
Exploring potential performance of wide PowerPC-based superscalar processors

J. H. Moreno
RISC Architecture and Analysis
IBM Thomas J. Watson Research Center

Topics

- Wide-issue out-of-order superscalar processor model
- Simulation environment
- Evaluation of potential performance

- On-going activity
 - Initiated early 1997
 - Basis for researching/evaluating new topics

- Methodology, infrastructure
- Trends among results instead of absolute values

Team

Mayan Moudgill
John-David Wellman
Jaime Moreno
Pradip Bose

Acknowledgements

Erik Altman, Al Chang, Dan Prener
Eric Kronstadt, Louise Trevillyan
Dave Meltzer, Chuck Moore, Mary Mosher

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Why yet another superscalar processor model?

- **Superscalar processors continue dominating the field**
 - No apparent likelihood of ending superscalar paradigm in near future
 - Continuing improvements in features and capabilities
 - Certain aspects getting easier due to number of transistors available
 - Existing programs (binary compatibility)
- **Need for evaluating new implementation challenges**
 - High frequency objectives:
 - few levels of logic per pipeline stage, relatively long wires
 - New structures and algorithms
- **Need to understand**
 - Potentials
 - Impact of new ideas
 - Impact of changes in characteristics of workloads

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Desirable capabilities in an "early modeling" environment

- Ability to assess impact of various features
 - performance
 - suitability for given contexts
 - client (workstation), server
 - scientific vs. commercial workloads

- Infrastructure to study contexts/requirements
 - performance trends
 - applications
 - microarchitecture

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Goals

- Ability for understanding the limits and potential of
 - out-of-order, speculative, highly concurrent superscalar processors
 - explore alternative features

- Do not focus on specific implementations

- ➔ *Get understanding for the future*

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Limitations in existing tools

- **Flexibility for modifying the microarchitecture**
 - models usually reflect a specific microarchitecture
- **Modeling aggressive out-of-order features**
 - beyond current state-of-the-art in implementations
- **Fast simulation capabilities**
 - millions of processor cycles/hour
- **PowerPC-based**
 - ...
- **Modeling of instructions executed speculatively**
 - usually not available

Exploring potential performance of wide PowerPC-based superscalar processors

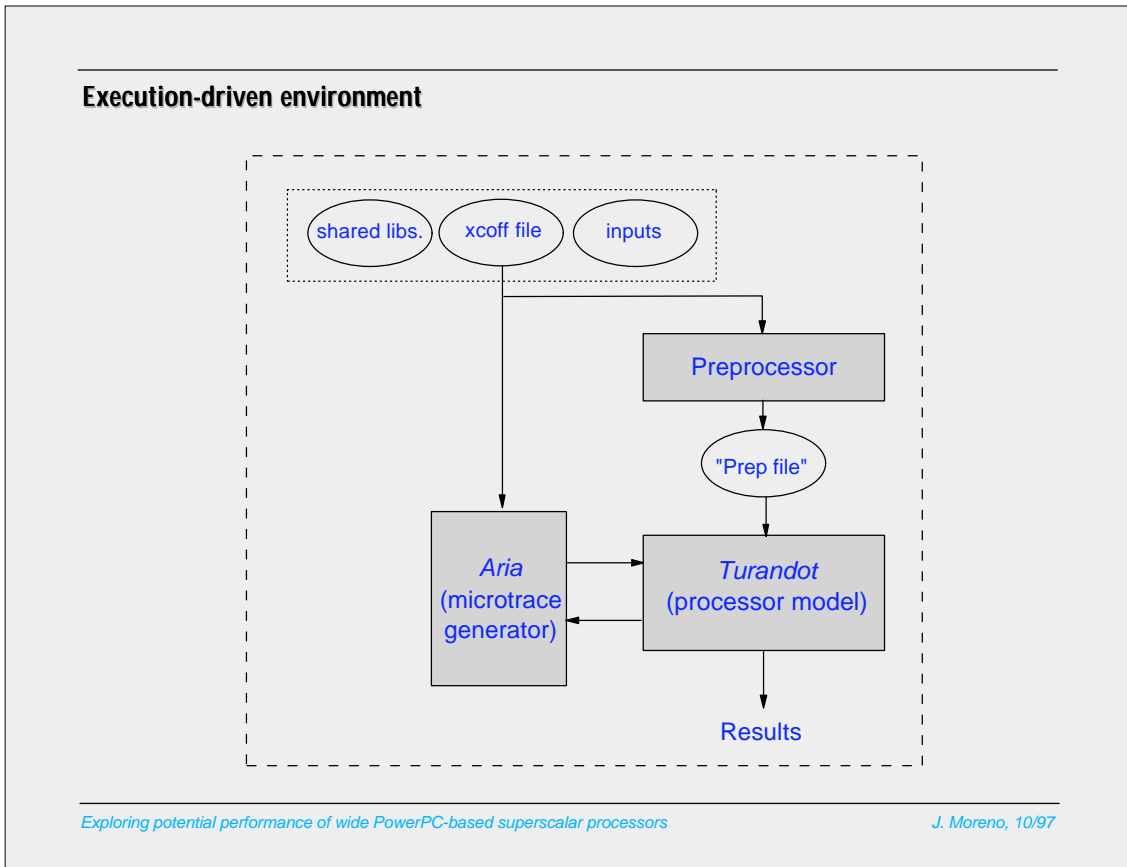
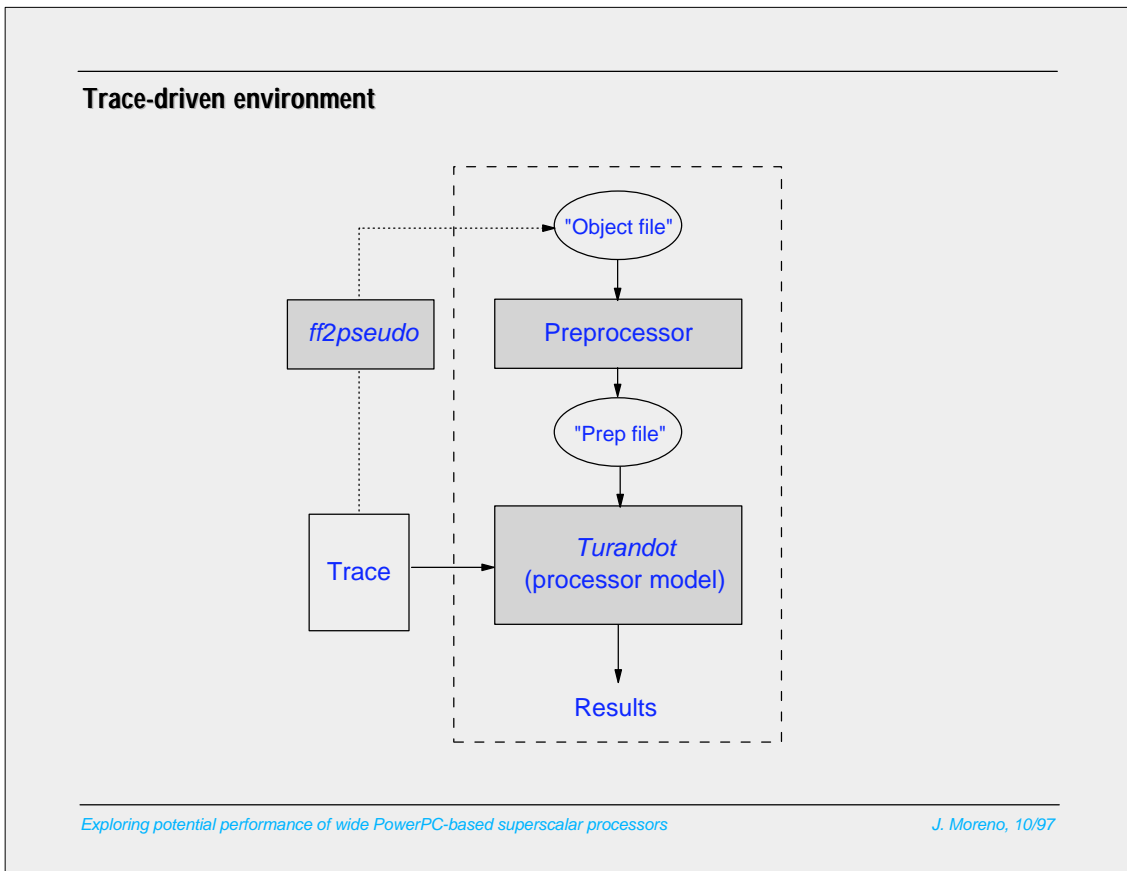
J. Moreno, 10/97

