**Modul Perkuliahan Sesi 13**

**Organisasi dan Arsitektur Komputer**

**Parallelism and Superscalar**

Apa itu Superscalar ?

* Salah satu jenis dari arsitektur, dimana superscalar adalah sebuah uniprocessor
* Suatu rancangan untuk meningkatkan kecepatan CPU
* Mengeksekusi intruksi umum (arithmetic, load/store, conditional branch) secara independen dan secara bersamaan dalam tahap pipeline yang berbeda
* Mengeksekusi dua kali atau lebih operasi scalar dalam bentuk paralel
* Dapat diterapkan untuk RISC dan CISC
* Dalam praktek, biasanya RISC





Gambar 1 : Superscalar vs Scalar Organization

Perbedaannya dengan adanya beberapa unit fungsional memungkinkan prosesor untuk mengeksekusi aliran instruksi secara paralel, satu aliran untuk setiap pipeline, sehingga meningkatkan kecepatan kerja

Superscalar vs Superpipelined

* Superpipelined adalah pendekatan alternatif untuk mencapai kinerja yang lebih besar
* Perbedaannya, Superpipelining mengeksploitasi fakta bahwa banyak tahapan pipeline melakukan tugas-tugas yang membutuhkan **waktu kurang dari setengah clock siklus.**
* Sedangkan superskalar mengijinkan proses untuk **bekerja secara bersamaan** pada saat **satu clock siklus** yang sama



Keterbatasan

Untuk meningkatkan instruksi level parallelism, perlu dilihat keterbatasan mendasarnya:

1. Kebenaran ketergantungan data (RAW)
2. Prosedural ketergantungan
3. Konflik sumber daya
4. (Output dependency) (WAW)
5. (Antidependency) (WAR)

Kebenaran ketergantungan data Read After Write

•ADD r1, r2 (r1 := r1+r2;)

•MOVE r3,r1 (r3 := r1;)

•Jika tidak ada ketergantungan, dua instruksi dapat diambil dan dieksekusi secara paralel

•Jika ada **ketergantungan**, semua instruksi (instruksi kedua) harus **ditunda** sampai semua nilai input telah diproduksi (instruksi pertama)

Prosedural ketergantungan

•Memiliki ketergantungan prosedural pada cabang dan tidak dapat dilaksanakan sampai cabang dieksekusi

•**Konsekuensi**: kehilangan kesempatan dalam jumlah yang lebih besar pada setiap delaynya

Konflik sumber daya

•Kompetisi dari dua atau lebih instruksi untuk memperoleh sumber daya yang sama pada saat yang bersamaan

•Contoh sumber daya: memori, cache, bus, register-file, port, dan unit-unit fungsional lainnya

•Contoh konflik:

–Instruksi: 2 instruksi aritmetik

•Diatasi dengan duplikasi sumber daya

–Solusi: 2 unit aritmetik, dimana unit fungsionalnya dipipelinekan



Dessign Issues

•1. Instruction- Level Parallelism and Machine Parallelism

•Instruksi level parallelism

–Terjadi jika Instruksi bersifat independen (tidak saling berhubungan)

–Dieksekusi secara paralel dengan overlapping (tumpang tindih)

–Ditentukan oleh frekuensi ketergantungan data yang benar dan prosedural ketergantungan dalam kode

•Contoh:

Kiri || Kanan

•LOAD R1 R2 || ADD R3 R3,”1”

•ADD R3 R3,”1” || ADD R4 R3, R2

•ADD R4 R2 || STORE [R4] R0

•Instruksi sebelah kiri independen, sebelah kanan tidak

•Machine parallelism

–Kemampuan prosesor dalam memanfaatkan paralelisme tingkat instruksi

–Ditentukan oleh jumlah pipeline yang paralel dan oleh kecepatan serta kecanggihan mekanisme yang menggunakan prosesor untuk menemukan instruksi-instruksi independen

•2. Instruksi isu kebijakan

•3 hal penting:

–Urutan dimana instruksi diambil

–Urutan dimana instruksi dieksekusi

–Urutan dimana instruksi memperbarui isi dari register dan memori lokasi

•Secara umum kebijakan instruksi superscalar dapat digolongkan dalam 3 kategori berikut:

– In-order issue with in-order completion

– In-order issue with out-of-order completion

– Out-of-order issue with out-of-order completion

In-order issue with In-order completion

•Isu instruksi mengeluarkan instruksi dalam urutan yang pasti yang akan didapatkan dengan eksekusi sekuensial (in-order-issue) dan menulis hasilnya dalam urutan yang sama (in-order-completion)

•Tidak begitu efisien (sebagai dasar untuk membandingkan pendekatan yang lebih canggih)

•Mampu mengambil lebih dari satu instruksi pada satu waktu

In-Order Issue In-Order Completion (Diagram)

Contoh: I1 membutuhkan dua siklus untuk mengeksekusi I3 dan I4 bersaing untuk unit eksekusi yang sama I5 tergantung pada nilai yang dihasilkan oleh I4 I5 dan I6 bersaing untuk unit eksekusi yang sama



In-Order Issue Out-of-Order Completion

•Prosesor akan mendekode instruksi hingga dijumpai ketergantungan atau konflik.

