Competing Technologies for Digital VLSI:
Silicon CMOS and ECL, Gallium Arsenide, and Silicon on Insulator

Stephanie Oberg.

Summary

Several technologies are capable of being used for custom, semi-custom, or programmable VLSI digital logic circuits and memories. This report compares current TMC technology (silicon CMOS or complementary metal-oxide-semiconductor) with several alternatives: silicon ECL (emitter-coupled logic), various gallium arsenide logic families, and silicon-on-insulator (SOI). The following factors are compared (where information is available): 1) inherent capabilities of the technology, 2) process maturity and trends, 3) current and projected future costs, and 4) system integration and compatibility with current technology, from design through assembly. None of these technologies will be pushing silicon CMOS out of competition in the next two years or so, although in gallium arsenide especially, the integration level, marketability, and producibility of circuits increased tremendously in 1990. When time permits, Thinking Machines systems and chip designers should begin active discussions with vendors of alternative technologies.

Purpose

This report is meant to describe the directions in which technologies are heading and recommend TMC actions. A company our size probably cannot independently decide to use a technology that "the market" has not accepted. This report covers development possible over the next ten years or so.

Overview of gallium arsenide as a VLSI material

Electron mobility in gallium arsenide (a compound semiconductor, known by its chemical symbol, GaAs) is about six times higher than it is in silicon (Si), which means simplistically that transistors can switch six times faster and makes gallium arsenide an attractive research target for high-speed devices. Gallium arsenide circuits probably should not directly replace any in use currently; rather, to quote one source, "future GaAs ICs will complement, not compete with, silicon VLSICs. The former will emphasize moderate gate counts and high gate speeds, while the latter will exploit slower individual gates but high gate counts. Complete advantage can be taken of gallium arsenide digital technology only if traditional signal-processor architectures are completely recast at the memory layout, logic design, arithmetic implementation, and system architecture levels." [1] Developments in 1990 and early 1991 have changed the situation; gallium arsenide gate counts are now comparable to those of CMOS gate arrays.

Silicon and gallium arsenide IC R&D began at roughly the same time (early 1960's). Silicon pulled ahead and has been developed to the level where scientific and engineering understanding of every step that goes into the production of ICs is thorough. Digital

gallium arsenide has taken decades to get off the ground because each new effort to develop it has met fresh frustrations: crystal growth problems, lack of understanding of Schottky energy barriers and surface states which are inherent to all gallium arsenide devices, difficulty in controlling dopants, the decision to use microwave gallium arsenide gold-based metallization technology which is not easily adaptable to high integration levels, etc. Until recently, gallium arsenide technology would not have been a good deal in terms of its "risk cost" and compatibility with TMC philosophy. The situation is now changing (see below).

Gallium arsenide has been widely used for analog microwave devices to operate at 0.5 to 40 GHz, where it makes sense because wavelengths are comparable to substrate sizes, the availability of a semi-insulating gallium arsenide substrate reduces signal losses and decreases process complexity, the required integration levels for analog circuits are low (usually less than twenty transistors per circuit) so lower yields don't hurt as much, and the military communications market has tended to tolerate high-priced, low yielding lab curiosities. Digital markets are much less tolerant of such demanding circuits and so gallium arsenide has not been able to get off the ground until, according to Vitesse and Gazelle, foundries adopted the approach of giving gallium arsenide the "look and feel" of silicon ECL or TTL (transistor-transistor logic) technology, particularly as regards the supply voltages and other design considerations (but with lower power consumption than CMOS).

Where gallium arsenide is discussed, this paper will focus on the capabilities of Vitesse, since it is by far the farthest along and I have the most info on it. Maturity level: 7

Frequently-mentioned attractions of gallium arsenide

1) *Faster switching times (for equivalent power) because of higher electron mobility.* True (see "Electron mobility" in table below). The question is whether faster switching speeds alone will significantly improve our product. The favorable inherent qualities of gallium arsenide cannot always be fully used in a practical implementation. For example, gallium arsenide devices are fastest with a supply voltage of 1 to 1.5 V, but 5 V (or later 3.3 V) is more practical in a systems application since those voltages will already be present. The new strategy is to use CMOS-compatible voltages and accept non-optimum (but still better than competing technologies') performance from the gallium arsenide.

2) *Larger temperature range of operation.* Gallium arsenide devices can operate at higher temperatures without being swamped by thermal generation of carriers. At high temperatures (e.g. channel at 150 °C), however, although short-term operation is not impaired, reliability is a concern. Also, its higher band gap means higher temperature operation (up to about 200°C) without swamping of extrinsic carriers (although not necessarily without reliability worries based on contact failures, etc.).

3) *Radiation hardness.* Not relevant until a CM gets installed in a satellite.

4) *Semi-insulating substrate.* Semi-insulating (SI) gallium arsenide is readily available and has lower parasitic capacitance for devices built on it. SI-silicon is not easily available, but two technologies are appearing that give the same result: Silicon on sapphire and silicon with buried oxide, collectively known as silicon on insulator (SOI). SOI process maturity has been estimated at 6 on a scale of 1 to 10 [10]. See more SOI discussion below.

5) *Light emission.* Having a direct bandgap, gallium arsenide is capable of emitting light and is a frequent choice for laser diodes. (silicon effectively cannot emit light.) This points the way toward optoelectronic devices, fiber optic interconnects, etc. Futuristic.

6) *Greater reliability*. Since there is no gate oxide, there is no possibility of punchthrough or breakdown (although the bulk material can experience breakdown at high voltages). Not sensitive to ionic contamination (unlike CMOS which is degraded by sodium in parts per billion quantities).

7) Simpler processing than straight CMOS

Not-so-frequently mentioned disadvantages

1) Higher raw material cost and offshore material supply

2) Smaller wafers and greater fragility and thus more expensive processing. Gallium arsenide wafers are 3" to 4" in diameter; silicon are 5", 6", or 8".

