three input cells, and the output cell polarises along the direction corresponding to the majority of the inputs, in order to minimise the electrostatic energy.

There are two main problems to be faced when implementing a QCA circuit: the need for individual adjustment of each cell and the limits on the operating temperature. Individual adjustment of each cell is needed because of fabrication tolerances, the presence of stray charges, and the

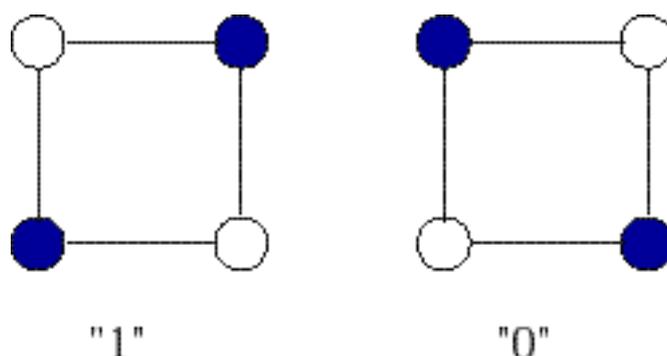


Figure 4.4:- The two polarisation states possible in a QCA cell.

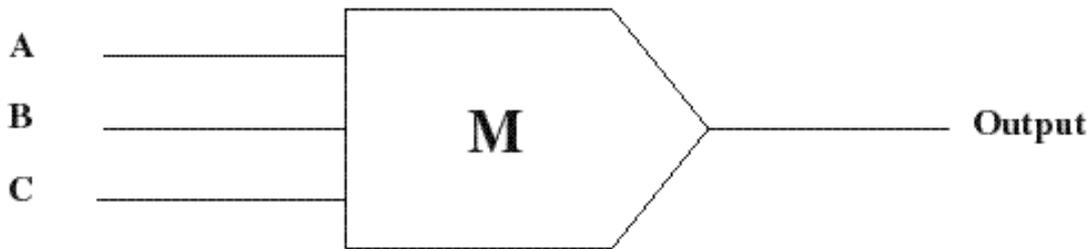
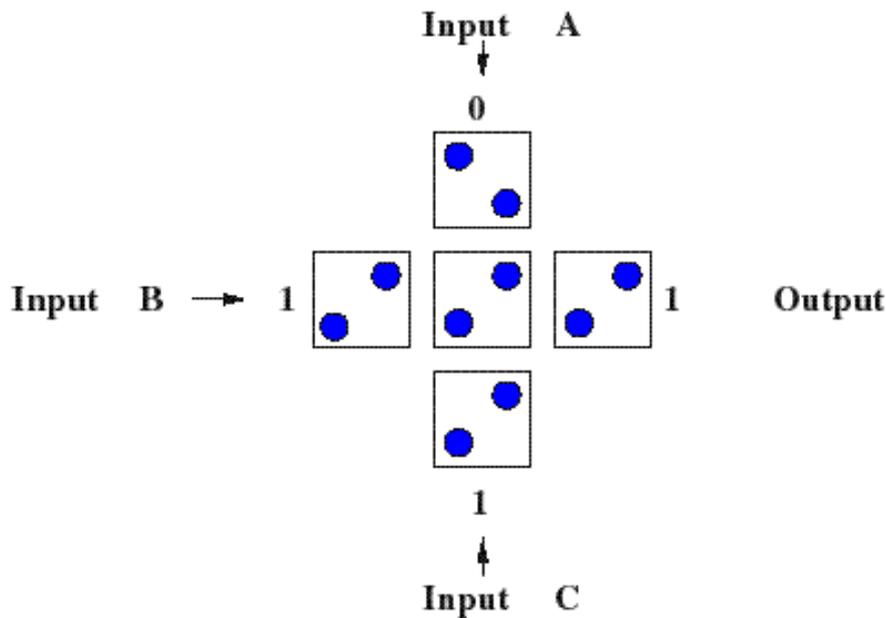


Figure 4.5:- An example of the logic functions that can be performed with QCA arrays is the majority voting gate

need for exactly $4N+2$ excess electrons in each cell. In order to make adjustment possible, leads are needed, reaching each single cell. Such leads are also required to load onto the quantum dots the required number of excess electrons. The presence of leads prevents straightforward lateral branching from a chain of cells, one of the basic features, which would be needed to create logic gates.

Limitations on the operating temperature are of a more fundamental nature and are due to the weakness of the dipole interaction between the cells, which must be significantly larger than kT . The maximum operating temperature for a chain of six cells in silicon-on-insulator (thick line) and in gallium arsenide (thin line) is plotted in Fig. 4.6, assuming a readout time long enough to filter out fluctuations and correctly detect the output for an average polarisation that is 20% of the maximum. Improvements could be achieved by shrinking the cell dimensions and by using materials with lower dielectric constant. Structures at the molecular level will be needed to approach room temperature operation.

The current state of the art in QCA devices is a majority voting gate based on metal-insulator tunnel junctions, operating at a few millikelvin [Amlani 1998]. The demonstration of a silicon-based cell with a potential for operation at a temperature above 1 K and of a cell defined by split gates on a GaAs/AlGaAs heterostructure is currently being pursued. Once the technology for a single cell has been developed this can be applied without substantial modifications to the realisation of chains of cells, while the implementation of logic gates will require a solution to the abovementioned problem of lateral branching.

From the architectural point of view, the main advantage of QCA systems is to get rid of interconnect lines since operation relies on the interaction between adjacent cells. In addition, the absence of charge transport allows low power consumption. Potentially QCA may consume very

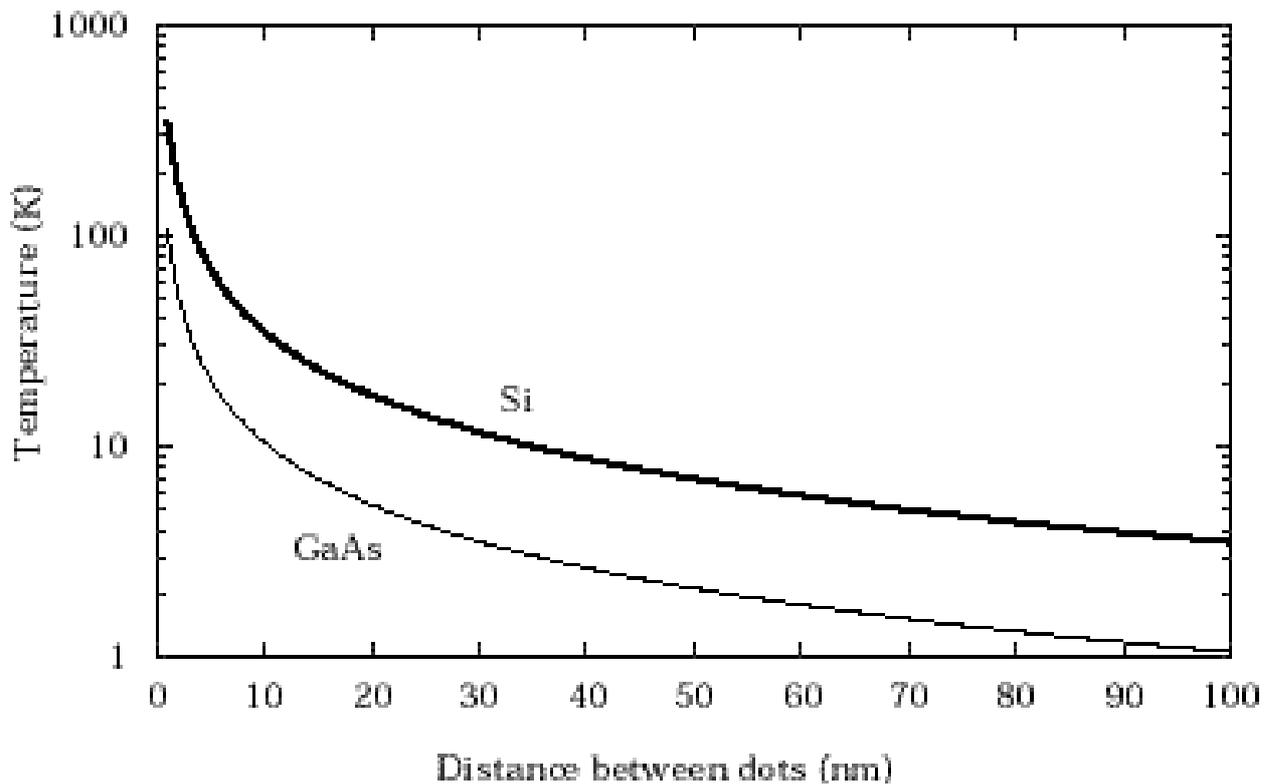


Figure 4.6:- Calculated operating temperature for a Si and a GaAs chain of 6 QCA cells as a function of the distance between the dots.

low power benefiting from a ground state computation paradigm: once a given polarisation is enforced onto the input cells, the system relaxes to the ground state and each output cell assumes the polarisation state which minimises the total energy. However, the time required for relaxation towards the ground state cannot be reliably predicted, and, if metastable states exist, the system could get stuck in one of them for a significantly long time. A solution out of this dilemma would be to perform a so-called “adiabatic switching” consisting of slowly varying the height of the potential barriers separating the quantum dots in a cell. In this way, the system evolves remaining always in its instantaneous ground state, as guaranteed by the adiabatic theorem [Lent 1997].

So far, the major obstacles hindering the production of a QCA architecture on a large scale integration are the need for individual adjustment of each dot and the severe limitations on the operating temperature due to the small value of the electrostatic splitting between the energies of the two polarisation states.

Operating temperatures reasonably close to room temperature require small cell dimensions, of the order of a few nanometers. Cells could be built upon single bistable molecules, i.e. molecules which in the absence of an external electric field some of the valence electrons can localise in two different groups of bonds with the same probability. By analogy to the four-dot cells, the presence of an external electric dipole field would make one of the two possible configurations energetically favorable. The problem arises how to arrange the molecules in structured arrays on some substrate. But if successful it also solves the problem of fabrication tolerances, because molecules are intrinsically precise and bound electrons would be sufficiently localised to act as “excess” electrons, without the need for external leads. Stray charges will still represent a major issue, unless it becomes possible to develop substrates virtually free from them. Overall, when device dimensions are pushed down to the molecular scale, the QCA concept appears promising, since it is capable of exploiting basic molecular properties.

4.3 Current Trends in Novel System Architectures

4.3.1 Starting Point: Systems on chip and Innovations in Microprocessor Designs

The superscalar RISC (Reduced instruction set complexity) architecture is currently the dominant microprocessor architecture. Due to its prevailing position in the commercial microprocessor market it will be undoubtedly the starting point for all innovations in the direction of complete systems-on-chip as well as for more speculative architectures. The characteristic feature is an instruction level parallelism based on multiple execution units and the capability to simultaneously fetch and decode 2-4 instructions with a fixed (32 bit) format [Hennessey 1990]. Pipelining of subsequent instructions increases the implementation efficiency and the data throughput. To avoid pipeline stalls during conditional branches modern RISC processors often contain a branch prediction unit. On chip instruction and data cache is part of the memory hierarchy and the classical method to solve the processor memory speed bottleneck.

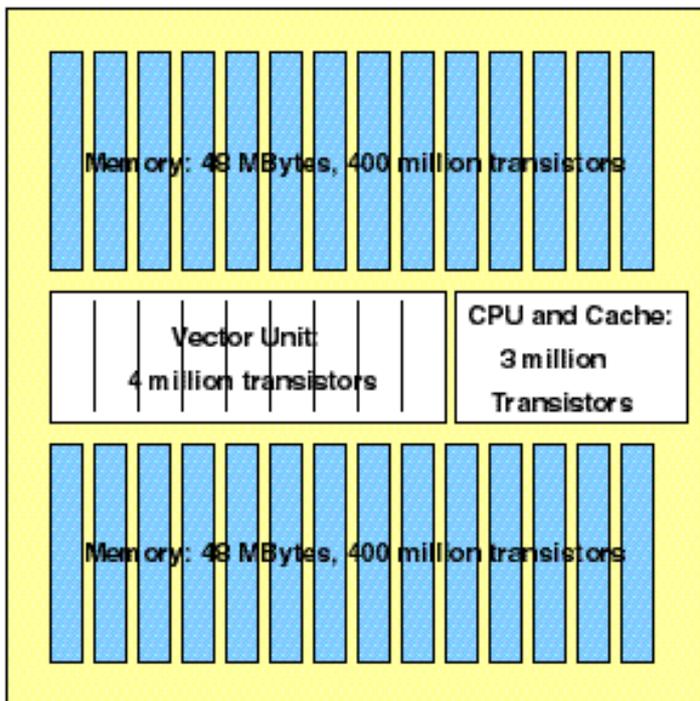


Figure 4.7: IRAM (Intelligent RAM) chip combining multiple vector processing units for multimedia applications, a conventional CPU and a large on-chip memory [Kozyrakis 1998].

