Cahners
MICROPROCESSOR REPORT
www.MPRonline.com
THE INSIDER'S GUIDE TO MICROPROCESSOR HARDWARE

IBM Paving the Way to 0.10 Micron

First Copper, Then SOI, Now Low-k and E-Beams

By Keith Diefendorff {5/1/00-01}

As it has many times in the past, IBM is once again blazing the trail to next-generation IC processing way ahead of the rest of the semiconductor industry. Two years ago (see *MPR 9/14/98-msb*, "IBM Delivers on Copper Promise With 750-400"), IBM rocked the industry with its leap to copper interconnects—a feat most other vendors are still scrambling to match. A year later, IBM made another startling announcement: it would move its mainstream logic processes to silicon-on-insulator substrates (see *MPR 8/24/98-02*, "SOI to Rescue Moore's Law"). The company has now made good on that promise by shipping an SOI-based PowerPC processor, code-named IStar, to its AS/400 group. Then, just last month, on April 3, IBM announced yet another giant technological leap, this time to a low-*k* process ($k < 3.0$) using a spin-on polymer dielectric, called SiLK by developer Dow Chemical. Copper, SOI, and SiLK will be the baseline materials for IBM's 0.13-micron-generation CMOS-9S process, which will enter production next year.

As if copper, SOI, and low-*k* weren't sufficient to prove its prowess, on March 2 IBM announced a breakthrough in electron-projection lithography (EPL). This development, which dramatically boosts e-beam-stepper throughput, could potentially render unnecessary the enormously expensive extreme-ultraviolet (EUV) optical steppers that are currently the odds-on favorite for next-generation lithography (NGL). This IBM development could lead to a commercial EPL stepper from partner Nikon by early 2003, opening the door to billion-transistor chips.

While leadership in any one of these technologies would be impressive, IBM's command of all of them is almost unbelievable. Only Motorola, which until last year was a partner of IBM, has so far managed to get copper processors into mass production (see *MPR 11/16/98-04*, "G4 Is First PowerPC With AltiVec"), but even Motorola is still well behind IBM on copper manufacturing. Other companies have claimed use of "low-*k*" dielectrics, but these companies are mostly referring to fluorine-doped silicon-dioxide materials with dielectric constants only about 10% lower than conventional SiO_2. A few companies have also claimed to be working on SOI, but none that we know of (besides IBM) is yet to the stage of seriously considering it for volume mainstream manufacturing. And while a few companies are funding industry consortia research into next-generation lithography, most will simply wait until NGL tools become broadly available from traditional equipment suppliers.

In Conscious Pursuit of a Risky Strategy

IBM could just be blowing smoke, tooting its technology horn more loudly than other semiconductor vendors to gain the appearance of a technology leader. But history does not support this theory. Over the years, IBM has demonstrated a clear pattern: invest heavily in research and development on aggressive new technologies; announce them when they're ready; ram them into volume production; then disseminate the technology to the rest of the industry while moving on to new technologies before the crowd catches up.

Bijan Davari, IBM Fellow, vice president of IBM's Semiconductor Research & Development Center in East Fishkill (NY), and the mastermind of IBM's semiconductor R&D strategy, admits this strategy involves some risks. For

one thing, the development of advanced processes is extraordinarily expensive. For another, proprietary processes are not consistent with low-cost manufacturing. On the one hand, IBM would like to maximize the return on its investment by keeping its technology to itself to use as a competitive weapon. On the other hand, it realizes that it cannot afford to be out on a technology limb by itself. IBM needs other semiconductor manufacturers to adopt its technology so that the equipment industry will invest in developing the reliable low-cost, high-throughput tools that IBM needs for high-volume chip production.

Davari's plan to resolve this dilemma is twofold: stay ahead and partner with other companies. If IBM can stay ahead of the industry, he argues, it opens a window of time during which the company can exploit an advanced technology before others catch up. During this period, Davari says that IBM Microelectronics garners a significant amount of business building for its customers' parts that simply cannot be built by any other vendor.

If IBM stays far enough ahead, then even after this period of exclusivity, its intellectual property will still have enough residual value to be licensed to close partners and, eventually, to the rest of the industry. IBM then plows these licensing revenues back into process development to fund its efforts to stay ahead. Also, IBM allows selected partners, such as UMC and Infineon, to pitch in to help defray development costs in return for earlier access to some of IBM's advanced technologies (see *MPR 2/14/00-02*, "IBM, Infineon, UMC Gang Up On 0.13").

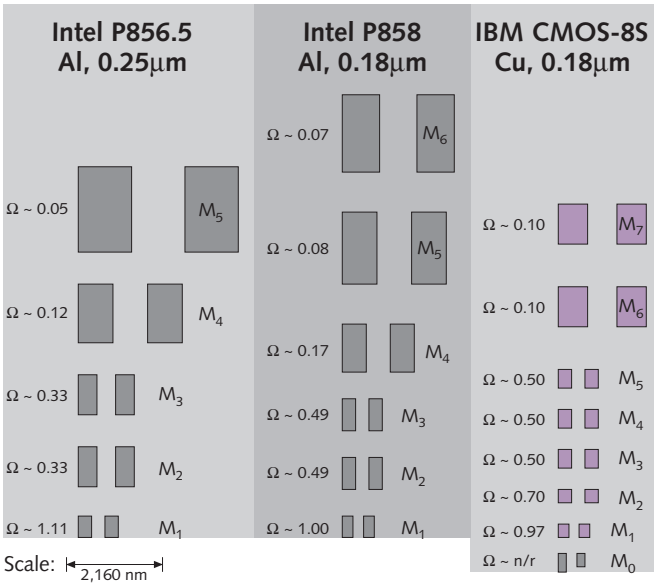

Figure 1. From this scale drawing of Intel's 0.25-micron P856.5 and 0.18-micron P858 it is clear that the wires get closer together, but the vertical dimensions do not shrink proportionally, which increases capacitance. In contrast, IBM's 0.18-micron CMOS-8S copper interconnect has much thinner layers—thus less capacitance—even though the wire resistance (including cladding) is similar to that of P858.

