Extremely Low-Power Logic

Organizer: Christian Piguet, CSEM, Neuchâtel Switzerland
Presenters: Jacques Gautier, CEA-LETI, Grenoble, France; Christoph Heer, Infineon
Technologies, Munich, Germany; Ian O'Connor, Ecole centrale de Lyon, Lyon,
France; U. Schlichtmann, Technical University of Munich, Germany.

Abstract

For extremely Low-power Logic, three very new and promising techniques will be described. The first are methods on circuit and system level for reduced supply voltages. In large logic blocks, interconnect becomes a main issue, that could be solved by on-chip optical interconnect. Nano-devices will also be presented, as a possibility to compute with nearly zero power, and compared to future 10 nanometers transistors.

1. Ultra Low-Voltage Logic (C. Heer, U. Schlichtmann)

The progress of silicon process technology marches on relentlessly. As predicted by Gordon Moore decades ago, silicon process technology continues improvements at an astonishing pace [MOO]. The number of transistors that can be integrated on a single IC approximately doubles every 2 years [ITR].

1.1 Power Consumption Becomes Critical

With the exponentially increasing number of transistors on a die, power consumption has emerged as a key challenge to product success. This becomes more critical as many of today's high-volume consumer products are battery-powered.

Depending on the type of end-product and its usual application, different aspects of power consumption are the primary concern. Reduction of leakage power consumption today is primarily a concern for products that are mostly operating in some type of standby mode, such as cell phones.

Additionally, as leakage power increases dramatically with newer process generations, it is becoming a significant contribution to overall IC power consumption even in normal operating mode. It's contribution varies very significantly with type of transistor being used (e.g. high V_{th} vs. low V_{th}) and operating condition (esp. temperature).

Reduction of active power is a concern for almost all IC products today. For battery-powered products, reduced power consumption directly results in longer operating time. Even for non-battery-powered products, reduced power consumption brings advantages, such as allowing cheaper packaging, thus reducing cost. Especially for highly complex communications system ICs or microprocessors, reduced power consumption is required to enable heat dissipation at all and to avoid performance reduction due to excessive temperatures.

1.2 Traditional Approaches to Power Reduction

Our main focus will be active power consumption. It is determined by the following well-known relationship: $P \sim f * C * V^2$

This simple formula immediately identifies the key levers to reduce power:

- Reduce operating frequency
- Reduce driven capacity
- Reduce supply voltage

Traditionally, reduction in supply voltage has been the most often followed strategy to reduce power consumption. Unfortunately, it has the side effect of reducing performance as well, primarily because gate overdrive is diminishing as supply voltages are scaled faster than threshold voltages. As supply voltages are driven below 1.0V, these reductions are more pronounced than previously. In addition, newer process technologies give significantly less of a performance boost than has traditionally been the case, therefore a further reduction in performance is highly undesirable. Finally, as shown in Figure 1, the power reduction achieved by moving to a new process generation has trended down over time.
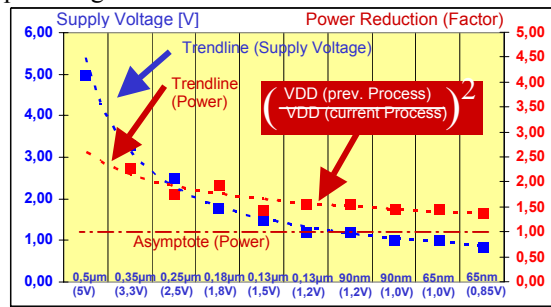

**Figure 1: Reduction of Supply Voltage**

Consequently, more advanced approaches are required. First, we will present an approach to ultra-low voltage logic and discuss its benefits and disadvantages. Then, we will discuss design techniques to reduce power consumption, which are partly already applied in leading-edge industrial ICs.