3) Certain chemical instabilities such as lack of a stable oxide

Comparison of advantages and disadvantages

Often, the litany of advantages seems to be rather frantic, as optimistic claims are made again and again. The simple-minded summary is that silicon is as close to an ideal material for ICs as is imaginable. Gallium arsenide is a better choice when the desire for better speed relative to power overrides considerations of cost, controllability, and risk.

Physical Properties at room temperature

	Si	GaAs	SiO_2
band gap	1.12 eV indirect	1.42 eV direct	-
dielectric constant	11.9	12.85	3.85
melting temperature	1415 °C	dissociates 600°C	1600°C
mobility - electrons (1)	1500 $cm^2\ V^{-1}sec^{-1}$	8500	-
- electrons (2)	700 $cm^2\ V^{-1}sec^{-1}$	4300	-
- holes (1)	475	450	-
saturated velocity	8×10^6 cm sec^{-1}	20×10^6	-
thermal conductivity	1.5 W cm^{-1} $°C^{-1}$	0.48	0.01
specific heat	0.7 J gm^{-1} $°C^{-1}$	0.35	
thermal diffusivity	0.92 cm^2 sec^{-1}	0.25	
thermal expansion	2.6×10^{-6} $°C^{-1}$	6.9×10^{-6}	10^{-6}
breakdown field	3×10^5 V cm^{-1}	4×10^5	$3 \times 10^6 - 10^7$

(1) Undoped semiconductor (2) Semiconductor doped to 10^{17} cm^{-3}

Materials supply

Four inch (100 mm) gallium arsenide wafers are the largest available; some foundries still use 3" (including Vitesse). (For comparison, silicon wafers are available up to 8", and our current vendors are using 5" and 6" wafers.) Gallium arsenide wafer sizes, constrained by the difficulties in keeping electrically active defect levels low, will probably increase slowly in the absence of demand for ICs in high volume, although Vitesse does plan to move to 6" gallium arsenide at an indefinite time. Wafer sizes are important because processing costs are usually more dependent on the number of wafers than on the total area.

Gallium arsenide crystal growth pullers are more expensive to build and operate because they usually run at several atmospheres of As overpressure. Wafers come mostly from Japan (more so than do silicon wafers).

Risks

Apart from the particular risks of a new technology (e.g. gallium arsenide materials supplies are uncertain, SOI has not yet been produced in large volumes, or whatever), the general risk is great. Reason: silicon processing is fully mature and there are multiple sources even for specific subsets of the business. I believe that it is a technology that will never "go bad" on us. Regardless of delays, yield problems, reliability problems, unresponsive suppliers, and everything else we experience with LSI Logic, Cypress Semiconductor, or whomever, we know that the technology is reasonably predictable and can somehow ultimately be made to work. This applies to our current and future (five to ten year) needs.

I don't feel that I can currently say the same about the alternative technologies I am discussing: although promising, they just don't have the history or the weight of an enormous industry behind them.

Processing

Silicon CMOS is a fully mature, high volume process. The base wafer is semiconducting, and is patterned by electron beam direct write or UV (ultraviolet) exposure in a stepper. Processing includes from 15 to 19 steps such as oxidation, plasma CVD (chemical vapor deposition), metal deposition by sputtering or evaporation, poly-silicon deposition, etching, ion implantation, and diffusion. The native oxide is adherent, durable, resistant to dielectric breakdown, and controllable.

Gallium arsenide processing equipment is usually converted from silicon use. Controllable diffusion and oxidation are impossible because of the reactive, binary (two elements, gallium and arsenic) nature of gallium arsenide; etching is used much less than in silicon for similar reasons. Gallium arsenide is susceptible to attack by almost all etchants and so must be processed with liftoff technology, which doesn't require the use of wet chemical etchants. Doping is accomplished by ion implantation. The process for creating the active channel gate may be self-aligned or manually aligned; the former is preferred to avoid the need for an extra, yield-reducing alignment step.

Gallium arsenide is fragile, which leads to lower yield because wafers break, and it cannot take temperatures much above 250 or 300 °C without losing arsenic to the atmosphere, so many silicon processing steps, such as high-temperature oxidation, diffusion, and deposition, cannot be used. The thermodynamics of a binary semiconductor adds one variable to metallurgy problems. Metallized contacts are gold-based, compatible with gold wire bonding (or aluminum-based, like CMOS, for Vitesse). The entire fabrication process is less complicated than ECL or CMOS, so 21 day turnaround is possible. Vitesse is the self-proclaimed pioneer in replicating the integration and processing success of silicon MOS technology in gallium arsenide.

Silicon ECL (emitter-coupled logic)

ECL technology is broadly familiar to most people as a fast and power hungry logic family chosen when the need for speed overrides the problem of cooling. Most of the advantages

of ECL have been reduced by gallium arsenide as a competitor; the biggest one that remains is the greater history and experience with silicon ECL.

ECL does not use TTL voltage levels and the power required is inherently so high that progress towards integration levels above 100,000 gates will be slow. Purely because of power considerations, ECL is not suited to large-scale, densely-packed use such as processing nodes in a Connection Machine. The use of ECL cannot be excluded from consideration in any parts of our system that are not highly dense or massively replicated, e.g. front ends.

Silicon-on-insulator (SOI): the basics

SOI attempts to maintain most of the good qualities of bulk silicon technologies while improving one of the drawbacks, a (semi-)conductive substrate that leads to parasitic capacitances and thus lower speeds. Essentially, one buys an expensive, complicated base wafer for the privilege of making much simpler device structures on it.