Capacitance, the Microprocessor's Worst Enemy

The transition time of a signal on a wire in an IC is proportional to the product of the wire's resistance (R) and its capacitance (C). Thus, lowering R and C reduces signal delay. Furthermore, the noise that a signal accumulates as it propagates through a wire is related to the degree of capacitive coupling to adjacent signals. Thus, reducing capacitance both reduces signal delay and improves signal integrity.

Unfortunately, capacitance does not scale with process shrinks. The capacitance a signal encounters is proportional to the area of adjacent parallel conductors and inversely proportional to the thickness of the insulator between them. As process dimensions shrink, wires get shorter, reducing C, but they also get closer together (which increases C) and narrower (which increases R). Thus, the net effect of process scaling is to leave the RC-delay component roughly the same, or to make it somewhat worse. So, as process dimensions shrink and transistors speed up, RC interconnect delay becomes an increasingly large component of overall circuit delay. Furthermore, capacitive coupling of noise among signal lines gets worse, because vertical wire thickness generally isn't reduced by the same scale factor as horizontal line widths and spaces (thickness is usually maintained to keep resistance to a minimum).

RC delay and noise coupling have not always been huge problems. In 0.25-micron and larger processes, transistors largely dominated circuit delays, and wires were far enough apart that only long parallel buses created serious noise problems. But at 0.18 micron, things change: interconnect delays and noise become more significant problems. And at 0.13 micron, unless something is done, these problems become serious obstacles to continued circuit-speed increases.

IBM made a step-function improvement in this situation for its 0.22-micron (CMOS-7S) and 0.18-micron (CMOS-8S) process generations when it introduced copper as the interconnect material. Copper has about 40% lower resistivity than the aluminum alloy used previously, a fact IBM exploits to build thinner interconnect layers—which have less capacitance—without increasing resistance. Figure 1, which compares IBM's 0.18-micron copper CMOS-8S interconnect system with Intel's 0.18-micron aluminum P858 system, clearly illustrates the advantage of copper in this respect.

Although this improvement is substantial, as dimensions shrink further, to 0.13 micron and beyond, and the wires get even closer, capacitance once again becomes a limiting factor. This time, however, no new conductor material will come to the rescue. Silver, the only material more conductive than copper (at normal temperatures), is only slightly more so (about 5%). Fortunately, manufacturers have one more handle on capacitance: the permittivity of the insulator, also called the dielectric constant.

Finding the Least-Worst Alternative

The interlayer dielectric (ILD) material used by most manufacturers today is silicon dioxide (SiO_2), which has many ideal physical properties for this purpose. As a glass, it is mechanically solid, allowing it to provide good support for the interconnect layers and to form a tight hermetic seal from the environment. Silicon dioxide is chemically inert and thermally stable, making it compatible with the silicon substrate, with all types of interconnect materials, and with high-temperature manufacturing steps. In addition, the material offers low leakage currents and high breakdown voltages. It also has excellent adhesion and is amenable to planarization using chemical-mechanical polishing (CMP).

Unfortunately, silicon dioxide doesn't have such ideal electrical properties. Pure SiO_2 has a dielectric constant (k) of about 4.0; including overcoats necessary in the manufacturing process, silicon-dioxide insulation typically delivers a k_{eff} in the range of 4.3–4.5. Some manufacturers, including IBM in CMOS-8S and Intel in P858 (see *MPR 1/25/99-06*, "Intel Raises the Ante With P858"), use a fluorine-doped silicon dioxide called FSG (fluorosilicate glass) or SiOF. FSG is attractive because it has manufacturing properties similar to pure SiO_2; unfortunately, it improves the k by only about 10%. The improvement in a copper-interconnect environment is even less (about 6%), because less fluorine must be used to remain compatible with copper.

While many materials have lower k than pure or fluorinated SiO_2, all other known insulating materials are inferior to SiO_2 with respect to their thermal, mechanical, or chemical properties, making them more difficult to use in manufacturing, or less desirable in the final product. It is an intrinsic property of low-k materials, for example, that they also have a low modulus—that is, they are soft. IBM spent several years identifying possible candidates, which are shown in Table 1, and deciding which had the fewest drawbacks—or at least had only problems IBM thought it could tackle.

Materials	k	Process
Silicon Dioxide	3.9–4.5	PECVD
Fluorosilicate Glass (FSG)	3.2–4.0	PECVD
Polyimides	3.1–3.4	Spin-on
HSSQ	2.9–3.2	Spin-on
Diamond-Like Carbon	2.7–3.4	PECVD
Carbon-Doped SiO_2	2.7–3.3	PECVD
Parylene-N	2.7	CVD
Benzocyclobutenes	2.6–2.7	Spin-on
Fluorinated Polyimides	2.5–2.9	Spin-on
MSSQ	2.6–2.8	Spin-on
Aromatic Thermosets	2.6–2.8	Spin-on
Fluorinated DLC	2.4–2.8	PECVD
Parylene-F	2.4–2.5	CVD
Teflon AF	2.1	Spin-on

Table 1. SiLK, a type of aromatic thermoset, has low dielectric constant (k) and has the advantage of being a spin-on material, which makes it equipment-set compatible with future lower-k dielectrics. PECVD = plasma-enhanced chemical-vapor deposition. (Source: IBM)

Another criterion IBM imposed on its search for a low-k material was the requirement that it be extensible. For example, IBM knows that in the future (below 0.10 micron) it will have to adopt porous insulating materials to get a k closer to 2.0. These advanced porous materials are likely to be "spin-on" materials, as opposed to being applied with a plasma-enhanced chemical-vapor-deposition process (PECVD), as is silicon dioxide. Porous materials, however, will not be ready for manufacturing for several years. Therefore, for this generation, IBM wanted a spin-on material that would be compatible with future tool sets, allowing a smooth transition to porous materials when the time arrives.