1.3 Zero-V_{th} devices

Developed in the mid 90-ies this concept overcomes the diminishing gate overdrive by radically setting the threshold voltage of the active devices to zero. Goal was an off-current of about 1/10 of the saturation current. Therefore the devices will never completely switch off. But from an overall power perspective the gain in active power consumption is tremendous.

On the other hand these devices become very susceptible to process and temperature variations. Significant yield is only achievable with back biasing via active well control. The ladder approach is not feasible for any mobile application. Therefore a more conservative approach with respect to zero-V_{th}, but still aggressive compared to current devices, had to

be chosen. An ultra-low V_{th} device with about 150mV threshold voltage proved to be the best compromise between zero-V_{th} and current low-V_{th} of about 300mV within a 0,13µm CMOS technology.
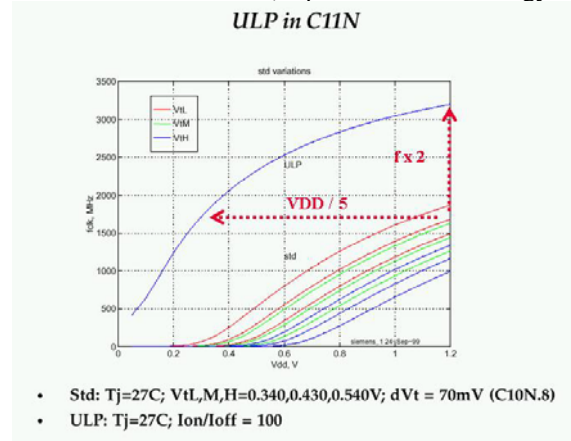
Figure 2: Zero-V_{th} concept

Fabrication of this ultra-low V_{th} device is possible, but affects some standard methods to overcome short-channel effects. The so-called halo- or pocket-implantation had to be removed to bring the threshold voltage down. Unfortunately short-channel effects are now heavily increased, leading to a very strong V_{th} roll-off with slight variations of the channel length. Finally this effect was prohibitive for the overall approach and led to cancellation of many zero-V_{th} projects in the industry.

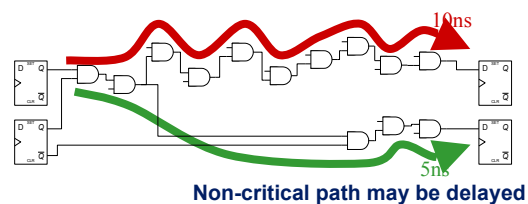**1.4 Design Approaches to Power Reduction**

Multi-Vdd design, as shown in Figure 3, recognizes that most logic paths in a design are not critical. They can be slowed down, often significantly, without reducing the overall system performance, but active power consumption.

While conceptually simple, the implementation is quite challenging. Different power supplies need to be managed. Level-shifters are required between different supply domains, increasing die area and necessitating tradeoffs between power reduction and area increase.

Another more essential technique to reduce leakage power is multi-V_{th} design. The idea is similar to multi-Vdd design: paths that do not need highest performance are implemented with special leakage-reduced transistors (higher V_{th} transistors, thicker gate-oxide). Design-tool support for this technique is also rudimentary at best. While it is becoming established to design different blocks with different transistors, it is very challenging to do this on the level of individual transistors.

But many large semiconductor companies already offer standard cell libraries with different base devices (low-leakage, regular and high performance devices). Replacement of cells within non-critical paths is easy as the same gate function has exactly the same footprint in all versions.

But also, manufacturing costs need to be taken into consideration as each new V_{th} or Tox requires typically two additional masks, driving costs up.
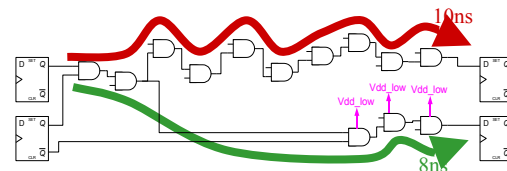
Figure 3: Multi-Vdd design

But often the largest power reductions are achieved by architectural approaches: partitioning the system such that large areas can be powered off for significant periods of time; partitioning memory systems such that large parts can be turned off in standby mode. Clock gating is a first step in that direction with local off-switching of non-active gates.

