

Overcoming the Barriers to 10Ghz Processors Bill Pohlman, Chairman, CTO and Founder, Primarion

> Microprocessor Forum Keynote Tuesday, Oct. 16, 2001





Good Morning!

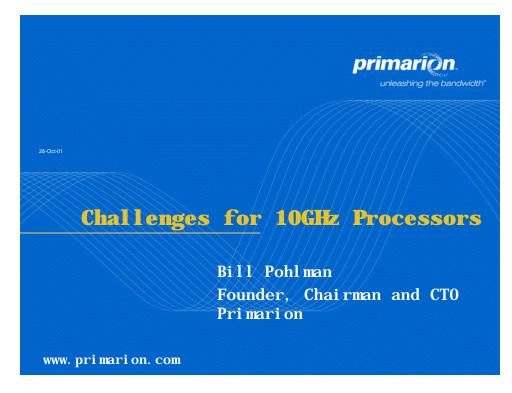
As a 'retired' Microprocessor Developer who participated in this technology since the early 70's, it's great to have a chance to give a keynote at the industry's premiere conference!

Sometime in my years at Intel, Andy Grove once characterized me as a "Maverick" in a positive way. Maybe that was because I built a RISC processor business inside Intel!? (the 960)

Well, I admit I do try to look at things differently. I suspect many of you can relate this!

So, in this keynote, I promise you at least a couple of controversial new ideas to engage. So, let's get started!!





The purpose of this keynote is to engage the issues of 10GHz processors. But first let me set some context before I cover my agenda slide.

I am sure as process technology advances, we can all project these clock rates for super-pipelined processors to occur within a few years, as we get below the 100nm technology node.

This progression of clock rate is, frankly, *really getting boring* to me, so I am not going to talk on this, per se, or the related process technology, which is beyond my expertise.

But, as Moore's Law advances, there is a number of related technology scaling issues impacting these systems, i.e. second order effects that we will need to deal with.

In some cases these will create inflection points where new technologies will emerge. That's exciting. That's worth talking about!





My first experience with Moore's Law was when I lead the 8086/88 program at Intel. The 8086 was introduced in early 1978.

The Marketing challenge at the time was to develop a processor 10X faster than the 8080/85. They needed it in 18 months to fill a gap in time until a real next generation processor could be architected and delivered.

It was clear that Marketing had never internalized all aspects of Moore's Law, i.e. transistor counts double about every two years, and its corollary that *performance doubles about every 18 months.* 

The project took 18 months all right and performance and integration doubled on Moore's Law. A couple of tiny benchmarks did hit 10x performance improvement but that was it.

Thinking back, no one realized what the impact of the 8086 would be on this industry at the time. In fact, it was planned as a one shot *"gap filler"*. That was one hell-of-a gap we filled!

I, for one, probably would have spent an extra week on the one-month product definition had I known! ... Perhaps to add a couple more segment registers. (Ha!)

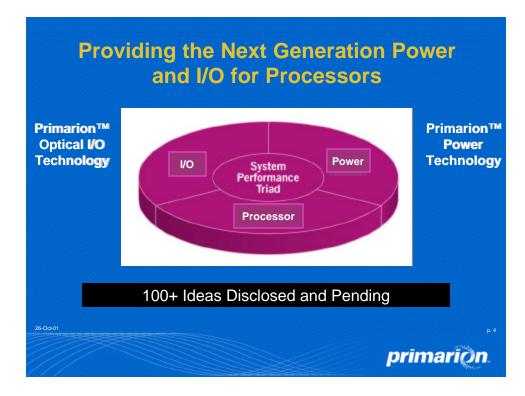


I wonder if Marketing still wants 10x performance improvement on every generation?

...Probably!

But, I suspect Moore's Law continues to bind our real engineering expectations.





OK, I really *did* retire, but that lasted just a couple of months until I was convinced to join a semiconductor start up - Primarion.

My task was to develop a technology and product strategy. I reflected a lot on Moore's Law during my short-lived retirement, on a cruise of the Greek Islands. My thesis is that, as it advances, technology inflection points will occur in some areas as existing technologies fail to keep up with the needs of multi-GHz processors. Typically, technology inflection points are periods in time where new, innovative solutions can or *must* enter the market.

Startups are great places to foster the innovations needed and in a highly focused way.

So Primarion is uniquely focused on two of these technology inflection areas: wideband power delivery for GHz processors and optical busses and I/O. Both areas, it seems to me, have had too little attention, leading to the issues and enabling the solutions I will address today. Power has been treated as a "given" by microprocessor designers, and anybody who has been on a microprocessor project knows the bus is typically the last unit to get attention.