Plastic Dielectric Is Smooth As SiLK

The dielectric material that IBM finally settled on for its 0.13-micron CMOS-9S process belongs to a class of materials known as aromatic thermosets, specifically an organic polyarylene-ether resin sold commercially by Dow Chemical under the brand name SiLK (see sidebar). Pure SiLK has a dielectric constant of 2.62; including overcoats, SiLK delivers a k_{eff} of around 3.0, about 25% better than FSG and more than 30% better than pure silicon dioxide.

Although Dow will sell SiLK to the industry, it will not be easy for other manufacturers to follow in IBM's footsteps. Ron Goldblatt and Jim Ryan, key contributors to IBM's low-k effort, point out that they had to develop a number of new techniques to integrate SiLK into IBM's copper process, which is shown in Figure 2.

One problem with SiLK is that, unlike SiO_2, it etches at the same rate as resist, a characteristic that makes it incompatible with the traditional copper dual-damascene process flow. To solve this problem, IBM developed a dual hardmask consisting of two dissimilar layers. The dual-damascene pattern is first etched into the hardmask layers, then transferred to the SiLK dielectric. Other techniques had to be developed

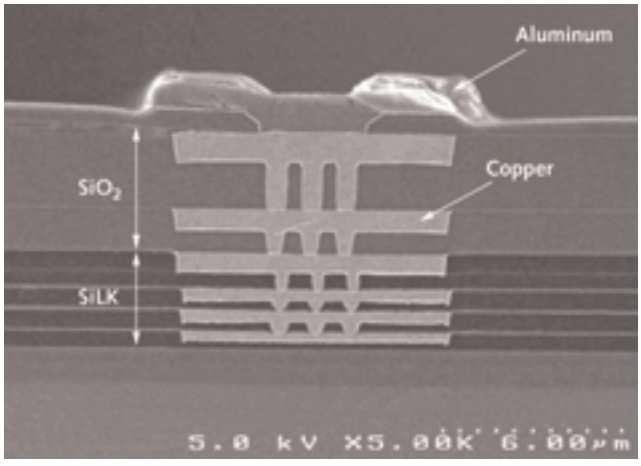

Figure 2. IBM had to develop a number of techniques to successfully integrate the soft SiLK dielectric with its copper dual-damascene process. For example, special structures had to be developed for supporting the upper interconnect layers and bond pads.

A Really Low *k*

IBM has announced that it will use SiLK resin from Dow Chemical in its next-generation 0.13-micron CMOS-9S process, which will enter volume production next year.

SiLK is spin-on aromatic hydrocarbon polymer with a dielectric constant of 2.62. SiLK is stable at temperatures of up to 450°C, allowing it to withstand the rigors of the semiconductor manufacturing process. The new material has an etch selectivity of 20:1 and can be etched with standard O_2/N_2 plasma. It is compatible with either aluminum or CVD- or electroplated-copper metal systems. With a toughness of only 0.62MPa-$m^{1/2}$, however, SiLK is softer and less adhesive than traditional silicon-dioxide interlayer dielectrics (ILDs), making it difficult to planarize with conventional chemical-mechanical polishing (CMP)—a problem IBM had to work around.

SiLK Property	Value
Dielectric Constant (k)	2.62
Leakage Current	3.3x10 $^{-10}$A/cm at 1mV/cm
Breakdown Field	4mV/cm
Glass Transition Temp (Tg)	>490°C
Thermal Stability	>450°C
Young's Modulus	2.7 GPa
Toughness	0.62 MPa-m $^{1/2}$
Film Stress	45 MPa
Moisture Uptake	0.25% @ 80% RH, 25°C
Thermal Conductivity	0.18 W/mK
Crack Growth Rate in Water	<10 $^{-11}$m/sec

Dow and IBM are now working together on ultra-low-*k* ($k \approx 2.0$) porous dielectrics for 0.10 micron and beyond as part of the National Institute of Standards and Technology's advanced technology program.

For more information on SiLK go to *www.silk.dow.com*.

to compensate for SiLK's low modulus (4% that of SiO_2) and poor thermal conductivity (15% that of SiO_2). IBM's techniques involve, among other things, special structures for supporting the interconnect layers and bond pads, changes in design rules to account for SiLK's different etch properties, and optimization of the barrier films to guard against copper contamination.

Solving this latter problem was one of the most challenging for the IBM team. Integrating a new dielectric material into a conventional aluminum metal system isn't an easy task, even though aluminum is chemically benign and its characteristics are thoroughly understood. But integrating a completely new nonoxide-based dielectric with copper—which is highly contaminating and understood much less well—is a far more challenging task. Motorola has previously disclosed progress toward integrating a porous inorganic dielectric ($k \approx 2.0$) with its copper-metal system (see *MPR 5/31/99-msb*, "Motorola Takes Capacitance to New Low"), but it admits that much work remains to be done to put that dielectric into production. IBM is the only company we know of that has cleared all the hurdles of integrating a low-*k* dielectric into a high-volume-production copper process.