The MET: Microarchitecture Exploration Toolset

- **Collection of tools for exploration of microarchitecture features**
 - Aria
 - Turandot
 - LeProf
 - ...
- **Trace-driven and execution-driven tools**
- **Fast simulation: ~100 Mcycles/hour**
- **Intended to support early exploration of processor organizations**
 - detailed model of generalized pipeline
 - trends among results instead of their magnitudes

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

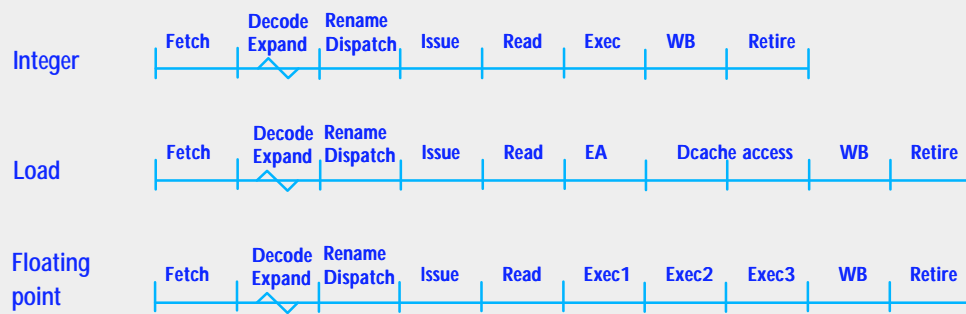


Processor organization

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Pipeline stages



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Other features of the processor model

- Extensive predecoding of input program
- Programming for low simulation overhead
 - macros instead of function calls
 - no pointer-linked data structures
 - single procedure
 - few branches
 - novel cache emulation technique
- No run-time parameters; recompilation required
- ...

