S

CS 211

Introduction to Explicitly Parallel Instruction Computing (EPIC) and Very Long Instruction Word (VLIW) Architectures

Bhagi Narahari

Static ILP: VLIW/EPIC Architectures

- Overview of key Explicit Parallel Instruction Computing (EPIC) concepts
 – speculation, predication, register files
- Very Large Instruction Word (VLIW) and EPIC:
 - VLIW architectures progressed to EPIC
 - Let's take a quick look at VLIW

CS 211

C

VLIW and EPIC VLIW architectures progressed to EPIC A quick look at "pure" VLIW approach .

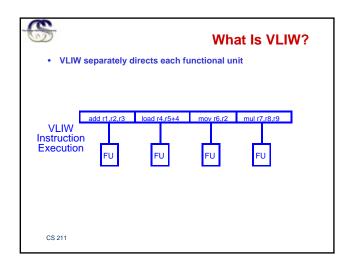
CS 211

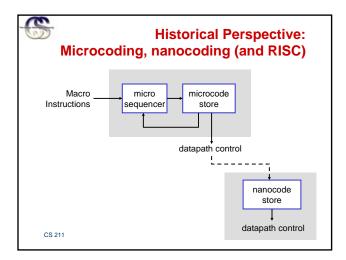
C

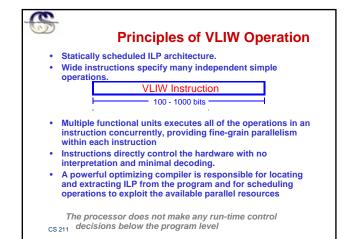
VLIW: Very Large Instruction Word Each "instruction" has explicit coding for multiple operations In IA-64, grouping called a "packet" In Transmeta, grouping called a "molecule" (with "atoms" as ops) Tradeoff instruction space for simple decoding

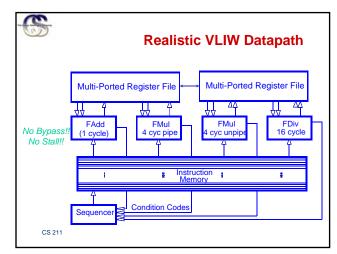
- The long instruction word has room for many operations
- By definition, all the operations the compiler puts in the long instruction word are independent => execute in parallel
- E.g., 2 integer operations, 2 FP ops, 2 Memory refs, 1 branch
 16 to 24 bits per field => 7*16 or 112 bits to 7*24 or 168 bits wide

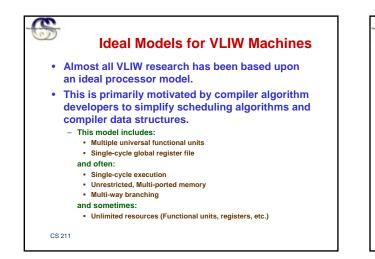
1





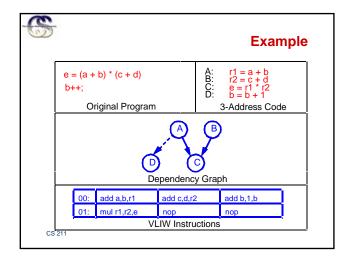


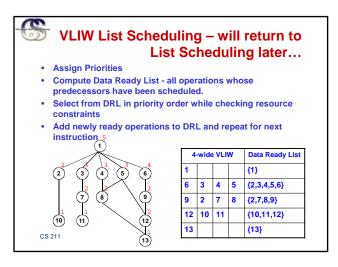




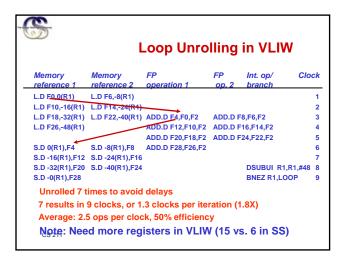


- The program is translated into primitive RISCstyle (three address) operations
- Dataflow analysis is used to derive an operation precedence graph from a portion of the original program
- Operations which are independent can be scheduled to execute concurrently contingent upon the availability of resources
- The compiler manipulates the precedence graph through a variety of semantic-preserving transformations to expose additional parallelism