While most of the industry investment today is centered on the processor's internal chip design, it is Primarion's thesis that these two additional technology areas – power and I/O – will increasingly capture more of the system value

<sup>© 2001</sup> Primarion, Inc.



going forward. This will come about directly from the significant incremental performance that they can deliver.

Furthermore, all three of the legs of this stool must be properly in place to deliver maximum value to the customer in multi-GHz designs.

These related gains are much like any new microarchitecture innovation, such as multithreading, a reportedly 18 percent to 30 percent performance enhancement. Both new GHz speed power and high bandwidth bus interconnect technologies should bring similar levels of overall system performance improvement and thus warrant similar levels of attention.

A more balanced approach will be necessary in the future as Moore's Law advances!

Now I would like to summarize my agenda for discussing these challenges.



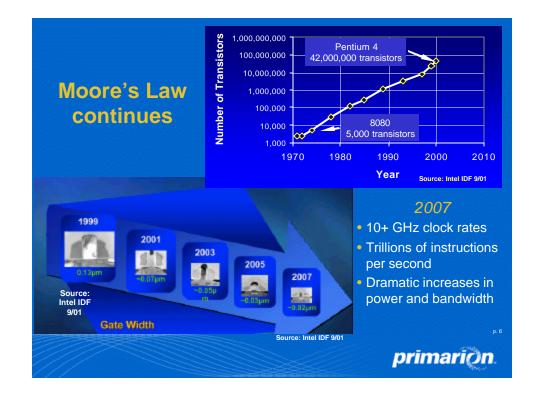


First, let's revisit Moore's Law quickly and then outline some basic attributes foreseen for a 10GHz processor. This will lay the foundation for a more detailed discussion on its Power and I/O challenges as Moore's Law continues. As, I said, failing to resolve these means we will be unable to bring the full value of a 10GHz CPU to the markets.

Finally, I will wrap the two concepts together with future microprocessors into what we at Primarion refer to as the encapsulated processor vision.

This is a way of thinking of processors of the future where a carefully structured environment allows optimum performance realization.





Here, I borrowed a couple of slides from Intel's recent Developer Forum.

We can see that both from a process scaling and transistor integration trend we are going to continue on Moore's Law through much of this decade.

New process nodes appear about every two years.

In Fall 2001, we have seen the start of the 2GHz processor and demos in the 3GHz range using Intel's Netburst\* super-pipeline execution architecture.

In 2001, at the 130nm lithography node we have also seen the first communication chips with 10GHz front-end circuits for OC-192 and 10G Ethernet.

Incidentally, don't confuse effective device gate channel width with the lithography node. This foil shows channel gate width.

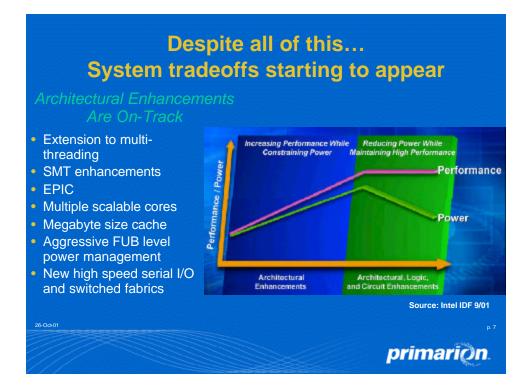
Somewhere in the 2005-2007 period 10GHZ clock rates should be in volume production.



Integration trends are also on track with this improved lithography for a 300-400 million-transistor device by then.

All looks good! So, where are the system trade-offs and real challenges starting to appear?





Power and thermal limits are impacting mobile computers as shown in another Intel IDF foil.

Moore's Law of performance improvement, the doubling every 18 months or so, is hitting the wall in Mobile designs due to power limits. Now the emphasis is moving to getting great performance at low and lower power extending battery life.

Yet, in desktop and servers, we are still seeing rapid advancement in micro architectures designed to extract more parallelism from instruction streams without much emphasis on power saving.

But, the emphasis of future microarchitectures must move to incorporating lower power approaches and circuit techniques while still providing performance advances.

These power-saving microarchitectures are well along in mobile today in processors like Banias from Intel and Crusoe from Transmeta, but by the era of 10GHz microprocessors, I suspect this will become the central focus of all future microarchitecture designs.

