**342 REDUCED INSTRUCTION SET COMPUTING**

# **REDUCED INSTRUCTION SET COMPUTING**

# **ARCHITECTURE**

The term *computer architecture* was first defined in the article by Amdahl, Blaauw, and Brooks of International Business Machines (IBM) Corporation announcing the IBM System/

computer seen by the machine language programmer as de- tions are actually used in the real programs. This evidence scribed in the *Principles of Operation.* IBM referred to the came from the analysis of the *trace tapes,* a collection of mil-Principles of Operation as a definition of the machine that lions of the instructions that were executed in the machine enables the machine language programmer to write function- running a collection of representative programs (10). It ally correct, time-independent programs that would run showed that for 90% of the time only about 10 instructions across a number of implementations of that particular archi- from the instruction repertoire were actually used. Then the tecture. **obvious question was asked: "why not favor implementation** 

machine that are observable by the program (3). On the other cycle and emulate the rest of the instructions?" The following hand, Principles of Operation are used to define the functions reasoning was used: ''If the presence of a more complex set that the implementation should provide. In order to be func- adds just one logic level to a 10 level basic machine cycle, tionally correct, it is necessary that the implementation con- the CPU has been slowed down by 10%. The frequency and forms to the Principles of Operation. performance improvement of the complex functions must first

architecture, which includes: tional cost'' (5). Therefore, RISC architecture starts with a

- 
- 
- 
- 
- All registers and memory locations that may be directly performance of the machine.
- 

tion enabled several embodiments of the same architecture to<br>be built. Operational evidence proved that architecture and<br>implementation could be separated and that one need not im-<br>ford MIPS project, used by Silicon Graphi them to produce the same result which defined the notion of *architectural compatibility.* Implementation of the whole line **RISC Performance**

RISC architecture has been developed as a result of the 801 chine organization and the machine architecture where the project which started in 1975 at the IBM Thomas J. Watson skills and experience of computer design are shown. RISC Research Center and was completed by the early 1980s (5). deals with these two levels—more precisely their interaction This project was not widely known to the world outside of and trade-offs. IBM, and two other projects with similar objectives started in The work that each instruction of the RISC machine perthe early 1980s at the University of California Berkeley and forms is simple and straightforward. Thus, the time required Stanford University (6,7). The term RISC (reduced instruc- to execute each instruction can be shortened and the number tion set computing), used for the Berkeley research project, is of cycles reduced. Typically the instruction execution time is

''the most important product announcement that this corpora- Development of RISC architecture started as a rather tion has made in its history." 
"fresh look at existing ideas" (5,8,9) after revealing evidence Computer architecture was defined as the attributes of a that surfaced as a result of examination of how the instruc-The architecture specification covers all functions of the of those selected instructions so that they execute in a short The Principles of Operation document defines computer overcome this 10% degradation and then justify the addismall set of the most frequently used instructions which de- • Instruction set termines the pipeline structure of the machine enabling fast • Instruction format • • Instructions in one cycle. If addition of a • Instruction format • Instruction increases the "critical path" (typi-• Operation codes cally 12 to 18 gate levels) for one gate level, then the new • Addressing modes cally 12 to 18 gate levels) for one gate level, then the new instruction should contribute at least 6% to 8% to the overall

manipulated or tested by a machine language program One cycle per instruction is achieved by exploitation of • Formats for data representation parallelism through the use of pipelining. It is *parallelism through pipelining* that is the single most important charac-Machine Implementation was defined as the actual system or-<br>ganization and hardware structure encompassing the major<br>functional units, data paths, and control.<br>Machine Boslization includes issues such as logic tophol based

Machine Realization includes issues such as logic technol-<br>ogy, packaging, and interconnections.<br>Separation of the machine architecture from implementa-<br>stream architectures today are of the RISC type. Those in-<br>tion enabl

of computers according to a common architecture requires un-<br>usual attention to details and some new procedures which are<br>described in the Architecture Control Procedure. The design<br>and control of system architecture is an **RISC Architecture RISC Architecture RISC Architecture** only in proportion to the amount of technology improvements; A special place in computer architecture is given to RISC. this is, more or less, available to everyone. It is in the ma-