There are three major approaches for producing a layer of device-quality silicon on an insulating substrate: 1) use MBE (molecular beam epitaxy) or MOCVD (metallo-organic chemical vapor deposition) to grow silicon on a quartz or sapphire substrate, 2) use heat to bond a thin silicon sheet to a silicon wafer through fusion of an insulating oxide film in between, or 3) use ion implantation to create an embedded insulating layer, then recrystallize silicon on top of it (SIMOX technology). Most companies developing the successful SOI wafers are Japanese. (A Boston area company, Ibis, is using approach 3 with some success.)

Approaches 1 and 3 are capital-intensive. Each ion implanter can cost upwards of $1M and takes many hours to implant a few dozen wafers. MBE is almost as expensive; MOCVD can be somewhat faster and accommodate larger batches, so its capital cost per wafer is lower, but still large compared to the cost of ordinary silicon wafers. Approach 3 has been the most successful to date.

SOI maturity level: 3.

Frequently-mentioned silicon-on-insulator advantages

1) *Simpler processing* (no need for isolation from the substrate). True if you ignore the processing necessary to make the SOI wafer in the first place

2) *Higher density theoretically possible*. This is because you don't need as many space-consuming isolation trenches

3) *Faster devices because of much smaller parasitic capacitance*

4) *Lower incidence of α-particle induced soft errors* (in memories) due to lower capture cross section, since only α-particles that are absorbed in the surface layer of silicon can cause error in memories

5) *Decreased susceptibility to latchup*

Designing with SOI devices

Devices have not been made on SOI wafers in commercial quantities. There are many reports of 64 K SRAMs made on SIMOX wafers with good yield and performance. Among others, Honeywell is developing a 256K SOI SRAM [27].

Device technology for digital ICs

Silicon MOS technology families (nMOS and CMOS) are based on the MOSFET (MOS field effect transistor), using n-channel or both n- and p-channel.

The MOSFET structure is impractical in gallium arsenide because of lack of a good native oxide, plus Schottky barrier considerations. Gallium arsenide complementary devices are also not practical because of much lower hole mobility. Therefore, gallium arsenide digital logic families are based on the MESFET (metal-semiconductor field effect transistor) and Schottky diodes and, to a lesser extent, the JFET (junction FET). Typical power consumption can be 0.25 to 2.0 mW / logical gate (GigaBit Logic) or 0.2 mW / gate (Vitesse). For isolation of gate from substrate, MESFETs rely on Schottky barriers, i.e.potential barriers of various energy heights that form on gallium arsenide when certain metals are deposited under favorable conditions.

Both enhancement-mode (normally non-conducting or "off") and depletion-mode devices are now available in gallium arsenide.

Designing gallium arsenide ICs

MESFETs are integrated based on gallium arsenide logic designs, such as: SBFL (super buffered FET logic), DCFL (direct-coupled FET logic, analogous to nMOS), SCFL (source-coupled FET logic), or SDFL (Schottky-diode FET logic), none of which closely resembles CMOS (see reference 6 for schematics of these logic families). Logic families have not yet settled down as well as they have in the silicon world and they probably won't until "the market" has a clear idea of what it wants to do with gallium arsenide. An effect of this is that voltage levels have not been determined or are adjustable in many cases. (Convex's C3800 series of all-gallium arsenide computers uses a -2 V supply.)

All foundries offer design assistance in the form of seminars, handbooks, application engineers, cell libraries, simulators, prototype tools, etc., and most claim to offer excellent turnarounds on prototypes and final designs. Relatively low-cost prototyping is also available from TriQuint through many users sharing a wafer. One change that TMC would need to make is to think of things below the logical gate level. We definitely would need designers experienced in gallium arsenide; several universities now offer courses in gallium arsenide digital logic design. Vitesse is working on creating and standardizing a library to make gallium arsenide logic design more closely resemble design in silicon.

A few extra layout constraints will keep densities slightly lower in gallium arsenide than in CMOS in the long run. One is that gates must all be parallel across the chip to avoid differential threshold voltage shifts caused by strain-induced piezoelectric charges in this non-centrosymmetric material. Another is the need for numerous ohmic contact areas.

According to Long and Butner [6], "virtually all of the problems...associated with using gallium arsenide in a system stem from the extremely fast signals that propagate between chips". Transmission line theory must be used and off-chip interconnects must be

impedance-controlled, using elements that are usually lower-density (i.e. coaxial cables and connectors are fatter than ribbons and PCB connectors). Crosstalk, signal attenuation, and propagation delay increase with frequency. I don't believe Long and Butner's statement...I would mention the following things as potential problems: a) interfacing from silicon to gallium arsenide, b) less reliability history, etc., c) more difficulty in cooling due to lower thermal conductivity, d) lack of suppliers, and e) shifting and evolving technology.

"RISC-type designs are so attractive for implementation in gallium arsenide that we are likely to witness such a component in the next few years." [1] Published in 1985, this prediction has not been fulfilled six years later, although a Texas company, Systems & Process Engineering Corporation, is working on designing a three-chip gallium arsenide SPARC set, one of which may go to foundry (at Vitesse) in 1991. They have simulated the gallium arsenide-based ALU (only) at a stand-alone speed of 300 MHz.

Gallium arsenide logic devices in production and used in computers

Analogues to the ASICs that TMC uses are available in gallium arsenide. Vitesse announced a 30 Kgate logic chip in 1990 and designs for 102 and 195 Kgates (raw) are being accepted as of early 1991, with about 50 to 70% of gates available. Designs for 323,000 raw gates will be accepted in mid- to late 1991.

The Cray-3 uses 500-gate equivalent gallium arsenide CPU and control logic chips from GigaBit Logic, assembled in 16-chip printed circuit boards (PCBs) using chip-on-board technology with wirebonding. The boards, each one inch square, are stacked into modules of sixty-four PCBs, and then immersed in Fluorinert, an inert fluid, for cooling [9]. The first exemplar of the Cray-3, made up of 208 of these modules (and thus 208 x 16 x 64 = 212992 gallium arsenide chips total), is planned to finally ship by the end of 1991, close to four years behind the original schedule. Some analysts expect the shipment to be delayed until 1992 [23]. The memory chips are of silicon. The CRAY-3 is projected to cost $29M, which implies a per chip cost of $136 — IF gallium arsenide chips were the only cost (clearly they are far from being that). Probably each gallium arsenide chip costs about $20 or so.