It's likely that the quality of all designs by then will be measured by how well they deliver performance within a given thermal environment or power limit.

<sup>© 2001</sup> Primarion, Inc.



MIPS/Watts will be the key metric from top to bottom.

It is also very likely that power limits may well be "only one of a number of issues" that could "bend" the performance improvement rate of Moore's Law as we continue technology scaling.

Let's examine some other areas...



	Moore's Law Trend Summary
•	<ul> <li>On track to 10 GHz</li> <li>Lithography, transistors, gate delays, density, integration,</li> <li>Bottlenecks</li> <li>Electrical buses, interfaces and interconnects</li> <li>Packaging and decoupling technology</li> <li>I/O and memory speed, latency</li> <li>Design complexity and productivity</li> <li>Show Stoppers</li> <li>Thermal management</li> <li>Power supply noise and transients</li> <li>I/O SNR, BER, transient errors</li> </ul>
26-Oct-01	primarion.

Here's my summary of how things are trending or tracking with on Moore's Law:

Obviously, as I mentioned before process and lithography and thus integration trends and device speed appear on track.

But behind the trend are some key areas, which are becoming bottlenecks to staying on the Performance Trend even though they are improving somewhat:

Electrical Interconnect and Buses are hitting practical speed limits. Even exotic electrical signaling technology, like multilevel approaches, is only delaying the day we will need a new medium.

Packaging technology, from an interconnect inductance point of view is slowly improving. I will chat on this more later.

And capacitive decoupling technology for power plane noise reduction is also improving too slowly.

I/O continues to be bottleneck in systems and we all know memory performance has not tracked with processor speed. Design complexity is certainly on track growing 58 percent a year but design productivity is only growing 21 percent a year.

<sup>© 2001</sup> Primarion, Inc.



I am very aware of this in that my last job at Intel was an assignment to fix this specific issue. In the best case, you might get back to a Moore's Law improvement rate but it seems impossible to catch up. Projects continue to get bigger and take much longer failing to even hold steady with the trend.

But, real performance improvement show-stoppers are also emerging.

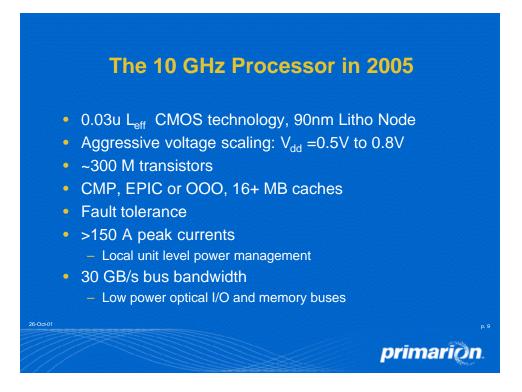
Pat Gelsinger, Intel's CTO, a good friend of mine, gave an eye-opening keynote on this at the Solid State Circuits Conference earlier this year where he projected microprocessor case temperatures would eventually be getting close to those of the surface of the sun. Obviously, this can't happen. Thermal management is emerging as the major show-stopper. And, packing devices closely together in a system as necessitated by electrical considerations makes this worse.

Thus, we must stop the growth in power consumed while still gaining performance: aggressive voltage scaling and functional unit clock gating are the only methods practical in the short-term horizon. System desegregation will help spread out the thermal management problem. But, then electrical noise issues in power delivery limit voltage scaling and also impact device performance.

Finally, at very low VCC voltages, reduced signal to noise ratios will lead to increases in bit errors and thus frequent transient processor failure.

So, in this context, let's look at some likely attributes of a 10GHz processor...





You can project the 10GHz processor will be available in 2005 based on the availability of some 2GHz processors today using the doubling of performance every 18 months or so.

But, we may be surprised with one sooner as the race to 10GHz is on today. We should see some yield to 10GHz as we push 90nm technology but wide availability will likely wait for the next process generation after that.

We can all extrapolate what we think a 10GHz processor will look like simply by projecting from today's position: more cache, more execution units, more extraction of instruction, thread and task level parallelism and so forth. Pick your favorite parallelism technique.

I will stay out of that religious pursuit except to say the focus must move "real MIPS per watt delivered" for all processors in this time frame, top to bottom.

Operating voltage must drop to limit power and reduce gate stress effects. Extensive use of FUB power gating, even in high-end processors, may mitigate how low we have to go. Nevertheless, with 4x more integration and 10GHz clock rates we will likely be well below a 1 volt Vcc supply. In fact, at these levels, I am sure power supply technology will likely become a source of competitive advantage. This is very hard technology to develop since we need low noise and lots of bandwidth at these processor clock rates.