			Stalls for Scalar	
	: L.D	F0,0(R1)		D to ADD.D: 1 Cycle
2	L.D	F6,-8(R1)		D.D to S.D: 2 Cycles
3	L.D	F10,-16(R1)		D.D 10 0.D. 2 0ycles
4	L.D	F14,-24(R1)		
5	ADD.D	F4,F0,F2		
6	ADD.D	F8,F6,F2		
7	ADD.D	F12,F10,F2		
8	ADD.D	F16,F14,F2		
9	S.D	0(R1),F4		
10		-8(R1),F8		
11	S.D			
12	DSUBUI	R1,R1,#32		
13	BNEZ	R1,LOOP		
14	S.D	8(R1),F16	; 8-32 =	-24

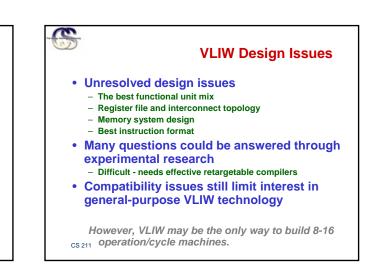


Enabling Technologies for VLIW

- VLIW Architectures achieve high performance through the combination of a number of key enabling *hardware* and *software* technologies.
 - Optimizing Schedulers (compilers)
 - Static Branch Prediction
 - Symbolic Memory Disambiguation
 - Predicated Execution
 - (Software) Speculative Execution
 - Program Compression

CS 211

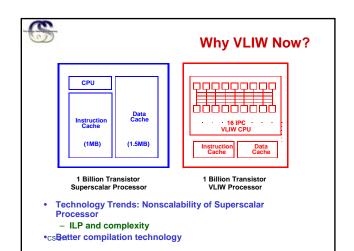
C

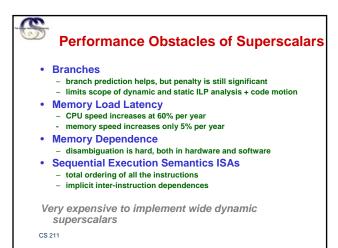


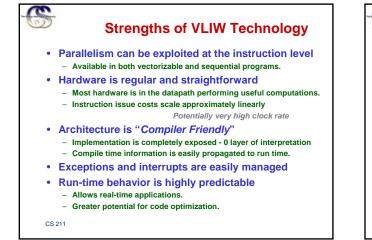
VLIW vs. Superscalar [Bob Rau, HP]				
Attributes	Superscalar	VLIW		
Multiple instructions/cycle	yes	yes		
Multiple operations/instruction	no	yes		
Instruction stream parsing	yes	no		
Run-time analysis of register dependencies	yes	no		
Run-time analysis of memory dependencies	maybe	occasionally		
Runtime instruction	maybe	no		
reordering	(Resv. Stations)			
Runtime register allocation	maybe	maybe		
	(renaming)	(iteration frames)		



- CStin MAJC (Microarchitecture for Java Computing)









- No object code compatibility between generations
- Program size is large (explicit NOPs) Multiflow machines predated "dynamic memory compression" by encoding NOPs in the instruction memory
- Compilers are extremely complex
 Assembly code is almost impossible
- Philosophically incompatible with caching techniques
- VLIW memory systems can be very complex
 Simple memory systems may provide very low performance
 Program controlled multi-layer, multi-banked memory
- Parallelism is underutilized for some algorithms.

The EPIC Model

- VLIW concept in terms of static ILP – Use compiler to extract parallelism
- Try to overcome limitations of VLIW

CS 211

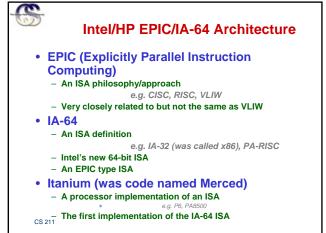
• Can we use additional H/W support to enhance S/W techniques?