Convex Computer Corporation's C3800 series supercomputer, announced in May 1991 and to be available in the fourth quarter, is the "first supercomputer to be fully implemented with gallium arsenide circuitry" and is air-cooled. It has eight processing nodes, uses a -2 V supply, clocks at 16.6 nsec (60 MHz), could reach 2 GFlops peak, and costs $8M. The Vitesse gallium arsenide ASICs have 45,000 gate equivalents and the whole system costs from $2M to $8M. Convex also has several lower-end models which use mixed silicon and gallium arsenide technology. [19, 23, 24] Various companies will be announcing workstations and personal computers using part gallium arsenide logic in 1992 [27].

Gallium arsenide memories available

Vitesse is the primary U.S. supplier of gallium arsenide-based SRAM. Their largest memory is 4K x 4 which will be available in October 1991, with ECL-compatible outputs. They choose to design new memories only according to specific customer requests because of their company's small size and the uncertainty of demand. Solbourne's Series 5 workstations use a 1 Kbit SRAM from Vitesse.

It is interesting to note that as of July 1991, Vitesse was continuing to operate on a largely custom-design and custom-build basis, for both logic and memory chips. Their technical ability to manufacture memories and logic devices exceeds the current demand, which raises the question of how progress can be sustained if the marketplace fails to embrace

their products. Vitesse is a long way from supplying gallium arsenide memory as a commodity.

Programmable devices

Programmable logic devices (PLD) that can be fully or partially programmed by the user after manufacture have become popular and are used in TMC designs. The base technology is usually CMOS or BiCMOS and programmability is achieved either reversibly through an EPROM (charge storage) mechanism or irreversibly through a fuse or antifuse. Vendors include Altera, Xilinx, and Actel. Widespread use of PLDs can ease the chip designer's task and reduce time to market, but there are also associated problems. The use of PLDs reduces the discipline needed to get a board designed and functioning with "mask-made" chips ("please don't throw things at me" — the author). Also, just about every vendor has a different user interface for programming, which wastes users' time. Still, EIA's industry-wide standard library scheme, EDIF/LPM or Electronic Design Interface Format / Library of Parametrized Macros, may provide the way to a single interface in the future [26]).

Only one gallium arsenide vendor, Gazelle, offers a PLD, and it is TTL-compatible. Programming is done irreversibly by fusing links according to a "standard JEDEC" file which may be created to the specifications of any silicon PLD vendor.

PLDs also exist in ECL technology.

Packaging, manufacturing, and assembly

The coefficient of thermal expansion about 3 times greater for gallium arsenide than for silicon, and thermal conductivity is about 3 times lower. Gallium arsenide is brittle, so chip pick and handling are tougher. GigaBit will deliver wafers or bare dice; Vitesse and TriQuint will deliver wafers, bare dice, or packaged units.

In general, packages developed for silicon chips can be used for gallium arsenide, as long as the lower thermal conductivity and greater thermal expansion and brittleness are taken into consideration. Vitesse offers packages that we are familiar with, such as PLCCs, and also a variety of TAB and exotic options.

Electrostatic discharge sensitivity

There is no particular concern about gallium arsenide in this respect. The ESD sensitivity of a finished and packaged chip is a function of the protective circuitry that surrounds it. Here are some examples:

Vendor	Maximum safe voltage
LSI	2000 V
Weitek	2000 V
GigaBit	100 to >1000 V
Vitesse	2000 to 3000 V

Reliability

To the extent that FIT rates can be compared, the projected lifetime of gallium arsenide devices is similar to that of devices from our better silicon vendors.

Vendor	Part type	Est. FIT rate	Assumed E_a
LSI Logic	LCA100K	405	0.7 eV
Weitek	3164 et al.	120	0.7
Vitesse	All	66	0.7
Vitesse	All	20	1.2
VLSI Technology	1.0 µm	49	0.7
GigaBit Logic	Many	42	1.4
GigaBit Logic	SC10000	25	1.4

Projected FIT rates depend strongly on several variables and assumptions, such as junction temperature, activation energy, and complexity of the test vehicle. The numbers above don't detail all relevant factors, but I've made sure that the numbers are comparable in a broad sense.

Integration levels

Gallium arsenide integration levels have been lower than those for silicon technologies, because of lower investment of time and money in process development. SOI integration levels are also lower than silicon technology integration levels. During late 1990, gallium arsenide DCFL and silicon CMOS technologies achieved rough parity in maximum feasible integration level for the first time, closing an enormous gap. See spreadsheet for more details.

Technology	Integration level (early 1991)
GaAs DCFL (Vitesse 1 M)	161 K (useable) gate array
CMOS (LSI LCA200K)	200 K (useable) gate array
MOS memory	4 M DRAM
SIMOX	64 K SRAM
GaAs	16 K SRAM

Integration level in any technology is constrained by:

1) *Printable linewidths*. Controlled by the achievements of photolithography and by the surface quality of the substrate; equivalent for silicon and gallium arsenide.

2) *Devices (transistors, passives, etc.) needed per logical gate.* Controlled by logic family.

Family	Gate	Devices needed
DCFL	3-input NOR	4 transistors
SBFL	3-input NOR	8 transistors
CMOS	3-input NOR	6 transistors
BiCMOS	3-input NOR	3 transistors, 1 resistor
ECL	2-input OR/NOR	6 transistors, 3 resistors

3) *Average size of devices and thus of logical gates.* Controlled by inherent material properties, need for uniformity in device properties (such as

threshold voltage), printable linewidths, technology advances, vendors' choices of power / density / speed tradeoffs, and scaling considerations. Advantage to silicon-based families and especially to SOI.