<sup>© 2001</sup> Primarion, Inc.

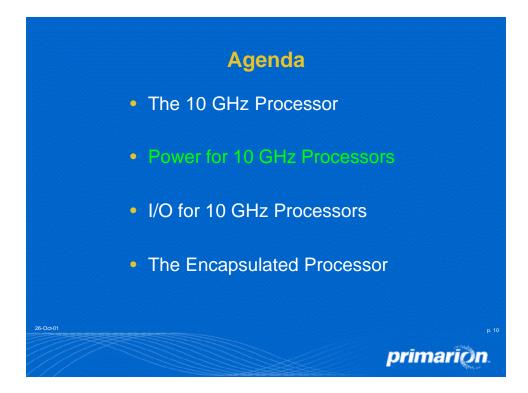


Supply currents will need to be well over 100 AMPS, but more importantly peak to average current ratios will grow much more rapidly as power management gets more aggressive. These create instantaneous load changes called di/dt events. Clearly, this also will be quite a challenge to power delivery systems by extending the load lines they must support and response bandwidth mandatory for much lower supply noise.

As transient error rates increase due to lower SNR associated with lower operating voltages and higher supply noise, we must develop more dynamic fault tolerant designs as well. You'll likely see extensive use of error correction and on-the-fly recovery techniques.

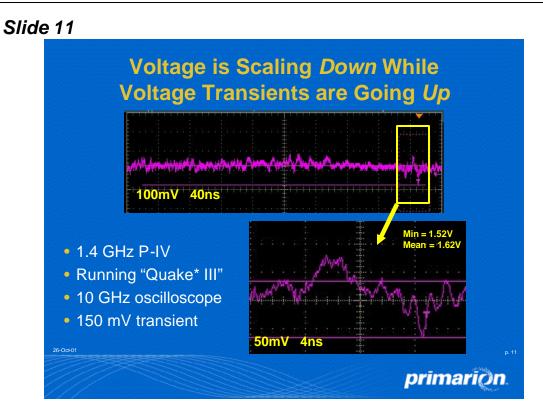
Finally, we need more bandwidth to feed these beasts! And more memory performance! My thesis is we will eventually be pushed to develop optical media interconnect. I will speak much more on this later. The alternative is LVDS signaling and the use of many more I/Os. Optical is attractive to me because bandwidth can be dramatically scaled up in the future with Moore's Law.





Next let's focus on the power delivery challenge for multi-GHz microprocessors!





Here's a look at a 1.4 GHz Pentium IV processor today, sensing at the core voltage sense pins with a super fast scope:

Droop events > 100mV occur on today's processors and have a typical duration < 15ns.

They are caused by complex computations, which need a huge amount of energy delivered instantaneously.

They *will* get much worse with faster clock speed.

They *will* limit yield and maximum processor clock rate as we need to derate specs for them.

Today, we try to mitigate these droops with high-performance decoupling capacitors with not much success. We use automatic voltage positing as well.

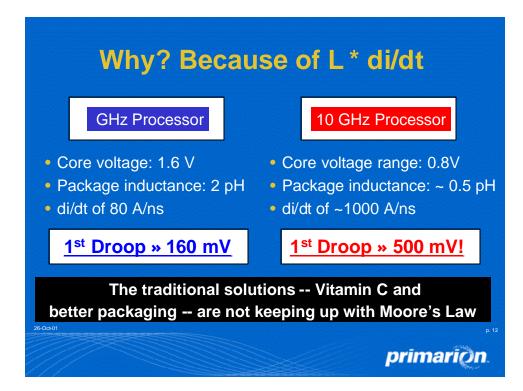
But current generation switching VRM's are inductively too far away and are way too slow to solve this problem.

So, as we scale device speed and device integration up and voltage down, would you expect this problem to get worse or better?

Yes, it gets much worse! Let's take a closer look at why...

© 2001 Primarion, Inc.





Here's why the noise and voltage droop. . . Inductance is killing us!

As I said, Transient Current Load during power hungry computations causes instantaneous need for large amounts of charge. But the power network inductance limits the ability to deliver the charge fast enough, thus leading to core voltage droops.

"Vitamin C" is our affectionate term for passive decoupling caps. They are limited by parasitic inductance as well and at low voltages store less charge. These droops can cause transient failure unless we guard band power specs. That means more power and slower specs.