This fact may shed light on IBM's strategy to make an early jump to copper in its 0.22-micron (CMOS-7S) and 0.18-micron (CMOS-8S) processes. The move to copper was criticized by many industry experts, who thought the move was unnecessarily aggressive. Intel, for example, argues that at 0.18 micron, it can achieve equivalent performance just by adding a low-*k* dielectric (SiOF) to its existing aluminum metal system. While that may be true, IBM now has two generations of copper-manufacturing experience under its belt and thus has a stable next-generation interconnect platform from which it can make the move to a true low-*k* material. By procrastinating, copper-naysayers will be facing a giant step up when they move to 0.13-micron lithography, copper interconnects, and a new dielectric material all in one generation.

IBM intends to deploy copper and SiLK across its entire process family, including its less-expensive foundry processes. The company has announced that in 3Q00 it will offer a design kit for Cu-11, a 0.13-micron CMOS-8SF ASIC with 40 million wireable gates. It expects to begin sampling the part in 1Q01 and be in full production by 3Q01. In this part, IBM will exploit the low resistance and capacitance offered by its copper/SiLK interconnect system to pack wires more closely together, doubling the number of wireable gates over the previous CMOS-7SF part. IBM says the embedded DRAM array in this part will be 40% denser and 25% faster than the embedded DRAM in its previous CMOS-7SF ASIC.

The embedded-DRAM cell in next-generation CMOS-9SF ASICs will be based on yet another IBM innovation: a vertical access transistor that is self-aligned with a buried strap into the trench capacitor. The vertical transistor eliminates the problems associated with continual shrinking of the gate length, thereby allowing a smaller cell size. The technique, which IBM described at the International Electron Devices Meeting (IEDM) last December, reduces the size of a DRAM cell by 25% compared with conventional cells.

Fast Interconnects And Fast Transistors

Copper interconnects and low-*k* dielectrics are all about reducing wire delay. And, to the extent that wire delay is a limiting factor in circuit speed, they do improve the situation significantly. To quantify the gain, IBM performed a complete 3D parametric extraction to simulate signal propagation through four different metal/dielectric systems. As Figure 3 shows, the simulation of a 200-micron M3 wire showed a 37% reduction in wire delay for copper/SiLK over aluminum/SiO_2—not including any indirect gains from reduced capacitive noise coupling (crosstalk). Because of

the conservative assumptions used in the simulation, IBM says it sees even better performance in real silicon than is predicted by the simulation: measurements indicate that copper alone provides up to 20% improvement rather than the 11% predicted by these simulations.

Of course, if wire delay isn't a limiting factor, then the gains predicted in Figure 3 will not result in faster overall circuits. Intel, in its campaign to defend its decision to forgo copper in P858 (see *MPR 2/28/00-02*, "Processors Penetrate Gigahertz Territory"), says it knows how to rebias the design to be transistor-delay dominated, eliminating potential gains from interconnect speedups. We find this argument unconvincing, however; while this unnatural technique may minimize wire delay, it is not clear that it results in faster circuits. In fact, Texas Instruments found that signal-propagation speed is optimal when gate delay and wire delay are balanced (within a clock cycle), and we estimate that in most 0.18-micron processors today, wire delays and gate delays contribute equally to circuit speed—notwithstanding Intel's techniques.

Moreover, since gate speed increases much more dramatically than interconnect speed when gong from one process generation to the next, wire delay will rapidly become the dominant delay term. By the time we reach 0.10 micron, or maybe even 0.13 micron, most of the 37% speed gain IBM predicts from copper and SiLK will manifest itself in higher processor frequencies. The remainder of the problem—gate delays—IBM is attacking aggressively with SOI and lithography.

"Industry Must Go to SOI," Says IBM

According to Ghavam Shahidi, manager of IBM's SOI program, scaling of bulk CMOS becomes extremely difficult below 0.13 micron, primarily due to short-channel effects. As transistor channel length shrinks, parasitic factors, which at long channel lengths were insignificant, become dominant. Loss of gate control (and transistor gain), high gate-overlap capacitance, subthreshold leakage, and tunneling, among other problems, conspire to eliminate the speed gains that have accompanied all previous process shrinks.

Although a few tricks remain at 0.13 micron to counteract some of these problems, at 0.10 micron and below they become unmanageable. Intel, for example, in a paper presented at last December's IEDM, described a notched-poly technique that undercuts the gate poly to reduce overlap capacitance. IBM says it shies away from such stopgap solutions, however, because they do not scale well to shorter channel lengths. IBM says that even at 0.18 micron, notched poly is more trouble than it's worth. The problem is that ultraprecise control over the etch is required to achieve consistent gate lengths, but such precise control is difficult because of factors such as the proximity of other structures, which create unavoidable local variations in the effectiveness of the etch.

Solutions to other short-channel problems are equally hard to find. At extremely tiny dimensions, manufacturing tolerances simply cannot be kept tight enough to adequately control source/drain doping profiles, for example. And some effects simply cannot be eliminated, even if manufacturing tolerances are perfect. For example, as transistors shrink, the critical charge required to upset SRAM cells and dynamic nodes is lowered. Below 0.13 micron, soft errors induced by charged particles become a big problem, putting a limit on how far these devices can be scaled. But, thanks to the isolation provided by its buried-oxide layer, SOI has a naturally immunity to such disturbances and thus has a much lower soft-error rate (SER) than short-channel bulk processes.

IBM's research into these issues has convinced Shahidi and Davari that there is simply no viable solution to scaling problems in general, save for one: silicon-on-insulator. SOI offers many advantages over bulk CMOS, which we detailed in our 1998 SOI article (see *MPR 8/24/98-02*, "SOI to Rescue Moore's Law"). The advantage Shahidi cites in defense of IBM's bold assertion that the industry must move to SOI, however, is that SOI offers another knob for controlling the shape of the channel. As Figure 4 shows, the silicon layer above the buried oxide—whose thickness can be precisely controlled—allows source/drain profiles that cannot be created otherwise, solving many of the short-channel problems. This extra knob also allows the creation of unique device structures with characteristics precisely matched to specific circuit needs.