1.5 Conclusion and Future Challenges

There is no single "silver bullet" to solve the challenge of power reduction. While ultra-low voltage logic is a conceptually very convincing concept, its widespread implementation today is hindered by manufacturing concerns. Today, design techniques are most promising to reduce power – both active and leakage. Dynamically adjusted supply and threshold voltages are promising concepts which still require more investigation.

2. On-Chip Optical Interconnect for Low-Power (Ian O'Connor)

The role of interconnect as a dominant factor in determining circuit performance is growing in importance. Timing is already no longer the sole concern with physical layout: power consumption, thermal control, crosstalk and voltage drop drastically increase the complexity of the trade-off problem. Even with the most optimistic estimates for RC time constants using low-resistance metals such as copper and low-k dielectrics, global interconnect performance required for future generations of ICs cannot be achieved with metal. The 2001 ITRS [ITR] (International Technology Roadmap for Semiconductors) states that before 2010, "conventional interconnect scaling will no longer satisfy performance requirements. Defining and finding solutions beyond copper and low-k will require material innovation combined with accelerated design, packaging and unconventional interconnect".

2.1. Above Chip Interconnect

A promising approach to the interconnect problem is the use of an optical interconnect layer. Such a layer could empower an enormous bandwidth increase, immunity to electromagnetic noise, a decrease in the power consumption, synchronous operation within

the circuit and with other circuits, and reduced sensitivity to temperature variations. Important constraints when developing the optical interconnect layer are the fact that all fabrication steps should be compatible with future IC technology and that the additional cost incurred remains affordable. Difficulties expected are obtaining a large enough optical-electrical conversion efficiency, reducing the optical transmission losses while allowing for a sufficient density of photonic waveguides on the circuit and reduction of the latency while operating above the 10GHz mark.

Hybrid (3D) integration of the optical layer on top of a complete CMOS IC (Fig. 4) is a realistic solution, with good scope for evolving. A microlaser such as that described in [FUJ] is coupled to the passive waveguide structure and provides a signal to an optical receiver (or possibly to several, as in the case of a broadcast function). At the receiver, the high-speed optical signal is converted to an electrical signal and subsequently distributed by a local electrical interconnect network. The source and detector devices are realised in III-V materials and bonded to silicon using advanced heterogeneous grafting technology [SEA].

To form a planar optical waveguide, silicon is used as the core and SiO_2 as the cladding material. Si/SiO_2 structures are compatible with conventional silicon technology and transparent for 1.3-1.55μm wavelengths. Such waveguides with high relative refractive index difference $\Delta \approx (n_1^2 - n_2^2)/2n_1^2$ between the core ($n_1 \approx 3.5$ for Si) and claddings ($n_2 \approx 1.5$ for SiO_2) allow the realisation of a compact optical circuit, with bend radius of the order of a few μm [SAK]. For a wavelength of 1.55μm, single-mode conditions impose a waveguide width of 0.3μm.
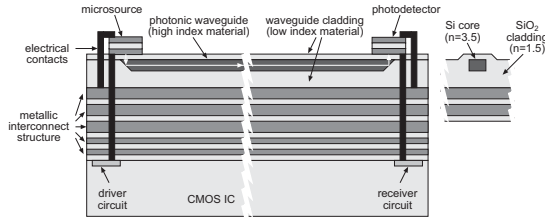

Figure 4. Cross section of hybridised interconnection structure

The optical process is completely independent from the CMOS process, which is appealing from an industrial point of view. Disadvantages of this approach include the complex electrical link between the CMOS subcircuits (source drivers and detector amplifiers) and inevitably more advanced technological solutions for bonding.