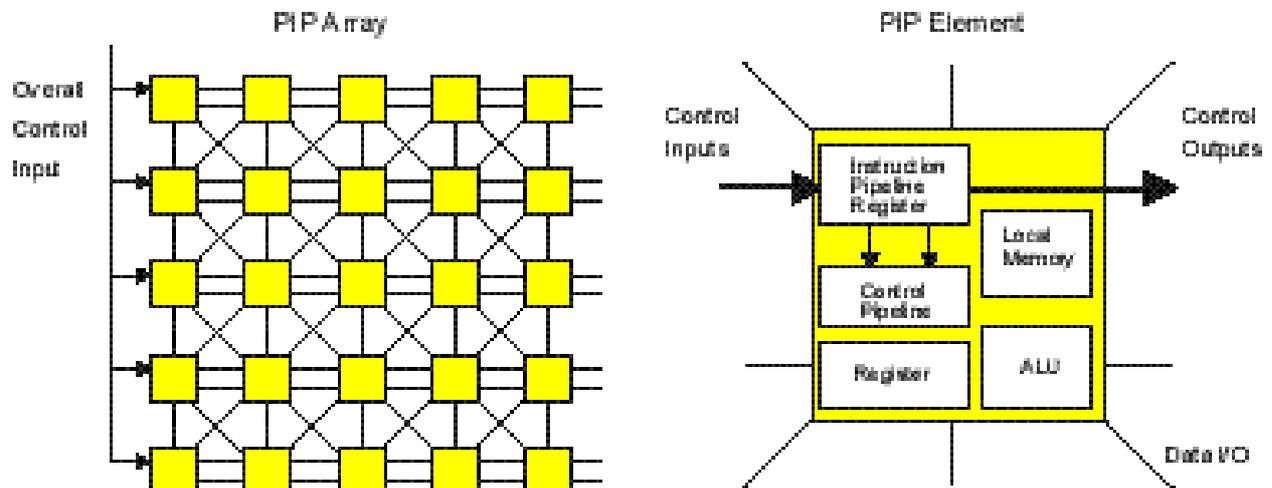


Figure 4.8: Propagated Instruction Processor [Fountain 1998a]

With the increasing relevance of multimedia data processing Very Long Instruction Word (VLIW) architectures and Single-Instruction Multiple Data type instructions are developed to enable an efficient real time computation of 8-16 bit audio and video data streams. Even for nanometer scaled CMOS more radical system architectures are being investigated to exploit the capability of integrating 1 billion transistors on a single chip. Based on technological innovations, such as systems on chip, RISC cores are combined with vector processing units and large on-chip DRAM. In that context the “Intelligent RAM (IRAM)” is a potential solution to limit the impact of the increasing DRAM-processor speed bottleneck and to increase the performance power dissipation ratio. In nanoelectronics the data-path-orientated, scalable design of multiple vector units is especially interesting because of its simple control logic and its modular design which ease the interconnection problem and the design complexity.

4.3.2 Parallel Processing

4.3.2.1 Propagated Instruction Processor

Among different concepts for parallel computing Single Instruction Multiple Data architectures (SIMD) are an attractive system architecture for nanoscale integration due to their regular design and local system-level interconnections. SIMD array processor is a well-known special purpose system architecture and has been primarily developed for image processing tasks where a simple mapping of the algorithm on hardware is possible. However, the classical SIMD architecture suffers from the drawback that the instructions for the different image processing tasks have to be transmitted in a global fashion to each processing unit.

The interconnection problem originating from the global instruction flow can be solved by a so-called Propagate Instruction Processor (PIP) [Fountain 1998a]. The PIP differs from a classical SIMD-processor in the sense that instructions are pipelined in a horizontal direction while the PIP units perform their computations (figure 4.8). Thus, multiple image processing algorithms are simultaneously solved in parallel. Without going into details, the pipelined operation would improve the overall system performance if high bandwidth communication with external components was available. To store the propagating instructions each processing unit contains a local instruction pipeline register. Further components are the local memory, a control pipeline, a register, and the arithmetic logic unit. The size and capabilities of a single processing unit depend on the complexity of the algorithms and the image resolution.

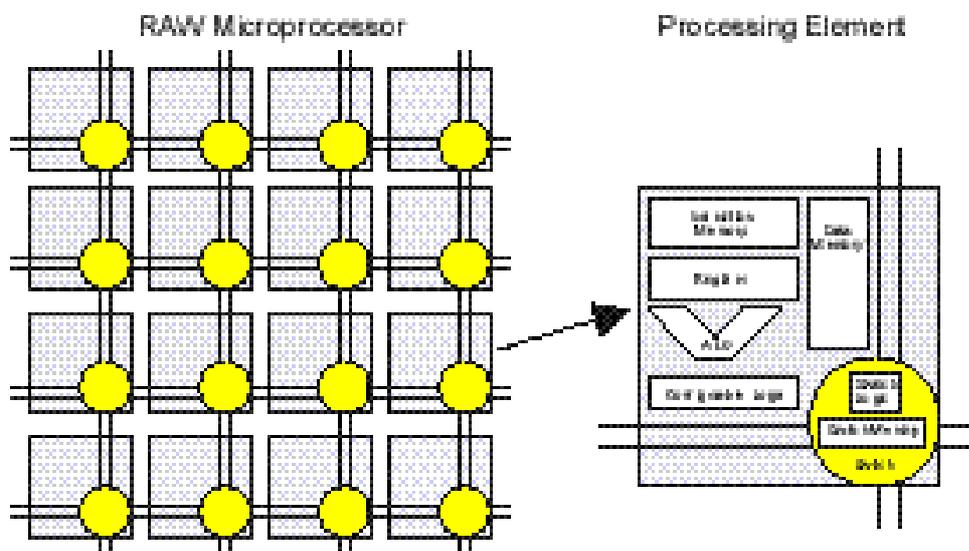


Figure 4.9: RAW machine: One-chip reconfigurable computer architecture. Each processing element contains a local instruction memory, data memory, registers, ALU, control logic. The SMEM block stores the configuration of the programmable crossbar switch.

4.3.2.2 Reconfigurable Hardware

The Reconfigurable Architecture Workstation (RAW) is a replicated architecture of identical processing elements. Each processing element contains a RISC-like pipeline, a local instruction and data memory [Waingold 1997]. The complexity of a processing element is comparable to a R2000 CPU and comprises about 2 million transistor equivalents. The interesting feature for nanoelectronics is that the programmer can also modify the logic functionality by means of the configurable logic blocks. In addition, the programmable crossbar switches allow it to implement different point-to-point interconnections. Since the configuration of the crossbar switches is stored in a memory block, the RAW machine is a dynamic and adaptive architecture. The implementation of these programmable interconnects requires about 30 percent of the total chip area.

Besides the advantage of the regular layout, the underlying idea of the RAW architecture is to transfer certain tasks from hardware to software. For example the amount of testing in a billion transistor chip could be reduced because reconfiguration and rerouting, in the case of defective elements, could be performed by a compiler. Furthermore for application specific tasks, a significant increase in speed over a pure software solution is achieved by exploiting the local configurable logic blocks.

4.3.2.3 Defect and Fault Tolerant Architectures

Defect tolerant architectures may be essential for computational nanoelectronics since this is probably the only way to build systems with billions of devices economically. Any computer with nanoscale components will contain a significant number of defects, as well as massive numbers of wires and switches for communication purposes. Such a concept, which differs significantly from the usual concepts of building complex computer systems, is the Teramac experiment [Heath 1998]: The Teramac is a defect tolerant custom configurable computer based on field programmable gate arrays (FPGAs). Although the idea is similar to the RAW machine, it differs in the sense that a computation is performed on the basis of programmable look-up tables instead of complex boolean logic circuits. Especially if a nanoscale circuit technology allows an efficient implementation of memory components while logic circuits are more difficult to fabricate, the look-up table approach of the Teramac could be an alternative to the traditional way of building a computer using logic circuitry.

More important, however, is the fundamental paradigm of the Teramac to build a computer cheaply and imperfectly, find the defects, configure the resources with software, compile the program, and then run the computer. Here, the underlying idea is a redundant interconnection structure, such as the fat tree architecture shown in figure 4.10. Based on this experimental architecture we can draw several important conclusions for nanoelectronic systems:

First, it is possible to build a very powerful computer that contains defective components, as long as there is sufficient communication bandwidth in the system to find and use the working components.

Second, the resources in a computer do not have to be regular but rather they must have a sufficiently high degree of connectivity and of technical intelligence

The third aspect addresses the issue that the most essential components for nanoelectronics may be the address lines and the crossbar switches that control the settings of the hardware configuration. For a defect tolerant system a self-testing architecture may be useful since the system itself can find out the defective components. To perform the reconfiguration of the logic interconnections (rerouting) in the case of defects, some additional hardware and a set of test vectors are required to identify defective components and to activate redundant hardware. Similar concepts are used to increase the yield of semiconductor memories by activating redundant memory cells. Therefore, the efficiency of the whole system can be improved by implementing technical intelligence.

Again, comparable to the RAW machine, the configuration software plays a dominant role, due to the paradigm of solving difficult hardware problems by means of software. The same strategy was the principal motivation for designing the first microprocessor and the continuation of this trend is very likely due to the existing knowledge and the economic success of the semiconductor industry.

In summary the Teramac suggests an entirely different approach to nanoelectronic research, it is a top-down approach, and not a bottom-up approach dominated by the technology. The key hardware components of this concept are wires, switches and configuration.

In addition to defect tolerant architectures, other techniques, such as triple modular redundancy, have to be considered to tolerate transient errors during operation, which is extremely important for QCAs and SETs in the presence of fluctuating offset charges [Ancona 1997]. At the current state of technology the detection of transient errors, that is fault-tolerance, seems to be a question of circuit design. In other words, the circuit implementation of the programmable switch in figure 4.10 should be robust against these transient errors even if the whole hardware has detected no defective components.

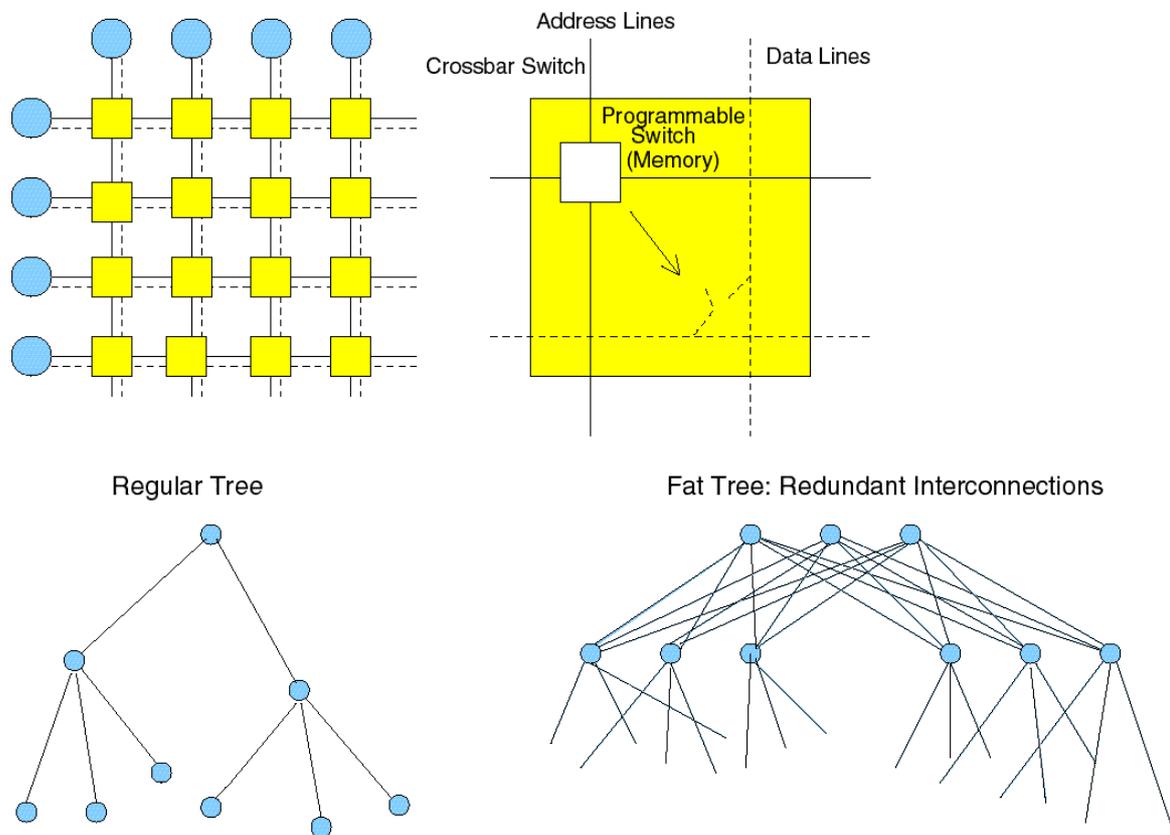


Figure 4.10: Redundant interconnections of the Teramac.

