Extended Abstracts of the 22nd (1990 International) Conference on Solid State Devices and Materials, Sendai, 1990, pp. 353-356

Invited

Low Temperature Microelectronics

Fritz H. Gaensslen
IBM T. J. Watson Research Center
Yorktown Heights, NY 10598

Richard C. Jaeger
Electrical Engineering Department
Auburn University, AL 36849

Past and present developments toward operating microelectronic circuits at low temperature are reviewed. To assess the future potential of this approach, its advantages and disadvantages are discussed. The fact that devices and materials behave generically better at low temperature will have some bearing on the ultimate attainable technology limits.

1. Introduction

Over the past twenty years, integrated circuit cost and performance improvements have been achieved with relative ease through continued miniaturization of device geometries. In the submicron regime, however, performance enhancements are becoming more and more difficult and expensive to attain solely through feature size reduction. In addition, the design of high performance computing systems is becoming less and less dominated by raw device switching speed. High circuit density combined with low per circuit power dissipation are greatly increasing in importance. Superior switching speed is offered by bipolar technology, but CMOS technology is superior in circuit density and power dissipation.

Operation of CMOS at Liquid Nitrogen Temperature (LNT: 77.3K) can directly yield a significant performance improvement exceeding that provided through down-scaling by a factor of two. Thus, CMOS technology limits are extended by one generation through LNT operation; therefore we expect low temperature operation to become increasingly competitive in the future. Because of its very attractive speed, density, power and leakage attributes, and its reliability potential, liquid nitrogen cooled CMOS represents a strategic technology alternative for large scale computers, an area which has traditionally been the exclusive domain of bipolar technology.

Contrary to the long held belief that bipolar devices are unusable at low temperature, recent work has shown that large bipolar current gains can be maintained at low temperature [1]. However, overall bipolar performance typically peaks at temperatures around 150K [2-4], and identifying proper methods for maintaining high values of f_T at LNT still represents an open question. It remains unclear whether a decrease in cutoff frequency or the increase in bipolar transconductance will be most important in determining the performance limitations of BICMOS at LNT.

2. Past Developments

Possible advantages of operating MOS devices in a computer system at low temperature began to be explored in detail during the early 1970's, when it became clear that the high device and circuit density and low power associated with future CMOS technologies would result in a computer system packaged in a small, readily-cooled volume [5-7]. The development of MOS device scaling theory [8] provided a methodology for systematic miniaturization of MOS devices. One important limitation to this theory is the lack of scaling in the subthreshold region. Device turn-off is extremely important for operation of dynamic random access memories. Lowering the ambient temperature provides the additional variable needed to enhance the subthreshold slope. This fact constituted the original impetus to look at low temperature MOS operation.

Proportionality between subthreshold slope and reciprocal temperature was demonstrated in [7], and it was also established that the classical threshold voltage theory still holds for enhancement-mode devices at LNT. Since miniaturization and low temperature operation are closely linked together, it is essential that geometry or size effects do not get worse at lower temperature. Device currents increase at 77K in accordance with the carrier velocity versus electric field relationships. Approximate drain current improvements for low and high fields range from 6 to 1.6 times. Similar improvements are experienced by n-channel and p-channel devices.

Leakage currents of pn-junctions, as a result of an exponential temperature dependance of the intrinsic carrier concentration, are sharply reduced at LNT. The theoretical change is about 30 orders of magnitude, while at least 5 to 6 can be realized in a practical environment. The junction breakdown voltage at LNT is about 10 to 15% lower due to a lengthened mean free path between collisions; this poses no problem.

Electrical and thermal conductivities also increase substantially (5-6x) at LNT. The improvements in mobility, subthreshold slope and other more subtle changes yield generically better MOS channels at LNT than at normal operating temperature. In addition, low-temperature operation of CMOS eliminates parasitic latchup as a concern [9,10].

Problems with impurity freezeout in common buried-channel depletion-mode devices [11,12], obviate their straightforward use at low temperatures. Exactly the same effects, related to the parallel existence of a surface and a buried channel, occur in p-channel enhancement-mode devices with n-doped poly gates, as currently fabricated in n-well CMOS processes.

Operation at LNT yields a more tightly controlled temperature environment and hence tighter distributions of electrical parameters enabling the design of better circuits. Most wearout failure mechanisms follow an Arrhenius relationship and decrease exponentially with temperature. Low temperature operation should greatly improve reliability, which is yet to be substantiated experimentally. The soft error rate, an important reliability detractor, has been reported to decrease significantly with temperature [13].