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Parameters in model

- Approx. 100 parameters
 - number/size resources
 - enable(disable) features
 - select among alternative policies

Model validation approach

- Derived from processor validation techniques
- Extensive cross-checking of data collected

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Aria, a "micro-trace" generation engine

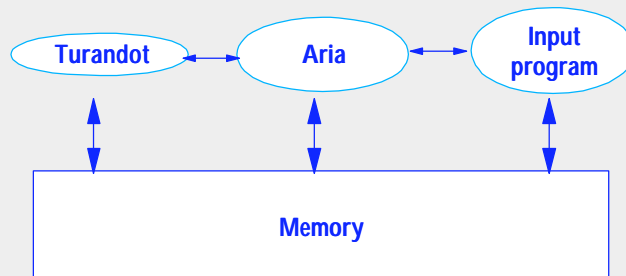
- Uses principles developed for binary translation
 - at first-time execution, translate basic block into instrumented version
 - same functionality
 - generates trace of execution
 - captures dynamically-linked libraries
- Two versions of each basic block
 - "normal" version: executed under normal conditions
 - "not-taken" version: executed in speculative manner
 - mispredicted paths
 - no changes to state of the program (memory)
 - load instructions are guarded (no segmentation faults)
 - illegal instructions replaced by no-ops
- Capable of emulating execution of instructions not in the ISA
 - translated into sequence of existing instructions
 - trace includes the non-architected instruction and its effects

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Aria/Turandot interaction

- Processor model and micro-tracing engine running concurrently
- Processor model requests trace for each basic block
 - normal or not-taken version
 - model provides state of the program to tracing engine (register state)
 - memory shared among model and tracing engine



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Exploration space (in this presentation)

- Issue policy
 - class-order
 - out-of-order
- Width
 - 4, 8, 12
- Cache size
 - 64K/2M, 128K/4M, infinite
- Branch prediction
 - simple, perfect

- Just some examples of exploration possibilities

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Workloads

Commercial	TPCC	PowerPC DB2 trace	170M
	SLIQ	Reduced version of data mining algorithm in Intelligent Miner.	350M
SPECint95	GCC95	Gnu C Compiler (program cc1)	500M*
	compress95	Compression algorithm	38M
	go	Game of Go	420M
	m88ksim	Motorola 88000 simulator	110M
Technical	TPP	Gaussian Elimination (1000x1000)	170M
	sparsemv	Sparse matrix vector multiplication	198M
Misc.	perl	Pattern Extractor/Recognizer	12M
	lex	Lexical Analyzer	100M
	yacc	Yet another compiler compiler	50M

> 2B

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Exploration dimensions

Widths	Units	Ports	Queues	Physical registers
Fetch/Dispatch/Retire	FX/FP/LS/BR	Data cache and TLB	Issue/Retire/IBuf	GPR/FPR/CCR/SPR
4/4/6	3/2/2/2	2	20(12)/128/24	80/80/32/64
8/8/12	6/4/4/4	4	40/160/48	128/128/64/96
12/12/16	8/4/6/4	6	60/160/72	128/128/64/96

Issue policy	Width	Cache size	Branch prediction
	Fetch/Dispatch/Retire	L1-I, L1-D, L2	
Class-order	4/4/6	64K, 64K, 2M	8192 entry BHT, 4096 BTAC
Out-of-order	8/8/12	128K, 128K, 4M	Perfect
	12/12/16	128K, 128K, Inf	
		Inf, Inf, Inf	

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

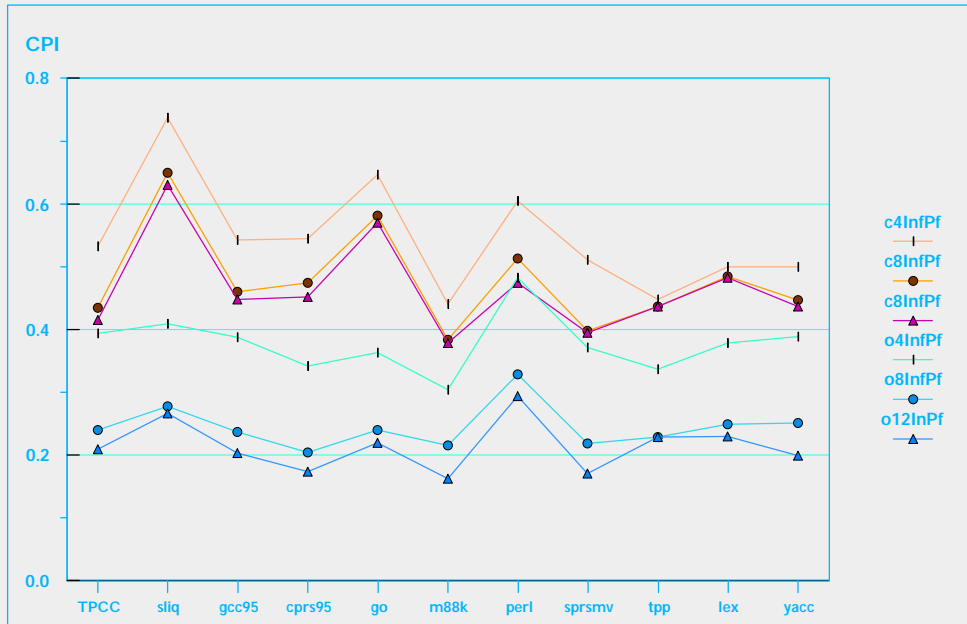
Other parameters (examples)

Maximum intrs. in flight	160	Miss queue, cast-out queue (entries)	8
I-prefetch buffer latency (cycles)	1	Store queue, reorder buffer (entries)	31
I-prefetch buffer (entries)	4	Cast-out overhead (cycles)	5
Latency from L2 to I-prefetch buffer at I-prefetch buffer hit (cycles)	8	D/I-TLBs (entries)	128
Latency from L2 to I-prefetch buffer at I-prefetch buffer miss, after L1 reload (cycles)	4	D/I-TLBs miss penalty (cycles)	4
BTAC (entries)	4096	TLB2 (entries)	1024
Next fetch address misprediction penalty	2	TLB2 miss penalty (cycles)	40
LR stack size (entries)	32	L1-I/D, L2-cache line size (bytes)	128
Branch history table (entries)	8192	L1-I/D cache miss penalty (cycles)	8.7
Page size (bytes)	4096	L2 cache miss penalty (cycles)	40