# **344 REDUCED INSTRUCTION SET COMPUTING**

as processing of one stage is finished, the machine proceeds  $T_0 = \text{clock period (ns)}$ with executing the second stage. However, when the stage becomes free it is used to execute the same operation that be- While CISC instruction will typically have less instructions

By overlapping the execution of several instructions in a pipe-<br>ine fashion (as shown in Fig. 1), RISC achieves its inherent  $\tau$ execution parallelism which is responsible for the perfor- as follows: mance advantage over the complex instruction set architec-

cycle per instruction (CPI = 1.0), which would be the case structions.<br>when no interruptions in the pipeline occurs. However, this  $\sigma$  PISC solid

efficient execution of the RISC pipeline.

The simplicity of the RISC instruction set is traded for **RISC MACHINE IMPLEMENTATION** more parallelism in execution. On average, a code written for RISC will consist of more instructions than the one written The main feature of RISC is the architectural support for the for CISC. The typical trade-off that exists between RISC and exploitation of parallelism on the instruction level. Therefore CISC can be expressed in the total time required to execute a all distinguished features of RISC architecture should be concertain task: sidered in light of their support for the RISC pipeline. In addi-

$$
Time (task) = I \times C \times P \times T_0
$$

divided into five stages, namely, machine cycles; and as soon  $P =$  number of clock periods/cycle (usually  $P = 1$ )

longs to the next instruction. The operation of the instruc- for the same task, the execution of its complex operations will tions is performed in a pipeline fashion, similar to the require more cycles and more clock ticks within the cycle as assembly line in the factory process. Typically, those five pipe- compared to RISC (11). On the other hand, RISC requires line stages are as follows: more instructions for the same task. However, RISC executes its instructions at the rate of one instruction per cycle, and IF: Instruction Fetch its machine cycle requires only one clock tick (typically). In ID: Instruction Decode addition, given the simplicity of the instruction set, as re-EX: Execute flected in simpler machine implementation, the clock period  $T_0$  in RISC can be shorter, allowing the RISC machine to run MA: Memory Access at the higher speed as compared to CISC. Typically, as of to-<br>WB: Write Back and MA: Maximum Maximum At the frequency day, RISC machines have been running at the frequency reaching 1 GHz, while CISC is hardly at the 500 MHz clock

The trade-off between RISC and CISC can be summarized

- tures (CISC).<br>The goal of RISC is to achieve an execution rate of one the program consisting of a fewer number of powerful inprogram consisting of a fewer number of powerful in-
- when no interruptions in the pipeline occurs. However, this is not the case.<br>
is not the case.<br>
The instructions and the addressing modes in RISC archives its performance advantage by having<br>
is not the case.<br>
The instruct

tion to that, RISC takes advantage of the principle of locality: spatial and *temporal*. What that means is that the data that was used recently is more likely to be used again. This justiwhere **fies the implementation of a relatively large general-purpose** register file found in RISC machines as opposed to CISC. Spa-*I* = number of instructions/task tial locality means that the data most likely to be referenced  $C =$  number of cycles/instruction is in the neighborhood of a location that has been referenced.



**Figure 1.** Typical five-stage RISC pipeline.



Figure 2. Pipeline flow of a Register-to-Register operation.

