

A 500 MHz 64 bit RISC CPU with 1.5Mbyte on chip Cache

Philip Barnes
Member of the Technical Staff
Fort Collins Microprocessor Lab
VLSI Technology Center
Hewlett-Packard Company
3404 East Harmony Road, Fort Collins, Colorado

Welcome to this presentation on the design and implementation of Hewlett-Packard's most recent 64 bit PA-RISC Microprocessor.

This design involved porting an existing processor core from a 0.5 micron CMOS process into an advanced 0.25 micron process, moving the large L1 caches on to the die and adding additional functionality.

This processor, tape released in the spring of 1998 and has been shipping in systems since late 1998.

Outline

- Program Goals
- Legacy Features
- Major Changes
- Electrical Modeling
- Clocking
- Technology Scaling Issues
- Dual Voltage Memory Bus Interface
- Manufacturing Test Support
- Mask Processing
- Total System Cost Reductions

Today's presentation will cover several areas including an overview of the program goals, a brief description of the legacy processor features and an overview of the major functional changes made to the design. I will follow that up with a discussion of the detailed electrical modeling strategy used to ensure high quality first silicon, an overview of the on-chip clocking strategy and a discussion of some of the more significant technology scaling issues encountered and how they were addressed with circuit redesign. Following that, I will go through the changes made in the main memory bus interface, (my part of the processor), and conclude with a brief discussion of manufacturing and product impacts of this design.

Program Goals

- Schedule Driven - 1999 Product Delivery
- Cost / Footprint / Power Reduction over previous design
- ~2X Incremental performance over existing products
 - Competitive application performance in commercial and technical markets.
 - Target 29 SPECint95
 - Target 37 SPECfp95
 - Full binary compatibility with previous designs

The processor development program was launched with clear goals - to deliver industry leading performance on an aggressive schedule, while reducing the total system cost, power dissipation and system foot print of its predecessor.

As many of you may know, performance leadership is a moving goal that requires balancing performance delivered with point of market entry. Missing targets on either front will leave you behind the competition.

The processor is targeted at both the technical and commercial markets, spanning the product space from the uniprocessor workstation to greater than 32-way scalable shared memory multiprocessors. Results coming in from our system development labs indicate we shall exceed our performance goals with performance in the range of 30 SPECint95 and 50 SPECfp95.

Legacy Functional Features

- 64 bit PA-RISC 2.0 architecture
- Four way superscalar out-of-order execution
- 2 load/store units
- 2 integer ALU's
- 2 integer shift/merge units
- 2 floating point multiply-accumulate (FMAC) units
- 2 floating point divide/square root units
- glueless snoopy MP support

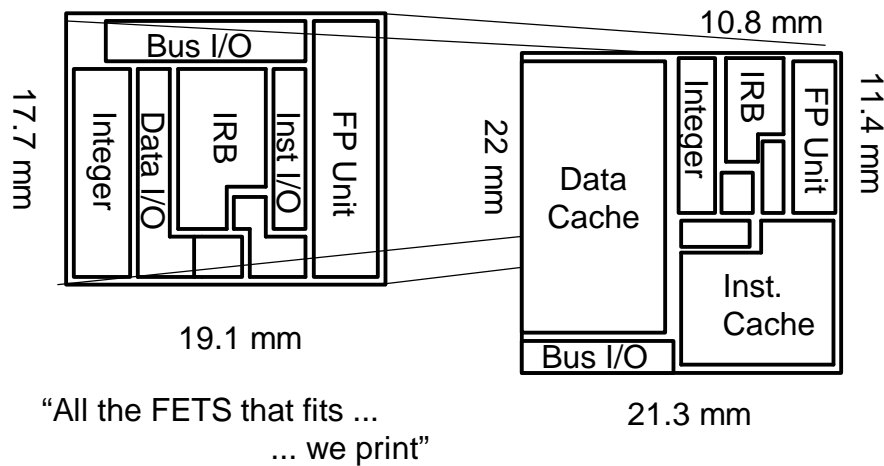
The processor retains the same rich functionality of its 0.5 micron predecessor including an aggressive four issue out of order superscalar core supported by a 56 entry instruction reorder buffer feeding dual integer ALU's, dual load/store pipes, dual shift merge units, dual floating point multiply/accumulate units and dual floating point divide / square root units. Looking beyond uniprocessor performance, the processor fully supports glue-less snoopy symmetric multiprocessing and scalable shared memory multiprocessing.

Major Functional Changes

- 4 way set associative on-chip caches.
 - 1MByte Data, 0.5MByte Instruction
 - 32 byte or 64 byte line size
 - elimination of about 600 I/O's.
- Expansion of Instruction Front End (4x Branch prediction cache 2048 entries, modified algorithms.)
- Increased TLB size from 120 to 160 entries (dual ported, fully associative.)
- 2X bandwidth asynchronous main memory bus interface

Major functional enhancements to the processor include four way set associativity on the L1 data and instruction caches and support for an increased 64 byte cache line size in addition to the legacy 32 bytes. The instruction front end has been enhanced with a four X increase in the number of entries in the branch prediction cache. The dual ported, unified, fully associative translation look aside buffer has been expanded to 160 entries. And finally, the asynchronous main memory bus interface delivers twice the bandwidth of the previous design while allowing bus and core frequencies to be set independently for maximum performance.