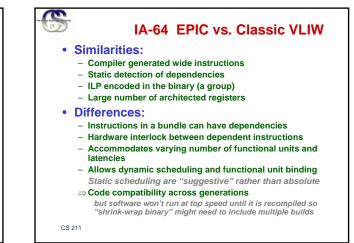
S

EPIC Concepts

- Explicitly Parallel Instruction Computing

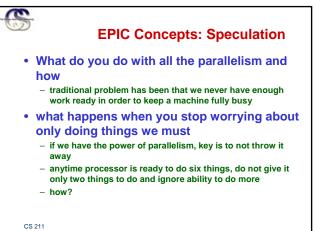
 unlike early VLIW designs, EPIC does not use fixed width instructions....as many parallel as possible!
- Programs must be written using sequential semantics
 - parallel semantics not supported
 - explicitly lay out the parallelism
 - eg: swapping of operands

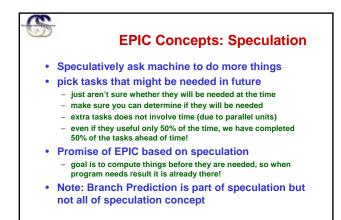




EPIC: Key Concepts Speculation

- Predication (and parallel compares)
- Large (Rotating) Register Files



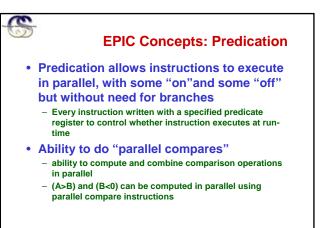




EPIC Concepts: Predication

- Branching is generally bad because it interferes with the ideal pipeline model of reading instructions while earlier inst is executed
- ideally, if we eliminate branches then this problem disappears
 - Can we lin instruction to a condition ?
- Predication is process by which branches are eliminated
 - Note: predication is not branch prediction!!

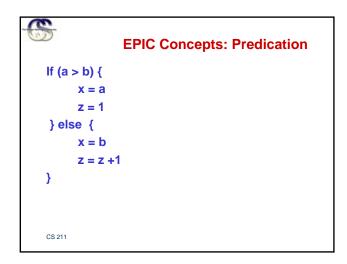
CS 211

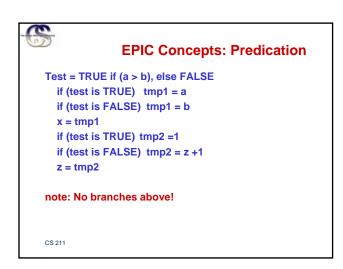


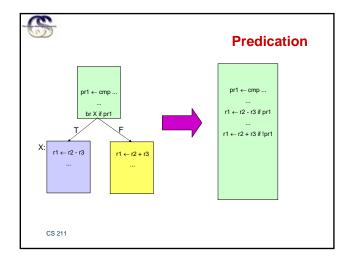


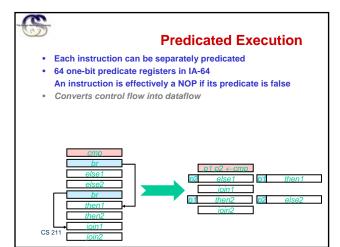
EPIC Concepts: Predication

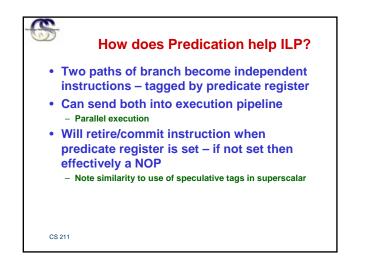
- EPIC provides predicated instructions
 every instruction can be executed in predicated manner
 - instruction execution tied to result of a predicate register
 - one predicate register hardwired to a 1; use this to always execute