It is not explicitly stated, but that implies the use of caches **Carefully Selected Set of Instructions**

Often, RISC is referred to as Load/Store architecture. Alter-<br>natively the operations in its instruction set are defined as<br>Register-to-Register operations. The reason is that all the<br>RISC machine operations are between th side in the General Purpose Register File (GPR). The result <br>of the operation is also written back to GPR. When restricting of the operation is also written back to GPR. When restricting between the locations of the operands to the GPR only, we allow for determinism in the RISC operation. In the other words, a po-<br>determinism in the RISC operat tentially multicycle and unpredictable access to memory has<br>been separated from the operation. Once the operands are<br>available in the GPR, the operation can proceed in a deter-<br>ministic fashion. It is almost certain that o the operands which can, nevertheless, be easily handled in a widespread misunderstanding that the main feature char-<br>hardware. The execution flow in the pipeline for a Register-<br>acterizing RISC is a small instruction set. hardware. The execution flow in the pipeline for a Register-<br>to-Register operation is shown in Fig. 2.<br>The number of instructions in the instruction set of RISC can

Memory Access is accomplished through Load and Store be substantial. This number of RISC instructions can grow<br>instructions only; thus the term  $Load/Store$  Architecture is until the complexity of the control logic begins to impo instructions only; thus the term *Load/Store Architecture* is until the complexity of the control logic begins to impose an often used when referring to RISC. The RISC pipeline is spec- increase in the clock period. In practice, this point is far beand memory access with equal efficiency. The various pipeline we have reached a possibly paradoxical situation, namely, stages of the Load and Store operations in RISC are shown that several of representative RISC machines known today in Fig. 3. have an instruction set larger than that of CISC.

The principle of locality is applied throughout RISC. The fact that only a small set of instructions is most frequently used, **Load/Store Architecture** was used in determining the most efficient pipeline organiza-

- 
- 
- 

Register operation is shown in Fig. 2. The number of instructions in the instruction set of RISC can<br>Memory Access is accomplished through Load and Store be substantial This number of RISC instructions can grow yond the number of instructions commonly used. Therefore



**Figure 3.** The operation of Load/Store pipeline.

118 instructions, while IBM RS/6000 (PowerPC) contains 184 instruction is encountered. instructions. This should be contrasted to the IBM System/ 360 containing 143 instructions and to the IBM System/370 **Simple Addressing Modes**

size of RISC instructions is also fixed to the size of the word<br>
(32 bits); however, there are cases where RISC can contain  $\frac{1}{2}$ . Base + Displacement<br>  $\frac{1}{2}$ . Base + Displacement two sizes of instructions, namely, 32 bits and 16 bits. Next is the case of the IBM ROMP processor used in the first commercial RISC IBM PC/RT. The fixed format feature is very important because RISC must decode its instruction in one Those three addressing modes take approximately over 80%<br>and the algorithment because RISC must decode its instruction in one of all the addressing modes according cycle. It is also very valuable for superscalar implementations of all the addressing modes according to Ref. 3: (1) 30% to (12). Fixed size instructions allow the Instruction Fetch Unit  $40\%$ , (2)  $40\%$  to 50%, and (3) guarantees only single I-TLB access per instruction. **Separate Instruction and Data Caches** One-cycle decode is especially important so that the out-

come of the Branch instruction can be determined in one cycle One of the often overlooked but essential characteristics of in which the new target instruction address will be issued as RISC machines is the existence of cache memory. The second well. The operation associated with detecting and processing most important characteristic of RISC (after pipelining) is its a Branch instruction during the Decode cycle is illustrated in use of the locality principle. The locality principle is estab-Fig. 4. In order to minimize the number of lost cycles, Branch lished on the observation that, on average, the program instructions need to be resolved, as well, during the Decode spends 90% of the time in the 10% of the code. The instruction stage. This requires a separate address adder as well as com- selection criteria in RISC is also based on that very same obparator, both of which are used in the Instruction Decode servation that 10% of the instructions are responsible for 90%

For example: IBM PC-RT Instruction architecture contains Unit. In the best case, one cycle must be lost when Branch

containing 208. The first two are representatives of RISC ar-<br>chitecture, while the latter two are not.<br>line. That is, in order to be able to perform the address calcu-<br>directure, lation in the same predetermined number of pipeline cycles **Fixed Format Instructions** in the pipeline, the address computation needs to conform to What really matters for RISC is that the instructions have a<br>fixed and predetermined format which facilitates decoding in<br>one cycle and simplifies the control hardware. Usually the<br>favors three relatively simple addressing