Physical Changes

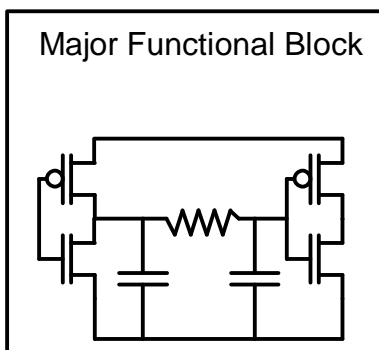


Changes to the CPU core include the removal of 460 I/O pads associated with the off chip cache memory interfaces and relocation of the bus control and I/O circuitry. The Bus control is moved to below the integer datapath, where it is better situated between the data cache and the address reorder buffer portion at the lower left of the instruction reorder buffer. This also puts it in close proximity to the translation look aside buffer and the instruction fetch unit.

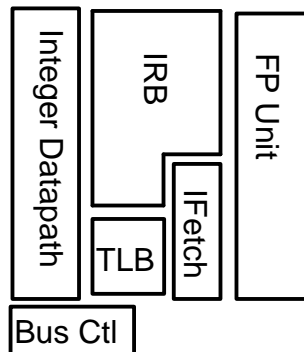
The entire core is shrunk by approximately a factor of two and moved to the upper right hand side of the die, making room for the large L1 data and instruction caches. Moving the L1 cache on chip allows the cache based critical paths to scale with processor frequency while maintaining a 1 to 1 frequency ratio with the core and 3 cycle load use penalty

Electrical Modeling Strategy

Graybox models based on FET's and parasitics generated for all major blocks



Grayboxes combined with Top level route parasitics to catch global issues



The predecessor 0.5 μM design made extensive use of aggressive full custom and structured custom CMOS circuit techniques. Successful migration to 0.25 μM technology could only be accomplished through the use of detailed electrical modeling. All major functional blocks are modeled at the FET level using extracted parasitic capacitance and resistance measurements. These block level models are used stand alone and are also abstracted into “gray box” models for use at the top level. Gray box models retain all the essential path and state information necessary to support static timing analysis at the next higher level of the design. The graybox information, along with extracted resistance and capacitance measurements from the global composition and route support a full static timing analysis of the core.

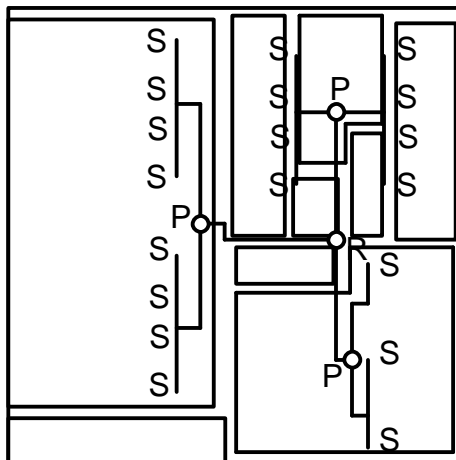
Detailed Electrical Analysis

- Hierarchical FET based model making extensive use of extracted parasitic resistance and capacitance data.
- Hierarchical Static Timing analysis taken to the full chip level to ensure all critical paths run at speed.
- Electrical robustness checking tool analyzes major functional blocks looking for electrical margin issues. Runs background simulations to gather more data on such issues as charge sharing and capacitive coupling.

In addition to supporting detailed static timing analysis, the electrical models are used by an in house developed “electrical robustness” checker. This rule based tool set looks through the design for instances of known potential problems. Over 100 checks are made on each major functional block. While tools may never be a substitute for good engineering judgement, they do provide a solid safety net and allow large expanses of custom circuitry to be covered quickly. By pointing out design anomalies for further detailed inspection by the circuit engineer we avoid common pitfalls while raising institutional awareness of some hard learned lessons from the past. After running the tools we often have both a better circuit and a better engineer, even if the two do not always agree.

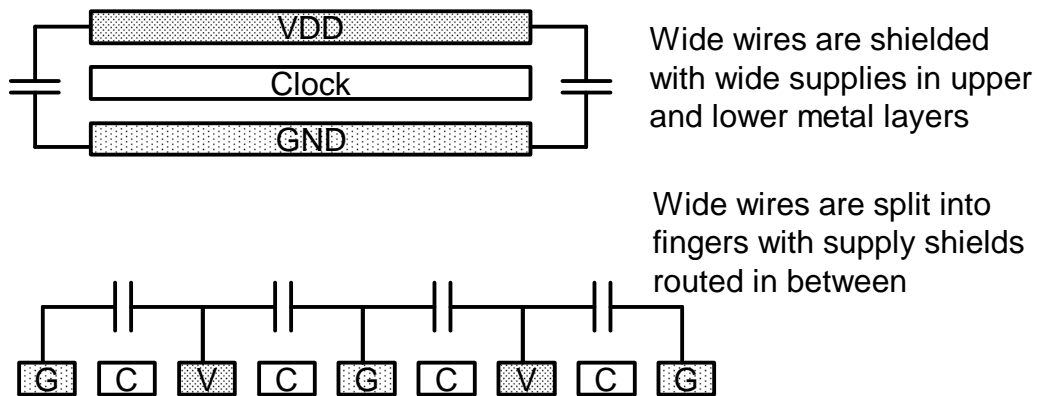
Clocking

- 1 Receiver/Buffer (R)
- 3 Primary Buffers (P)
- 19 Secondary Buffers (S)
- Balanced Tree distribution
- Secondary Buffer outputs all tied together to form CK net
- CK locally qualified to generate Phase clocks for logic



This design retains and extends the open loop clock distribution used in its predecessor. At speed differential clock inputs are received at a central receiver just below the core. True and complimentary outputs are routed out to three primary buffers - one in the center of the core, one in the data cache and one in the instruction cache. Each of these primary buffers fan out in balanced tree structures to the secondary buffers (19 in total). These secondary buffers all connect to one globally distributed “CK” clock net. This net is tapped by local clock load buffering circuits (called clock gaters in the local vernacular) to generate qualified phase clocks for the functional circuitry. The clock network is extensively analyzed with field and circuit simulators to understand the transmission line effects. The final analysis shows less than 80 picoseconds of skew across the entire CK net, with local skew being more in the 25 picosecond range.