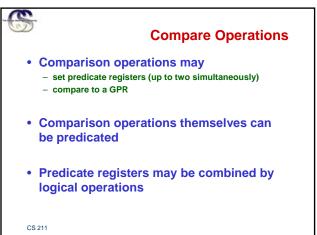


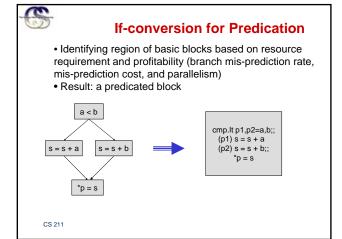


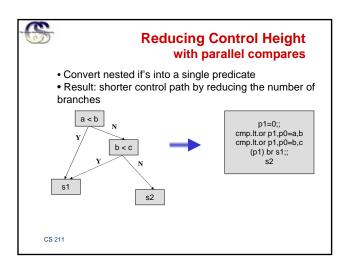


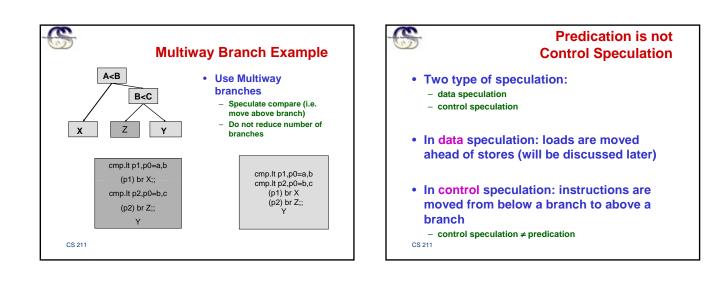


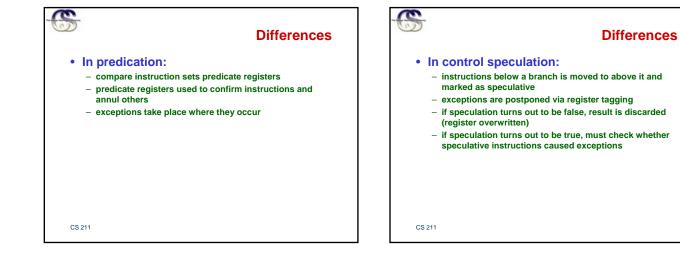


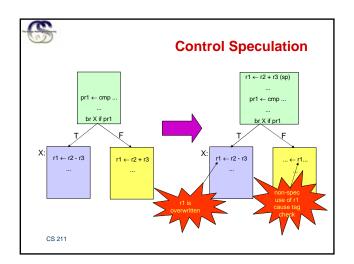


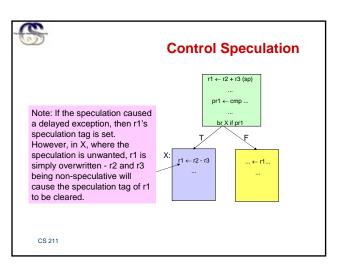


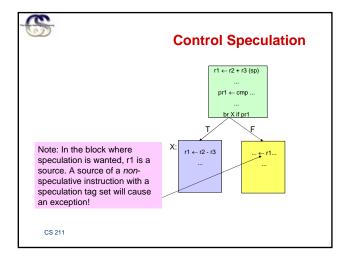


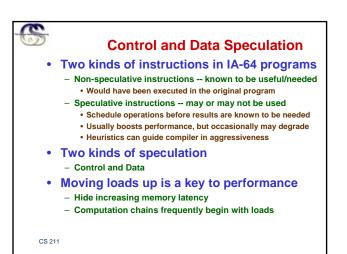




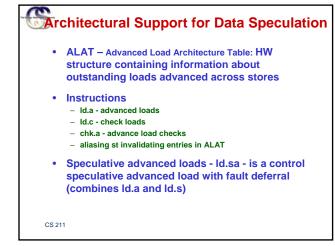


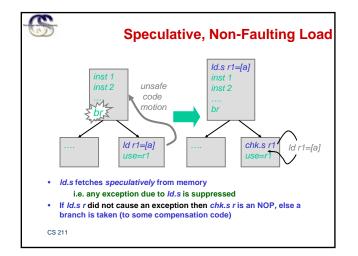


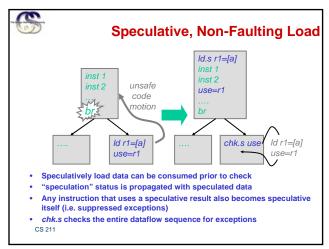


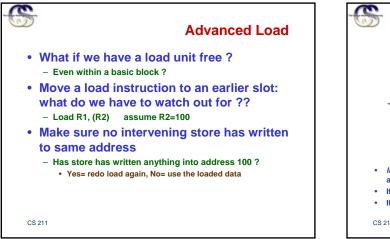


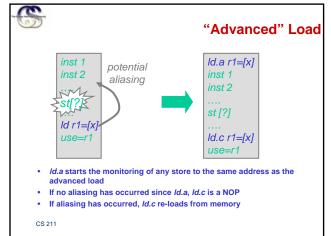


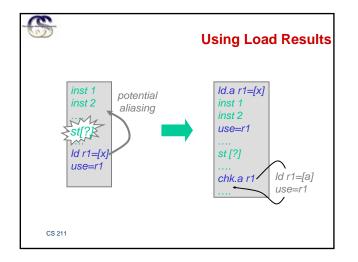


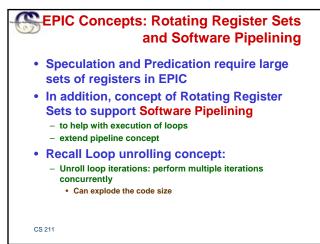


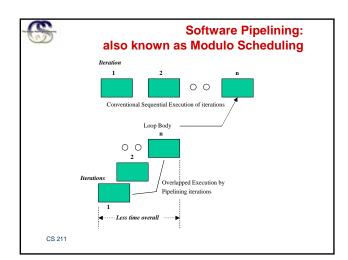


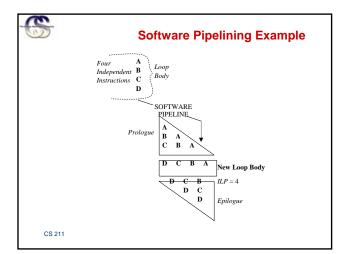


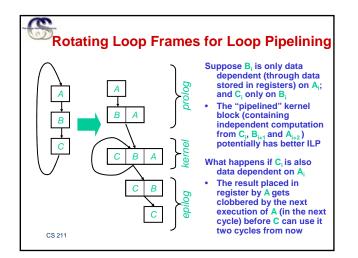


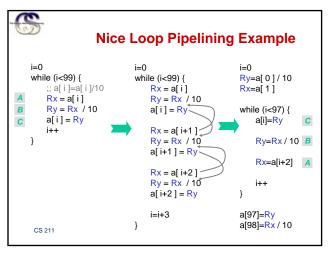


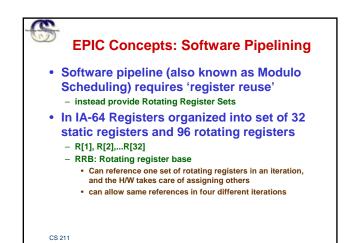


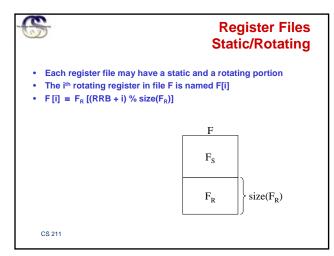