- 
- 
- 3. Base  $+$  Index

### **REDUCED INSTRUCTION SET COMPUTING 347**



a 90–10 rule (13). not collide. It is from there that the separation of instruction

ral. *Spatial locality* means that the most likely location in feature for RISC. the memory to be referenced next will be the location in the **Branch and Execute Instruction** neighborhood of the location that was just referenced previously. On the other hand, *temporal locality* means that the Branch and Execute or Delayed Branch instruction is a new most likely location to be referenced next will be from the set feature of the instruction architecture that was introduced<br>of memory locations that were referenced just recently. The and fully exploited in RISC. When a Br of memory locations that were referenced just recently. The

The RISC machines are based on the exploitation of that principle as well. The first level in the memory hierarchy is the general-purpose register file GPR, where we expect to find the operands most of the time. Otherwise the Register-to-Register operation feature would not be very effective. However, if the operands are not to be found in the GPR, the time to fetch the operands should not be excessive. This requires the existence of a fast memory next to the CPU—the Cache. The cache access should also be fast so that the time allocated for Memory Access in the pipeline is not exceeded. One-cycle cache is a requirement for RISC machine, and the performance is seriously degraded if the cache access requires two or more CPU cycles. In order to maintain the required one- **Figure 5.** Pipeline flow of the Branch instruction.

of the code. Often the principle of the locality is referred to as cycle cache bandwidth the data and instruction access should In case of the cache, this locality can be spatial and tempo- and data caches, the so-called Harvard architecture, is a must

cache operates on this principle.<br>The RISC machines are based on the exploitation of that This is illustrated in Fig. 5.





**Figure 6.** Lost cycle during the execution of the load instruction.

The subject instruction can be found in the instruction stream preceding the Branch instruction, in the target instruction stream, or in the fall-through instruction stream. It is the task of the compiler to find such an instruction and to fill-in this execution cycle (14).

Given the frequency of the Branch instructions, which varies from 1 out of 5 to 1 out of 15 (depending on the nature of the code), the number of those otherwise lost cycles can be substantial. Fortunately a good compiler can fill-in 70% of those cycles which amounts to an up to 15% performance improvement (13). This is the single most performance contributing instruction from the RISC instruction architecture.

However, in the later generations of superscalar RISC machines (which execute more than one instruction in the pipeline cycle), the *Branch and Execute* instructions have been abandoned in favor of *Brand Prediction* (12,15).

The Load instruction can also exhibit this lost pipeline cycle as shown in Fig. 6.

Program to calculate:



**Figure 7.** An example of instruction scheduling by compiler.

			$EX   MA   IF   WB   IF   D   EX   MA  $	WB

Total of cycles for two instructions

**Figure 8.** Instruction execution in the absence of pipelining.

The same principle of scheduling an independent instruction in the otherwise lost cycle, which was applied for in Branch and Execute, can be applied to the Load instruction. This is also known as delayed load.

An example of what the compiler can do to schedule instructions and utilize those otherwise lost cycles is shown in Fig. 7 (13,14).

### **Optimizing Compiler**

RISC architecture solves the lost cycle problem by intro-<br>ducing Branch and Execute instruction (5,9) (also known as<br>Delayed Branch instruction), which consists of an instruction<br>pair: *Branch* and the *Branch Subject* in





*<sup>a</sup>* IBM PC-RT Instruction architecture contains 118 instructions, while IBM RS/ 6000 (PowerPC) contains 184 instructions. This should be contrasted to the IBM System/360 containing 143 instructions and IBM System/370 containing 208. The first two are representatives of RISC architecture; the latter two are not.