Two Clock Wiring Strategies



This slide shows two routing strategies employed in the clock distribution.

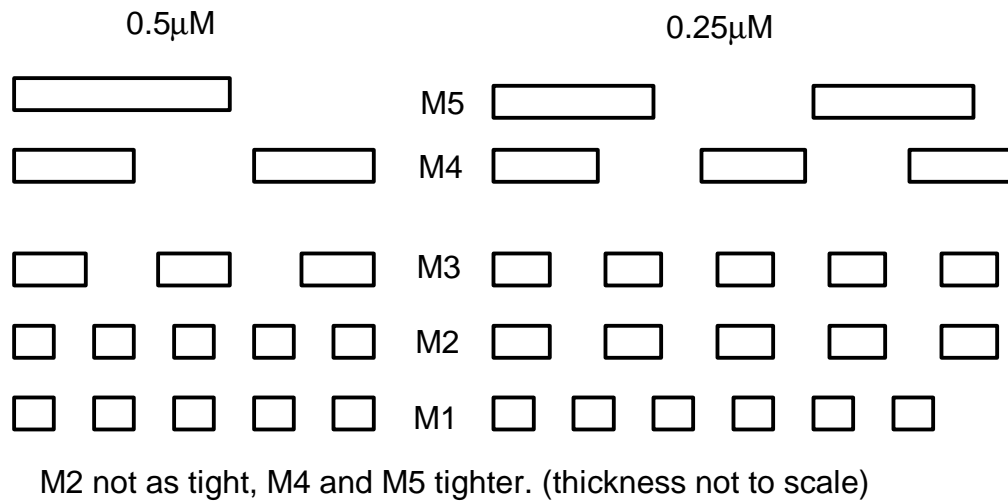
These strategies keep the signal and return current paths close by, thereby addressing signal integrity problems, reducing inductance effects and lowering AC resistance to near the calculated DC resistance. The first scheme, employing wire widths up to 80 microns, is used when metal is available to afford such an approach. The second approach is used when distributing the global clocks on a single layer. Both schemes make use of the extensive on chip bypassing capacitance between the supplies so that high frequency return currents may pass through both supply rails.

IC Process Migration

Feature	0.5 μM	0.25 μM
Metal 1 Pitch (μM)	1.4	0.64
Metal 2 Pitch (μM)	1.4	0.93
Metal 3 Pitch (μM)	2.4	0.93
Metal 4 Pitch (μM)	3.2	1.6
Metal 5 Pitch (μM)	5.0	2.56
Leff (μM)	0.28	0.14
Tox (Angstroms)	80	40
Supply Voltage	3.3	2.0
Frequency (MHz)	240	500

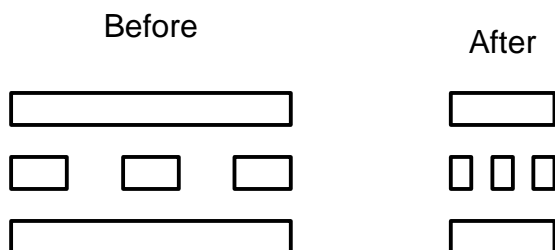
This table offers a side by side comparison between the predecessor's half micron IC process and this processor's quarter micron technology. You will notice that in the half micron process, metal 1 and metal 2 had the same pitch, with metal 3 being more coarse than metal2. In the new quarter micron technology, metal 1 has a relatively tighter pitch, whereas metal 2 is relatively coarse - comparable to the metal 3 pitch. These changes in wiring pitch offered compositional challenges to wire limited designs. In addition to pitch changes, resistance was as significant factor. The quarter micron IC process's tight metal one pitches came at the expense of significantly increased resistance.

Relative Metal Pitch Changes



This slide attempts to illustrate graphically the relative pitch changes that occurred. The vertical dimensions are not drawn to scale. The relatively tighter pitch in the metal five layer afforded opportunity to route more signals in this upper layer to address critical timing paths in longer routes.

Interconnect Scaling Issues



Horizontal shrink is greater than vertical shrink, resulting in an increase in lateral capacitance with respect to vertical (plate) capacitance. Requires greater attention to cross-coupling and increased noise margin on receivers for long routes

Another interesting effect occurs in the wiring as the design is shrunk. The new quarter micron process pushes wiring pitches to the point where the wires may be taller than they are wide. While this allows for greater flexibility in routing, it also offers greater risk of coupling problems between signals and their nearest neighbors. While most ratioed static CMOS circuits have adequate noise margin, pass fet logic and ratio-less dynamic gates receiving signals from long routes are at risk and require more careful attention.