2.2 Clock Distribution (One Point to Multipoint)

One application of optical interconnect technology is in clock distribution networks (CDN) [FRI]. By replacing the electrical clock distribution tree by an optical one, the need for repeaters or clock multiplier circuits would be eliminated, thus reducing power consumption and clock skew. However, it would be illusory to believe that the optical clock signal could be routed down to the single-gate level: optoelectronic interface circuits are of course necessary and consume power. An example system realizing a clock distribution function, shown in Fig. 5, requires a single photonic source coupled to a symmetrical waveguide structure routing to a number of optical receivers. At the receivers the high-speed optical signal is converted to an electrical one and provided to local electrical networks. The number of clock distribution points is a particularly crucial parameter in the overall system.
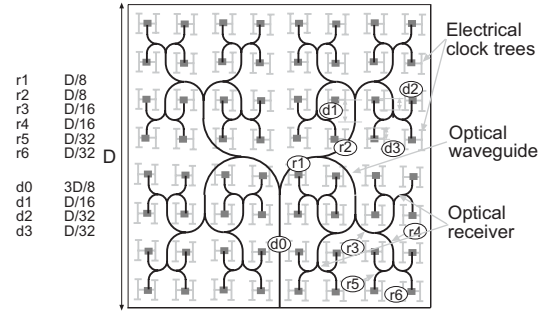
Figure 5. Optical H-tree network with 64 output nodes.

The global optical H-tree was optimised to achieve minimal optical losses. The radius of curvatures are designed to be as large as possible. For 20mm die width and 64 output nodes in the H-tree at the 70nm technology mode, the smallest radius of curvature (r_5, r_6 in Fig. 5) is 625μm, which leads to negligible pure bending loss.

2.3. Power Consumption Comparison

The power dissipated in optimised clock distribution networks was analysed for both electrical and optical systems at the 70nm technology node. Key parameters for the optical system are summarised in Table 1.

In order to calculate the power dissipated in the electrical system we used an internally developed simulator, which allowed us to model and calculate the electrical parameters of clock networks for future technology nodes [TOS]. Berkeley BSIM3v3 [CAO] parameters were used to model the transistors used in the buffers. The power dissipated in the system is extracted from transistor-level simulations.

|	
Wavelength λ [μm]	1.55
Waveguide core index (Si)	3.47
Waveguide cladding index (SiO_2)	1.44
Waveguide thickness [μm]	0.2
Waveguide width [μm]	0.5
Transmission loss [dB/cm]	1.3
Loss per Y-junction [dB]	0.2
Input coupling coefficient [%]	50
Photodiode capacitance [fF]	100
Photodiode responsivity [A/W]	0.95

Table 1. Optical system parameters

To estimate the electrical power dissipated in the optical system we used the methodology shown in Fig. 6. We assumed the use of an external VCSEL

(vertical cavity surface emitting laser) source, rather than an integrated microsource.
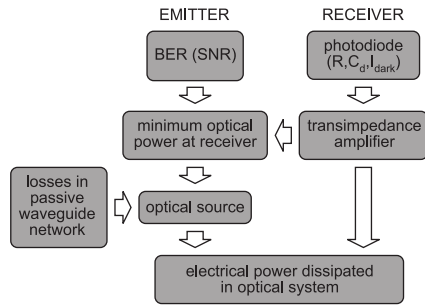
Figure 6. Methodology used to estimate the electrical power dissipation in an optical clock distribution network

Based on given photodiode parameters (parasitic capacitance C_d, responsivity R and dark current I_{dark}), the method described in [OCO] is used for the automated design of a maximum bandwidth transimpedance amplifier. Next, for a given system performance (BER, bit error rate) and for the noise signal associated with the photodiode and transimpedance circuit we calculate the minimum optical power required by the receiver to operate at the given error probability [MOR] in the preamplifier noise calculations. This power figure, coupled with the losses incurred through the passive optical waveguides enables us to estimate the optical power that needs to be emitted by the VCSEL. The electrical power dissipated in optical clock networks is the sum of the power dissipated by the number of optical receivers and the energy needed by the VCSEL to provide the required optical power. For a BER of 10^{-15} the minimal power required by the receiver is -22.3dBm (at 3GHz).
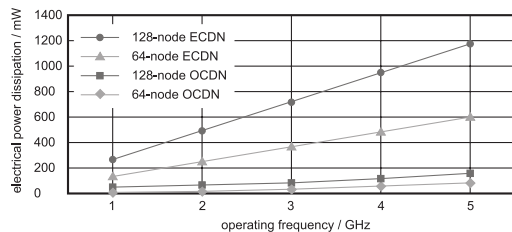
Figure 7. Electrical power dissipated by electrical (ECDN) and optical (OCDN) clock distribution networks for varying operating frequency