4) *Power dissipation per gate.* Gallium arsenide generally dissipates less power than silicon CMOS; both run cooler than ECL.

5) *Maximum permissible size of chip.* Controlled by mechanical considerations, power dissipation, typical defect density, etc. Advantage to silicon CMOS because of stronger material and lower defect densities than gallium arsenide; again, ECL has to stay smaller than either of the former because of higher power dissipation (how do you package a 50 W chip?). SOI may be at a a disadvantage.

6) *Need for interconnections* and the area they occupy. Equivalent for silicon and gallium arsenide.

Overall, with a best guess at effect of the above factors, silicon CMOS and gallium arsenide are roughly equivalent with respect to maximum integration level achievable, possibly with a slight advantage to gallium arsenide DCFL because of its lower power consumption. ECL will continue to be limited in achievable integration levels, and SOI has integration capability similar to silicon CMOS.

Outlook for improvement in gallium arsenide maturity, volume, scale

Predictions of the size of the digital gallium arsenide market have been grossly optimistic in the past. Silicon has a tremendous background of practical experience and highly detailed research, and all silicon processing steps have been studied to death. There is some room for optimism in that Vitesse, in particular, seems to have broken out of the "handmade lab curiosity" mode into the realism of making gallium arsenide chips easier to use.

Figure 1 shows historical and projected integration levels for several technologies. Vitesse's achievement of scaling up rapidly in integration levels is impressive and puts their technology on a par with CMOS (see graph). Gallium arsenide costs are higher and so are speeds. See attached spreadsheet for details. In general, I will protect myself by being "guardedly optimistic" about Vitesse's future.

Three dimensional integrated circuits

Three dimensional integrated circuits are fabricated by stacking successive layers of SOI devices and interconnects. After the first layer is processed, an insulating layer is deposited on top of the devices and silicon is grown on top of that. More devices are then fabricated. Maturity level: 2.

The advantages are potentially much greater packing density and higher speed communications. The disadvantages, besides the overriding feasibility concerns, are that power still needs to be removed and crosstalk becomes worrisome. In fact, crosstalk is expected to be the factor that limits the density. The major players are Japanese (Mitsubishi and others).

Business

Digital gallium arsenide merchant sales, worldwide, in 1988 were $50M and are projected to be somewhere between $125M and $145M in 1990. There is also a large military market for analogue gallium arsenide circuits, mostly microwave. Vitesse is purely digital, GigaBit pushes digital logic and memories, TriQuint makes digital, analogue, and mixed technology on a single chip, and Gazelle concentrates specifically on programmable logic devices.

Company	1990 Revenues	1991 Revenues (proj.)
Vitesse	$17M	$28M
GigaBit Logic	$12M	na
TriQuint	$21M	na
TriQuint / GigaBit / Gazelle	na	$40M

Of the above, Vitesse conforms the best to TMC philosophies and business practice. They have reached a second source product license and royalty agreement with Fujitsu, a major player in the gallium arsenide microwave field, and Thomson-CSF in France is a second fab. Within the past three years, Ford Microelectronics, Harris Semiconductor, and several other companies have left the digital gallium arsenide business because of the small market.

Both GigaBit and Gazelle have agreed to merge with TriQuint, the former in March 1991, the latter in May 1991 (for now, all product lines will continue as before). These mergers, when made final, will make TriQuint the largest commercial U.S. supplier of gallium arsenide ICs, while Vitesse remains the largest supplier of digital gallium arsenide circuits.

The gallium arsenide industry has two major segments: digital / commercial on one hand, as represented by Vitesse, and microwave / military on the other, represented by TriQuint et al. and several smaller captive manufacturers. In addition to the similarities between them, such it is important to realize the vast differences

Feature	Digital	Microwave
Metallization	Aluminum	Gold
Integration level	100 Kgates	100 transistors and passives
Customers	Mostly commercial	Military
Wafer size	3" or 4"	2" or 3"
Building blocks	DCFL gate	FETs, HEMTs, passives
Challenge	Defects, gate density	Noise, speed
Technology	Planar	Recessed gate or planar

The point is that the military suppliers would have a long way to go before they could be considered competition or a second source for Vitesse. Most or all of them would be unable to compete.

Costs

As an example of the cost of finished circuits, Vitesse gallium arsenide gate arrays cost $0.03 to $0.05 / gate in 1990, e.g. up to $700 for a 14,000 gate array. The price for CMOS gate arrays was about one-tenth the cost of gallium arsenide in 3Q90, or $0.003 to 0.005 / gate. The attached spreadsheet has more details and includes generic figures of merit based on minimization of power usage, cost, area, and gate delay.

Future costs for gallium arsenide ICs will be determined by how fast (or if) the market develops. Over the next ten years or so, take the per-gate cost for CMOS as a floor for per-gate gallium arsenide cost; the actual cost of gallium arsenide gate arrays may be about two to three times that of CMOS. The move from 3" to 6" wafers at Vitesse should mean a steady-state gallium arsenide IC cost reduction of about a factor of two (not four), and miscellaneous improvements will probably account for the rest of the expected reduction.

Predicting the future

Past predictions of yield, integration level, cost, die size, wafer size, etc. have usually erred on the side of conservativism, since people have assumed that the problems they saw blocking them could only be solved by being rammed head-on. Since the head-on approach was clearly doomed, they assumed that progress would stop. As an example, consider optical wavelength lithography for chip manufacture. It was predicted to lose its usefulness below 1 micron; linewidths of 0.6 micron are routinely drawn by optical lithography, thanks to creative new techniques. Farther back in time, consider the VLSI studies in the early 1960s, where it was "proved" that random defects would limit the achievable size of VLSI chips. Experience showed that defects were not randomly distributed and that their density could be drastically decreased. The result is the integrated circuit industry.