Transistor Scaling Issues

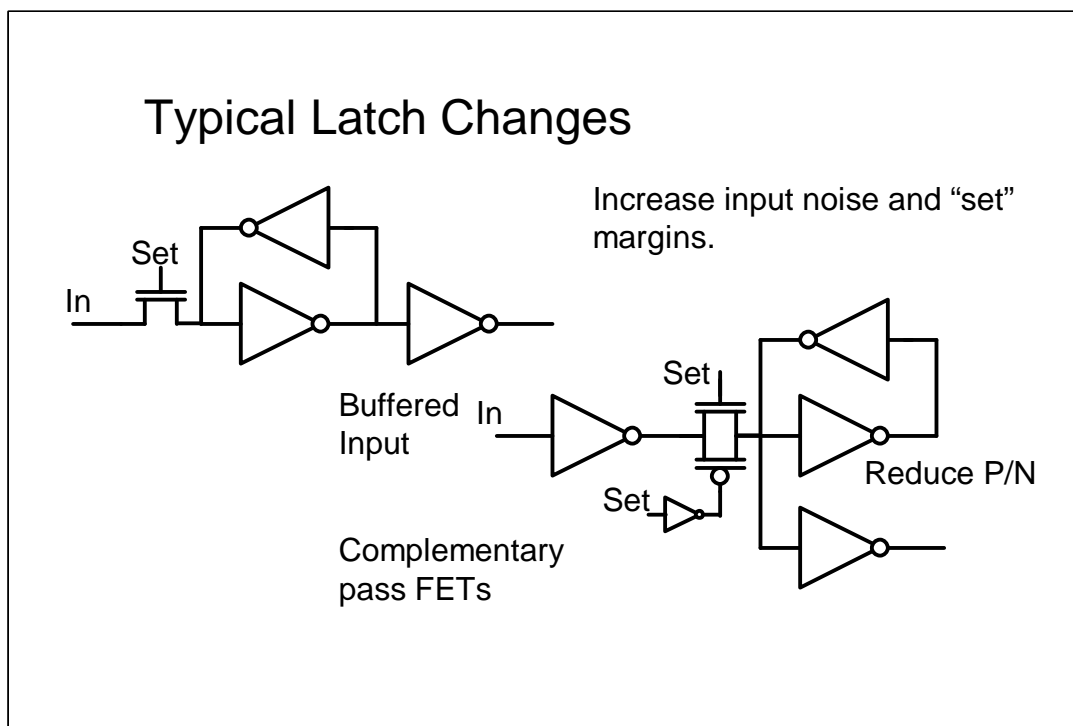
- Transistor threshold voltage (V_t) is a larger percentage of supply. Significantly degrades pass NFET logic.
- Source and drain capacitance increases more than gate capacitance. Creates additional charge sharing concerns.
- PFET performance increases more than NFET performance, resulting in a change in gate ratioing and an upward shift in gate trip points.

As you can imagine, our circuit designers were rather pleased to be turned loose with the fast new quarter micron FETs. However, the new FETs were not without their challenges.

In comparison to the half micron process, the transistor thresholds are a greater fraction of the supply voltage. This has a direct negative impact on pass transistor steering logic prevalent in many of our latch designs.

Increases in the relative magnitude of source and drain capacitance with respect to gate capacitance aggravated the ever present charge sharing problems inherent in dynamic logic designs.

And lastly, the PFET performance increases more than the NFET performance, resulting in an upward shift of ratioed logic trip points.

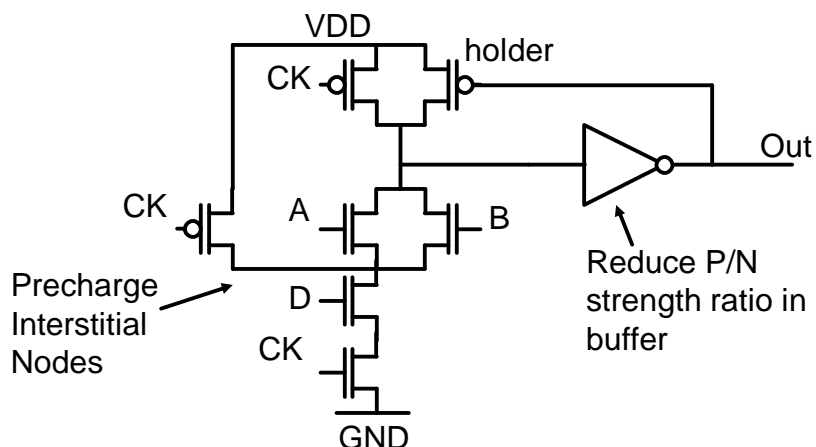


This slide illustrates the classic transparent latch “make over” applied extensively throughout the design. First, the input is buffered to provide a locally referenced logic level to the pass fet logic. (Adverse line to line coupling can cause the input signal in the original design to be booted below the local ground and cause loss of a stored ‘1’ even when the Set signal is low.)

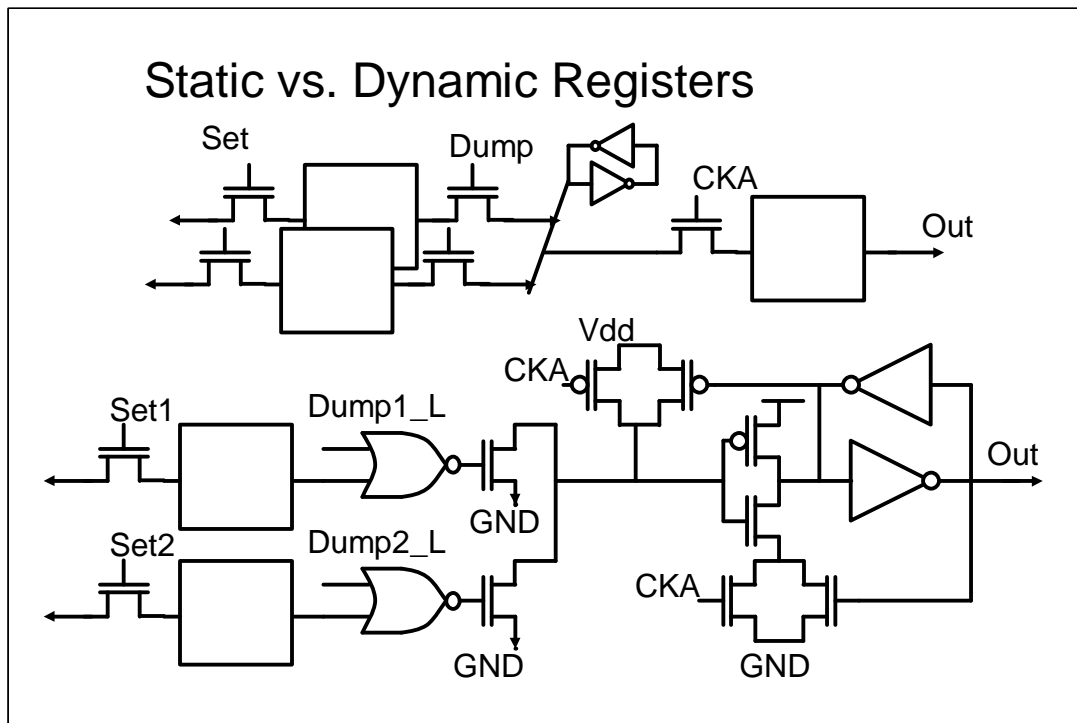
The locally buffered input also provides a known input drive strength and more constant capacitive load to the previous driving stage.