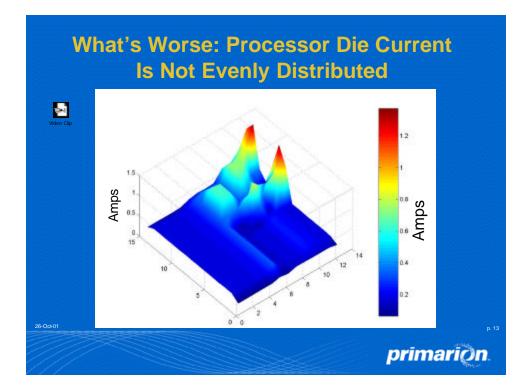
Much before 10GHz clock rates are reached these issues will drive us to new and innovative power technology. We need an active, high-speed charge delivery technology. Obviously, we must be able to regulate out droops as high as 500 mVolts. Ouch!

The alternative is to create microarchitectures that present a smooth continuous load to the power system. But, aggressive clock gating to reduce power is not that direction.

Moore is meeting Maxwell here! Moore's Law is meeting Maxwell's aw of Electromagnetism here!

<sup>© 2001</sup> Primarion, Inc.





Worse yet, this is a spatial transient power delivery problem! Look at the variation across an advanced GHz processor die.

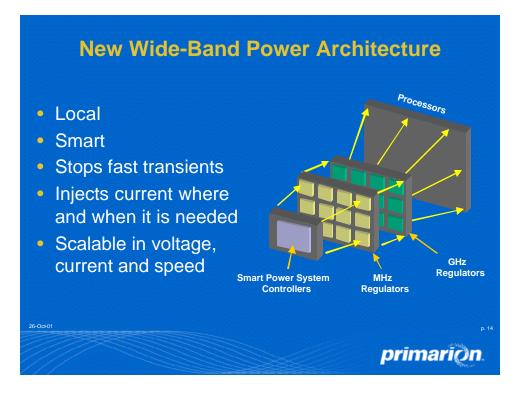
Any new power technology needs to deal with that as well.

Note the high current transients under execution units; floating point units are especially problematic in this regard. Let's take a look at the power plane in animated simulation of a high di/dt event... [run animation]

And I suppose you expect your processor to keep running when this happens?

Dream on!





Primarion is developing such wideband power technology. It requires a level of transient regulation that can actually respond beyond the processor's clock rate. This multi-tiered technology can deal with very high di/dt events both temporally and spatially in a closely integrated response. Packaging is the key to success here as we must minimize parasitic inductance.

This unique technology uses SiGe BiCMOS to deliver huge amounts of charge in a nano second or two, and also responds in less than a nano second to high di/dt events. The bandwidth of the regulation technology is a multiple of the processor's GHz clock rate. This is what you need to handle a GHz changing load.

It is a fully scalable technology. It will be available next year.

We believe the technology will be necessary to enable 10GHz general purpose microprocessors to run reliably below 1 volt core voltage.

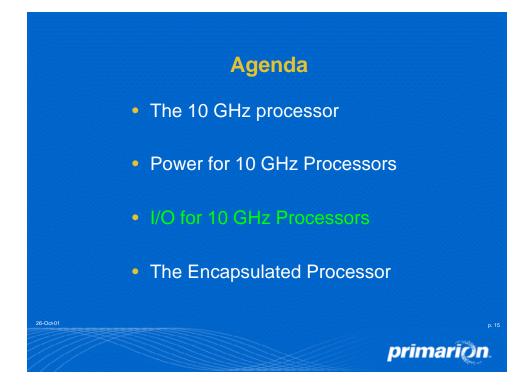
But, today its value proposition is clear: Wideband power delivery will enable much tighter power specs having 4-7x smaller voltage droops. This will translate to adding back higher speed bins. I predict we will recapture at least 20% more clock rate with this wideband power based on our extremely detailed simulations. This is very comparable to many of today's microarchitecture



performance advances. This will translate to competitive advantage for the early adopter.

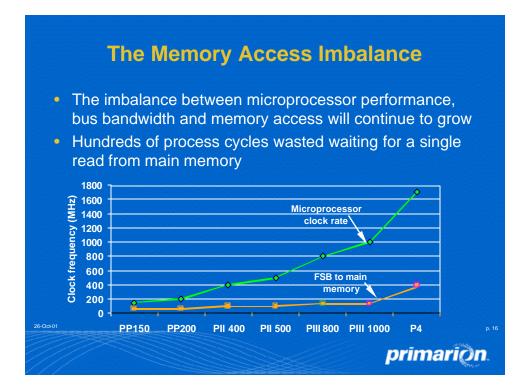
By the way, as we announced yesterday, Intersil has partnered with Primarion to help develop the technology and bring it to market.