The comparison in terms of dissipated power between the optical and electrical global clock distribution networks is shown in Fig. 7. It can be seen that the power dissipated by the electrical system is highly dependent on the operating frequency, while in the optical system, it remains almost the same. The difference between the power dissipated in both systems is clearly higher if we increase the frequency and number of nodes in H-trees. For a classical 64-node H-tree at 5GHz frequency the power consumption in the optical CDN should be 5 times lower than in an electrical network.

2.4. Future of On-Chip Optical Interconnect

New research into the possibility of on-chip wavelength division multiplexing (WDM) is currently underway at LEOM. For example, a single waveguide could be used to replace a 64-bit bus, where each individual signal makes use of a distinct wavelength. Reconfigurable networks could also be realised in the optical domain using compact micro-resonators, capable of selecting and redirecting a signal based on its wavelength. Such networks would be entirely passive; i.e. no power would be required to transport the data, whatever the communication route necessary, leading to power reduction and higher integration density. However, such a scheme would imply a shift in the routing paradigm from a centralised arbiter acting on the switch boxes, to one acting on the block interfaces to select the wavelength(s) to be used. Also, tunable and thermally stable microlasers would be required.

3. Nano-Devices for Low-Power (Jacques Gautier)

In addition to packing-density increase and speed improvement, the scaling down of technologies comes with a reduction of the power consumption per function. However this gain is offset by the tremendous increase in the number of transistors per chip. A possible solution is to go further towards nano-scale devices where a lower amount of charge is needed to code a bit. This is the basis of what is known as single electronics. Using molecules could be a realistic way to fabricate these tiny devices and other useful nanostructures.

3.1 Single electronics

The dynamic power consumption in CMOS circuits, $P=a.C.V^2.f$, can also been expressed as $P=a.Q.V.f$ (1), where Q is the amount of charge in transistors and interconnects for coding a bit of information. So it is clear that reducing it would result in power saving. Looking at the transistor contribution, mainly the channel charge, a value of about only ten electrons is expected for 10nm MOSFET's from the extrapolation of the historical trend and from the ITRS roadmap anticipation [ITR]. This is much lower than the thousands of electrons in current devices. It is possible to go still further, towards only one electron, using what is call a single electron transistor or SET [GRA]. However this gain will be effective only if the capacitance of interconnect is accordingly reduced. Another factor in expression (1) is the electrostatic potential to which the charge Q is brought. At present, there is a strong incentive for reducing it. Whereas the supply voltage of current High Performance circuits is in the range 1.2-1.8V, operation at only 0.3V on experimental circuits has already been demonstrated [DOU], which is close to the bottom limit anticipated by the ITRS. For a lower value the device is never in well defined On or Off states which results in either leakage or poor performance. What can be expected from SET's ?