The complimentary pass FETs provide for a higher voltage ‘1’ value on the storage node, assuring that the storage node flips as expected. Reducing the P to N ratio on the forward inverter also aids the AC performance of the latch.

Domino Gate Changes due to increased Source/Drain Capacitance



This slide illustrates changes to a typical dynamic domino gate overcome the transistor scaling concerns. Greater use of prechargers to charge up interstitial nodes helps to offset the increased charge sharing potential. PFET holders are beefed up to account for increased NFET leakage. And lastly, the P to N ratio of the buffer is backed down slightly. All of these changes provide for a more robust circuit at the expense of a bit of performance. Library circuits are given the full robustness treatment, whereas custom applications are left up to the engineer's judgement.



The circuits on this slide help to illustrate the trade-offs made in adjusting to the new process. Both circuits represent valid options for register transfer from multiple sources into a destination register. In the top circuit, we have a widely used static implementation, where data from the source register is conditionally dumped onto a tristate bus, from which it can be loaded into a destination register. The bus is equipped with a bus holder to ensure that it does not float. While this scheme looks safe (and static) it is actually quite risky as transistor thresholds go up. After the direct shrink, many circuits of this type had difficulty flipping the destination register due to degraded “one” levels and the need to fight two feedback circuits - one in the bus holder and one in the destination register.

The circuit below illustrates a robust alternative. While at first glance it appears more complicated and perhaps riskier due to its dynamic nature, its benefits far outweigh the risks. In this scheme, the bus is precharged high while CKA is low. The bus evaluates when CKA goes high. At this time, one of the logic low “DUMP_L” signals goes low, enabling the associated pull down NFET to discharge the bus should the source register be holding a low value.

Variants of this approach are widely applied in large fan-in multiplexor applications in place of pass-fet steering logic.

125 MHz Memory I/O Bus

- 2X Data Transfer Mode
- 3.3 V Series terminated mode - legacy, bring up.
 - receiver protection circuits
 - Analog biased Cascode output stage
 - Controlled impedance push/pull.
 - Level shifting PFET pre-driver
- 1.5 V open drain mode - new systems.
 - Controlled impedance pull-up termination.
 - Variable slew rate pull down drivers.

Several enhancements were made to the main memory and I/O bus to ensure the core has adequate data and instruction bandwidth. The bus enhancements were made while ensuring full backward compatibility with existing product bus structures.

The primary functional enhancement was the addition of a 2X data mode, allowing 64 byte cache line transfers to complete in the same amount of time taken by the legacy 32 byte cache line transfers. A second functional change allows the bus to run asynchronously from the processor core - this allows each part of the system to run at its maximum rate free from any ratio limitations.

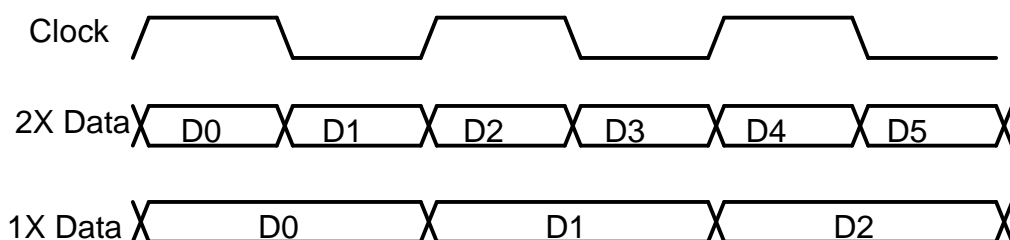
To support this higher speed signaling protocol, we opted to move from the legacy 3.3V push/pull series terminated bus topology to a 1.5 volt parallel terminated open drain topology.

The new 1.5V topology better supports incident wave switching and is more accommodating to the decreasing supply voltages used in deep submicron IC processes. All new products are being designed to the 1.5V standard.

By maintaining backward compatibility to legacy bus chip sets, we were able to make extensive use of existing turn-on systems and could debug the processor silicon independently from the new bus chipset turn on effort. An upside of this strategy was the capability to take advantage of extremely high quality processor silicon early in the turn on effort to support early introduction of the new processor in some legacy product configurations.