•Tidak akan ada instruksi lainnya yang akan didekode sampai konflik teratasi

•Contoh:

•Output dependency

–R3:= R3 + R5; (I1)

–R4:= R3 + 1; (I2)

–R3:= R5 + 1; (I3)

–Intruksi I2 tidak dapat dijalankan sebelum instruksi I1data dependency

–Jika eksekusi I3 selesai sebelum I1 , maka nilai I1 akan salah akibatnya I3 harus menunggu hingga I1 menghasilkan output yang benar output dependency

In-Order Issue Out-of-Order Completion (Diagram)



Out-of-Order Issue Out-of-Order Completion

•Decouple tahapan-tahapan dekode dan eksekusi pipeline

•Setelah selesai pendekodean instruksi disimpan di jendela instruksi

•Ketika buffer belum penuh, prosesor dapat terus mengambil dan mendecode instruksi hingga buffer penuh

•Kemudian ketika unit fungsional tersedia maka instruksi dari jendela instruksi dikeluarkan untuk eksekusi

Out-of-Order Issue Out-of-Order Completion (Diagram)



Antidependency

•Write-write dependency

–R3:=R3 + R5; (I1)

–R4:=R3 + 1; (I2)

–R3:=R5 + 1; (I3)

–R7:=R3 + R4; (I4)

–I3 tidak bisa dikerjakan sebelum eksekusi I2 dimulai, dan I2 membutuhkan nilai R3 dari I1dan I3 mengubah R3

–**Antidependency kendalanya mirip dengan kendala dari dependensi data yang benar, namun dibalik**.

Register Renaming

•Digunakan untuk mengeliminasi WAW dan WAR

•WAW dan WAR muncul karen register tidak dapat lagi merefleksikan nilai-nilai dari aliran program

•Dengan adanya register renaming maka yang menentukan lokasi adalah nama, jadi bisa memiliki banyak lokasi daripada nama

•Contoh:

•I1: R1R2/R3 pembagian membutuhkan waktu yang lama untuk penyelesaiannya

•I2: R4R1+R5 RAW dependency dengan I1

•I3: R5R6+R7 WAR dependency dengan I2

•I4: R1R8+R9 WAW dependency dengan I1

•Lalu direnaming

•I1: R32R2/R3 pembagian membutuhkan waktu yang lama untuk penyelesaiannya

•I2: R33R32+R5 masih RAW dependency dengan I1

•I3: R34R6+R7 tidak lagi WAR dependency dengan I2

•I4: R35R8+R9 renaming kedua dari R1



Gambar 2. Speedups of Various Machine Organizations without Procedurel Dependencies.

Teknik dalam Superscalar

•Branch Prediction

•Program yang terdiri dari kelompok perintah bercabang sering digunakan dalam pemrograman.

•Pada CPU yang mendukung perintah pencabangan, CPU membutuhkan cukup banyak clock cycle

•Contoh: 80486 fetch keduanya pada instruksi rangkaian selanjutnya setelah cabang (branch) , dan target instruksi cabang (branch). Memberikan 2 cycle delay jika cabang (branch) diambil



Gambar 3. Superscalar Processing

SUPERSCALAR

**1. Pengertian**

**Superscalar** adalah sebuah unitprocessor yang dapat mengeksekusi dua atau lebih operasi scalar dalam bentuk paralel. Superscalar merupakan salah satu rancangan untuk meningkatkan kecepatan CPU.

Kebanyakan dari komputer saat ini menggunakan mekanisme superscalar ini. Standar pipeline yang digunakan adalah untuk pengolahan bilangan matematika integer (bilangan bulat, bilangan yang tidak memiliki pecahan), kebanyakan CPU juga memiliki kemampuan untuk pengolahan untuk data floating point (bilangan berkoma). Pipeline yang mengolah integer dapat juga digunakan untuk mengolah data bertipe floating point ini, namun untuk aplikasi tertentu, terutama untuk aplikasi keperluan ilmiah CPU yang memiliki kemampuan pengolahan floating point dapat meningkatkan kecepatan prosesnya secara dramatis.

Peristiwa menarik yang bisa dilakukan dengan metoda superscalar ini adalah dalam hal memperkirakan pencabangan instruksi (brach prediction) serta perkiraan eksekusi perintah (speculative execution). Peristiwa ini sangat menguntungkan buat program yang membutuhkan pencabangan dari kelompok intruksi yang dijalankankannya.

Program yang terdiri dari kelompok perintah bercabang ini sering digunakan dalam pemrograman. Contohnya dalam menentukan aktifitas yang dilakukan oleh suatu sistem berdasarkan umur seseorang yang sedang diolahnya, katakanlah jika umur yang bersangkutan lebih dari 18 tahun, maka akan diberlakukan instruksi yang berhubungan dengan umur tersebut, anggaplah seseorang tersebut dianggap telah dewasa, sedangkan untuk kondisi lainnya dianggap belum dewasa. Tentu perlakuannya akan dibedakan sesuai dengan sistem yang sedang dijalankan.