**Figure 9.** Main branches in development of computer architecture.

mizing Compiler" was introduced in RISC (5,9,14). This com-<br>piler was capable of producing a code that was as good as the<br>piler was capable of producing a code that was as good as the<br>code written in assembler (the hand-c

interruption to the pipeline, Loads and Stores that cannot be due to the clock skews and clock jitter. This could very soon



scheduled, and finally the effect of finite size caches, the number of "lost" cycles adds up, bringing the CPI further away from 1. In the real implementations the CPI varies and a  $CPI = 1.3$  is considered quite good, while CPI between 1.4 to 1.5 is more common in single-instruction issue implementations of the RISC architecture.

However, once the CPI was brought close to 1, the next goal in implementing RISC machines was to bring CPI below 1 in order for the architecture to deliver more performance. This goal requires an implementation that can execute more than one instruction in the pipeline cycle, a so called *superscalar* implementation (12,16). A substantial effort has been made on the part of the leading RISC machine designers to build such machines. However, machines that execute up to four instructions in one cycle are common today, and a machine that executes up to six instructions in one cycle was introduced in 1997.

## **Pipelining**

**One Instruction per Cycle Cycle Cycle Cycle Parallelism**, we may increase the RISC machine performance further. However, this idea does not lead to a simple and The objective of *one instruction per cycle* (CPI = 1) execution straightforward realization. The increase in the number of was the ultimate goal of RISC machines. This goal can be pipeline stages introduces not only an o pipeline stages introduces not only an overhead in hardware theoretically achieved in the presence of infinite size caches (needed to implement the additional pipeline registers), but and thus no pipeline conflicts, which is not attainable in prac- also the overhead in time due to the delay of the latches used tice. Given the frequent branches in the program and their to implement the pipeline stages as well as the cycle time lost

**Figure 10.** History of RISC development.

### **Table 2. Some features of RISC Processors**



*<sup>a</sup>* No cache.

bring us to the point of diminishing returns where further mour Cray, who is by many given the credit for the invention increase in the pipeline depth would result in less perfor- of RISC. mance. An additional side effect of deeply pipelined systems is hardware complexity necessary to resolve all the possible **History of RISC**

sudden development. It was rather a long and evolutionary process in the history of computer development in which we learned how to build better and more efficient computer sys- **BIBLIOGRAPHY** tems. From the first definition of the architecture in 1964 (1), there are the three main branches of the computer architec-<br>ture that evolved during the years. They are shown in Fig. 9. The IBM System/360, IBM J. Res. Develop., 8: 87-101, 1964.

The CISC development was characterized by (1) the PDP-<br>11 and VAX-11 machine architecture that was developed by<br>11 and VAX-11 machine architecture that was developed by<br>11 and Examples, Advanced Computer Science Series, Digital Equipment Corporation (DEC) and (2) all the other New York: McGraw-Hill, 1982.<br>architectures that were derived from that development. The  $\alpha$  C  $\Delta$  Blazuw and E P Broo middle branch is the IBM 360/370 line of computers, which is *IBM Syst. J.*, 3: 119–135, 1964.<br>characterized by a balanced mix of CISC and RISC features. A R P Case and A Paders, Arch The RISC line evolved from the development line character- *Commun. ACM,* **21**: 73–96, 1978. ized by Control Data Corporation CDC 6600, Cyber, and ulti-<br>mately the CRAY-I supercomputer. All of the computers be-<br>search Center. Rep. RC 9125, 1981; also in SIGARCH Comput. longing to this branch were originally designated as *Archit. News,* **10** (2): 39–47, 1982. *supercomputers* at the time of their introduction. The ulti- 6. D. A. Patterson and C. H. Sequin, A VLSI RISC, *IEEE Comput.* mate quest for performance and excellent engineering was a *Mag.,* **15** (9): 8–21, 1982. characteristic of that branch. Almost all of the computers in 7. J. L. Hennessy, VLSI processor architecture, *IEEE Trans. Com*the line preceding RISC carry the signature of one man: Sey- *put.,* **C-33**: 1221–1246, 1984.

conflicts that can occur between the increased number of in-<br>structions residing in the pipeline stages is mainly determined by the type of the number search Center under the name of the 801. 801 is the<br>struction core (th work was directly transferred to MIPS (17).