2X Data Transfer Mode

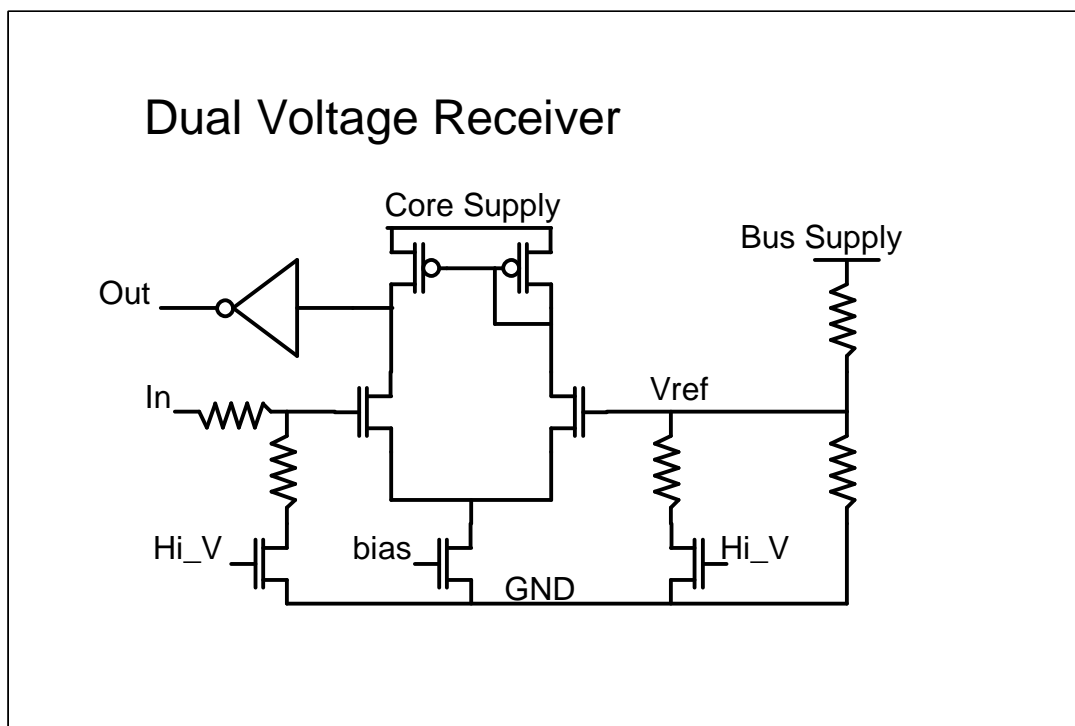
- Both 32 byte and 64 byte cache line support
 - Uses four 2X Data cycles on 64 byte line transfers.
 - Uses four 1X Data cycles on 32 byte line transfers.



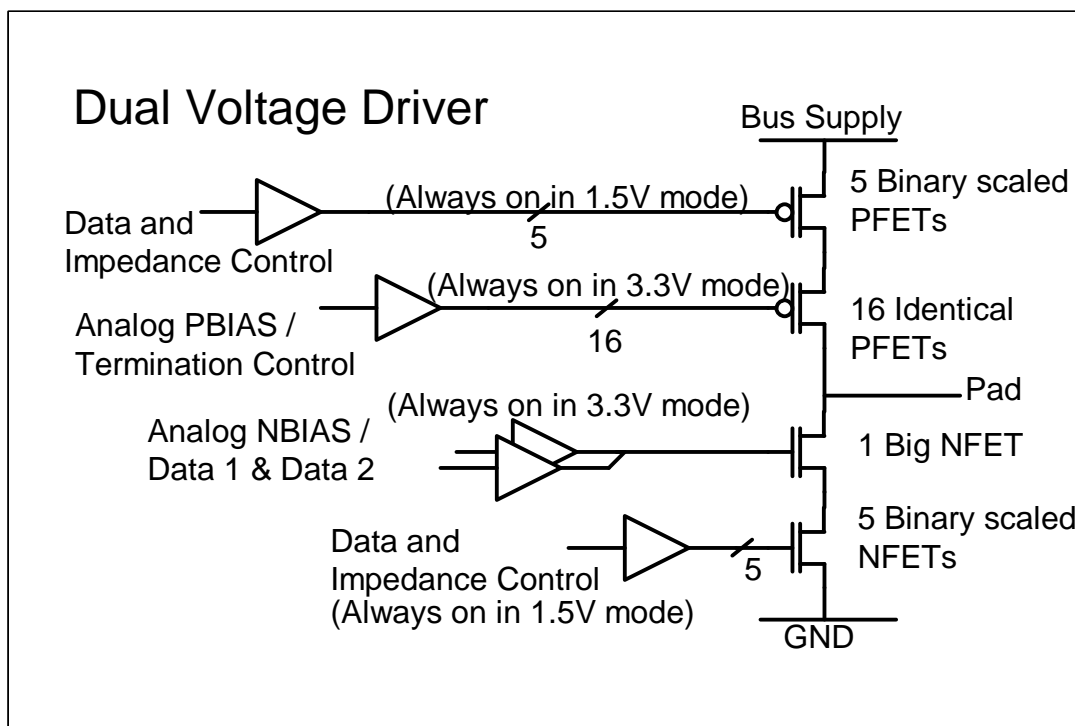
This timing diagram illustrates the 1X and 2X data transfer modes. Note how in 2X mode, data is transferred on both the rising and falling edges of the bus clock. Address cycles and non-cache line data cycles are transferred in 1X mode, driving new values on the bus only on the rising edge of the bus clock.

This approach was taken to allow the bus interface to be reasonably ignorant of the 32/64 byte line mode differences, in terms of bus arbitration protocol and bus cycle consumption while maintaining an addressing and data bandwidth match with both the new and legacy bus chipsets.

The bus arbitration and ownership algorithms support the conditional insertion of dead states on ownership transfer, allowing the processor to accommodate a wide variety of bus sizes and loading options.