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

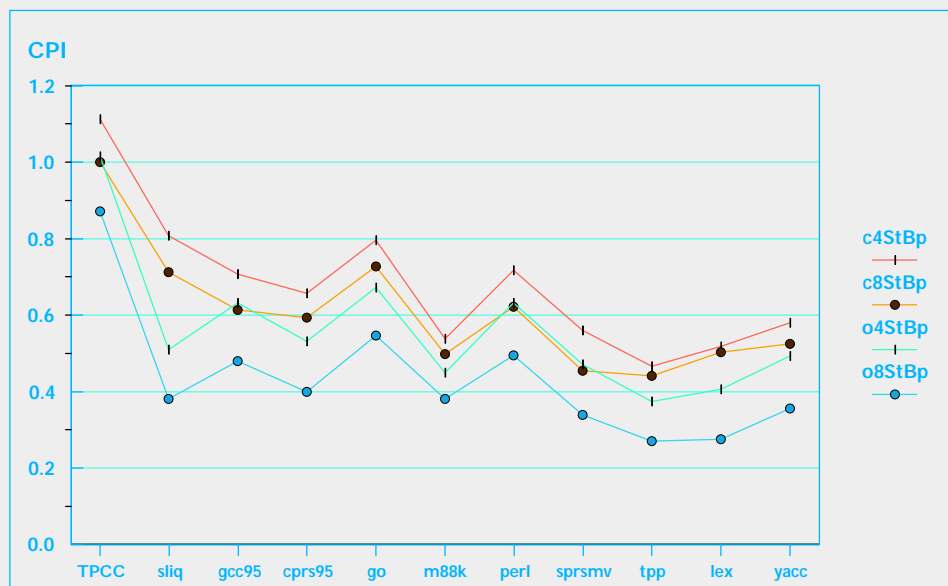
CPI with infinite cache and perfect branch prediction



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

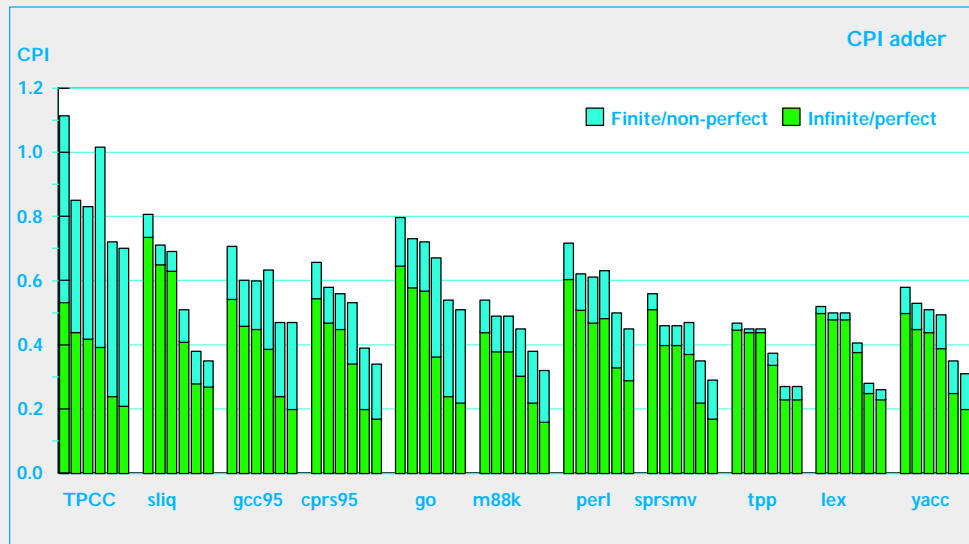
CPI with finite cache and branch predictor



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

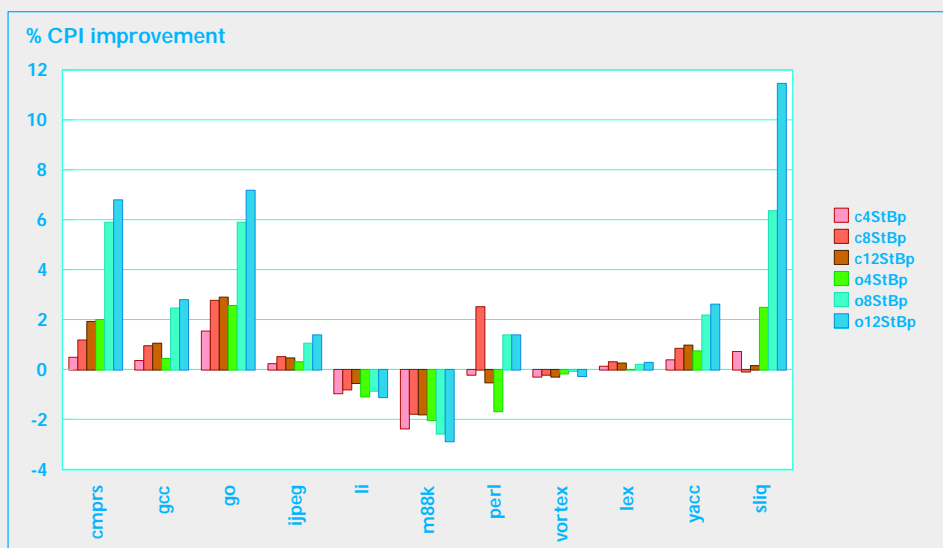
CPI adder



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Effects of instructions from mispredicted paths