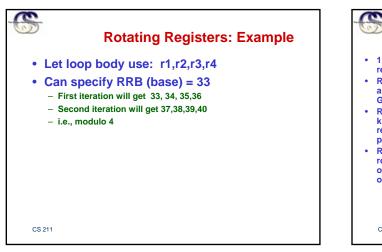


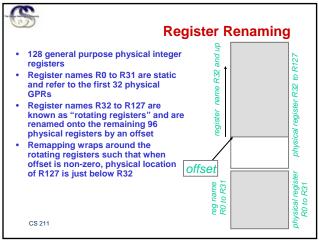


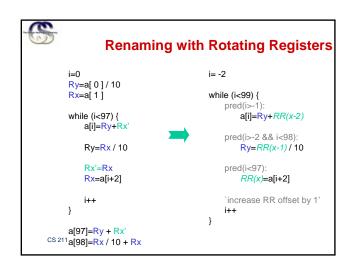


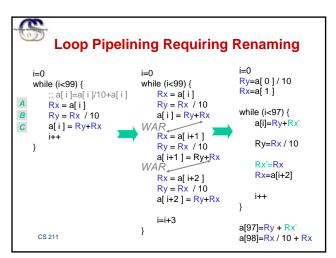


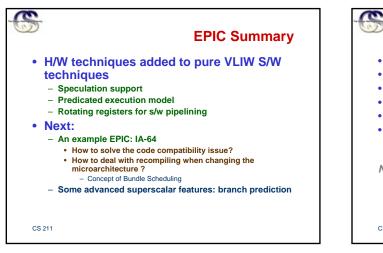






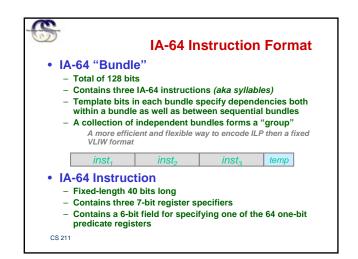


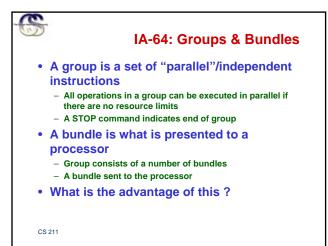






- 128 general-purpose registers
- 128 floating-point registers
- Arbitrary number of functional units
- · Arbitrary latencies on the functional units
- Arbitrary number of memory ports
- Arbitrary implementation of the memory hierarchy
- Needs retargetable compiler and recompilation to achieve maximum program performance on different IA-64 implementations





Cool Features of IA64

- Predicated execution
- Speculative, non-faulting Load instruction
- Software-assisted branch prediction
- Register stack
- Rotating register frame
- Software-assisted memory hierarchy

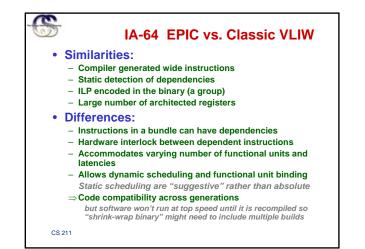
Mostly adapted from mechanisms that had existed for VLIWs

CS 211

0

Itanium Specifics

- 6-wide 10-stage pipeline
- Fetch 2 bundles per cycle with the help of BP into a 8-bundle deep fetch queue
- 512-entry 2-level BPT, 64-entry BTAC, 4 TAR, and a RSB
- Issue up to 2 bundles per cycle some mixes of 6 instructions e.g. (MFI,MFI) or (MIB,MIB_h)
- Can issue as little as one syllable per cycle on RAW hazard interlock
 or structural hazard (scoreboard for RAW detection)
- 8R-6W 128 Entry Int. GPR, 128 82-bit FPR, 64 predicate reg's
 4 globally-bypassed single-cycle integer ALUs with MMX,
- 2 FMACs, 2 LSUs, 3 BUs
- Can execute IA-32 software directly
- Intended for high-end server and workstations
- You can buy one now, finally.



S

EPIC and Compiler Optimization

- EPIC requires dependency free "scheduled code"
- Burden of extracting parallelism falls on compiler
- success of EPIC architectures depends on efficiency of Compilers!!
- We provide overview of Compiler Optimization techniques (as they apply to EPIC/ILP)