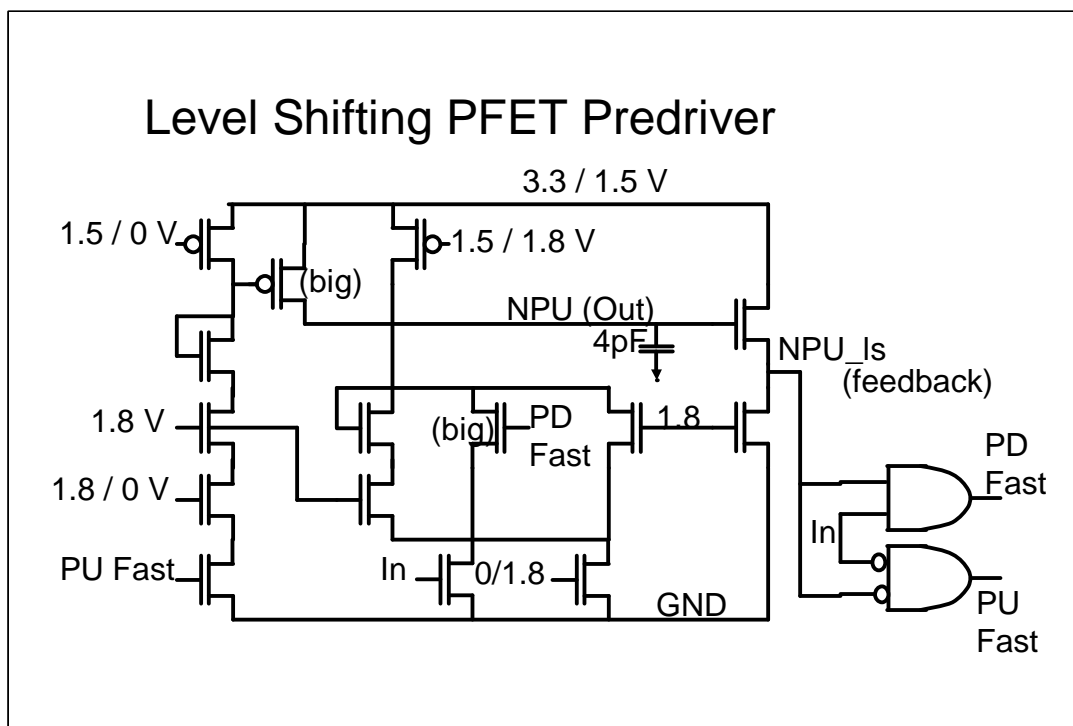
This simplified receiver schematic illustrates how one receiver is configured to operate at two significantly different bus voltages. In both cases, the receiver is operating from the core supply voltage, while the on-chip reference is derived off the bus voltage. By including switched voltage dividers on both the input and the reference signal, operated by the HI_V signal, the signals seen by the receiver FETs can be level shifted down to a more accommodating operating range during 3.3 volt operation.



This diagram illustrates the basic scheme used in the off-chip driver to support the 3.3V push/pull and 1.5V open drain operation.

When operating as a 3.3V push pull driver, the inner FETs of the cascode output stage are biased with intermediate analog values. These protection FETs shield the outer FET's from the full swing of the I/O pad, since the 3.3V swings would be damaging to the 40 angstrom oxide quarter micron devices. The outer drive FETs are arranged in binary weighted fingers to support process/voltage/temperature (PVT) compensation of the drive strength. This helps control reflections and signal integrity on the bus. The gates of the outer NFETs are driven between ground and core VDD, while the gates of the outer PFETs are driven between the bus high voltage (3.3V) and an intermediate voltage of approximately BUS VOLTAGE minus CORE VOLTAGE. This prevents high voltage oxide stress in these devices.

When operated as a 1.5 volt open drain driver, the outer FETs in the cascode are turned on, while the inner FETs are switched for data signalling. The inner NFET is fed by dual tristating predrivers, one for each phase of data in the 2X mode. The inner PFET is fingered in sixteen segments to support process/voltage/temperature compensation of the pull-up termination.



This is a simplified schematic of one of the level shifting predrivers used to control the outer PFETs in the 3.3V mode of operation. The goal of this circuit is to swing the output signal (NPU) between 3.3 volts and a lower intermediate voltage of approximately 1.5V. This circuit drives a substantial load (that being the gate load of up to half the pull-up PFETs - roughly 1000 microns of FET).

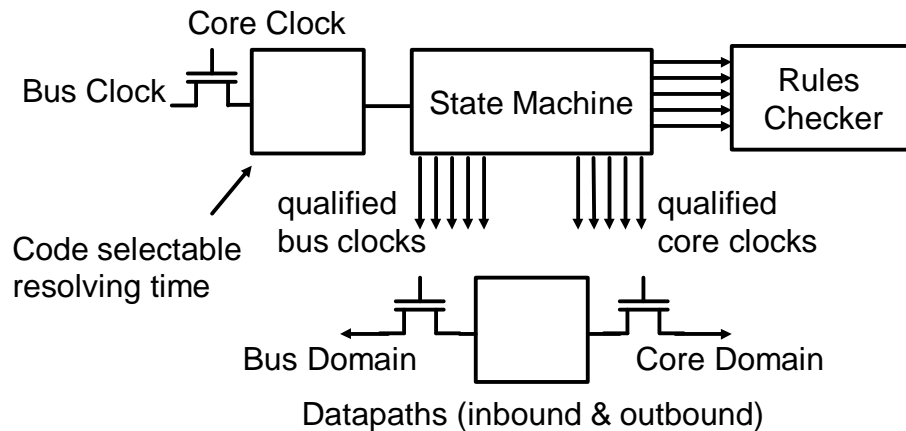
The basic approach is to use an inverting common source circuit driven with the IN signal. This circuit uses cascode protection of the driving NFET and a DC biased PFET load. This circuit provides adequate DC levels, but burns excessive current if sized for reasonable speed driving the 4pF load. To improve switching time while reducing DC current, the predriver employs large booster FETs (labelled “big” in the diagram.) These booster FETs are controlled via a feedback loop that combines a level shifted version of the NPU output signal with the Input signal to generate PULL DOWN FAST (PD_FAST) and PULL UP FAST (PU_FAST).