The logarithmic extrapolations for density, gate count, feature size, etc. will continue to hold true, with some qualifications. Over the past two years or so, Vitesse in particular has changed the slope of its developmental curve steeply upward, at a rate of increase in integration that cannot be maintained. They are playing a skillful game of catch-up and will slow when they are at as high a level of integration as the silicon folk. I am afraid that they may already have outrun their market, cutting them off from the things that drive a technology: market share and money.

Recommendations

Thinking Machines should begin actively exploring digital gallium arsenide vendors, particularly Vitesse, and should look at its own architecture and chips to see if and where digital gallium arsenide would make sense for use several years out. We should continue to concentrate current design efforts on established CMOS technology, while staying informed on the development of alternative technologies, particularly with respect to volume, price, and marketplace acceptability. If further investigation of Vitesse in particular suggests that we could work together, we might want to use our influence to e.g. encourage DARPA funding for them, preparatory to the time when we will be able to consider designing a system that incorporates Vitesse's technology.

Vendors

Gazelle Microcircuits
2300 Owen Street
Santa Clara, CA 95954
(408) 982-0900

GigaBit Logic
1908 Oak Terrace Lane
P. O. Box 2518
Newbury Park, CA 91320
(805) 499-0610

Systems & Process Engineering Corporation (SPEC)
Austin, TX
(512) 385-0318

TriQuint Semiconductor, Inc.
P.O. Box 4935
Beaverton, OR 97076
(503) 644-3535

Vitesse
741 Calle Plano
Camarillo, CA 93010
(805) 388-3700
(617) 239-8075

References and bibliography

The company affiliation of the author is noted in parentheses after the citation when it is known.

1. Gallium Arsenide Technology, ed. David K. Ferry. Howard W. Sams & Co., Inc.: Indianapolis, 1985. Chapter 1 is good for an overview of gallium arsenide digital ICs.
2. VLSI Fabrication Principles, Sorab K. Ghandhi. John Wiley & Sons: New York, 1983. Recipe approach (without emphasis on basic understanding) to various steps in manufacture of ICs. Occasionally misleading.
3. VLSI Technology, ed. S. M. Sze. McGraw-Hill: New York, 1983. Excellent explanation of IC manufacture steps, plus discussions of yield, reliability, packaging, process simulation and diagnostics, and device integration. This book has dated much more slowly than most because of Sze's understanding and foresight.
4. J. S. Blakemore, "Semiconducting and other major properties of gallium arsenide", J. Appl. Phys. **53** (10), October 1982, p. R123. Detailed survey of literature on gallium arsenide as a raw material.
5. Gallium Arsenide Computer Design, ed.V. M. Milutinovic and David A. Fura. Computer Society of the IEEE, Computer Society Press, 1988. Reprints of articles from several IEEE publications, dating from 1981 through 1987.
6. Gallium Arsenide Digital Integrated Circuit Design, Stephen I. Long and Steven E. Butner. McGraw-Hill: New York, 1990. [Barker TK7874.L66 1990] What the name implies: thorough explanation of gallium arsenide logic design, plus materials and processing issues.
7. Gallium Arsenide Technology and its Impact on Circuits and Systems, IEEE Circuits and Systems Series. ed. D. Haigh and J. Everard. Peter Peregrinus Ltd.: London, 1989. [Precatalogued E15281 Barker] This book, though useful, seems to be strangely out of date on many points. Its reference in Chapter 6 for "current" gallium arsenide integration levels is from 1982. It refers to obsolete mesa isolation as a standard technique, although it has been supplanted totally by implantation isolation for ICs.
8. Vitesse Semiconductor Corporation, Vitesse 1991 Product Data Book. Extensive info on Vitesse products, plus application notes on packaging, DCFL ASIC design, etc.
9. D. Kiefer and J. Heightley, "CRAY-3: A GaAs Implemented Supercomputer System", IEEE GaAs Symposium Technical Digest, 1987, p.3 (Cray)
10. Ravi Shankar and Eduardo B. Fernandez, VLSI Electronics Microstructure Science, Vol. 20: VLSI and Computer Architecture. Academic Press: Boston, 1989.

Interesting, but annoyingly dated; the most recent references are from 1987 and graphs and compilations from 1979 (!) are used to illustrate trends over the period 1970 to 2000. [Barker TK7874.V56 v.20]