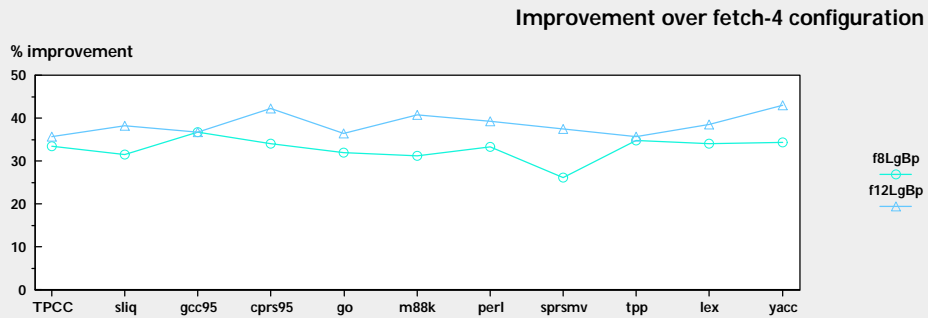


Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Observations

- Starting from fetch-4 configuration, there is "room to grow" by
 - Adding more units
 - Enlarging caches
 - Improving branch prediction
- More leverage in out-of-order organizations than in-order organizations
- Mispredicted paths might actually improve performance



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Evaluation of an OLTP workload

- Trace-driven instead of execution driven
 - difficulties in tracing OS-intensive applications
 - trace allows reproducibility of results
- Limitations in trace-driven evaluation
 - sample size
 - no mispredicted instructions/addresses

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Workload: PowerPC 601 trace

Length	172 M instructions, user and kernel space
Branch instructions	18.9 %
Branches taken	44.3 %
Instrs. in kernel space	22.1 %
Memory access instructions	34.8 %
Load/store multiple instructions	1.6 %
String instructions	1.4 %
Load/store w/update instrs.	1.7 %
Average block size	5.3 instrs.

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

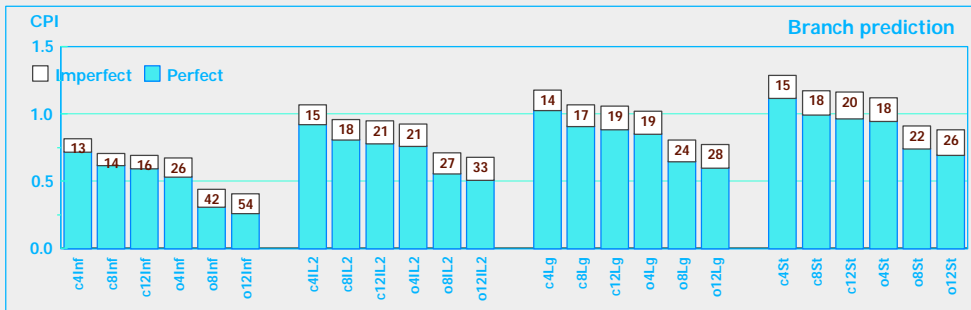
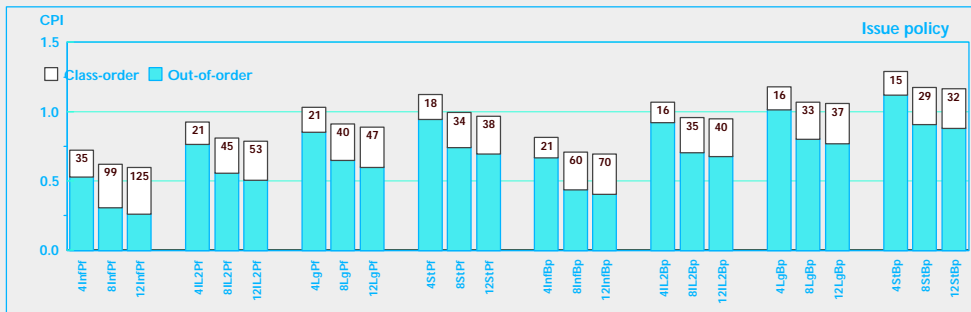
CPI results

Issue policy	Width	Bp: 2-bit branch history table (8192 entries)				Pf: Perfect branch predictor			
		Inf	IL2	Lg	St	Inf	IL2	Lg	St
c: Class-order	4	0.82	1.07	1.18	1.29	0.72	0.93	1.03	1.12
	8	0.71	0.96	1.07	1.18	0.62	0.81	0.91	1.00
	12	0.70	0.95	1.06	1.17	0.60	0.79	0.89	0.97
o: Out-of-order	4	0.67	0.93	1.02	1.12	0.53	0.77	0.86	0.95
	8	0.44	0.71	0.81	0.91	0.31	0.56	0.65	0.75
	12	0.41	0.68	0.77	0.88	0.27	0.51	0.60	0.70