4.3.2.4 DNA Computing

Computations by chemical or biological reactions overcome the problem of parallelism and interconnections in a classical system. In DNA computing [Adleman 1994] where the similarities between mathematical operations and biological reactions are used to perform calculations, the key idea is to find the similarities between DNA (Deoxyribo Nucleic Acid) – the basic genetic information – and well-known digital computers. If a string of DNA can be put together in the right sequence, it can be used to solve combinational problems. The calculations are performed in test

tubes filled with strands of DNA and gene sequencing is used to obtain the result. Adleman calculated the travelling salesman problem to demonstrate the capabilities of DNA computing.

DNA computing on parallel problems potentially provides 10^{14} MIPS while using less energy and space than conventional supercomputers [<http://www.clearlight.com/~morph/dna>]. While CMOS supercomputers operate 10^9 operations per Joule, a DNA computer could perform about 10^{19} operations per Joule (10^{10} times more efficient). Data could potentially be stored on DNA in a density of approximately 1 bit per nm^3 while existing storage media as DRAMs require 10^{12}nm^3 to store 1 bit.

4.3.3 Artificial Neural Networks

Popular concepts of parallel and adaptive computing are the various forms of artificial neural networks [Haykin 1994]. Again biological information processing motivates this approach, but in contrast to DNA computing the learning algorithms are formulated as complicated sets of mathematical equations. Mostly, these equations are implemented in conventional hardware or in software if flexibility is more important than execution time of the neural algorithm.

Due to the fundamental relevance of the adaptive interconnection, these large arrays of simple neural processing elements show features such as association, fault tolerance, and self-organisation. Here, it has to be considered that on the one hand the complexity of a neural processing element is much lower than the complexity of the processing elements in the Teramac or RAW machine discussed above. On the other hand the interconnection density of neural processing elements is orders of magnitudes higher, so that the question - if artificial neural networks are appropriate for nanoelectronics - remains open.

As a starting point, however, a neural nanoelectronics architecture with a low interconnection density, such as cellular neural networks (CNNs) and associative matrices, are interesting [Goser 1997] even if their capabilities in regard to adaptation and generalisation are limited. CNNs are regular processing arrays where only the neighbouring cells are connected. Often these processing elements are implemented in analogue circuit techniques so that low power applications, such as intelligent sensor pre-processing are preferred applications. To a certain degree they can be regarded as low-level SIMD architectures and therefore suffer from the same disadvantage of a global instruction distribution. The associative matrix is a concept for a distributed memory and is based on digital input and output states. The learning rule is strictly local and only binary weights are used. Due to the distributed storage principle and a certain degree of fault-tolerance, the associative matrix could be interesting for a multi-Gbit RAM made of single electron technology.

The phenomenon of self-organisation is similar to self-structuring in technology [Feldheim 1998] aiming at producing nanometer scale structures without extensive processing steps from outside. The self-organising neural network maps the input data on a two-dimensional area without a teaching help from outside [Kohonen 1995]. This property would be useful in nanoelectronics for structuring an initially homogeneous array of cells.

Most of the present neural networks are static networks. In a first step they map the data according to learning algorithms. In the second step the adaptive interconnections are fixed and novel, unknown data is classified or associated based on the internal configuration. Alternative approaches are time dependent, biology inspired networks in which the input data is processed in a dynamic way using dynamical phenomena, such as spiking, oscillations, bifurcations, and local excitations. Local adaptation and excitations occur according to non-linear integral-differential equation for the time variant weight connections and the cell polarisation (excitation). These dynamic networks, known as resonant neural systems, not only show outstanding new features of artificial intelligence similar to small brains but also need a tremendous number of devices if an implementation is intended. At this point it should be noted that neural networks are a supplement and not a replacement of traditional data processing if a useful application can be found. In this way, a certain self-adaptation of neural networks can compensate the inherent drawbacks of nano-

electronic devices, for example fluctuating background charges of SETs are presently under investigation [Goossens 1998].

4.3.4 Quantum Information Processing (QIP)

Quantum computing is a fundamentally different way to process information which is not based on the Turing machine. The basic concept of QIP relies up so called Quantum-Bits (Qubit) which are a superposition of different eigenstates of a quantum system, like two-state spin systems [Steane 1998]. The states of different qubits can become “entangled”, that is, there exists a coherent interaction between two qubits, in a controlled manner and this can be exploited for calculations. The natural parallelism of superposition and entanglement enables, in principle, computations to be carried out that are not practical for classical systems, such as the computation of NP-problems. For example the factorisation of large integers would have a revolutionary impact on cryptography. Therefore, quantum communication may be one of the first applications of QIP. A recent discovery is the ability to extract a single entry from a large database of unordered information, which would yield an associative memory for speech recognition.

Although QIP systems have a huge potential, they might replace classical processors only if the system offers indisputable benefits over the existing devices. For this to happen, QIP systems would require registers for ten or hundreds of qubits. A medium-term objective are would be to develop an elementary scaleable quantum processor. Although the practical importance of an elementary (e.g. 4-qubit) quantum processor may be limited in itself, it may represent an important step towards larger-scale quantum computing. Quantum systems capable of handing a few qubits have been demonstrated, for example with NMR of small molecules or quantum photon interferometry. At the present stage, it is impossible to select one specific technology for the achievement of quantum computation, since all existing approaches are affected by rather severe limitations, but the winning technology must comply with the following factors: 1. Ability to control decoherence and perform fault-tolerant operations. 2. Scaling properties in terms of number of qubits, time per gate, and physical size. 3. Chances to become a low-cost, low-power technology potentially able to enter the market.

Solid state demonstrators do not exist, but various devices are investigated for this purpose, such as quantum dots, single flux superconducting devices or nanoscale NMR and ESR based devices. Theoretical models how these systems might work have been proposed [Kane], but for a successful implementation fundamental progress has to be made in technology of nanoscale devices. For example, the current purity of materials is good enough for microelectronics and partially adequate for some nanoelectronic devices, but for producing a solid state quantum processor the purity must be better by several orders of magnitude. In view of the large technological difficulties, it is likely that a solid state demonstrator will not be available in before 2003.

4.4 Comparison of the Various System Architectures

A rough overview of different architectures and circuits discussed in this chapter is shown in figure 4.11. Classifying different popular computational paradigms, some characteristic features that become relevant in nanoelectronics are the amount of wiring, the degree of parallelism, and the range of data transfer (figure 4.12). The difference between wiring and data transfer is that the first uses physical interconnections, that is, metallic wires, whereas data transfer may be realised without direct physical wiring. In this case, long-range data transfer is done by sending a signal from module to module, analogous to systolic structures. It is obvious that, today, parallel operation and locally interconnected modules are of less significance in traditional microprocessor-memory ar-

chitecture. In contrast, the basic elements of cellular automata and artificial neural networks process the given input data fully parallel. Additionally, figure 4.12 specifies the region of highly parallel operating and locally interconnected nanoelectronic systems. Since the MEL-ARI nanoscale integrated circuit cluster primarily investigates novel devices, the question of suitable architectures can only be considered indirectly. One important result, however, of the short review of the different systems is that each architecture relies on an efficient implementation of several basic circuit functions. In this connection table 4.3 summarises the corresponding key components and suggests an appropriate technology as well as a prototype circuit of lower complexity. For example the RAW machine and Teramac make intensive use of programmable crossbar switches and thus a prototype SET or RTD implementation would demonstrate that nanoscale technology is able to deliver such a key component. By combining experimental simulation results it then possible to extrapolate the performance of larger circuits and systems.

To handle the increased diversification and to cope with design complexity it is expected that intelligent design tools will become more and more important within a hierarchical design flow which is characterised by interfaces with design rules and standardised functional blocks.

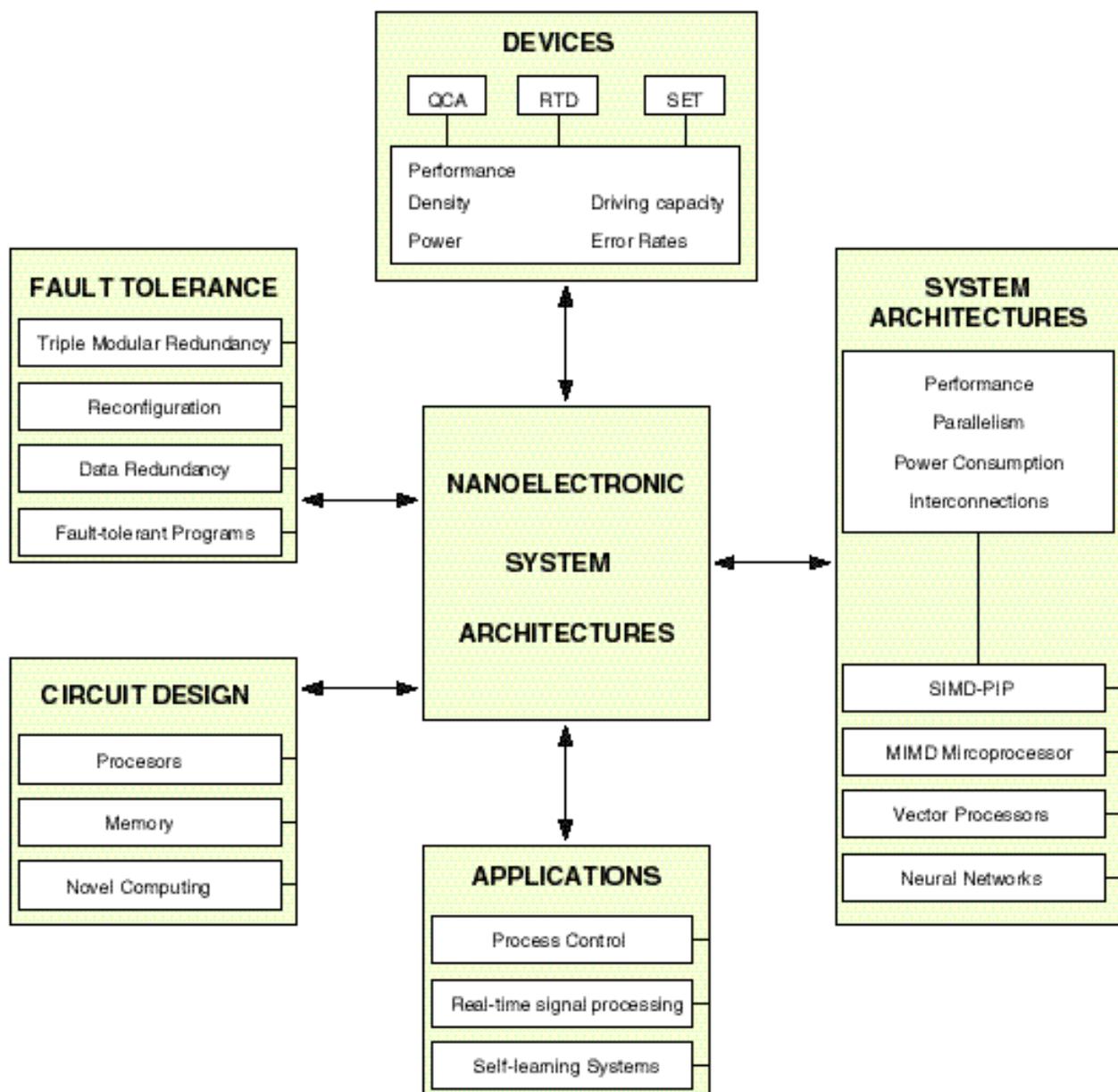


Figure 4.11: Aspects of different architectures for nanoelectronic systems [Fountain 1998b]

Table 4.3: Key components of possible nanoelectronic system architectures and prototype implementation of basic circuits for different technologies.