To answer this question, their properties and mode of operation are briefly recalled. A SET is a device which comprises two Source and Drain reservoirs of electrons and a control gate, like in a MOSFET. In

between, there is an island where carriers should be confined. A common solution to do that, is to insert tunnelling or potentials barriers between the reservoirs and the island. This is the main structural difference from MOSFET's, but it is essential for the operation of SET's. Due to this confinement, there is always an integer number of electron in the island. When the external biasing is varied, there are electrostatic conditions to meet to change this number, provided that just one more or less electron has a significant effect on the electrostatic energy of the island compared with kT. This is the Coulomb blockade effect. The corresponding condition is:

$$E_C = \frac{e^2}{2C_\Sigma} >> kT \,,$$ where C_Σ is the total

capacitance of the island ($C_\Sigma=2C_j+C_G$). For room temperature operation, it should be less than 0.3 aF, which requires an island smaller than a few nm. The condition related to the confinement is that the resistance of tunnel barriers exceeds the resistance quantum $R_K=h/e^2\sim25.8$ kΩ.

As shown in figure 8, the characteristics of SET's are very different from those of MOSFET's. In the $I_D(V_G)$ curves there are periodic oscillations of current and the output characteristics look like a resistance with a low drain voltage domain where the device is periodically Off/On as a function of V_G. The period of Coulomb blockade oscillations, CBO, is given by e/C_G, where C_G is the gate to island capacitance.

One can notice that room temperature operation implies a low value of Cg which leads to a period higher than 0.5V. Consequently, for low V_{DD} circuits, only a part of a period can be exploited. In the case of logic gates, the current peak voltages of pull-up and pull-down devices should be tuned to make a complementary of action and to obtain an effect equivalent to the existence of PMOS and NMOS transistors. A solution is to add a second gate to tune the phase of CBO [TAK]. Such a gate can also be used to balance the shift of CBO resulting from the influence of possible parasitic (background) charges in the neighbourhood of the SET.
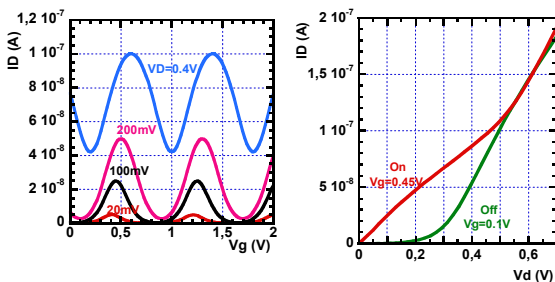
Figure 8: simulated $I_D(V_G)$ and $I_D(V_D)$ characteristics of SET's. The main parameters are: Cj=0.1aF, Cg=0.2aF, Rj=1MΩ, T=300K

An example of transfer characteristic is given in Figure 9 for different temperatures. Parameters have been chosen for V_{DD}=0.2V, including control gate

biases. At 300K the curve is not so ideal as for lower temperatures. The reason is the all the more soft transition between On and Off states as the temperature increases, which is equivalent to the subthreshold current effect of MOSFET's. This can be clearly seen in Figure 10 where the static current in the inverter has been plotted. A quasi-CMOS operation inverter at a supply voltage as low as 20mV has been demonstrated by NTT [TAK], but in this case the temperature was only 27K.
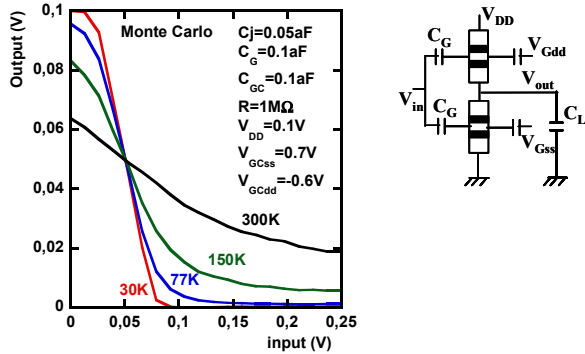
Figure 9: transfer characteristics of a SET inverter