IBM has been building SOI-based microprocessors for some time now, and through that effort it has gained considerable insight into SOI's properties. This experience, according to Davari, has given IBM increasing confidence that SOI is the right strategic path. Simple experiments, such as rendering the same PowerPC design in both bulk CMOS-7S

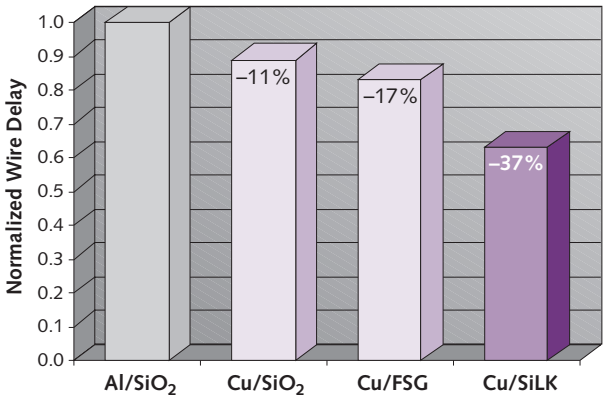

Figure 3. This graph shows the relative speed of a 200-micron M3 interconnect wire for four metal/dielectric systems. All three copper wires were the same thickness, and the aluminum wire was scaled to the same sheet resistance. Al/SiO_2 was used in IBM's 0.27-micron CMOS-6S2 generation, while Cu/SiO2 and Cu/FSG were used in 0.22-micron CMOS-7S and 0.18-micron CMOS-8S, respectively. Cu/SiLK will be used in 0.13-micron CMOS-9S. (Source: IBM)

and 7S-SOI, have demonstrated a raw speedup of more than 20% across a 7-sigma variation in channel lengths. Other experiments indicate that redesign to utilize the variable-threshold voltages (V_t) and deeply stacked gates made possible by SOI (and impossible in bulk CMOS) can achieve speed gains of 50%, and sometimes more. If these results carry through to volume production, which IBM says they will, just on the basis of SOI alone (independent of copper and low-*k*), IBM could be one full generation ahead of the industry in process speed while using the same lithography.

IStar, PA-8700 Debut in SOI

Proving that it isn't kidding about its move to SOI, IBM quietly revealed that it is shipping production 540MHz CMOS-7S-SOI processors, code-named IStar, to its AS/400 group. (IStar is a PowerPC-compatible processor with modifications for use in AS/400s.) The company did not say when IStar-based E-Server systems would be available, but historically it takes several months to put server systems into production, indicating availability early in the second half of this year.

IStar, which IBM first described at ISSCC in February of 1998, is essentially the same design as its predecessor, Pulsar, which operates at 450MHz in bulk CMOS-7S. A direct comparison between IStar and Pulsar provides powerful evidence in support of IBM's claim of 25% speed boost due strictly to SOI, without redesign.

In fact, this comparison may underestimate the gain from SOI. Since Pulsar has been in production for some time, its L_{eff} is probably being pushed more aggressively than that of the new IStar. If true, IBM probably still has enough headroom to push IStar's speed closer to 600MHz, making it 33% faster than Pulsar. Whether the company will make this move depends on how quickly it intends to follow with a CMOS-8S2 version. (CMOS-8S2 is an SOI-only process.) According to IBM's data, shown in Figure 5, 8S2 is 20% faster than 7S-SOI at nominal channel lengths and 33% faster at aggressive L_{eff}. Thus, an 8S2 version of IStar should easily coast to 700MHz.

In an announcement that shocked everyone, including IBM, HP disclosed on April 11 the details of an 800MHz PA-8700 processor, which will be available in systems by 1H01.

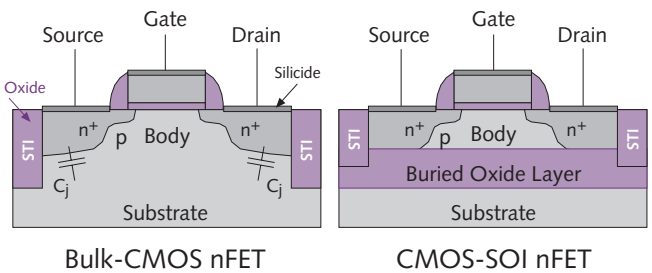

Figure 4. In addition to the beneficial effects of reducing capacitance, SOI provides more control over the source/drain doping profile, thereby providing an additional handle on the short-channel effects that threaten device scalability. STI = shallow-trench isolation.

While the 8700 announcement was expected, the disclosure that it would be built in a copper SOI process was a surprise. Although HP didn't officially announce the fab for the 8700, IBM is the only vendor on the planet with a production-worthy copper SOI process. Thus, the mystery of who is building PA-RISC chips these days is now pretty much settled. In fact, the HP-IBM linkage is so transparently obvious that IBM execs are probably more than mildly upset with HP for preempting their official SOI AS/400 announcement.

HP is not the only company looking to IBM for process technology. Sun recently confirmed our suspicions that its MAJC-5200 (see *MPR 10/25/99-04*, "Sun Makes MAJC With Mirrors") will be built by IBM rather than by its long-time UltraSPARC partner, Texas Instruments. The 5200 is now entering production in 0.22-micron copper CMOS-7S, but it will soon move to 0.18-micron copper CMOS-8S, and eventually to SOI.

Also, at Microprocessor Forum last October, Compaq said that its 21464 would be constructed in a 0.13-micron copper low-*k* SOI process (see *MPR 11/15/99-msb*, "Alpha 21464 targets 1.7GHz in 2003"). Furthermore, rumors persist that Compaq is on the verge of announcing a deal with IBM to produce copper Alphas, probably the 21264, probably in CMOS-8S. Given Compaq's Microprocessor Forum statements, we suspect it is also negotiating for access to SOI-based CMOS-8S2 and CMOS-9S for its 21364 and 21464. Such a deal would be a good move for Compaq and would give us a more favorable outlook on the future of Alpha.