**HISTORICAL PERSPECTIVE** The chronology illustrating RISC development is illustrated in Fig. 10.

The architecture of RISC did not come about as a planed or a<br>sudden development. It was rather a long and evolutionary shown in Table 2.

- 
- 
- 3. G. A. Blaauw and F. P. Brooks, The structure of System/360,
- 4. R. P. Case and A. Padegs, Architecture of the IBM System/370,
- search Center, Rep. RC 9125, 1981; also in *SIGARCH Comput.*
- 
- 

### **REFLECTOMETERS, TIME-DOMAIN 351**

- 8. J. Cocke and V. Markstein, The evolution of RISC technology at IBM, *IBM J. Res. Develop.,* **34**: 4–11, 1990.
- 9. M. E. Hopkins, A perspective on the 801/reduced instruction set computer, *IBM Syst. J.,* **26**: 107–121, 1987.
- 10. L. J. Shustek, Analysis and performance of computer instruction sets, PhD thesis, Stanford Univ., 1978.
- 11. D. Bhandarkar and D. W. Clark, Performance from architecture: Comparing a RISC and a CISC with similar hardware organization, *Proc. 4th Int. Conf. ASPLOS,* Santa Clara, CA, 1991.
- 12. G. F. Grohosky, Machine organization of the IBM RISC System/ 6000 processor, *IBM J. Res. Develop.,* **34**: 37, 1990.
- 13. J. Hennessy and D. Patterson, *Computer Architecture: A Quantitative Approach,* San Mateo, CA: Morgan Kaufman.
- 14. H. S. Warren, Jr., Instruction scheduling for the IBM RISC System/6000 processor, *IBM J. Res. Develop.,* **34**: 37, 1990.
- 15. J. K. F. Lee and A. J. Smith, Branch prediction strategies and branch target buffer design, *Comput.,* **17** (1): 1984, 6–22.
- 16. J. Cocke, G. Grohosky, and V. Oklobdzija, *Instruction control mechanism for a computing system with register renaming, MAP table and queues indicating available registers,* U.S. Patent No. 4,992,938, 1991.
- 17. G. Kane, *MIPS RISC Architecture,* Englewood Cliffs, NJ: Prentice-Hall, 1988.

### *Reading List*

- D. W. Anderson, F. J. Sparacio, and R. M. Tomasulo, The IBM 360 Model 91: Machine philosophy and instruction handling, *IBM J. Res. Develop.,* **11**: 8–24, 1967.
- *Digital RISC Architecture Technical Handbook,* Digital Equipment Corporation, 1991.
- V. G. Oklobdzija, Issues in CPU—coprocessor communication and synchronization, *EUROMICRO '88, 14th Symp. Microprocessing Microprogramming,* Zurich, Switzerland, 1988, p. 695.
- R. M. Tomasulo, An efficient algorithm for exploring multiple arithmetic units, *IBM J. Res. Develop.,* **11**: 25–33, 1967.

VOJIN G. OKLOBDZIJA Integration Corporation

- **REDUNDANT SYSTEMS ANALYSIS.** See RELIABILITY OF REDUNDANT AND FAULT-TOLERANT SYSTEMS.
- **RE-ENGINEERING.** See BUSINESS PROCESS RE-ENGI-NEERING; SOFTWARE MAINTENANCE, REVERSE-ENGINEERING AND RE-ENGINEERING; SYSTEMS RE-ENGINEERING.

**REFLECTANCE.** See GONIOMETERS.

**REFLECTION MEASUREMENT.** See STANDING WAVE ME-TERS AND NETWORK ANALYZERS.