Another possibility to reduce the power consumption is based on the use of SET's to build logic gates with increased functionality. A first example has been the RT demonstration by Toshiba [UCH] of an hybrid SET-MOSFET gate in which a SET operates either in a positive or in a negative transconductance region, depending on a charge stored in the neighbourhood. This way, the gate can be programmed to be inverting or non-inverting. Compact logic gates have also been designed by NTT with multiple inputs SET's [TAK]. In both cases, for a given targeted logic function, the transistor count is reduced compared to a CMOS design. Moreover, SET's can be used for memory and analog applications.
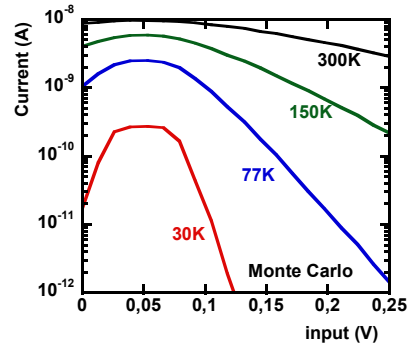
Figure 10: current in a SET inverter

3.2 Molecular electronics

For the fabrication of SET's, any kind of conducting material can be used. Whereas the basic research was done on metallic SET's [GRA], circuit demonstrations are performed mainly on silicon [TAK, UCH], for complementary with MOSFET's, but it could be advantageous to use molecules for real applications. The reason is the size requirement

discussed previously. Also, the load capacitance should be very low for power consumption and speed considerations which implies short interconnects. The most promising way to achieve that is the bottom-up approach, using naturally formed tiny structures or self-assembled methods. The best example is the carbon nanotube (CNT) which can be used to fabricate MOSFET's [JAV], SET's [MAT], interconnects [LI] and even non volatile memory arrays [NAN].

CNT's are long cylinder of carbon atoms consisting of rolled-up sheets of graphite. For Single Wall CNT's the diameter is as small as 1-5 nm. Depending on their chirality, they are semiconductor or metallic materials. Their mobility is higher than the one of silicon, and a ballistic transport has been demonstrated for lengths less than a few hundreds of nm, but the subthreshold characteristics of CNFET's are not better than those of MOSFET's. Worldwide, several teams are conducting research on the selective growth or deposition of CNT's that would have the right chirality and on the evaluation of CNFET's as potential candidate to replace MOSFET's in the future.

Different kinds of molecules are also currently investigated to make nanometer scale electronic components and circuits, but a single molecule transistor has not yet been obtained. To date, one of the most advanced achievements is a $1\mu m^2$ 64 cells crossbar matrix fabricated by HP Labs. [CHE]. The switching units are bundle of rotaxane molecules.

True single molecule device will require interconnects at a similar scale. This is also essential to reduce parasitic capacitances and the power consumption. Since the needed resolution is far beyond the possibility of lithographic tools, including NGL, the solution will come from the bottom-up approach. An example is the realization by Caltech of a Pt nanowire lattice with widths and pitches of 8 nm and 16 nm respectively [MEL]. Biology can also come to the rescue for the self-assembly of nano-circuits [FAI].

3.3 Discussion

For SET's or CNT devices, as well as for nano MOSFET's, the supply voltage reduction is dependent upon the effects of leakage mechanisms. Although it is not yet clear if they could achieve a lower V_{DD}, there are other candidates, like the resonant tunnelling devices (RTD). Their operation is based on electron transport via discrete energy levels in double barrier quantum well structures, which results in the existence of a negative differential resistance. This implies a perfect control of the geometry and the fabrication in suitable materials. A promising approach is their implementation along semiconductor nanowires [BJO], but they are also studies for a molecular version.

The most important feature of the nanodevices, especially the molecular ones, is their size. That offers the possibility to reduce the power consumption by parallel processing. For instance, consider two blocks of low capacitance molecular devices doing the same task as one block of conventional devices, but at half the clock frequency. Going further, molecular Quantum-dot Cellular Automata is an attractive way, yet speculative, to reduce the power consumption, since there is no flow of current but only Coulomb interaction [LEN].

References

[MOO] Moore, G., Cramming more components onto integrated circuits. Electronics, Vol. 38, No. 8, 1965.