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

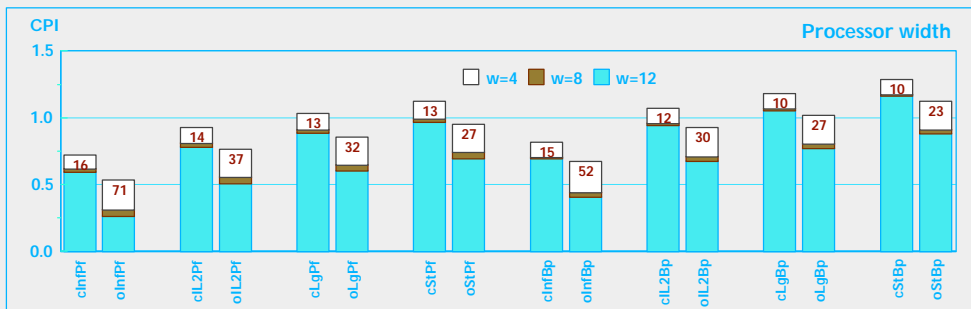
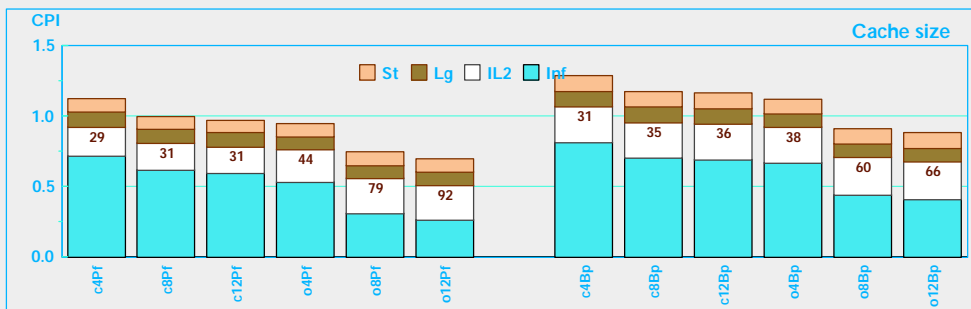
CPI addrs



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

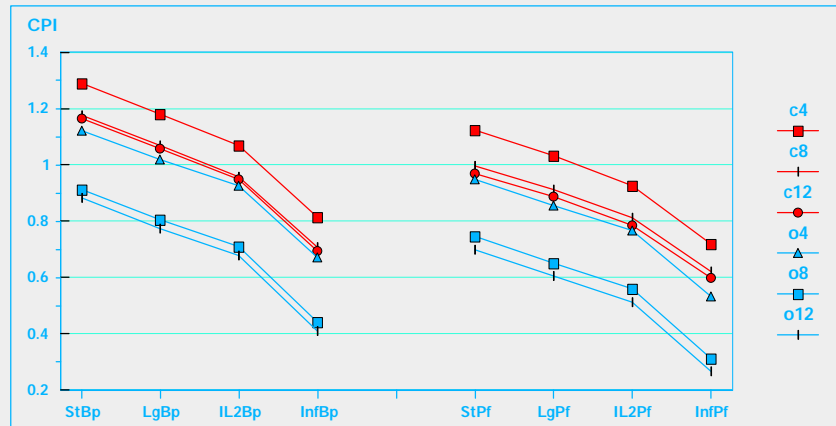
CPI addrs (cont.)



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

CPI for all configurations



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

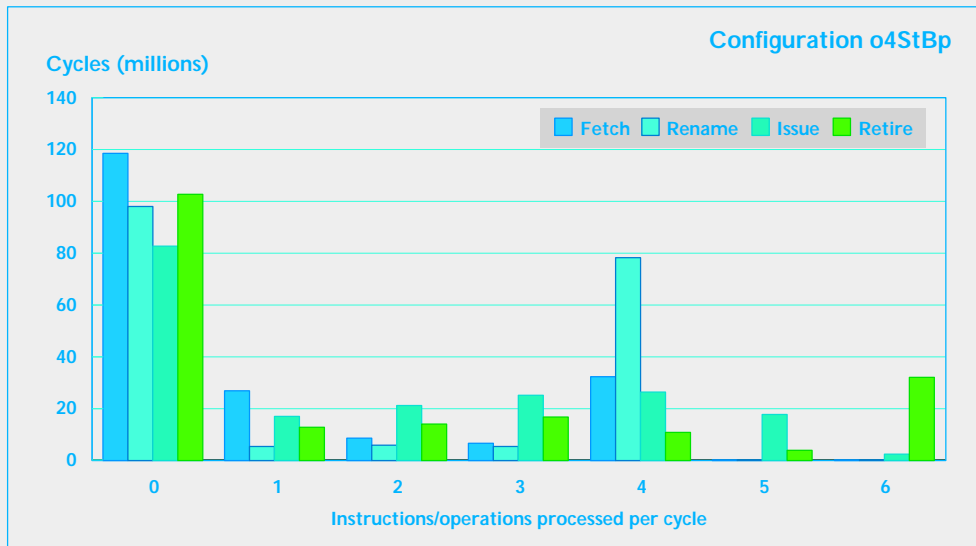
Observations

- With respect to least-aggressive out-of-order configurations
 - 15 to 32% degradation due to class-order issue
 - more severe degradation expected for in-order policy
 - 18 to 26% degradation due to imperfect branch predictor
 - 23% improvement when doubling resources
 - same branch predictor, same cache size
 - 10% additional improvement when doubling cache size
- Diminishing benefits beyond dispatching eight operations per cycle
- Still plenty of issues to investigate in detail

Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

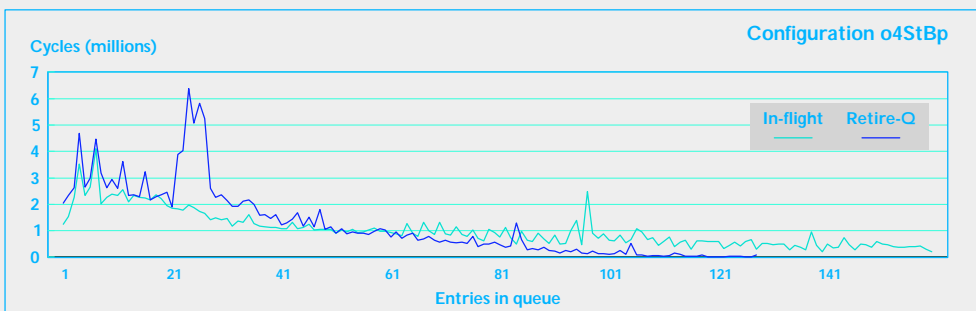
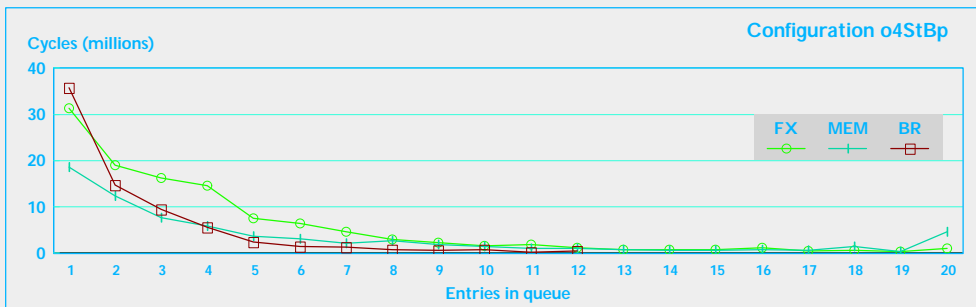
Utilization of pipeline stages



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

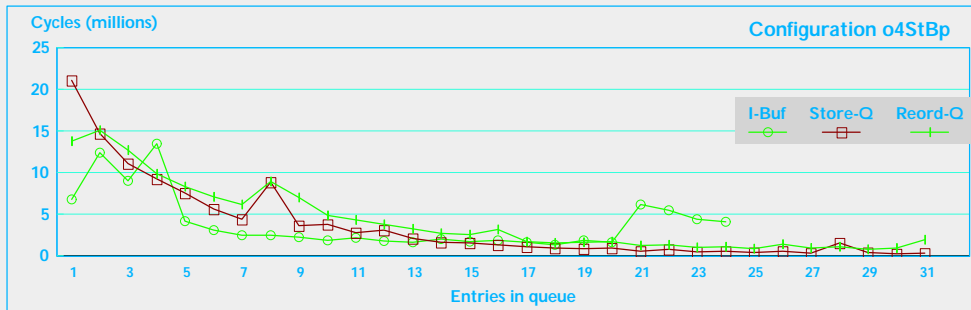
Utilization of queues



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

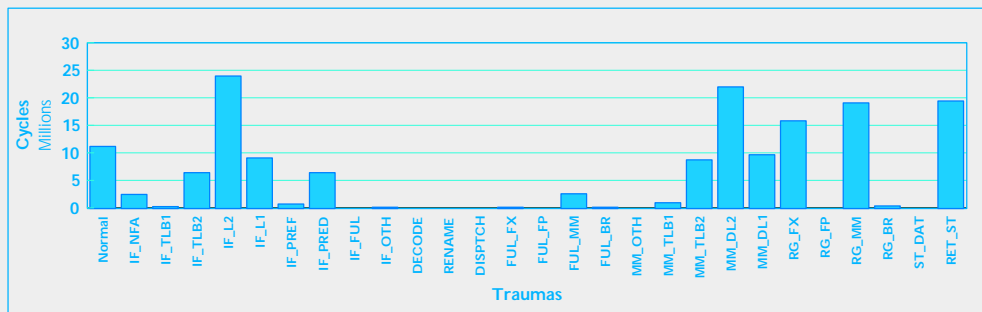
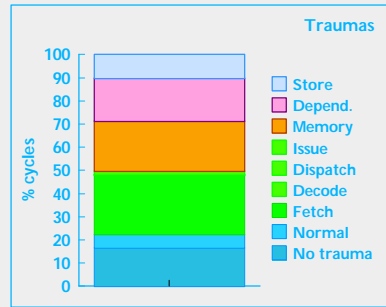
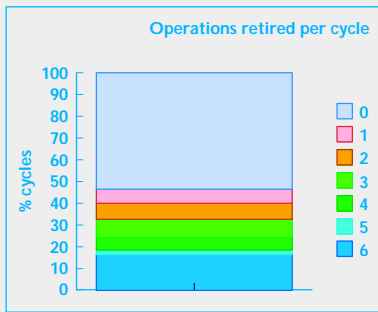
Utilization of queues (cont.)



Retirement's perspective

- Reasons for not retiring maximum number of operations
 - "traumas"
 - associated to each operation as it flows through the pipeline
 - only one trauma recorded per operation (last trauma)
- Identify trauma of first instruction that cannot be retired in a given cycle

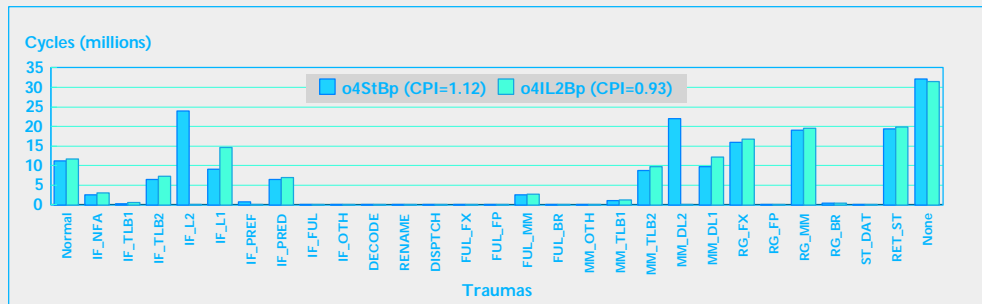
Retirement's perspective in o4StBp (CPI=1.12)