Now lets move on to busses and I/O.



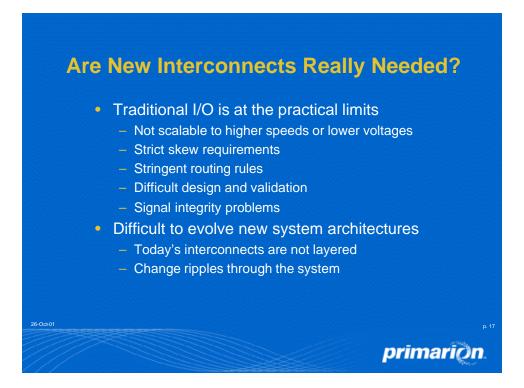


If we project forward current bus technology, we can see bigger and bigger gaps between processing rates and bus bandwidth. If you add memory access to this as well, we will waste hundreds of clock cycles on cache misses. Obviously, front side bus bandwidth is not scaling with Moore's Law in that the throughput cannot keep up with core execution rates. At current scaling rates this will become a huge issue at 10GHz clock rates. The industry will likely go to LVDS front side buses to try to mitigate this but the cost will be higher as pad counts will skyrocket.

To me, this suggests we are approaching a technology inflection point.

We need a better approach!





The case has been building for the last few years for a new system interconnect.

I could have filled this foil with press headlines of very expensive product delays and recalls caused essentially by signal integrity issues.

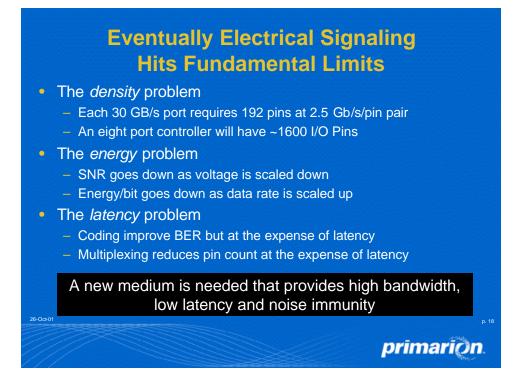
Today we are seeing even more complex multilevel electrical schemes emerging. We think that these will just add to the industry's signal integrity problems. Also they will add to the growing communication latency created by encoding and reliable recovery of multilevel signals.

Electrical bus throughput will just not scale with Moore's Law performance trends even with the use of massive numbers of I/O pins and lots of wasted power.

Thus, the case for new interconnect medium is becoming very compelling.

I hope that within the time horizon of 10GHz processors we will see a new optical interconnect technology to emerge. It must!





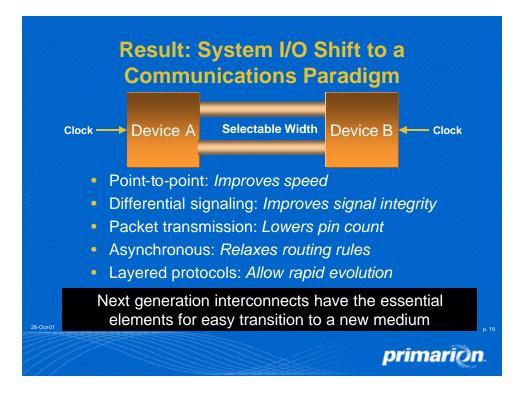
Again, the fundamental problem with electrical interconnect is that it is just not scaleable with Moore's Law. It is another place where Moore is meeting Maxwell.

I could have also added the distance problem to the list here. High-speed electrical interconnects cause the "densification of systems," concentrating thermal loads into small volumes. This makes thermal management a bigger problem.

Clearly, the wrong direction.

A good example is in server systems wherein a dual microprocessor system must have its two processors so close they face each other on opposite sides of the motherboard. Lack of signal integrity is driving this tight configuration. I will not belabor these points in the interest of time here.