These revelations by Compaq, HP, and Sun represent strong votes of confidence from the industry's top performance leaders for IBM's copper/SOI/low-*k* process roadmap.

Seeking Unlimited Resolution

While IBM pushes hard on the materials front with copper, low-*k*, and SOI, it is not ignoring the lithography front.

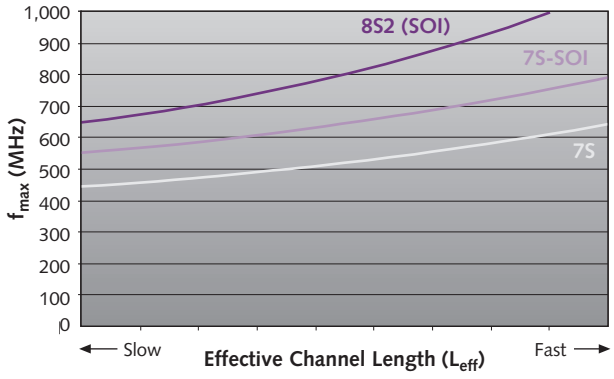

Figure 5. At nominal channel lengths, IBM's next-generation 0.18-micron CMOS-8S2 (an SOI-only process) is about 20% faster than 0.22-micron CMOS-7S, which is about 24% faster than the same process in bulk CMOS. CMOS-8S2, however, speeds up more than CMOS-7S as channel lengths are pushed toward the fast end of the processes. At aggressive L_{eff}, 8S2 is almost 50% faster than 8S2 at nominal L_{eff} and more than 30% faster than 7S-SOI at aggressive L_{eff}. (Source: IBM)

Today, for 0.18-micron processes, nearly all manufacturers rely on optical projection lithography using deep ultraviolet (DUV) light at a wavelength of 248nm. But this wavelength is just adequate to image the smallest features on a 0.18-micron chip while maintaining adequate depth of field for high-yield, high-volume production.

To go below 0.18 micron requires a number of resolution-enhancement techniques (RETs), such as off-axis illumination (OAI), strong-phase-shift masks (PSMs), optical proximity correction (OPC), and increased numerical-aperture lenses. Using these techniques, 248nm optical lithography can be pushed to serve the 0.13-micron generation—barely. The 1999 *International Technology Roadmap for Semiconductors* (ITRS99) calls for a transition to 193nm steppers during the 0.13-micron generation, which, with RETs, will suffice down to 0.10 micron—again, barely. For the 0.10-micron generation, the ITRS99 calls for another wavelength reduction, to 157nm. This time, RETs will allow 157nm steppers to serve down to 0.07 micron, but beyond that DUV isn't workable because, among other factors, lenses just become too opaque.

Therefore, during the 0.07-micron generation, the ITRS99 calls for a transition to a next-generation lithography (NGL) approach. There are four basic candidates for NGL: extreme-ultraviolet lithography (EUVL), X-ray lithography (XRL), electron-projection lithography (EPL), and ion-projection lithography (IPL). Although there is no industry consensus on which is the best approach, the majority of activity and investment over the past few years has been on EUVL, which, at a wavelength of 13.4nm, is suitable for as long as anyone reading this article is likely to care.

Intel has been the primary driving force behind EUVL, and it has formed an industry consortium, called the LLC, to help develop the technology. Three of the major national laboratories—Lawrence Livermore, Sandia, and Lawrence Berkeley—carry out the majority of the work for the LLC, which, surprisingly, includes AMD and Motorola. Sematech also contributes to the LLC's efforts.

This lithography roadmap, however, is not without problems. Chief among them is cost. Today, a single-column 248nm optical stepper costs $8 million to $12 million, and a large fab typically has a couple dozen of them. Replacing this equipment with 193nm steppers will be enormously expensive—not to mention the additional cost of RETs, which is also high. Some industry analysts believe that optical lithography will simply be too expensive for 0.10-micron processing, due both to the cost of equipment and to the poor yields that some expect as a result of narrower and narrower process windows. To turn around and repeat this exercise for 157nm DUV just a couple of years later would be staggering.

At one time it was hoped that EUVL would be ready for the 0.07-micron generation, possibly eliminating the need for the intermediate 157nm DUV step. This, however, does not appear to be feasible. The progress on EUVL has been excruciatingly slow, and the cost of the EUVL systems is likely to be higher than originally projected.

E-Beams to the Rescue

Meanwhile, IBM has been plugging away with its EPL research. For many years, the company used e-beam direct write (EBDW) to quickly turn bipolar chips for its mainframes. Initially, its Gaussian-beam EBDW steppers, which raster scan the circuit pattern directly onto the wafer at a rate of one pixel per flash, had lousy throughputs of 0.01 wafers per hour. During the 1980s—when feature sizes were 2 microns and there were fewer than 10^{10} pixels on a 5 mm wafer—IBM coaxed throughput upward to 20 wafers per hour with several-hundred-pixel-per-flash shaped-beam projectors.

The writing speed of these EBDW tools, however, did not keep pace with the Moore's Law rate of pixel growth, and it became clear that throughput would never be adequate for today's high-volume production, which will soon require writing 10^{13} pixels on a 200mm wafer. (Today's DUV steppers routinely achieve throughputs of 80–100 wafers/hour, and EUVL steppers—which are similar except for their use of mirrors rather than lenses—should have similar throughput.)