A summary of the advantages for CMOS microelectronics operated at LNT is shown below:

* Steeper Subthreshold Slope
* Higher Transconductance
* Well Defined Threshold Voltage Behavior
* No Degradation of Geometry Effects
* Dramatic Junction Leakage Reduction
* Enhanced Electrical Line Conductivity
* Improved Thermal Conductivity
* Bulk CMOS Latchup Suppression
* Reduced Soft Error Rates
* Improved Reliability
* Improved Digital Circuit Performance
* Compatible Analog Circuits

Because of low power consumption and a high level of integration, CMOS is the preferred VLSI technology; operation at LNT can directly yield a 2.5 to 3 times improvement in switching speed, when compared to operation at 358K (85C). This performance improvement exceeds that achieved through down-scaling by a factor of 2. LNT operation extends technology limits at least by one generation. Most performance data available so far present results of cooling room temperature technologies to 77K; optimized LNT technologies will yield substantial improvements. Here are some steps towards optimization. A new low temperature threshold voltage design is needed to compensate for a temperature dependent rise and to exploit the much steeper subthreshold slope. To force a clean surface p-channel device not affected by freezeout, a p+ polysilicon gate is required [14]. Use of $<110>$ silicon wafer material instead of $<100>$ can strike a better balance between p- and n-channel device performance [15]. A novel MOS temperature scaling concept described in [16] forces the channel carrier distribution to stay unchanged with temperature. The argument of a Fermi-Dirac distribution can be rendered invariant by changing temperature and potential simultaneously to maintain a constant ratio.

In addition to an optimized semiconductor technology, cooling and packaging technologies are required. Low temperature packaging also benefits from some of the above mentioned advantages. Due to higher mobilities hot electron device degradation was considered a potential problem area, but recent work indicates that adequate design margins can be maintained in LNT devices of proper design [17-19].

3. Memory and Logic at Low Temperature

A reduction in both subthreshold and junction leakage by several orders of magnitude occurs at LNT; this fact suggests a "pseudo-static" mode of operation of otherwise dynamic circuitry. When operated at low temperature, we can expect DRAM memories to retain information for periods of days or possibly more than a week [20] without being refreshed.

Three-device dynamic random access memory cell circuits are also possible [21]. Essentially they consist of a one-device cell with added sense circuitry. Nondestructive sensing is achieved through a d.c. path with a sense device in series with a read-select device. At low temperature multiple read operations are possible without requiring a rewrite operation - a "pseudo-static" mode of memory operation. In such a circuit the sense energy is not limited to merely the energy stored on the capacitor, as in a one-device cell. The sense energy is supplied from outside of the array

area making such a cell relatively insensitive to the specifics of the array design and to further scaling. With a sense path similar to those found in static cells, access is fast with no danger of upsetting the stored information. With overall power dissipation very low, the three device cell is very promising for a fast, low temperature memory design. An interesting approach based on a two device cell has been reported recently [22]; all the above mentioned advantages are retained with an improved current mode sensing scheme also included.

At room temperature static logic is dominant, since dynamic logic requires phase clocks with their respective distribution networks. Because of disappearing leakages a pseudo-static mode of dynamic logic also becomes possible at LNT as demonstrated in [23].

4. Computer Systems Applications

Low temperature CMOS technology allows more computing power per given volume, a most important consideration for end users. Evidence for the technical viability of low temperature operation was provided when the first liquid nitrogen cooled CMOS computer system, the ETA 10, was shipped at the end of 1986 by ETA Systems Inc., a subsidiary of Control Data Corporation [24]. Regular Honeywell CMOS hardware not optimized for LNT operation was used for this system. Since this hardware is functional from LNT to 85C (358K), high temperature systems with reduced performance can share the same technology. Furthermore, functional and margin testing for LNT can be done at room temperature to LNT specifications. By technology optimization, some of these test capabilities might get compromised.

Raw technology switching speed is no longer the dominant consideration for the design of high performance systems; the level of integration and power dissipation also have an important impact. The table below compares some of the attributes of an eight processor system with approximately 3.5 million circuits per processor, implemented in one micron technology (ETA 10), with those of a future 0.25 micron process optimized for LNT. The achievable performance as well as the reduction in overall system power and volume are impressive. Based on a liquid nitrogen cooled quarter micron CMOS technology, chips with several hundred thousand circuits and a gate delay below 100 picoseconds will be available to design future digital systems.

Although large systems have used liquid cooling (water or Freon) for a long time, there remains a widely perceived difficulty with liquid nitrogen cooling. Economical cooling systems are possible for the power levels needed at LNT with better then 10% efficiency [25]. Such a cooling system, though certainly not small nor free, occupies no more volume than the water cooling systems of competitive ECL machines. As a coolant, liquid nitrogen is inert and inexpensive, and its boiling temperature provides an excellent compromise between improved system performance and cooling costs. Large scale success of LNT systems, nevertheless, requires development of dedicated refrigeration equipment.

5. Low Temperature Research Activities

Recently, activities in low temperature electronics have greatly increased throughout the world. This growing attention manifests itself by the number of publications on the subject [26], of regular and panel sessions at major electronics conferences, and of special Low Temperature Electronics issues by scientific periodicals [27,28].