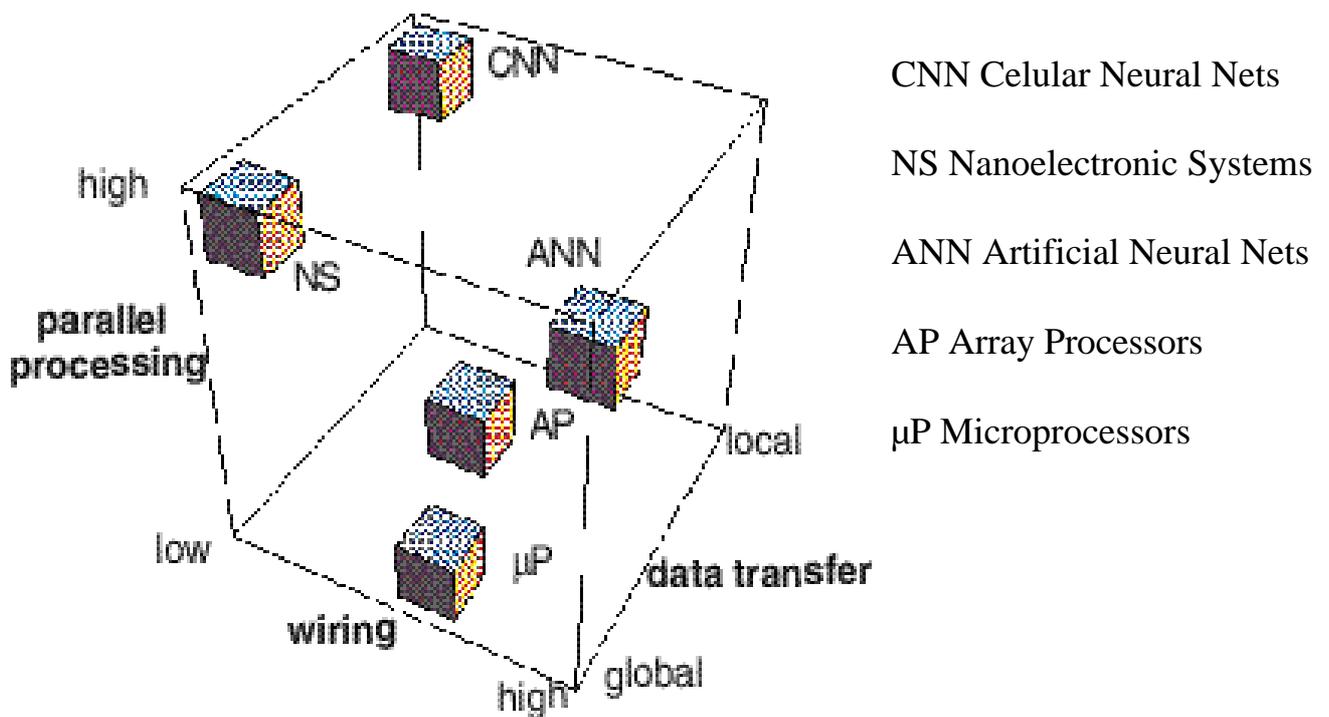


Figure 4.12:- Comparison between different architectures with respect to data transfer, wiring and parallel processing.

5. Conclusions

5.1 Introduction

In the past 15 years, nanoelectronics has developed as a versatile and fruitful playground for many innovative device concepts. Only a few of the concepts developed have been described in this Roadmap. Those working in the field have certainly witnessed that many novel ideas have not survived a first critical examination of their applicability on an industrial scale. The concepts we discuss here are for the largest part those that could eventually make it to applications, although some others may, within a few years, turn out to be simply the fashion of 1999. Notwithstanding these observations, we have compiled, to the best of our knowledge, a common wisdom shared among the majority of researchers in the field from the MELARI / NID projects.

One thing, which has become clear from this exercise, is that Si MOSFET technology still has many years to go. Even if downscaling would eventually slow down over the next decade, generations of performance improvement are still to be obtained from utilising more advanced architectures. However, there are also limits to MOSFET technology, especially considering issues such as power consumption, economics and system complexity. For these reasons, the novel devices discussed in this Roadmap deservedly receive strong attention from the research community. RTDs and spin devices are on the verge of commercialisation. SET derived devices in the form of nano-flash devices seem a most logical successor to several RAM technologies. RSFQ could provide powerful computing capability, provided one is prepared to pay the price of cryogenic operation. Molecular electronics is still at an immature stage but deserves vigorous investigation because of its huge potential.

The fabrication technologies developed for producing nanodevices of course provide a much more direct spin-off to the electronics industry, since they can be equally well applied to MOSFET fabrication. In fabrication it is remarkable that the work generated by the semiconductor industry, in this case optical lithography, has shown a much longer life than predicted even by its greatest advocates five years ago. Bottom-up approaches, closely linked as they are to the field of molecular electronics, are elegant, cheap and possibly enormously powerful techniques for future mass replication, but their applicability remains limited until total control over the emerging structures in terms of wiring and interconnections can be obtained. It is clear that new architectures are required for such bottom-up fabrication approaches.

Similar to fabrication, architecture development has also direct bearings for the Si electronics industry and is a field where much ground remains to be gained. At the same time, the specific quantum mechanical properties of some of the novel nanodevices require radically novel architecture approaches, and it is gratifying to see that such approaches are now starting to be developed by the architecture experts. Issues like fault tolerant architecture, parallel processing, local architectures and neural nets are directly translatable into Si hardware, whereas more advanced concepts like non-dissipative and quantum computing will almost certainly require the implementation of novel quantum mechanical nanodevices.

5.2 Devices

The list of devices discussed in this Roadmap has of necessity been limited to those presently under investigation as possible nanoelectronic circuit elements. We stress here that of course, as time progresses, other device concepts may have to be added to those in this list while some others may disappear.

SET at present shows promise for memory, probably not in the form of a single electron transistor but rather as a nano-flash memory. As such it seems a natural extension for conventional flash type memory devices and may even take over from SRAM, possibly bridging the gap between more standard CMOS and the actual single electron transistor. Whether the actual single electron transistor itself will ever make it to the market place will strongly depend on future developments in device fabrication, both in respect to feature size (2 nm range) and uniformity, along with the control of background charges. The charge sensitivity of the devices also imposes strong limitations on the allowable electrostatic interaction between different devices in a ULSI circuit.

RTDs are the most mature of all the device technologies reviewed in the present Roadmap. III-V devices are expected to be on the market in the near future. Research on Si-based devices still requires further development before marketable devices can be realised. The major concern remains wafer uniformity, especially for large integration levels. First applications will most probably be high frequency DACs or ADCs and transmitters, along with ultra low power memory for portable applications. It should be noted that the intrinsic bistability of RTDs also enables the implementation of compact and elegant architectures for logic circuitry, although these areas are not as mature as the lower integration level applications.

RSFQ has potential in niche applications, where high speed and computational power are important and the cost of cryogenics may be tolerated. One of the major problems is interfacing the high-speed circuits with conventional CMOS for input / output, especially when the large difference in circuit speeds is considered. At present the technology for low temperature superconductors seems mature, while for HTS, although considerable progress has been achieved in recent years, large-scale integration has not been demonstrated yet. The market presence of RSFQ will also depend on the availability of reliable, low-cost cooling engines.

Molecular electronics is at present taking off as a branch of transport physics. Its potential is large but there are formidable obstacles, which must be overcome. Here it seems of the utmost importance that chemists, biologists, physicists and engineers develop an interdisciplinary platform for communicating the needs of the electronics industry in one direction and the possibilities of chemical synthesis and self-assembly concepts in the other.

Spin devices in the form of tunnel junction MRAMs will be on the market in the near future. Recording heads based on the tunnel junction exchange biasing mechanism are already existing applications. A problem for widespread use in the semiconductor industry is that the metals used in the devices fabricated to date are not compatible with CMOS lines. Spin injection in semiconductors has considerable potential, for example, in quantum computing, although formidable obstacles in manufacture still have to be overcome.

A comparison between all the device technologies mentioned above and the CMOS developments predicted by the SIA Roadmap has been attempted and is included in Annex I. While the forecasts for CMOS are based on extrapolations of the last thirty years of progress along with recent research results as predicted by the SIA Roadmap, the nanoelectronic devices are best guesses provided by those participating in the MEL-ARI / NID projects. We hope that the numbers on nanodevices will serve as a guideline for a discussion basis to monitor the progress in the nanoelectronic field.

5.3 Fabrication

To date most of the gate and circuit downscaling depends on the ability to push optical lithography to smaller and smaller dimensions. Given the existing investments in optical lithography, it will remain the preferred manufacturing technique for CMOS as long as economically or physically possible. It seems reasonable to expect that serial techniques will not be suitable for mass fabrication at high integration levels, although some may still be applicable to metrology and test-

ing. For the next one or two generations beyond optical lithography, there are a number of options that are being investigated and still need further research such as EUV, X-ray, imprint, projection electron-beam or projection ion-beam lithography. These technologies are now all under active evaluation and are all compatible with the devices discussed in this Roadmap. To date X-ray and projection electron-beam lithography have had the greatest investment in research. Once structures of nanometer size dimensions and ultra-tight tolerance are required (such as SET), it seems unlikely that there will be another manufacturing approach apart from the bottom-up technique. As was stressed before, however, in the bottom-up approach a considerably stronger degree of control on the architecture by the engineers is necessary to enable actual circuit fabrication. The first applications of the bottom-up technique are likely to be small circuits fabricated on top of top-down fabricated chips.

5.4 Architectures

One of the issues we encountered in preparing this Roadmap is the importance of an early collaboration between device physicists and architecture designers. It is important for device physicists to realise what consequences a gainless device (for instance) has on the circuit architecture that must be implemented if the device is going to be used in applications. Another issue often missed by device people is the rapid increase in the timing (clock frequency) when dimensions are reduced and speed is increased. The arrival of local architectures, parallel computing and defect- and fault-tolerant architectures are a direct consequence of this drive towards miniaturisation and apply as well to CMOS as to the nanoelectronic devices discussed in this Roadmap. As long as nanoelectronic devices operate in a classical regime, the architectures could be very similar to those for CMOS devices of similar dimensions. A development towards embedded systems combining logic and memory on one chip, quite natural seems moving away from the standard CPU / memory architecture. As soon as quantum mechanical concepts are required at the device level, however, different approaches to system architecture are needed. For RTDs this leads to reduced system complexity while for the QCA implementation of SET this leads to quite involved but additionally quite compact circuit architectures. Totally novel architectures will be needed for relatively young concepts such as quantum computing, where at present research focuses mainly on the basic device and circuit level.

5.5 Outlook

While CMOS continues to dominate the semiconductor industry, it is apparent that several nanoelectronic devices originally conceived as successors to CMOS, are now finding their way into niche markets. It is also clear that the Moore's Law exponential increases in density and performance that CMOS has enjoyed for over thirty years cannot be maintained for ever. No exponential increase can go on for ever - eventually the increase in density, power consumption and volume of silicon chips will require all the energy in the universe to allow operation! If at some point where the exponential increase flattens out further increases in performance are required, one or some of the devices discussed in this Roadmap (or possibly another nanoelectronic concept) could find itself in the dominant position in the market place. If the nanoelectronic field wants to mature to this stage, there is a necessity to bring novel devices more on a par with CMOS by developing the necessary fabrication processes, simulation tools and design rules that are required for any industrial electronic manufacturing process. Nanoelectronics is now coming of age, and it appears to be the correct time to develop these production tools. This Roadmap is part of such a process of professionalisation of the field.