[AMA] Amann, M.C., Ortsiefer, M. and Shau, R., Surface-emitting Laser Diodes for Tele-communications, in *Proc. Symp. Opto- and Microelectronic Devices and Circuits*, Stuttgart, 2002

[CAO] Cao, Y. et al., New paradigm of predictive MOSFET and interconnect modeling for early circuit design, in *Proc. Custom Integrated Circuit Conference*, 2000

[FRI] Friedman E.G., Clock distribution networks in syn-chronous digital integrated circuits, *Proc. IEEE*, 89, 665, 2001

[FUJ] Fujita, M., et al. Ultrasmall and ultralow threshold GaInAsP-InP microdisk injection lasers: Design, fabrication, lasing characteristics and spontaneous emission factor, *IEEE J. Sel. Topics in Quantum Electronics*, 5, 673, 1999

[ITR] Semiconductor Industry Association, International Technology Roadmap for Semiconductors 2001 Edition, http://public.itrs.net, 2001

[MOR] Morikuni, J.J. et al., Improvements to the standard theory for photoreceiver noise, *IEEE J. Lightwave Techn*, 12, 1174, 1994

[OCO] I. O'Connor et al., "Predictive Design Space Exploration of Maximum Bandwidth CMOS Photoreceiver Preamplifiers", in *Proc. IEEE Intl Conf. on Electronics, Circuits and Systems*, 2003

[SAK] Sakai, A., Hara, G. and Baba, T., Propagation Characteristics of Ultrahigh-Δ Optical Waveguide on Silicon-on-Insulator Substrate, *Jpn. J. Appl. Phys. - Part 2*, 40, 383, 2001

[SEA] Seassal, C. et al., InP microdisk lasers on silicon wafer: CW room temperature operation at 1.6µm, *IEE Electron. Lett.*, 37, 222, 2001

[TOS] G. Tosik et al., "Optical versus metallic interconnections for clock distribution networks in new VLSI technologies", in *PATMOS 2003*, pp. 461-470, 2003

[DOU] T. Douseki et al., "A 0.3V 3.6GHz 0.3mW frequency divider with differential ED-CMOS/SOI circuit technology", in Proc. ISSCC, February 2003

[GRA] H. Grabert and M. Devoret, "Single charge tunnelling Coulomb blockade phenomena in nanostructures", Vol 294 of NATO ASI Series B, Plenum Press New York / London, 1992

[TAK] Y. Takahashi et al., "Silicon Single-Electron Devices for logic applications", in Proc. ESSDERC September 2002, Florence, pp. 61-68

[UCH] K. Uchida et al., "Room-temperature operation of multifunctional Single-Electron Transistor logic", in Proc. IEDM, pp.863-865, 2000

[LI] J. Li et al., "Bottom-up approach for carbon nanotube interconnects", APL Vol. 82, N°15, pp. 2491-2493, 2003

[MAT] M. Matsumoto et al., "Single-Electron Transistor with ultra-high Coulomb energy of 5000K using position controlled grown carbon nanotube as channel", JJAP Vol.42 Part 1 N°4B, pp. 2415-2418, 2003

[CHE] Y. Chen et al., "Nanoscale molecular-switch crossbar circuits", Nanotechnology 14 (2003) 462-468

[JAV] A. Javey et al., "High-K dielectrics for advanced carbon-nanotube transistors and logic gates", Nature Materials, Vol 1, pp. 241-246, December 2002

[MEL] N. Melosh et al., "Ultrahigh-density nanowire lattices and circuits", Science, Vol. 300, pp.112-115, 4 April 2003

[BJO] M. Björk et al., "Nanowire resonant tunneling diodes", APL, Vol. 81, N°23, pp. 4458-4460, December 2002

[LEN] C. Lent et al., "Molecular Quantum-dot Cellular Automata", J. Am. Chem. Soc. 125, pp. 1056-1063, 2003

[FAI] P. Fairley, "Germs that build circuits", IEEE Spectrum, pp. 37-41, November 2003

[NAN] http://www.nantero.com/pdf/Press_Release%205_%2003.pdf