New superconducting materials have also given additional focus to low temperature research. We do not, however, expect these materials to have an immediate effect on IC chip design because on-chip switching action will still be limited by the non-superconducting characteristics of MOS devices [29]. At 77K the standard on-chip metal line pattern

3.5 MILLION CIRCUIT PROC.:	TODAY (ETA 10)	FUTURE (OPTIMIZED)
Effective Channel Length:	1.0 μm	0.25 μm
Gate Delay:	400 psec	70 psec
Number of Circuits/Chip:	20K	320K
Number of Chips/Processor:	240	15
Power per Circuit/Chip/Proc./8 Proc.Total:	0.1mW/2W/500W/4kW	6.25μW/2W/30W/240W

already offers vastly improved conductivity and resistance to electromigration. On-chip compatibility between superconducting materials and highly sensitive MOS technologies is a critical question. To go along with superconducting interconnections, there is an acute need for a three-terminal, superconducting switch. Long term stability and the capacity to support sufficiently high current densities for high T_c films are assumed. At the packaging board level, however, superconducting interconnections may well have an important effect on the distribution of power and signals over long distances with better performance and/or noise immunity as a result.

6. Summary

Low temperature operation of CMOS technology can economically combine bipolar speeds with MOS yields, circuit densities and power levels, resulting in compact, power-efficient and fast systems [24,30]. Miniaturization and low temperature operation help each other in cooling, performance and reduction of parasitics. Functional and margin testing for low temperature can be done conveniently at room temperature to low temperature specifications. In addition, recent work has demonstrated the functionality of CMOS linear circuits at LNT [31]. Recent advances in superconductors will increase the thrust toward low temperature systems, although revolutionary changes are not expected to occur overnight.

The main reason for not going to low temperature operation in the past has probably been a simple lack of need. Throughout the last two decades, performance improvements have been achieved with relative ease through continued miniaturization of device geometries. Future enhancements, however, are becoming more and more difficult and expensive to realize through feature size reductions as we reach to well below one micrometer. This fact together with all the mentioned advantages makes low temperature operation increasingly attractive and ultimately necessary.

7. References

[1] J. D. Cressler et al., IEEE Trans., ED-36, pp.1489-1502, 1989.

[2] J. M. C. Stork et al., IEEE Trans., ED-36, pp.1503-1509, 1989.

[3] W. C. Dillard et al., IEEE Trans., ED-34, pp.139-142, 1987.

[4] M. Chrzanowska-Jeske et al., IEEE Trans., ED-36, pp.1475-1488, 1989.

[5] F. H. Gaensslen et al., IEEE IEDM, pp.43-46, 1975.

[6] R. L. Maddox, IEEE Trans., ED-23, pp.16-21, 1976.

[7] F. H. Gaensslen et al., IEEE Trans., ED-24, pp.218-229, 1977.

[8] R. H. Dennard et al., IEEE IEDM, p.168, 1972.

[9] F. H. Gaensslen, IEEE ICCD, pp.450-452, 1980.

[10] J. G. Dooley et al., IEEE, EDL-5, pp.41-43, 1984.

[11] F. H. Gaensslen et al., Sol.-St. Electr., vol.22, pp.423-430, 1979.

[12] R. C. Jaeger et al., IEEE Trans., ED-26, pp.501-508, 1979.

[13] W. H. Henkels et al., IEEE Trans., ED-36, pp.1414-1422, 1989.

[14] J. Y. Sun et al., IEEE Trans., ED-34, pp.19-27, 1987.

[15] M. Aoki et al., IEEE Trans., ED-36, pp.1429-1433, 1989.

[16] K. Masu et al., Proc. LT Semicond. Electronics Workshop, Burlington,VT, pp.104-108, 1989.

[17] M. Itsumi, J. Appl. Phys., vol.54, pp.1930-1936, 1983.

[18] D. Lau et al. IEEE IEDM, pp.565-568, 1985.

[19] A. K. Henning et al., IEEE Trans., ED-34, pp.64-67, 1987.

[20] W. Link et al., Arch. f. El. und Ubertr., vol.33, pp.229-235, 1979.

[21] W. M. Regitz et al., IEEE J. Sol. St. Circ., SC-5, pp.181-186, 1970.

[22] T. L. Blalock et al., Proc. VLSI Circ. Symp., pp.13-14, 1990.

[23] S. Hanamura et al., IEEE Trans., ED-34, pp.94-100, 1987.

[24] T. Vacca et al., VLSI Syst. Des., vol.8, pp.80-88, 1987.

[25] R. C. Longsworth et al., IEEE Trans., ED-34, pp.4-7, 1987.

[26] R. K. Kirschman, Editor, IEEE PRESS, NewYork, 1986.

[27] F. H. Gaensslen, R. C. Jaeger, Editors, Spec. Iss. IEEE Trans., ED-34, 1987 and ED-36, 1989.

[28] L. Clayton, Editor, Spec. Iss. Cryogenics, to appear Oct.,1990.

[29] J. T. Watt et al., IEEE IEDM, pp.393-396, 1987.

[30] G. Gildenblat et al., IEEE IEDM, pp.268-271, 1985.

[31] R. M. Fox et al., IEEE Trans., ED-34, pp.114-123, 1987.