We believe that research in nanoelectronics is important and rewarding and consider it vital for the community to vigorously encourage this economically valuable line of research.

References

- L. Adleman "Molecular computation of solutions to combinatorial problems" *Science* 266, pp1021-1024 (1994)
- M. Akazawa, "Binary-decision-diagram logic systems using single-electron circuits and single-flux-quantum circuits" preprint 1998 (akazawa@sapiens-ei.eng.hojudai.ac.jp)
- I. Amlani, A. O. Orlov, G. L. Snider, C. S. Lent, and G. H. Bernstein, *Appl. Phys. Lett.* 72, 2179 (1998)
- M.G. Ancona, "Systolic processor design using single electron digital circuits", *Superlattices and Microstructures*, 20(4) pp. 461-471 (1996)
- M. Ancona "Single electron digital circuits: possibilities and problems", *Phasdom 97*, Aachen 10-13 March (1997a)
- M.G. Ancona, "Bit errors in single-electron digital circuits", *Proc. Int. Workshop on Quantum Functional Devices*, NIST Gaithersburg VA, Nov (1997b)
- R.P. Andes, T. Bein, M. Dorogi, S. Feng, J.I. Henderson, C.P. Kubiak, W. Mahoney, R.G. Osifchin and R. Reifengerger, "Coulomb staircase at room temperature in a self-assembled molecular nanostructure" *Science* 272, pp1323-1325 (1996)
- S. Asai and Y. Wada, "Technology challenges for integration near and below 0.1 μ m", *Proc. IEEE*, 85, pp505-520 (1997)
- G.J. Ashwell, J.R. Sambles, A-S. Martin, W.G. Parker and M. Szablewski, *J. Chem. Soc. Chem. Commun.* 1374 (1990)
- A. Aviram, *J. Am. Chem. Soc.* 110, 5687 (1988)
- A. Aviram and M.A. Ratner, "Molecular rectifiers" *Chem. Phys. Lett.* 29 277 (1974)
- Behr, "The Lock and Key Principle" John Wiley and Sons, (1994)
- D. Berzon & T.J. Fountain, "A memory design in QCAs using the SQUARES formalism", 9th Great Lakes Symposium on VLSI, March (1999) *IEEE*, p. 166 Eds. R.J. Lomax, P. Mazumder
- H.A. Biebuyck et al., "IBM J. Res. Develop". 41, page 159 (1997)
- R.R. Birge and R.B. Goss "Biomolecular optoelectronics" in "Introduction to Molecular Electronics" Edited by M.C. Petty, M.R. Bryce and D. Bloor, Edward Arnold, London, (1995)
- R.R. Birge, *Annual Review Phys. Chem.* 41, pp683-733 (1990)
- H. Boeve, B. Dutta, J. De Boeck, G. Borghs, R. Van De Veerdonck, R. Coehoorn, J. S. Moodera, to be published.
- S. Borkar, "Technology Trends and Design Challenges for Microprocessor Design", *Proc. of the 24th European Solid-State Circuits Conference*, The Hague, The Netherlands, 22-24 Sept, pp. 7-8 (1998)
- S.G. Boxer, J. Stocker, S. Franzen and J. Salafsky in "Molecular Electronics - Science and Technology" Eds. A. Aviram (Am. Inst. Phys., New York, 1992)
- E. Braun, Y. Eichen, U. Sivan and G. Ben-Yoseph, "DNA-templated assembly and electrode attachment of a conducting silver wire" *Nature*, 391, pp775-778 (1998)
- T.P.E. Broekaert, B. Brar, J.P.A. van der Wagt, A.C. Seabaugh, F.J. Morris, T.S. Moise, E.A. Beam III and G.A. Frazier, "A monolithic 4-bit 2Gsp/s resonant tunnelling analogue-to-digital converter" *IEEE J. Solid-State Circuits* 33, pp1342-1348 (1998)
- T.P.E. Broekart, B. Brar, F. Morris, A.C. Seabaugh & G. Frazier, "Resonant tunneling technology for mixed signal and digital circuits in the multi-GHz domain", 9th Great Lakes Symposium on VLSI, March 4-6, 1999, Ann Arbor Michigan, p. 123 Eds. R.J. Lomax, P. Mazumder
- J. Brown and A. Pohm "1-Mb Memory Chip Using Giant Magnetoresistive Memory Cells", *IEEE Trans on Comp, Pack and Manufact Techn. A*, 17(3), pp.373 (1994)
- W. H. Bruenger, et al. "Resolution Improvement of Ion Projector with a Low Energy Spread Multicusp Ion Source", *MNE 98*, Leuven, to be published in *Microelectron. Eng.* (1999)

-
- L.A. Bumm et al. *Science* 271, 1705 (1996)
- W.P. Burlison, M. Ciesielski, F. Klass & W. Lu, "Wave-pipelining: a tutorial and research survey", *IEEE Trans. VLSI Syst.* 6, 464-473, (1998)
- J. Canning "Potential and challenges for lithography beyond 193 nm optics" *J. Vac. Sci. Technol. B15*, pp2109-2111 (1997)
- J. Canning "International SEMATECH Lithography activities" Presentation at XEL (1998)
- K.J. Chen, K. Maezawa, M. Yamamoto, "InP-Based High Performance Monostable-Bistable Transition Logic Elements (MOBILE's) Using Integrated Multiple-Input Resonant Tunneling Device". *IEEE Electron Device Letters*, 17(3), pp.127-129, Mar. (1996)
- T. Chen (ed) "Highlights of statistical signal and array processing", *IEEE Signal Proc. Mag.* September 1998, pp 21-63.
- S.Y. Chou and P.R. Krauss, "Imprint lithography with sub-10nm feature size and high throughput", Phasdom 97, Aachen 10-13 March (1997)
- D. Crawley, "An analysis of MIMD processor node designs for nanoelectronic systems", UCL IPG Report 97/3, available via <http://ipga.phys.ucl.ac.uk/reports/html>
- J.D. Cressler "SiGe HBT technology: a new contender for Si-based rf and microwave circuit applications" *IEEE Trans. Microwave Theory Techniques* 46, pp572-589 (1998)
- M.T. Cuberes, R.R. Schlittler and J.K. Gimzewski, "Room temperature repositioning of individual C60 molecules at Cu steps: operation of a molecular counting device" *Appl. Phys. Lett.* 69, pp3016-3018 (1997); see also <http://www.zurich.ibm.com/News/Abacus>
- M.T. Cuberes, R.R. Schlittler and J.K. Gimzewski, "Room temperature supramolecular repositioning at molecular interfaces using scanning tunnelling microscope", *Surface Sci. Lett.* 371, ppL231-L234 (1997)
- T. Dracopoulos, *Evolutionary learning Algorithms for Neural Adaptive Control*, Springer-Verlag (1997)
- N. Evers et al. *IEEE Electron Dev. Lett.* 17, p443 (1996)
- L. Feldheim & C.D. Keating, "Self-assembly of single electron transistors and related devices", *Chem. Soc. Rev.* 27, pp1-12, (1998)
- D.B. Fogel (ed), 'Evolutionary Computation: The Fossil Record', IEEE Press, (1998) (ISBN 0-7803-3481-7)
- T.J. Fountain, *Parallel Computing: Principles and Practice*, Cambridge University Press, (1994)
- T.J. Fountain, M.J.B. Duff, D.G. Crawley, C.D. Tomlinson, and C.D. Moffat, "The use of nanoelectronic devices in highly parallel computing systems" *IEEE Trans. On Very Large Scale Integrated Systems*, 6, 31-38, March (1998a)
- T.J. Fountain, D.G. Crawley, M.R.B. Forshaw, S. Spagocci, D. Berzon, "Novel processor arrays for nanoelectronics", Ultra electronics program review, Estes Park, Colorado, 18-21 Oct. 1998, Digest p.34-38. (1998b)
- FPGA-based Computing Machines website (1998): http://www.io.com/~guccione/HW_list.html
- A Fujiwara et al., 'Silicon double-island single-electron device', pp 163-167, Technical Digest, Int. Electron Devices Meeting (IEDM97), IEEE Inc., 445 Hoes Lane, Piscataway NJ 08855, USA, (1997)
- W.J. Gallagher et al. "Microstructured magnetic tunnel junctions", *J. Appl. Phys.* 81, 3741 (1997)
- R. Galvin, "Science Roadmaps" *Science* 280, 803 (1998)
- D. Goldhaber-Gordon, M.S. Montemerlo, J. C. Love, G.J. Opiteck and J.C. Ellenbogen, "Overview of nanoelectronic devices" *Proc. IEEE* 85, pp521-540 (1997)
- K. Goser, C. Pacha, A. Kanstein, and M.L. Rossmann "Aspects of systems and circuits for nanoelectronics" *Proc. of the IEEE*, 85, 558-573, April (1997)
- M.J. Gossens, "Analog Neural Networks in Single Electron Tunneling Technology", Ph.D Thesis, TU Delft, The Netherlands, (1998)
- Granly and Hurst, Presented at the 1996 Int'l NonVolatile Memory Technology Conference (IEEE)

-
- Digest pp.138 (1996)*
- E. Greenbaum, *J. Phys. Chem.* 96, pp514-516 (1992)
- P.E. Gronowski, W.J. Bowhill, R.P. Preston, M.k. Gowan & R.L. Allmon, "High-performance microprocessor design", *IEEE Journ. Solid-State Circ.* 33, 676-686, (1998)
- C.W.Gwyn, R. Stulen, D. Sweeney, and D. Attwood "Extreme ultraviolet lithography" *J. Vac.Sci.Technol.* B16, 3142 (1998)
- P. Hadley & J.E. Mooij, "Quantum nanocircuits: chips of the future?", *TU Delft reprint*, (1999) (<http://vortex.tn.delft.nl/~hadley/publications/qdevices99/pearsal.html>)
- P. Hadley, E.H. Visscher, Y. Chen and J.E. Mooij, "An offset charge independent single electronics RS flip flop" *Proc. International Conference on Quantum Devices and Circuits*, p224, Alexandria, Egypt, June 4-7, (1996)
- Hameroff et al. *Computer* 25(11), pp30-39, (1989)
- S. Hamilton, "Taking Moore's Law Into the Next Century", *IEEE Computer*, Jan. 1999, pp. 43-48 (1999)
- H. I. Hanafi, S. Tiwari and I. Khan, *IEEE Trans. Electron Devices* 43(9), 1553—8 (1996)
- L.R. Harriot "Scattering with angular limitation projection electron beam lithography for sub-optical lithography" *J. Vac. Sci. Technol.* B15(6), pp2130-2135 (1997)
- S. Haykin, *Neural Networks*, Macmillan, NY, (1994)
- J.R. Heath, P.J. Kuekes, G.S. Snider & R.S. Williams, "A defect-tolerant computer architecture: opportunities for nanotechnology", *Science* 280, pp1716-1721, (1998)
- J.L. Hennessy, D.A. Patterson, *Computer architecture: a quantitative approach*, Morgan Kaufman, San Mateo, Cal, USA, (1990)
- S. Hillenius and W.F. Brinkman "Microelectronics beyond roadmaps" *The Future of Microelectronics: Off the Beaten Path*, Ile des Embiez, France, June (1998)
- Hopfield, *Nature conference on Nanotechnology*, Tokyo, (1992)
- Hou and J.M. Philips. *Appl. Phys. Lett.* 67, pp286-288 (1995)
- H. Ishikuro and T. Hiramoto, "Fabrication of Si point contact MOSFETs acting as single electron transistor at room temperature", *Silicon Nanoelectronics Workshop*, Kyoto, June 8-9 (1997)
- K. Itoh et al., "Trends in Low-Power RAM Circuit Technologies", *Proc. IEEE* 84, pp524-543(1995)
- H. Iwai, "Future of LSI: Lithography, Interconnects, etc.", *The Future of Microelectronics: Off the Beaten Path*, Ile des Embiez, France, June (1998)
- G.J. Jeong and M.K. Lee, "Design of a Scalable Pipelined RAM System", *IEEE J. Solid State Circuits* 33, pp 910-914 (1998)
- C. Joachim and J.K. Gimzewski "An electromechanical amplifier using a single molecule" *Chem. Phys. Lett.* 265, pp353-357 (1997), see also <http://www.zurich.ibm.com/>
- S.J. Jou, C.Y. Chen, E.C. Yang, and C.C. Su, "A Pipelined Multiplier-Accumulator Using a High Speed Low Power Static and Dynamic Full Adder Design", *IEEE J. Solid State Circuits* 32, pp. 114-118 (1997)
- T.A. Jung, R.R. Schlittler and J.K. Gimzewski, "Conformational identification of individual adsorbed molecules with the STM", *Nature* 386, pp696-698 (1997)
- B.E. Kane, "A silicon-based nuclear spin quantum computer", *Nature* 393, 133-137, (1998)
- E. Kasper (1998) in addition to S. Luryi et al. *IEEE Trans. Elec. Dev.* ED31, 832 (1984)
- H. Kawaura, T. Sakamoto, T. Baba, Y. Ochiai, J. Fujita, S. Matsui and J. Sone, "Transistor operation of 30nm gate-length EJ-MOSFETs", *IEEE Elec. Dev. Lett.* 19, pp74-76 (1998)
- G. Klimeck et al., "Design and Implementation of Resonant Tunnelling Devices into Circuits and Applications", *Phasdom 97*, Aachen, Germany, March (1997)
- Kohonen, *Self-organising maps*, Springer-Verlag, Berlin, (1995)
- I. Koren & Z. Koren, "Defect tolerance in VLSI circuits: techniques and yield analysis", *Proc. IEEE* 86, 1819-1836, (1998)
- Koronkov, Chen and Likharev, *J. Appl. Phys.* 78, pp2520-2530 (1995)