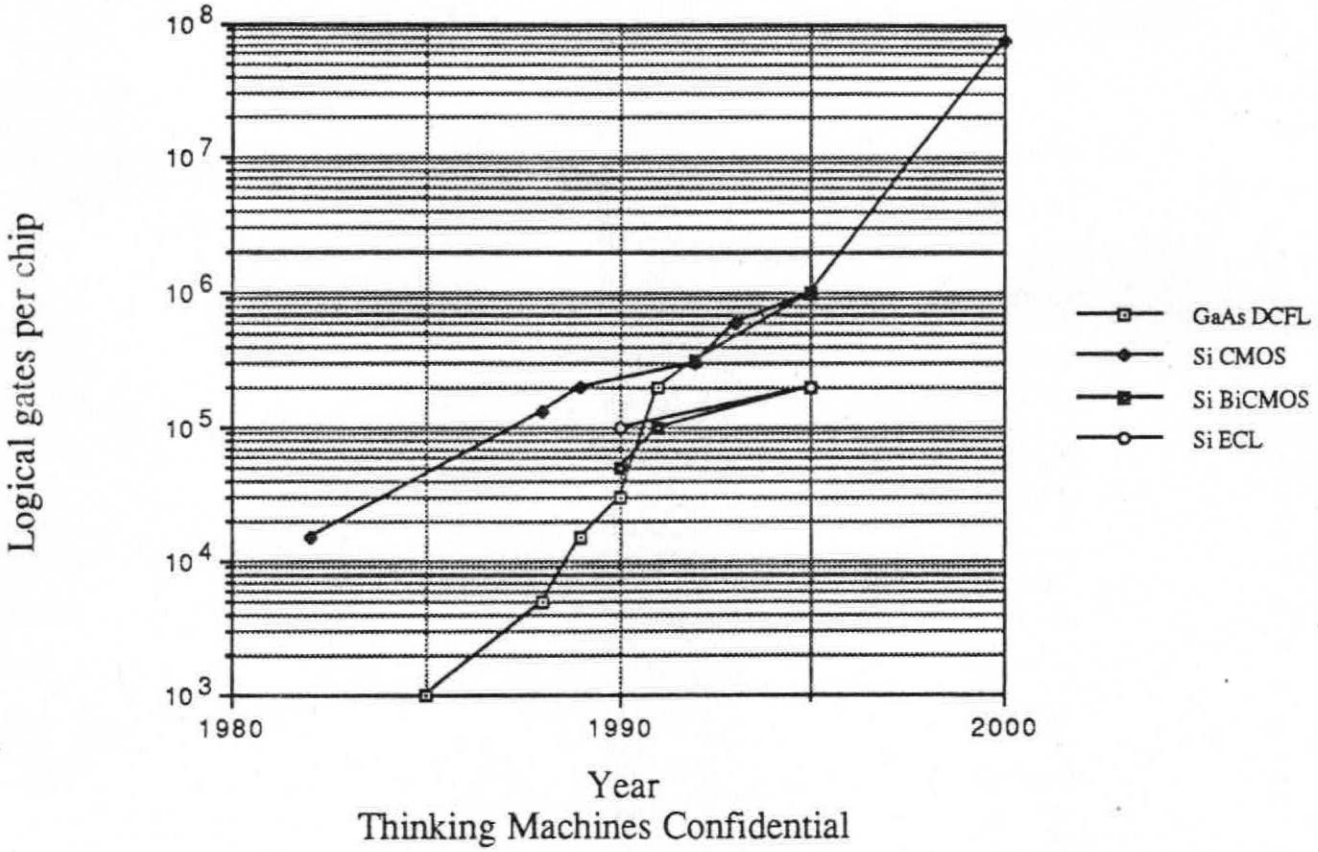
11. Takao Abe & John H. Matlock, "Wafer Bonding Technique for Silicon-on-Insulator Technology", Solid State Technology 33 (11), November 1990, p. 39 (Shinetsu Handotai)
12. P. K. Vasudev, "Ultrathin Silicon-on-Insulator for High Speed Submicrometer CMOS Technology", Solid State Technology 33 (11), November 1990, p. 61 (SEMATECH)
13. Harold H. Hosack et al., "SIMOX Silicon-on-Insulator: Materials and Devices", Solid State Technology 33 (11), November 1990, p. 61 (Texas Instruments)
14. Bill Groves, "1991: What will it mean for ASICs...?", ASIC Technology & News, 2 (9), January 1991, p. 1
15. Larry Walker, "Is GaAs ready for the mass market?", ASIC Technology & News, 2 (11), March 1991, p. 1. This interesting article is biased, because it uses, as authoritative, graphs on gallium arsenide vs. ECL and CMOS that were prepared by Vitesse.
16. GigaBit Logic, GigaBit Logic 1991 GaAs IC Data Book & Designer's Guide, August 1990.
17. Masaharu Watanabe, "Technical Trends in Large Diameter Silicon Wafers: Part I", Solid State Technology , 34 (3), March 1991, p. 69 (Toshiba / Japan Electronic Industry Development Association)
18. Robert Lipp, "New capabilities drive growth of CMOS ASICs", ASIC Technology & News, 2 (7), November 1990, p. 24
19. Darrell Dunn, "Convex's Supercomputer Push", Electronic Buyers' News, May 13, 1991, p. 12
20. Omar Wing, Gallium Arsenide Digital Circuits, Kluwer Academic Publishers: Norwell, Massachusetts, 1990. Simple-minded in the extreme and containing little of value.
21. Cray Computer Corporation report to stockholders, November 1989.
22. Shirou Horiuchi, "VLSI Technology Advances as Demand Increases: Parts I, II, and III", J. of Electronic Engineering, May, June, and July 1990.
23. Loring Wirbel, "Cray Computer details GaAs-based Cray-3". Electronic Engineering Times, May 20, 1991. p. 16
24. Terry Costlow, "Convex takes 'passive' action in C3800". Electronic Engineering Times, May 20, 1991. Impedance problems, board layup, connectors, and passive components.
25. Wesley R. Iversen, "GaAs up for another turn at the plate". Supercomputing Review, March 1990. p. 38
26. Larry Waller, "Will opposing FPGA standard camps create a stalemate?". ASIC Technology & News, 3 (4), August 1991, p. 1
27. "Honeywell Demos Faster High-Performance SRAMs". Microelectronics Manufacturing Technology, August 1991, p. 12
28. Larry Waller, "New era of high performance". ASIC Technology & News, 3 (5), September 1991, p. 22
29. "Circuits provide a uniform method for comparing high speed ASIC vendors". ASIC Technology & News, 3 (5), September 1991, p. 26
30. Justin Rattner, "On the horizon: massive parallelism", Electronic Design, August 23, 1990, p. 58 (Intel)
31. A. K. Goel & A. Kalia, "Comparison of nMOS, CMOS, and GaAs Technologies", IEEE Transactions 89CH2785-4/90/0000-1238$01.00 Disappointing.
32. Ron Cates, "High-Speed Gate Arrays: Considering the Options", VLSI Systems Design, Semicustom Design Guide, 1988. p. 24 (Vitesse)

33. Ron Cates, "Gallium arsenide finds a new niche", IEEE Spectrum, April 1990. p. 25 (Vitesse) Excellent overview.

A note about the attached spreadsheet

Boiling a technology down to numbers and squeezing them into a grid tends to obscure subtleties. I hope that this hasn't happened to the point of distortion, but comments and criticism are always welcome. Some of the numbers are moving targets, particularly price.