But IBM did not give up on e-beams. The company's latest breakthrough is the development of a practical e-beam projection-lithography (EPL) system, which uses mask projection analogous to that used in optical lithography. EPL is attractive as an NGL candidate because its resolution, for all intents and purposes, is unlimited. Both EPL and EUVL are capable of being extended to the 0.035-micron generation and beyond. EPL, however, has never been used with any success in semiconductor manufacturing because of practical limitations, primarily that of limited field size.

One source of problems, says IBM Fellow Hans Pfeiffer, is that electrons are charged particles, and they repel each other (Coulomb interactions). This effect tends to blur the image at high beam intensity. Moreover, while an EPL projected field can be larger than that of a EBDW system, it is still much smaller than most chips, requiring the field to be scanned over a considerable distance to cover the chip. Deflecting the beam very far, however, introduces off-axis aberrations that defeat attempts to contain Coulomb interactions.

To increase throughput in spite of these problems, IBM had to find a way to apply massively parallel pixel projection across a large field without creating distortion. For this feat it developed a novel magnetic lens system that minimizes off-axis aberrations by electronically shifting the optical axis of the lenses in sync with the beam. As Figure 6 shows, this creates a variable curvilinear axis for which the system is named PREVAIL (projection reduction exposure with variable-axis immersion lenses).

IBM is no longer the only company that believes in e-beams. It was apparently able to convince Nikon, the largest supplier of optical steppers today, that its system was a viable

NGL contender. Together, the two companies have constructed a proof-of-concept EPL system, shown in Figure 7, that employs a high-emittance, high-numerical-aperture e-beam source along with a silicon stencil mask and a proprietary distortion-correction system. The prototype system, which currently delivers a 12.8μA beam current during each 100μs pulse, has been used to successfully demonstrate 0.08-micron lithography over a 5mm-wide field without significant loss of resolution, as Figure 8 shows.

IBM expects to coax its PREVAIL alpha-tool performance to a 15μA beam current, delivering 10 million pixels per flash over a 7mm-wide field, which would support a throughput of 35 wafers per hour. On the strength of this prototype system, Nikon says it will build a commercial stepper for deployment in 2003.

Although Pfeiffer admits that EUVL systems will have some advantages over EPL systems, he says that production EPL steppers can be delivered earlier than production EUVL systems with competitive throughput, and that EPL steppers could cost even less than today's DUV optical steppers. If this is true, it would certainly make a compelling case for EPL as the industry's next-generation lithography system. IBM is currently investigating methods for extending EPL to 0.05 and 0.035 micron without sacrificing throughput.

Firing on All Cylinders

IBM has always been recognized by the industry as a technology leader. But other semiconductor companies have come to realize that technology is an incredibly important weapon in the microprocessor business—no architectural, microarchitectural, or circuit design innovation is likely to have even close to the impact of a half-generation lead in semiconductor technology. And conversely—nothing is likely to be as devastating as a half-generation technology lag. With such high stakes, other companies have also been investing heavily in advanced semiconductor process development, making us wonder just how long IBM could maintain its preeminent position at the top of the IC-process totem pole.

Despite heavy investment by other companies, however, IBM recently seems to be pulling even further ahead. The string of announcements over the past two years has been truly impressive. While other companies nibble around the edges of next-generation process problems, IBM takes giant bites out of them. Copper, SOI, plastic dielectrics, and e-beam lithography are big bites. But each move the company makes, while unquestionably aggressive, seems to be well justified.

Moreover, they are synergistic. While each technology is valuable in its own right, the combination is awesome. Together, copper, SOI, and SiLK support new design methods capable of producing chips that are easily twice as fast as could be built with a conventional bulk aluminum/SiO_2 0.13-micron process. Other companies will eventually follow

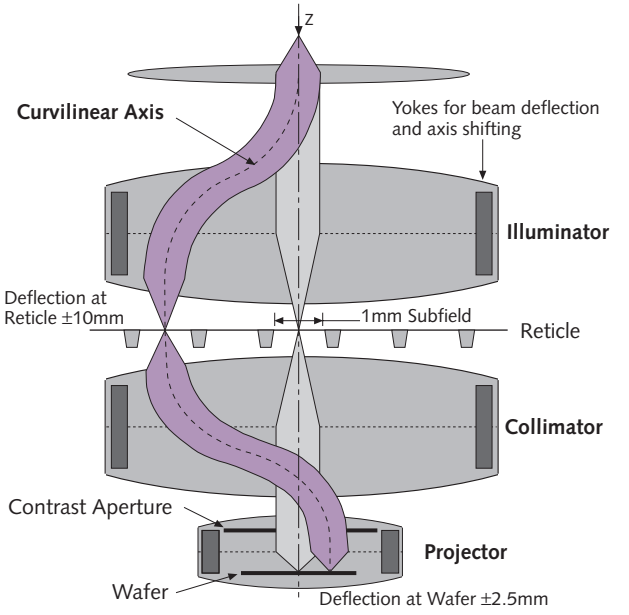

Figure 6. In IBM's EPL system, the optical axis of the lens system is electronically shifted by ±10mm in sync with the electron beam, thereby minimizing off-axis aberrations and allowing a larger electron-beam subfield and a higher beam current.