-
- M. Korkin, H.de Garis, F. Gers & H. Hemmi, "CBM (CAM_BRAIN MACHINE, <http://www.genobyte.com/GP97full.html>)
- C.E.Kozyrakis, D.A.Patterson. *A new direction for computer architecture research*, *IEEE Computer*, 31, 24-32, Nov.(1998)
- K. Kurihara, H. Namatsu, M. Nagase and Y. Takahashi "Fabrication and process for room temperature operated single electron transistor using electron beam nanolithography", *Silicon Nanoelectronics Workshop*, Kyoto, June 8-9 (1997)
- Lehn "Supramolecular Chemistry: Concepts and Perspectives" VCH, Weinheim, (1995)
- C. S. Lent, "Quantum Computation and its perspective," *A report to the Research and Development Association for Future Electron Devices*, *FED Journal*, April 1997
- C.S. Lent and P.D. Tougaw "A device architecture for computing with quantum dots" *Proc. IEEE* 85, pp541-557 (1997)
- K.K. Likharev, *Private Communication*, (1997)
- K. K. Likharev, *MEL—ARI meeting Lille (1998)*; K. K. Likharev, *GOMAC98 Proceedings* pp. 395—8 (1998)
- D. Mange, E. Sanchez, A. Stauffer, G. Tempesti, P. Marchal & C. Piguët, "Ekmbyronic: a new methodology for designing field-programmable gate arrays with self-repair and self-replicating properties", *IEEE Trans. VLSI Syst.* 6, 387-399, (1998)
- D. Matzke, "Will physical scalability sabotage performance gains?," *IEEE Computer* 30, 4, 37-39, (1997)
- P. Mazumder, S. Kulkarni, B. Bhattacharya, J.P. Sun, and G.I. Haddad, "Digital Circuit Applications of Resonant Tunneling Devices", *IEEE Proc.* 86, pp 664-686(1998)
- J. Melngailis, A.A. Mondelli, I.L. Berry III and R. Mohondro, "A Review of Ion Projection Lithography", *J. Vac. Sci. Technol.* B16(3) pp927-957 (1998)
- L. Melo, L. Rodrigues and P.Freitas "Novel Spin-Valve Memory Architecture" *INESC*, presented at *MMM conference*
- R.M. Metzger, B. Chen, U. Hopfner, M.V. Lakshmikantham, D. Vuillaume, T. Kawai, X.L. Wu, H. Tachibana, T.V. Hughes, H. Sakurai, J.W. Baldwin, C. Hosch, M.P. Cava, L. Brehmer, G.J. Ashwell "Unimolecular electrical rectification in hexadecylquinolinium tricyanoquinodimethanide" *J. Amer. Chem. Soc.* 119, pp10455-10466 (1997)
- S.C Minne, S.R. Manalis, G. Yaralioglu, A. Atalar and C.F. Quate, "Increasing the Throughput of the Atomic Force Microscope", *Phasdom 97*, Aachen 10-13 March (1997)
- C.A. Mirkin, R.L. Letsinger, R.C. Mucic and J.J. Storhoff, "A DNA-based method for rationally assembling nanoparticles into macroscopic materials" *Nature* 382, pp607-609 (1996)
- H. Mizuta and T. Tanoue "The Physics and Applications of Resonant Tunneling Diodes" *CUP* (1995)
- H-O. Müller, K. Katayama and H. Mizuta "Effects of disorder on the blockade voltage of two-dimensional quantum dot arrays" *J. Appl. Phys.* 84, pp5603-5609 (1998)
- A. Nakajima et al. *Appl. Phys. Lett.* 70(13), pp1742-4 (1997)
- K. Nakazato et al.. *IEDM97 Proceedings* (1997)
- H. Nambu, K. Kanetani, K. Yamasaki, K. Higeta, M. Usami, Y. Fujimura, K. Ando, T. Kusunoki, K. Yamaguchi and N. Homma, "A 1.8 ns access, 550 MHz, 4.5 Mb CMOS SRAM" *IEEE J. Solid State Circuits* 33, pp1650-1658 (1998)
- A. O. Orlov et al "Realisation of a Functional Cell for Quantum-Dot Cellular Automata", *Science* 277, pp 928-930 (1997)
- C. Pacha, K. Goser , A. Brennemann, and W. Prost, "A Threshold Logic Full Adder Based on Resonant Tunneling Transistors", to be published at *ESSCIRC'98*, The Hague, NL, September (1998)
- C. Pacha, O.Kessler, K. Goser , W. Prost, A. Brennemann, U. Auer, F.J. Tegude, "Parallel Adder Design with Reduced Circuit Complexity Using Resonant Tunneling Transistors and Thresh-

- old Logic”, to be published in: *Analog Integrated Circuits and Signal Processing, Special issue on Circuit Design in Nano and Quantum Device Technology*. A previous version was presented at the 2nd Workshop on Innovative Circuits and Systems for Nanoelectronics, TU Delft, (1997)
- C. Pacha, U. Auer, P. Gloesekoetter, A. Brennemann, W. Prost, F.-J. Tegude & K. Goser, “Resonant tunneling transistors for threshold logic circuit applications”, *Proc. of the Great Lakes Symposium on VLSI, March* (1999)
- F. Patten and S. Wolf (DARPA), Presented at the 1996 Int’l NonVolatile Memory Technology Conference (IEEE) Digest pp.1-2 (1996)
- D.J. Paul, “Silicon germanium heterostructures in electronics: the present and the future” *Thin Solid Films* 321, pp172-180 (1998)
- A. Pohm et al, *IEEE Trans. Magn.*, 30 (6), pp. 4650 (1994)
- A. Pohm, J. Daughton, J. Brown and R. Beech “The Architecture of a High Performance Mass Store with GMR Memory Cells”, *IEEE Trans. on Magn.*, 31 (6), pp.3200 (1995)
- A. Pohm et al. Presented at the 1996 Int’l NonVolatile Memory Technology Conference (IEEE) Digest pp.113-115 (1996)
- R. Porter et al., *Proc. Fifth Int. Conf. InP and Related Materials*, 1993, 37 (1993)
- M. Rabey, “Digital Integrated Circuits - A Design Perspective”, Prentice Hall, Chapter 7 (1996)
- I. Ranmuthu, K. Ranmuthu, C. Kohl, C. Comstock et M. Hassoun “A 512K-bit Magneto Resistive Memory with Switched Capacitor Self-Referencing Sensing”, *IEEE Trans on Circuits and Systems-II*, 39(8), pp.585 (1992)
- M.A. Reed, C. Zhou, C.J. Muller, T.P. Burgin and J.M. Tour, *Science* 278, 252 (1997)
- M.A. Reed, “Molecular electronics - Future Prospects and Reality” *The Future of Microelectronics: Off the Beaten Path, Ile des Embiez, France*, 1998
- S.L. Rommel, T.E. Dillon, P.R. Berger, R. Lake, P.E. Thompson, K.D. Hobart, A.C. Seabaugh and D.S. Simons “Si-based interband tunnelling devices for high speed logic and low power memory applications” *IEDM 98* (1998)
- A. Seabaugh et.al., “Resonant Tunneling Circuit Technology: Has it Arrived?”, *Techn. Digest of the 19th Gallium Arsenide Integrated Circuit Symposium* pp. 119-122 (1997)
- U. Schubert et al, *Chimia* 51 Nr 7 p 434 (1997)
- SIA Roadmap:- “The National Technology Roadmap for Semiconductors”, available from the Semiconductor Industry Association, 4300 Stevens Creek Blvd, San Jose, CA 95129. The document can also be downloaded at the web site: <http://www.sematech.org/public/roadmap/>
- J. Silbermann et al., “A 1.0-GHz single-issue 64-bit PowerPC integer processor”, *IEEE Journ. Sol. State Circ.* 33, 1600-1607, (1998)
- J.P. Silverman “Challenges and progress in X-ray lithography” *J. Vac. Sci. Technol. B*16, 3137 (1998)
- R.A. Soref “Silicon based optoelectronics” *Proc. IEEE* 81, pp1687-1706 (1993)
- A. Steane, “Quantum computing”, *Rep. Prog. Phys.* 61, 117-173, (1998)
- S.M. Sze, “Non-volatile memories” *The Future of Microelectronics: Off the Beaten Path, Ile des Embiez, France*, (1998)
- Y. Takahashi, A. Fujiwara, K. Yamazaki, H. Namatsu, K. Kurihara and K. Murase, “A multiple-gate single-electron-transistor and its application to an exclusive-OR gate” *IEDM 98* (1998)
- S. Takeda, G. Soda and H. Chihara “Solid neopentane C(CH₃)₄ as studied by nuclear magnetic resonance: A detailed examination of methyl and molecular reorientation in the low temperature phase” *Mol. Phys.* 47, pp501-517 (1982)
- S. J. Tans et al. “Individual Single Wall Carbon Nanotubes as quantum wires” *Nature*, 386, 474 (1997)
- S.J. Tans, A.R.M. Verschueren and C. Dekker, “Room temperature transistor based on a single carbon nanotube” *Nature* 393, pp49-52 (1998)