Figure 1. Gate array integration levels:
current and projected



Thinking Machines Confidential

Comparison of ASIC logic technologies (gate array, standard cell, hybrid, and full custom)

Technology	Si CMOS	Si CMOS	Si CMOS	Si CMOS	Si CMOS	GaAs DCFL	GaAs DCFL	GaAs DCFL	GaAs DCFL
Vendor	LSI Logic	LSI Logic	VLSI Tech.	VLSI Tech.	VLSI Tech.	Vitesse	Vitesse	Vitesse	Vitesse
Process / product	LCA100K	LCA200K	VGT300	VGT350	VGT353	VSC30K	VGFX200K	VGFX350K	1M
Type	gate array	gate array	gate array	gate array	gate array	gate array	gate array	gate array	gate array
Data sheet in hand		Apr-91	Jun-89	May-91	May-91				
TMC example	Miata	-	?	?	?	-	-	-	-
Gate length (µm)	1.0	0.7	1.0	1.0	1.0	0.8	0.6	0.6	0.6
Cost / gate	$0.003	$0.005	$0.005	$0.005	$0.005	$0.050	$0.010	$0.010	$0.013
Cost / max. no gates ($)	300	1000	370	342	342	750	1000	1750	2019
Power (µW / gate•MHz)		4.3	6.5	6.5	6.5			0.0	
Power / gate (mW) @ 25 MHz	0.48	0.11	0.16	0.16	0.16	0.40	0.14	0.25	0.20
Gate delay (psec)	460	400	630	405	350	90	130	130	100
Cost * delay ($ psec)	1.38	2.00	3.15	1.98	1.71	4.50	1.33	1.30	1.25
Delay * power (fJ)	221	43	102	66	57	36	18	33	20
Area / gate (sq µm, raw)	773		1030	1030	1030		545	626	
Cost*delay*power ($ fJ)	0.66	0.22	0.51	0.32	0.28	1.80	0.19	0.33	0.25
Cost*delay*pwr*area ($ fJ sq µ	512	0	527	331	286		102	203	
Maximum raw gates	238095	307500	246500	162791	162791	30000	195000	350000	323000
Usage factor	0.42	0.65	0.30	0.43	0.43	0.50	0.50	0.50	0.50
Maximum useable gates	100000	200000	73950	70000	70000	15000	97500	175000	161500
Voltage level (V)	5	3.3, 5		5	5	-1.1	-1.1	-1.1	-1.1
Interfaces with...	TTL, CMOS	TTL, CMOS				ECL, TTL	ECL, TTL	ECL, TTL	ECL, TTL
Process maturity (1-10)	9	9	9	9	9	4	4	4	4
Mask steps								13	
Logic complexity	4	4				3	3	3	3
Reliability (FIT)	405			26	26				
Begins shipping	current	Jun-91	current	current	current	current	Jun-91	Mar-92	?
Chip area (sq cm)									
Estimated chip cost									
Other features		36 KRAM 160 KROM 2 or 3 layer			3 layer metal				

Table 1 Thinking Machines Confidential 9/9/91

Comparison of ASIC logic technologies (gate array, standard cell, hybrid, and full custom)

CONT.

Technology	GaAs	Si ECL	GaAs SCFL	GaAs	Si CMOS	Si CMOS	Si CMOS	Si CMOS	Si CMOS	Si CMOS
Vendor	GigaBit	National	TriQuint	as used by	LSI Logic	VLSI Tech.	VLSI Tech.	VLSI Tech.	Weitek	Texas Inst
Process / product	SC10000	ECL ASIC	QLSI / QED	Cray	LCB007	VSC100	VSC320	VSC370	CMOS 34	EPIC II
Type	std cell		std cell	Cray-3	std cell	std cell	std cell	std cell	full custom	hybrid
Data sheet in hand	Aug-90		Jun-90				Nov-89			Jul-91
TMC example	-	-	-	-	Phoenix	Boxer	?	FAX-LP		DASH
Gate length (µm)	0.8		1.0		1.0	1.5	1.0	1.0		0.8
Cost / gate	$0.025	$0.003	$0.030	$0.040	$0.004	$0.002		$0.005		$0.007
Cost / max. no gates ($)	0	0	300	0	398	30				1500
Power (µW / gate•MHz)						20	3.6			0.5
Power / gate (mW) @ 25 MHz	1.30	0.40	0.01						0.80	12.50
Gate delay (psec)	120	120	141			740	780			500
Cost * delay ($ psec)	3.00	0.36	4.23		0.00	1.59				3.41
Delay * power (fJ)	156	48	2		0	0				6250
Area / gate (sq µm, raw)	4940		9316	29421	4066	11290		622		233
Cost*delay*power ($ fJ)	3.90	0.14	0.06	0.00		0.00				42.61
Cost*delay*pwr*area ($ fJ sq µ	19266	0	563	0		0				9929
Maximum raw gates	15000	100000	15000							293333
Usage factor			0.67							0.75
Maximum useable gates	0	0	10000	500	96000	14000	180000			220000
Voltage level (V)	-5		text -5, +5	text -5, +5	5	5	5	5	-5	5
Interfaces with...	ECL, TTL		ECL/TTL/CMOS		TTL, CMOS					
Process maturity (1-10)	3	8	3 to 4		9					
Mask steps		20								11
Logic complexity	3	8	3		4					
Reliability (FIT)	42				405			26	139	
Begins shipping	current		current		current	current	current	current	current	
Chip area (sq cm)			0.93	0.15						2.25
Estimated chip cost				$20.000	$398.200	$30.000				
Other features										