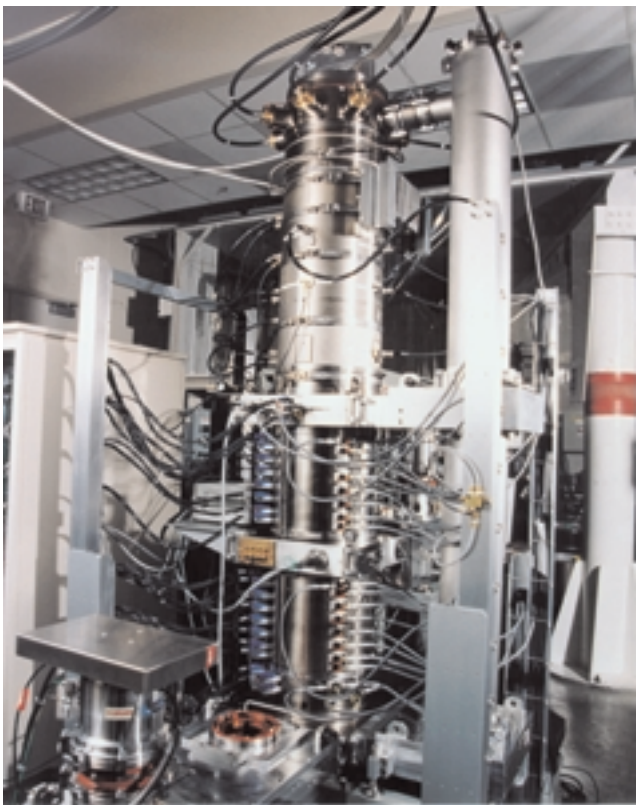

Figure 7. IBM and Nikon's proof-of-concept PREVAIL system is currently operating at a beam current of 12.8μA and has demonstrated 0.08-micron lithography over a 5mm-wide field without significant loss of resolution. (Source: IBM)

in IBM's footsteps, some willingly, some not. At this point, however, unless other companies are being incredibly secretive, IBM appears to be a good two years ahead of the rest of the industry.

IBM's technology lead is not going unrecognized. Nearly every major semiconductor vendor is actively trying either to license technology from IBM or to emulate it. UMC and Infineon, for example, have just entered into a major technology agreement with IBM (see *MPR 2/14/00-02*, "IBM, Infineon, UMC Gang Up On 0.13"). Motorola and AMD have joined forces to develop copper HIP6 and future processes that are likely to include SOI and low-*k* dielectrics (see *MPR 8/3/98-msb*, "Motorola, AMD Swap Technology"). We expect that even Intel, although it is forced to go slow because of its enormous volumes, will eventually follow IBM's lead, as it has done previously on such IBM innovations as shallow-trench isolation.

Moreover, nearly all high-performance processor design houses (except Intel, Motorola, and AMD) are beating down IBM's door to gain access to its advanced processes. Plans by Intel-partner HP for the PA-8700, TI-partner Sun for MAJC, Samsung/API-partner Compaq for Alpha, and startup Transmeta for Crusoe (see *MPR 2/14/00-01*, "Transmeta Breaks x86 Low-Power Barrier") are all strong endorsements of IBM's semiconductor technology. Every company in the world that is building a performance- or power-critical microprocessor or SOC knows instinctively that IBM is the place to look for the best technology. They also know, however, that it is the most expensive place to look. IBM is proud of its technology and is not ashamed to ask a premium price for it.

One next-generation technology on which IBM has been notably silent is the issue of 300mm (12-inch) wafers. John Kelly, the general manager of IBM Microelectronics, has stated that IBM doesn't intend to be the first company to use the foot-wide wafers. It's not a surprise, however, that IBM would be slow to adopt 300mm wafers. Although 300mm wafers are important from a fab-capacity point of view (300mm wafers carry 2.5 times as many chips as 200mm wafers), they do not directly contribute to performance, power, logic density, or reliability, which are IBM's primary concerns. Besides, the company does not intend to lag far behind the industry on 300mm. Davari says IBM will begin the transition to 300mm wafers during the 0.13-micron CMOS-9S and CMOS-8SF generations, putting it only slightly behind leaders Intel (see *MPR 6/21/99-msb*, "Intel Commits to 300-mm Wafers") and UMC (see *MPR 1/24/00-04*, "Hitachi, UMC Jump on 12" Wafers").

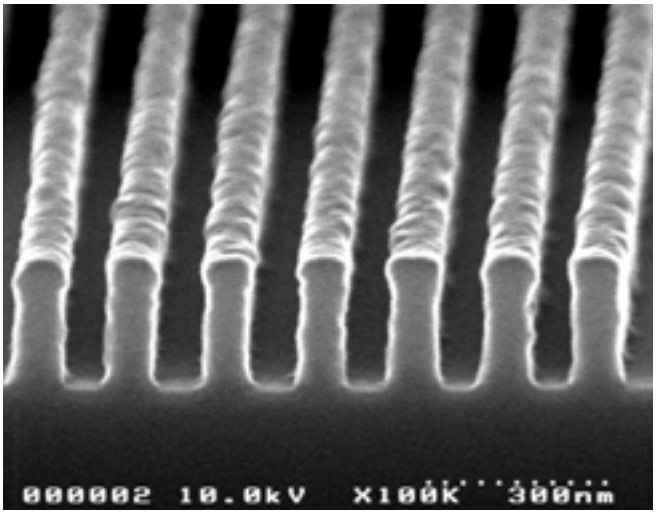

Figure 8. This photomicrograph shows 80nm lines and spaces (in resist) imaged by the PREVAIL proof-of-concept system at the edge of a 5mm field. The pattern shows almost no degradation when compared with patterns imaged at the center of the field. (Source: IBM)

IBM's strategy to stay ahead of the rest of the industry on technology is a bold one, if not an extremely risky one. To stay on this fast-moving treadmill, IBM cannot afford to stumble. A single misstep, such as falling into a losing-technology rat hole, could easily throw IBM off the treadmill, which runs too fast to get back on. To guard against such risks, IBM is attempting to follow a very well thought out long-range roadmap and to distribute the risks by working in parallel on multiple technology fronts. So far, the strategy is working, but it will require extreme vigilance to continue this strategy ad infinitum. IBM is silent on when or from where its next process advancement will come. But given its strategy and its past performance, it is a safe bet that East Fishkill researchers have something up their sleeve. ◇