-
- Y. Taur, D.A. Buchanan, W. Chen, D.J. Frank, K.E. Ismail, S-H. Lo, G.A. Sai-Halasz, R.G. Viswanthan, H-J. C. Wann, S.J. Wind and H-S. Wong "CMOS scaling into the nanometer regime" *Proc IEEE* 85, pp486-504 (1997)
- A. Terfort, N. Bowden and G.M. Whiteside, "Three-dimensional self-assembly of millimetre-scale components" *Nature* 386, pp162-164 (1997)
- T. Toffoli & N. Margolus, *Cellular automaton machines*, MIT Press, Cambridge, MA, (1985)
- S. Tiwari et al. *Appl. Phys. Lett.* 69(9), pp1232-1234 (1996)
- P. D. Tougaw, C. S. Lent, and W. Porod, *J. Appl. Phys.* 74, 3558 (1993)
- K. Tsukagoshi, B.W. Alphenaar and K. Nakazato, "Operation of logic function in a Coulomb blockade device" *Appl. Phys. Lett.* 73, pp2515-2517 (1998)
- ULTRA Electronics Review Meeting Proceedings, Estes Park, CO, 18-21 Oct (1998)
- Y. Umeda, K. Osafune, T. Enoki, H. Ito and Y. Ishii, "SCFL static frequency divider using InAlAs/InGaAs/InP HEMTs" in *Proc. 25th European Microwave Conference* pp222-228 (1995)
- Y. Wada, T. Uda, M. Lutwyche, S. Kondo and S. Heike, "A proposal of nanoscale devices based on atom/molecule switching" *J. Appl. Phys.* 74, pp7321-7328 (1993)
- C.H. Van Berkel, M.B. Josephs, S.N. Nowick, "Applications of Asynchronous Circuits", *Proc. of the IEEE*, Vol. 97, No. 2, Feb. 1999, pp. 223-233 (1999)
- J.P.A. van der Wagt et al., *54th Annual Device Research Conference Digest (IEEE, Inc., NJ, 1996)*, p168
- P. van der Wagt, A.C. Seabaugh and E.A. Beam III, "RTD/HFET low standby power SRAM gain cell" *IEEE Elec. Dev. Lett.* 19, 7 (1998)
- J. Villasenor & B. Hutchings, "The flexibility of configurable computing", *IEEE Signal Proc. Mag.* September 1998, pp 67-84 (1998)
- E. Waingold et al., "Baring it all to software: the raw machine", *MIT Lab for Computer Sci. Report TR-709*, Mar. (1997)
- Z. Wang and Y. Nakamura, *IEEE Trans. Magn.*, 32 (2), pp. 520 (1996)
- Y. Watanabe et al. "Monolithic integration of InGaAs/InAlAs resonant tunnelling diodes and HEMT for single transistor cell SRAM applications" *Tech. Dig. IEDM 92*, pp475-478 (1992)
- A. Wieder, "Paradigm shifts in processing, circuits, systems and hardware / software solutions", *The Future of Microelectronics:- Off the Beaten Path*, Ile des Embiez, France (1998)
- V.A. Wilkinson, M.J. Kelly and M. Carr "Tunnel devices are not yet manufacturable" *Semicond. Sci. Technol.* 12, pp91-99 (1997)
- W. Williamson et al. "12 GHz clocked operation of ultralow power interband resonant tunnelling diode pipelined logic gates" *IEEE J. Solid State Circuits* 32(2), pp222-231 (1997)
- K. Yano et al. *IEEE Trans. Electron Devices* 41(9), pp1628-38 (1994)
- K. Yano et al. *ISSCC98 Proceedings*, pp344-5 (1998)
- K. Yano, T. Ishii, T. Sano, T. Mine, F. Murai, T. Kure and K. Seki, "Status of single-electron memories" *IEDM98* (1998)
- M. Yasunaga, I. Hachiya, K. Moki & J.H. Kim, "Fault-tolerant self-organising map implemented by wafer-scale integration", *IEEE Trans. VLSI Syst.* 6, 257-265, (1998)
- L. Zhuang, L.J. Guo, S.Y. Chou "Silicon single-electron quantum-dot transistor switch operating at room temperature" *Appl. Phys. Lett.* 72, pp.1205-1207 (1998)

Annex I:- Comparison Between Technologies

One of the major objectives of the Nanoelectronic Roadmap is to allow comparisons between different technologies. In this section, an attempt is made to compare the technologies described in the previous sections with present production devices. In all cases, the present experimentally demonstrated performance in 1998 will be compared with predictions for 2006 and 2012.

The first section compares memory technologies. The second section assesses the similarities and differences of devices intended for logic and applications. The third section demonstrates the performance of single device nanoelectronic technologies and research device performance and the fourth looks at the circuit performance of different technologies. It is clear that it is difficult and subjective to compare certain parameters especially in technologies whose devices operate in a completely different way than conventional transistors, CMOS, etc.. In addition, changing the supply voltage or optimising for a single parameter may also vary the performance of individual transistors. Therefore, some standard guidelines will be used to attempt fair comparisons between differing technologies.

Notes:

1. Results in each column will be from a single device and hence the maximum f_T , f_{max} etc. from different devices are not quoted. This should give better comparisons between technologies, as one may always maximise a single parameter to the detriment of many others.
2. The power, switching time and power-delay product for devices will always be at the normal supply voltage of the circuit i.e. V_{dd} or equivalent.
3. The operating voltage will be loosely defined as the standard supply voltage required to operate the device or circuit. For DRAM, V_{dd} is an appropriate voltage, while for flash the program voltage V_{pp} is more appropriate.
4. The minimum noise figure quoted for logic and ASICs will be at 2 GHz, as this is one of the most commonly quoted parameters in the literature at present, due to the mobile telecom applications.
5. Transconductance, g_m or gain, b will always be the maximum value of a device.
6. Both lithography and feature size are quoted for e.g. logic transistors where the effective gate length can be significantly smaller than the width defined by lithography and especially for vertical tunnelling devices such as RTDs where the critical dimension for device operation is defined by epitaxial growth and not the lateral lithographic length. For DRAM the lithographic length will be the half pitch.

	CMOS			Single Electron Tunnelling			RTD			MRAM			Molecular
Technology	DRAM	Flash	SRAM	Yano-Type	Multidot nano-flash	Singledot nano-flash	1T HFET TSRAM	2T HFET SRAM	1T Si TSRAM	TJ	GMR	HHall	
Reference	SIA 97 ITR 99	SIA 97 ITR 99	Nambu 1998	Yano 1998	Nakazato 1997	Zhuang 1998	Van der Wagt 98	Pacha 1998	Pacha 1998	de Boeck 1998	De Boeck 1998	De Boeck 1998	
No. of devices	1GB	256MB	6 MB /cm ²	128MB	1	1	16		No demo	No demo	1 MB	No demo	No demo
Circuit Speed	100 MHz	100 MHz	200 MHz	50 kHz			500 MHz	1000 MHz	40 MHz	~20 MHz	10 MHz		
Feature size	180 nm	180 nm	200 nm	250 nm	100 nm	10 nm	500 nm			250 nm	350 nm		
Access time	10 ns	10 ns	1 ns	20 μs	10 ns		<1ns	1 ns	25 ns	≤ 50 ns	< 100 ns	8 ns	
Write time	10 ns	1 ms		10 μs	100 ns	1 μs	10 μs			<10 ns	100 ns	8 ns	
Erase time	0	1 ms		10 μs	1 ms					N/A	N/A	N/A	
Retention Time	128-256 ms	10 years	infinite	1day	7 days	5 s	static (to be tested)	static (to be tested)	static (to be tested)	Infinite (to be tested)	Infinite		
Endurance Cycles	Infinite	10 ⁵	infinite	10 ⁷	10 ⁹		10 ⁹	10 ⁹	10 ⁹	Infinite			
Operating Voltage (V)	1.8-1.5 V _{dd}	8 V _{pp}	1.8-1.5 V _{dd}	15 V	5 V	10 V	1-1.5V	< 1V	< 1V	3 - 5V			
Voltage to switch states	0.2 V	5 V	0.2 V	0.5V	0.65 V	0.1 V	0.38V	0.65V		50 mV			
Cell size	12F ² /bit 0.4μm ²	9F ² /bit		2F ² /bit	9F ² /bit	9F ² /bit	100 μm ²		9F ² /bit	2F ² /bit	9F ² +T		
No. of electrons	10000	1000	>10000	1000	2	1	10 ³ -10 ⁴	10 ³ -10 ⁴	10 ³ -10 ⁴	N/A	N/A	N/A	
Standby Power (W/Mbit)	10 ⁻⁷		<200				5	0.01	10 ⁻⁹	None	None		

I.I Comparison of memory devices in 1999

	CMOS			Single Electron Tunnelling		RTD			MRAM	Molecular
Technology	DRAM	Flash	SRAM	Multidot nano- flash	Single-dot nano-flash	RTD 1T HFET TSRAM	RTD 2T HFET SRAM	RTD 1T Si TSRAM	TJ	Mechanical C ₆₀ +STM
Reference	SIA 97 ITRS 99	SIA 97 ITRS 99	SIA 97 ITRS 99	Hitachi Cambridge	Hitachi Cambridge Likharev	Raytheon/ Pacha	Raytheon / Pacha	Raytheon / Pacha	de Boeck 1998	Welland/ Gimzewski
No. of devices	16GB	16GB	100MB/cm ²		16GB	~1 MB/cm ²	~1 MB/cm ²	20 MB/cm ²		100
Circuit Speed	125 MHz	125 MHz	554 MHz			500 MHz	> 1GHz	150 MHz	100 MHz	10 GHz
Feature size	100 nm	100 nm	100 nm	100 nm	100 nm	200 nm	200 nm	200 nm	100 nm	50 nm
Access time	8 ns		1.8 ns	10 ns		5 ns	~1 ns	~10 ns	10 ns	<1 ns
Write time	10 ns			100 ns	100 ns	5 ns	~1 ns	~10 ns	10 ns	<1 ns
Erase time	10 ns			100 μs		~ 5ns	< 1ns	~10 ns	N/A	
Retention time	256-512ms	10 years	infinite	1 year	5 s	infinite	infinite	infinite	infinite	<1 s
Endurance cycles	infinite	10 ⁵	infinite	10 ¹⁰		infinite	infinite	infinite	infinite	infinite
Operating Voltage (V)	0.9-1.2V _{dd}	7V _{pp}	0.9 - 1.2V _{dd}	5	3-5	0.5	0.5	0.5	2	< 1
Voltage to switch between states				1	0.1	0.2	0.2	0.2	<50mV	<1
Cell size	12F ² /bit 0.046 μm ²	9F ² /bit	0.21μm ²			0.04 μm ²			9F ² /bit	
No. of electrons				1	1					~1000
Standby Power (W/Mbit)										10 ⁻⁵

I.II Predictions of memory devices in 2006

	CMOS			Single Electron Tunnelling	Resonant Tunnelling			MRAM	Molecular
Technology	DRAM	Flash	SRAM	SET memory	RTD 1T HFET TSRAM	RTD 2T HFET SRAM	RTD 1T Si TSRAM	MRAM	molecular network
Reference	SIA 1997	SIA 1997	SIA 1997	Hitachi Cambridge	Raytheon Pacha	Raytheon Pacha	Raytheon Pacha	de Boeck 1998	Editorial Guess
No. of devices	256GB	256GB	180MB/cm ²		2 MB/cm ²	1-2 MB/cm ²	50 MB/cm ²		> 256GB
Circuit Speed	150 MHz	150 MHz	913 MHz		500 MHz	> 1GHz	150 MHz	100 MHz	<< 1 GHz
Feature size	50 nm	50 nm	35 nm		200 nm	200 nm	200 nm	50 nm	3-4 nm
Access time	10 ns	10 ns	1.1 ns	10 ns	5 ns	~1 ns	~10 ns	10 ns	~1 ns
Write time	10 ns	10 μs		10 ns	5 ns	~1 ns	~10 ns	10 ns	~1 ns
Erase time	< 1 ns	10 μs		< 1 ns	~ 5ns	< 1ns	~10 ns	N/A	~1 ns
Retention time	2 - 4 s	10 years		1 s	infinite	infinite	infinite	infinite	~1 ms
Endurance cycles	infinite	10 ⁵	infinite	10 ¹⁴	infinite	infinite	infinite	infinite	infinite
Operating Voltage (V)	0.5-0.6 V _{dd}	5 V _{pp}	0.6-0.5 V _{dd}	1	0.5	0.5	0.5	2	< 1
Voltage to switch state	0.2 V	5 V		0.2 V	0.2 V	0.2 V	0.2 V	<50mV	~1 mV
Cell size	2.5 F ² /bit 0.005 μm ²	2 F ² /bit		2 F ² /bit	0.04 μm ²			2 F ² /bit	> 2 F ² /bit
No. of electrons									< 10
Standby Power (W/Mbit)	10 ⁻⁷				0.01		10 ⁻⁹		

I.III Predictions of memory devices in 2012

Technology	CMOS	Si bipolar	SiGe HBT	GaAs HBT	GaAs MESFET	GaAs HEMT
Reference	SIA 1997	Cressler 1998	Cressler 1998	Haramé 1996	Sze 1990	
Lithography	250 nm	500 nm	500 nm	2000 nm	550 nm	150 nm
Feature size	200 μ m	90 nm	90 nm			150 nm
Switching time (gate delay per stage)	50ps unloaded 0.3 ns with 200 fF load	30ps ring oscillator	20ps ring oscillator		25 ps active load	
f_T	20,6 GHz	18 GHz	47 GHz	50 GHz	32 GHz	100 GHz
f_{max}	25,7 GHz	24 GHz	69 GHz	70 GHz	60 GHz	
g_m	347 (130) mS/mm					
Gain, β		100	113			
Power dissipation	21 μ W per stage	5 mW	5 mW		1 mW	
Power-delay product	0.1 fJ	150 fJ	100 fJ		20 fJ	
Minimum Noise Figure @2GHz	5.0 dB	2.0 dB	0.5 dB	2.0 dB	1.5 dB	<0.5 dB 1.8dB@8GHz
No of electrons						
Cost/transistor	0.1mcent					