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

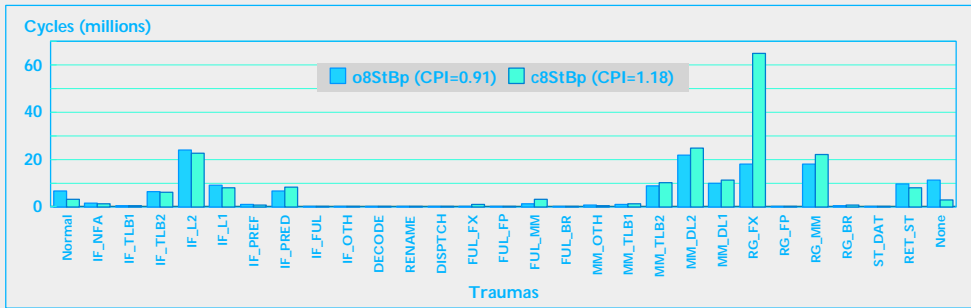
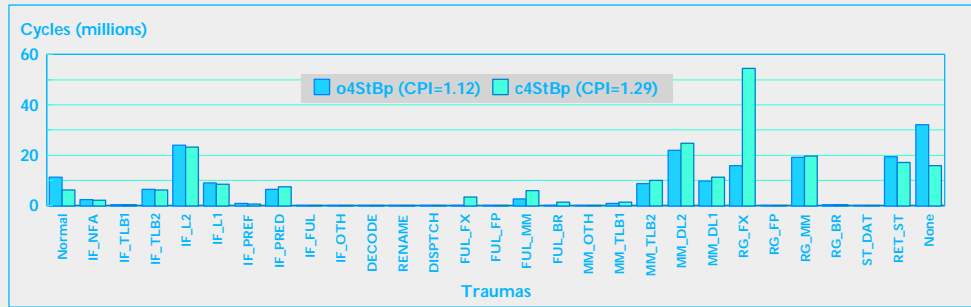
Effects of L2 cache



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

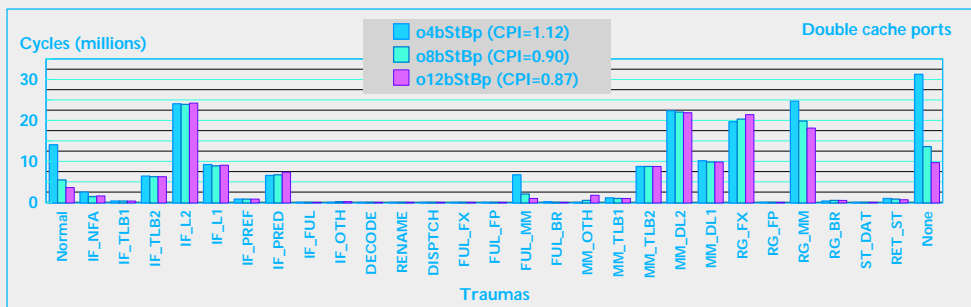
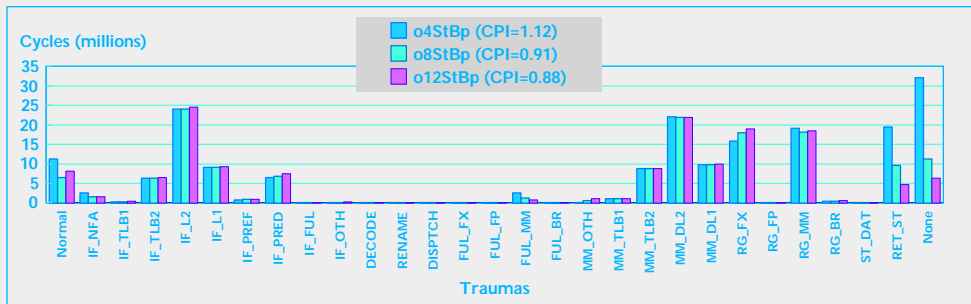
Effects of issue policy



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Effects of issue width



Exploring potential performance of wide PowerPC-based superscalar processors

J. Moreno, 10/97

Effects of other microarchitecture features

Feature	o4StBp		o8StBp	
	CPI	%	CPI	%
Original	1.12	-	0.91	-
No NFA prediction	1.16	-3.6	0.95	-4.4
No early branch resolution	1.18	-5.4	0.97	-6.6
Double I-fetch bandwidth	1.10	1.8	0.90	1.1
One fewer cycle in load operations	1.11	0.9	0.89	2.2
One additional decode stage	1.14	-1.8	0.93	-2.2
Two additional decode stages	1.15	-2.7	0.95	-4.4
Larger TLBs (4x)	1.08	3.6	0.87	4.4
Larger caches (2x)	1.02	8.9	0.81	11.0

Exploring potential performance of wide PowerPC-based superscalar processors
J. Moreno, 10/97

Observations

- **Bursty processor activity**
 - idle at times, quite busy at others
- **Limited instruction-level parallelism in the trace**
- **Small gains from various features**
 - cache size and early branch resolution most beneficial
- **Better leverage in out-of-order policy**
- **Potentially 30% improvement over decode/dispatch=4**

Exploring potential performance of wide PowerPC-based superscalar processors
J. Moreno, 10/97

Concluding remarks

- **Environment for early exploration**
 - fast flexible
 - trends among aggressive superscalar organizations

- **Basis for contrasting with other paradigms**

- **Aggressive superscalar seems able to outperform other organizations**
 - based on results reported in the literature
 - buildable?
 - need to quantify potential performance from realizable implementation
 - need to identify/develop features that provide better return

- **Continuing need for research on superscalar features**
 - considering constraints/possibilities arising from technology
 - understand interactions and tradeoffs among new features