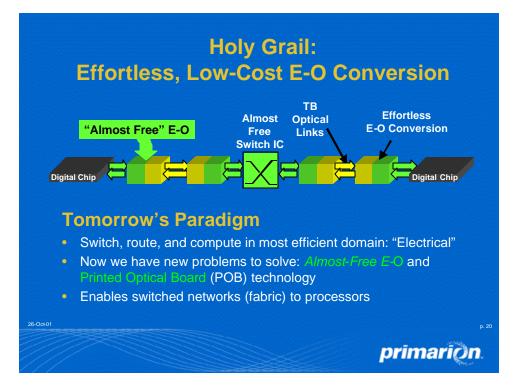




In the system I/O area we are shifting to new technology like InfiniBand, 3GIO as adopted by the PCI SIG, and HyperTransport. The protocols and physical layer of these new standards will support a fairly straight-forward transition to optical media since they are simple in nature.

Further, as we move beyond the 2.5GHz bit rates, these technologies are laying the ground work for an easy transition to 10 GHz optical bit rates per communication lane. At 2.5 GHz you can cover about 24" of FR 4 depending on the quality of the transmitter and receiver pair. But, take the data rate up to 10GHz and the distance is reduced by about 1,000x. If you need to go any practical distance off chip at 10GHz, you must consider optical media.





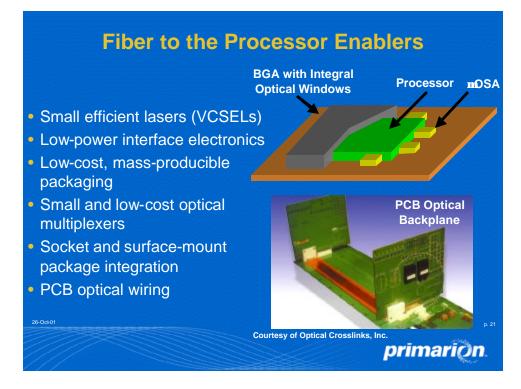
The primary barrier to extensive use of optical medium is the cost of conversions from electrical to optical and back again.

As E/O conversions become very economical, a new design paradigm will likely emerge wherein silicon chips are bused together with fiber/optical wave guides to build very high-performance systems. Silicon will get used where it makes sense but optical will be used for bus transport over distance.

In this new paradigm, systems may be built with switched fabric attached directly at the edge of microprocessors themselves. Thus, many conventional electrical communication bottlenecks will be mitigated or eliminated completely.

So, the Holy Grail is low-cost E/O conversions as this will enable low-cost optical interconnect and busing. This technology is coming to us quickly over the next few years. It will be extended to support CWDM capabilities making it incredibly scaleable in bandwidth as well!





We see enablers for this new paradigm emerging today.

Semiconductor lasers and detectors at 10Gbit/sec are available and being designed into new communication devices. They are incredibly tiny die and will drop in cost dramatically as volume ramps.

Also, low-power interface electronics are being developed and deployed. These are also tiny chips that can move to CMOS integration.

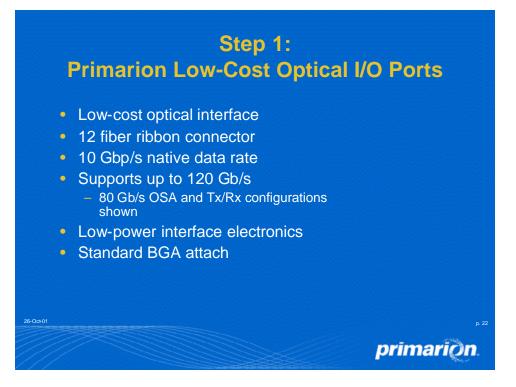
Low cost Plastic Packaging is being developed which will allow tiny optical assemblies to be co-packaged next to complex VLSI devices. I suspect that optical wave guides printed on PCB will emerge in the next five years.

Here's an example of an optical back plane offered by Optical Crosslinks.

All the pieces are coming into place to bring optical interconnect directly to our processors in the future. We call this Fiber to the Processor Technology.

Here the term "Fiber" includes use of optical wave guides!





In this Primarion technology prototype we can get a real feel for how small the transceivers for 12 fiber ribbon interfaces will become. At the top is an optical subassembly (OSA) for 10Gbit/sec bit rate on up to 12 fiber ribbon. It includes an optical port, VCELS, Detectors, Drivers and TIAs.

It would be easy to package this OSA next to a microprocessor in the same package

Below is a complete MT-RJ connectorized transceiver with 10Ge capability including SERDES and CDR and the OSA included.

This technology is designed for very low cost as it supports fully automated assembly and uses semiconductor high-volume packaging technology.