I.IV Comparison of locig and high frequency devices in 1998 (in production)

Technology	SiGe n-MODFET	SiGe p-MODFET	InAs RTD	SiGe RTD	SET logic	Molecular
Reference	Ismail 1995	Arafa 1997	Brown 1991	Ismail 1991	Tsukagoshi 1998	Tans 1998
Lithography	400 nm	100 nm	1.800 nm	100 μ m	200 nm	100 nm
Feature size	400 nm	100 nm	2 nm	5 nm	10 nm	5 nm
Switching time			~205 fs		~1 μ s	
f_t	40 GHz	70 GHz				
f_{max}	56 GHz	55 GHz	1.24 THz	150 GHz	~ 200 MHz	
g_m	420 mS/mm	233 mS/mm	60 mS/mm			
Gain, β					1	< 1
Power dissipation @ V_{dd}			0.3 μ W per gate		~ 20pW Excluding cooling	
Power-delay product						
Minimum Noise Figure @2GHz						
No. of electrons					100 - 1000	<100

I.V Comparison of logic and high frequency devices in 1998 (in research)

1. IV Predictions of logic and high frequency devices in 2006 (from research results)

Technology	SiGe n-MODFET	SiGe p-MODFET	InAs RTD	SET logic	Molecular
Reference	O'Neill	Arafa			
Lithography	100 nm	100 nm	200 nm	100 nm	100 nm
Feature size	70 nm	70 nm	2 nm	10 nm	5 nm
Switching time (gate delay per stage)			200 fs	500 ns	1 ns
f_T	150 GHz	70 GHz	1 THz	~ MHzs	500 MHz
f_{max}			1 THz		
k_n	950 mS/mm	237 mS/mm			
Gain, β				1	~ 1
Power dissipation @ V_{ds}	10s μ W		μ W	10 pW excluding cooling	~1 μ W
Power-delay product	200 aJ		~ 1 aJ		1 fJ
Minimum Noise Figure @ 2GHz					
No. of electrons	1000s	1000s		<100	< 100

I. VII Predictions of logic and high frequency devices in 2012 (from research results)

Technology	SiGe n-MODFET	SiGe p-MODFET	InAs RTD	SET logic	Molecular
Reference	Konig				
Lithography	50 nm	50 nm	~ 100 nm	50 nm	50 nm
Feature size	30 nm	30 nm	~2 nm	< 10 nm	< 5 nm
Switching time (gate delay per stage)	3 ps	5 ps	200 fs	100 ns	10 ns
f_T	320 GHz	200 GHz	1 THz	~10 MHz	~100 MHz
f_{max}			1 THz		
μ_n	>1000 mS/mm	~ 800 mS/mm			
Gain, β				~ 1	~ 1
Power dissipation @ V_{ds}	30 μ W	20 μ W		~1 pW	1 nW
Power-delay product	100 aJ	100 aJ	~1 aJ		10 aJ
Minimum Noise Figure @2GHz					
No. of electrons	500	500		<20	<10

I.VIII Comparison of circuit performance in 1998

Technology	CMOS	InAsRTD/HEMT	RSFQ LTS	RSFQ HTS	QCA	SET logic	Molecular
Reference	SIA 97	Maezawa 1998	Humphreys Niemeyer	Humphreys Niemeyer	No demo	Tasukagoshi 1998	No demo
Feature size	200 nm	700 nm HFET2x3µm RTD	3500 nm	2000 nm		10 nm (200nm lith)	
No. of active elements ^(a)	11M	2	5 10 ³	10		10	
Circuit Speed	750 MHz	35 GHz	1 GHz	30 GHz		1 MHz	
Events / chip / s	8.25 10 ¹⁵	7 10 ¹⁰	10 ¹² (b)	6 10 ¹⁰ (b)		10 ⁷	
Power supply, V _{dd}	2.5 V	0.7 V				0.1 V	
Power dissipation	70 W total chip	2 mW per gate	10 ⁻³ W (c)(d) total chip	10 ⁻⁶ W (c)(d) total chip		~20pW (d) per gate	
Cost/transistor	3 mcent						
Temperature	400 K	300K	4 K	40 K		4K	

Notes for tables I.VIII to I.X:

- (a) Active elements: transistors for semiconductor devices and junctions for superconducting devices.
- (b) An average gate contains about 5 junctions.
- (c) The power dissipation per gate is estimated to be ~ 1µW for LTS (Nb technology) and ~ 0,1µW for HTS including the bias current system containing the resistive load.
- (d) Excluding cooling
- (e) fan-in/out=3, Complex pipelined logic
- (f) fan-in/out=1, D-flip-flop
- (g) estimated from the difference between the highest occupied molecular orbital (homo) to the lowest unoccupied molecular orbital (lumo)
- (h) estimated that the resistance per device / molecule is in the order of 100 Ω.

Technology	CMOS	SiGe HCMOS	InAs RTD/HFET (10kA/cm ² current density)	InAs RTD/HFET (10kA/cm ² current density)	RSFQ Nb	RSFQ HTS	QCA	SET logic	Molecular
Reference	SIA 1997	Ismail 1995	Pascha 1998	Pascha 1998	Humphreys 1998	Humphreys 1998	Editorial Guess	Mueller 1998	Editorial Guess
Feature size	70 nm	70 nm	200nm HFET 200nm RTD	200nm HFET 200nm RTD	1 μm	1 μm	2 nm	100 nm	~ 3-4 nm
No. of devices	2 10 ⁶	>2 10 ⁶	10 ⁶	10 ⁶	10 ⁶	10 ⁴	4	100	10
Circuit Speed	3.5 GHz	5.25 GHz	6.25 GHz fan-in/out=3 (Complex pipelined logic) 18.75GHz fan-in/out=1 (D-flip flop)	22.6GHz fan-in/out=3 (Complex pipelined logic) 67.8GHz fan-in/out=1 (D-flip flop)	100 GHz	200 GHz	~10 kHz	5 MHz	~10 kHz
Events / chip / s	7 10 ¹⁷	1.05 10 ¹⁸	6.26 10 ¹⁵ (pipelined) 1.875 10 ¹⁶ (D-flip flop)	1.16 10 ¹⁵ (pipelined) 6.78 10 ¹⁵ (D-flip flop)	10 ¹⁸	10 ¹⁴	4 10 ⁴	5 10 ⁸	10 ⁵
Power supply, V _{DD}	0.9 - 1.2V	0.5 V	0.6V	0.6V	N/A	N/A	0.1 mV	1mV	~1V
Power dissipation	160 W	80 W	10μW per gate 10W total chip	150μW per gate 10W total chip	3 mW excluding cooling	100 μW excluding cooling	~ 1pW excluding cooling	< 1 μW	0.1 W (100Ω / dev)
Temperature	400 K	400 K	300 K	300 K	4 (10) K	40 K	4 K	77 K	300 K

Technology	CMOS	SiGe HCMOS	InAs RTD/HFET (10kA/cm ² current density)	RSFQ Nb	RSFQ HTS	QCA	SET logic	Molecular
Reference	SIA 1997	Ismail 1995	Pascha 1998	Humphreys	Humphreys	Maccuci	Mueller 1998	Editorial Guess
Feature size	30 nm	30 nm	50 nm HFET 50 nm RTD	500nm	500 nm	10 nm	50nm lith 2 nm	~ 2 nm
No. of Devices	2.34 10 ⁹	4.68 10 ⁹	10 ⁷	10 ⁷	10 ⁸	10 ⁸	10 ⁷	10 ¹⁰
Circuit Speed	10 GHz	15 GHz	6.25 GHz fan-in/out=3 (Complex pipelined logic) 18.75GHz fan-in/out=1 (D-flip flop)	200 GHz	500 GHz	100 GHz	10 MHz	~ 1 MHz
Events / chip / s	2 10 ¹⁹	7 10 ¹⁹	6.25 10 ¹⁶ (pipelined) 1.875 10 ¹⁷ (D-flip-flop)	2 10 ¹⁸	5 10 ¹⁶	10 ¹⁴	10 ¹⁸	10 ¹⁶
Power supply, V _{ds}	1V	0.5V	0.6V	N/A	N/A	0.1 mV	0.1V	1 V (HOMO - LUMO)
Power dissipation	70 W	17.5 W	1μW per gate 10W total chip	0.6 W excluding cooling	0.05W excluding cooling	excluding cooling	<1μW excluding cooling	10 ⁶ W (100Ω / dev)
Temperature	400 K	400 K	300 K	4 K	50 K	4 K	77 K	300 K

Annex II: MEL-ARI / NID Projects

The groups of the following research projects have contributed to this document. Abstracts of the projects and the list of the participating groups can be found at the MELARI homepage <http://www.cordis.lu/esprit/src/melari.htm>

CHARGE	The Coulomb Blockade Applied to the Realisation of Electronics (http://qt.tn.tudelft.nl/CHARGE/)
LASMEDS	Fabrication of elementary Molecular Electronic Devices
SIQUIC	Silicon Quantum Integrated Circuits (http://www.sp.phy.cam.ac.uk/~dp109/SIQUIC.html)
FASEM	Fabrication and Architecture of Single Electron memories
NANOWIRES	Conductance characteristics and mass fabrication of Nanoscale integrated circuit nanowires (http://www2.eng.cam.ac.uk/~nano-www/nanowire.htm)
QUEST	Quantum Electronics using Scanning Tunnelling microscopy based lithography (http://www.isen.fr/isen-lci/quest/)
SPIDER	SPIIn Dependent Nano-Electronics
QUADRANT	Quantum Devices foR Advanced Nano-electronic Technology (http://quadrant.iet.unipi.it/)
RSFQ-HTS	High Temperature Superconducting Rapid Single Flux Quantum Logic
LOCOM	Logic Circuits with Reduced Complexity based on Devices with Higher Functionality (http://www-be.e-technik.uni-dortmund.de/~pacha/locom/locom2d.html)
SPINUP	Semiconductor Processing by ImpriNt of Ultrasmall Patterns
NANOTECH	Development of Nanoimprinting Techniques suitable for large area mass production of nm-scale patterns
ANSWERS	Autonomous Nanoelectronic Systems With Extended Replication and Signalling (http://www-be.e-technik.uni-dortmund.de/~pacha/answers/answers1.html)
Q-SWITCH	Heterostructure defined electron wave guides for quantum based switching applications.

The following persons have particularly contributed:

Jo De Boeck, IMEC, Belgium

Vu Thien Bihn, University of Lyon 1 - CNRC, France

Michel Brillouet, France Telecom, Crolles, France

Wilhelm Brünger, Fraunhofer Institute, Berlin, Germany

Ramón Compañó, DGXIII European Commission, Belgium

Emmanuel Dubois, ISEM Lille, France

Terence Fountain, University College London, UK

Nikos Glezos, IMEL Demokritos, Greece

James Gimzewski, IBM Zürich Laboratories, Switzerland

Peter Hadley, Technical University of Delft, Netherlands

David Haviland, Chalmers University, Sweden

Richard Humphreys, DERA, UK

Jurriaan Huskens, University of Twente, Netherlands

Giuseppe Iannaccone, University of Pisa, Italy

Rainer Kaesmaier, Siemens AG, München, Germany
Michael Kelly, University of Surrey, U.K.
Dieter Kern, Tübingen University, Germany
Huguette Launois, L2M, CNRS, Paris, France
Massimo Macucci, University of Pisa, Italy
Günther Mahler, University of Stuttgart, Germany
Laurens Molenkamp, RWTH Aachen, Germany
Lars Montelius, University Lund, Sweden
Pietro Morales, ENEA, Italy
Heinz-Olaf Mueller, Hitachi Cambridge Laboratory, UK
Jürgen Niemeyer, Physikalisch Technische Bundesanstalt, Germany
Christian Pacha, University of Dortmund, Germany
Douglas Paul, University of Cambridge, U.K.
Werner Prost, University of Duisburg, Germany
Lars Samuelson, University Lund, Sweden
Pier Sazio, University of Cambridge, UK
Georg Schmidt, RWTH Aachen, Germany
Peter Vettiger, IBM Zürich, Switzerland
Mark Welland, University of Cambridge, U.K.