The feedback logic is such that if IN is high, and the output is still, high, the booster pulldown NFET is turned on, and the NPU output node slews down fast. When IN goes low, if the NPU output node is still low, then PU_FAST goes high resulting the big PULL UP PFET turning on. Once NPU has been pulled up to a high value, PU_FAST goes low again.

Additional circuitry in the predriver ensures that NPU is pulled all the way to GND when the predriver is operating as part of the 1.5V open drain scheme.

Asynchronous Bus Interface

- reactive synchronizer operates in parallel to data path



In order to increase the flexibility in processor and bus operating frequencies, while eliminating the need for tight control in routing the at speed processor clocks on the circuit, we opted for an asynchronous bus interface. This reactive synchronizer design involves sampling the bus clock with the core clock and feeding the results into a core clock domain state machine. The state machine then generates a series of qualified bus and core clocks to orchestrate data transfer through both the inbound and outbound datapaths. The datapaths include parallel staging latches at each boundary that can be multiplexed into the main bus.

The sampling circuitry has code selectable resolving time (through cascaded latch stages) allowing the processor to power up in a conservative mode and shift to a more aggressive lower latency mode under firmware control. The overall latency achieved in practice compares favorably to a fully synchronous design while providing timing flexibility in the bus to core interface.

As an added sanity check, a rules checker operates in parallel to the synchronizer to ensure that certain basic behaviors are adhered to. Rules include things like “make sure every bus cycle is transferred to the core clock domain once and only once.”, “make sure every core clock cycle is transferred once and only once to the bus clock domain.” If the checker detects a violation, it signals a high priority machine check to ensure the error does not result in silent data corruption.

Manufacturing Test Support

- eliminating all core CK I/O's enables speed testing at wafer sort without too much tester black magic.
- Cache Built In Self Test (BIST) and redundancy keeps test time low and yields high.
- Pseudo-Boundary scan at core interface supports core CK domain "pin vectors" for scan based functional testing.
- At speed multistep scan based interconnect testing supported in I/O's

This design offers significant advantages over its predecessor when it comes to manufacturing and test support. Significant benefit is received by eliminating all core speed I/O from the die. The previous design required close to 500 hundred at speed tester channels to provide data and assertions to the wide L1 cache interfaces at manufacturing test. This was prohibitively expensive at wafer sort, requiring such broadside speed testing to be performed after packaging. This design has only the bus interface to drive from the tester, allowing at speed functional testing to be conducted at wafer sort.

In addition, extensive built in self test (BIST) is used to cover the SRAM arrays. These features combine to significantly reduce the test time and test costs of this processor from its predecessor despite the increase in device complexity.

Additional test features added to the processor include a pseudo boundary scan at the clock domain crossing boundary, allowing diagnostic scan vectors to exercise the core in without imposing constraints on the bus clock frequency. And lastly, AC interconnect testing is supported in the I/O circuits through additional scan registers in the boundary scan.

Mask Processing

- tape out concerns addressed in parallel with processor design. (Staffed from day 1)
- 40 GB design processed in days. (Previous process would have taken months.)

Processing a design of this size required significant innovation in dataset processing. Early estimates using our previous methodology indicated it would take us several months to DRC the 40 gigabyte database, generate the 2 gigabytes of GDS data and fracture the mask set. Fracturing from our first mock tape release took over four weeks before we killed the job and started searching for a better way.

After modifying the tool flows to take advantage of as much parallelism as we could reasonably find and upgrading the hardware in the critical paths, we were able to complete the entire tape out process from a fully invalid data base to fractured mask data in under two weeks. Using hierarchical DRC and a partially valid database (such as the case with chip spins,) these times improve significantly.

Total System Cost Reduction

- Total system cost reduction was a major goal.
- Inexpensive 544 pin LGA package
- Fewer than 100 system interface signals to route on board (versus 768 in previous design)
- Power reduced by 50 percent while performance doubled and CPU cost reduced by more than 75 percent.

And in conclusion, I would like to point out the overall impact on system design of this processor in comparison to its predecessor. The elimination of two 128 bit wide off-chip cache ports and their associated address and tag pins allow the processor to be packaged in a reasonable inexpensive 544 pin land grid array package. With fewer than 100 system interface signals to route and no off chip cache wiring, the simplified system interface translates directly to a lower cost circuit board design for uniprocessor applications and greater processor packing density for multiprocessors. CPU power dissipation is reduced by over fifty percent while performance is doubled, providing a four fold increase in performance per watt. This is accompanied by a seventy five percent reduction in CPU cost resulting from the elimination of all off-chip high speed cache SRAM and the lower cost packaging. This design has successfully leveraged the tremendous intellectual property investment that went into designing, implementing and verifying its predecessor's aggressive out-of-order superscalar core and allowed that investment to be amortized over a new generation of processors that will continue to provide leadership RISC performance well into the future.

Thank you for your attention and interest in our design. At this point I would also like to thank my fellow engineers and managers at Hewlett-Packard's Fort Collins Microprocessor Lab for their contributions to this design and to this presentation.