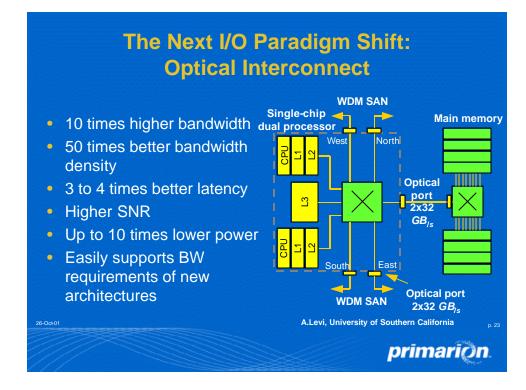
Alignment is ensured to be correct by construction design. This is the kind of technology that will exploit semiconductor cost learning curves. It will be adaptable PCBs with optical wave guides as well.

A conversion from electrical to optical and back again takes only 200ps with this design. This is much less time than it takes to drive an I/O bus line off a .13u CMOS chip.

How might microprocessor system designers scale these to buses and use such technology?

© 2001 Primarion, Inc. Primarion™ and the Primarion logo are trademarks of Primarion, Inc. Other names and brands are the property of their respective owners.





In the future, we can imagine the use of optical interconnect technology in new system paradigms. This is a system concept from Dr. Tony Levi of USC, a former Bell Labs Optical Guru who has written some 180 papers on optical technology and systems. He is a member of Primarion's technical advisory council. His numbers here are very compelling.

In the future, this kind of high-bandwidth switched optical fabric could touch directly to the microprocessor die itself. With a latency of just 200ps to go across an inter-chip optical link, the programming model will appear very much like all the processors in such an optical fabric are literally on one humongous die. Never have we seen a system where "computers can communicate at the same speed they can compute."

This cross-over point marks the inflection point. I can't even imagine the kind of applications that will be enabled with such technology. Can you?

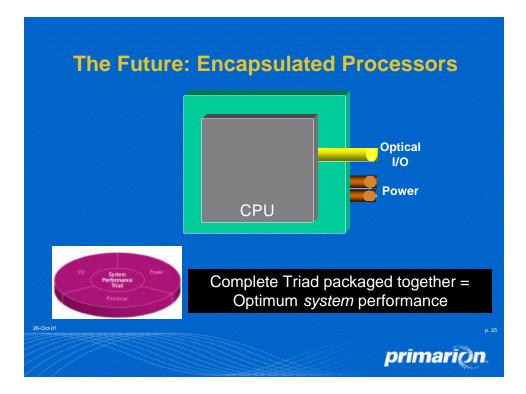
Certainly, it will change the whole industry if it materializes.





OK, let's put the power technology and optical I/O together into one vision, the 'Encapsulated Processor'.





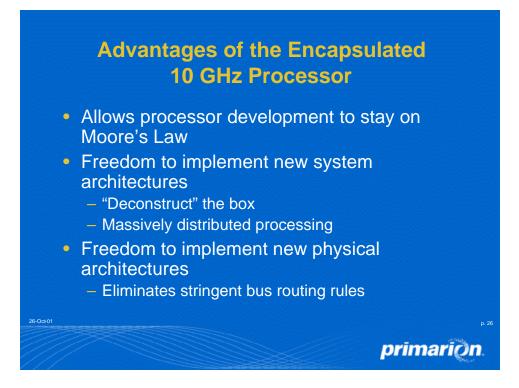
For a moment, try to imagine a 1billion-transistor microprocessor system on a chip running well beyond 10GHz at .5 volt or so.

Clearly, we will need to provide it an environment that is carefully powered and isolated. We will need to use all known techniques to minimize its power. Noise isolation between on-die subsystems may be key. This may necessitate fine grain power domains to prevent unit-to-unit interference. Maybe the floating point unit has it own power delivery so that its high di/dt events do not cause adjacent units to fail by stealing their energy.

Beyond, future 10GHz processor systems will require that power elements and optical interconnect be designed as a system and packaged together for maximum performance, reliability, availability, scalability and manageability. Electrical interconnect running at 10GHz speeds may well be limited to a CM or so on a multi-chip substrate.

In the encapsulated processor, you'll bring in pre-regulated power that then gets carefully conditioned and distributed, and you'll use optical wave guide buses or fiber ribbons to communicate to other subsystems with great SNR and speed and capacity.





So to summarize my views: The Encapsulated processor will bring a whole new degree of freedom allowing us to deconstruct the box in the same way switched fabric Data Centers today are being deconstructed by function. This will allow us to better manage increasing thermal loads, stay on Moore's Law, and build new classes of system architectures that allow...

...computers to communicate at the same speed they can compute!

Thank You!

###