Lalu apa yang dilakukan oleh CPU untuk hal ini? Komputer akan membandingkan nilai umur data yang diperolehnya dengan 18 tahun sehingga komputer dapat menentukan langkah dan sikap yang harus diambilnya berdasarkan hasil perbandingan tersebut. Sikap yang diambil tentu akan diambil berdasarkan pencabangan yang ada.

Pada CPU yang mendukung perintah pencabangan ini, CPU membutuhkan lumayan banyak clock cycle, mengingat CPU menempatkan semuanya pada pipeline dan menemukan perintah berikutnya yang akan dieksekusinya. Sirkuit untuk branch prediction melakukan pekerjaan ini bekerja sama dengan pipeline, yang dilakukan sebelum proses di ALU dilaksanakan, dan memperkirakan hasil dari pencabangan tersebut.

Jika CPU berfikir bahwa branch akan menuju suatu cabang, biasanya berdasarkan pekerjaan sebelumnya, maka perintah berikutnya sudah dipersiapkan untuk dieksekusi berikut data-datanya, bahkan dengan adanya pipeline ini, bila tidak diperlukan suatu referensi dari instruksi terakhir, maka bisa dilaksanakan dengan segera, karena data dan instruksi yang dibutuhkan telah dipersiapkan sebelumnya.

Dalam hal speculative execution, artinya CPU akan menggunakan melakukan perhitungan pada pipeline yang berbeda berdasarkan kemungkinan yang diperkirakan oleh komputer. Jika kemungkinan yang dilakukan oleh komputer tepat, maka hasilnya sudah bisa diambil langsung dan tinggal melanjutkan perintah berikutnya, sedangkan jika kemungkinan yang diperkirakan oleh komputer tidak tepat, maka akan dilaksanakan kemungkinan lain sesuai dengan logika instruksi tersebut.

Teknik yang digunakan untuk pipeline dan superscalar ini bisa melaksanakan branch prediction dan speculative execution tentunya membutuhkan ekstra transistor yang tidak sedikit untuk hal tersebut.

Sebagai perbandingan, komputer yang membangkitkan pemrosesan pada PC pertama yang dikeluarkan oleh IBM pada mesin 8088 memiliki sekitar 29.000 transistor. Sedangkan pada mesin Pentium III, dengan teknologi superscalar dan superpipeline, mendukung branch prediction, speculative execution serta berbagai kemampuan lainnya memiliki sekitar 7,5 juta transistor. Beberapa CPU terkini lainnya seperti HP 8500 memiliki sekitar 140 juta transistor.

Superscalar ini mampu menjlankan Instruction Level Parallelism dengan satu prosesor. Superscalar dapat diaplikasikan di RISC dan CISC, tapi pada umumnya RISC.

**2. Alasan desain Superscalar**

Sebagian besar operasi menggunakan besaran/nilai skalar Operasi ini memungkinkan peningkatan kinerja sistem hingga level tertentu  Superscalar Implementation.

Proses fetch dari beberapa instruksi secara bersamaan. Logika untuk menentukan ketergantungan sebenarnya yang meliputi nilai register Mekanisme untuk mengkomunikasikan nilai tersebut. Mekanisme untuk menginisialisasi instruksi paralel. Tersedianya sumber untuk eksekusi paralel dari beberapa instruksi. Mekanisme processing instruksi dengan urutan yg sesuai.

**3. Contoh Instruksi**

|  |  |
| --- | --- |
|

|  |
| --- |
| **Add R1, R2, R3** |

 |

Misalkan, instruksi

yang akan menambahkan isi register R1 dan R2 dan menempatkan jumlahnya dalam register R3. Isi dari register R1 dan R2 mula-mula akan ditransfer ke Unit aritmetika dan logika. Setelah operasi penambahan dilakukan, hasil penjumlahan tersebut akan ditransfer ke register R3. Prosesor dapat membaca instruksi selanjutnya dari memori, sementara operasi penambahan dilakukan. Kemudian jika instruksi tersebut juga menggunakan ALU, operand-nya dapat ditransfer ke input ALU pada waktu yang sama dengan hasil instruksi **Add** ditransfer ke register R3.

Pada kasus ideal, jika semua instruksi ditumpuk ke derajat yang maksimum yang mungkin dilakukan, maka eksekusi dilanjutnkan pada kecepatan penyelesaian satu instruksi dalam tiap siklus detak prosesor. Instruksi individual, mungkin masih memerlukan beberapa siklus detak agar selesai dilakukan. Tetapi untuk tujuan perhitungan, prosesor superskalar umumnya mampu melakukannya dalam tiap siklus.

Prosesor superskalar umumnya menggunakan beberapa unit fungsional, menciptakan jalur paralel di mana berbagai instruksi yang berbeda dapat dieksekusi secara paralel. Dengan pengaturan tersebut, maka dimungkinkan untuk memulai eksekusi beberapa instruksi secara paralel tiap siklus detak. Tentu saja, eksekusi paralel harus mempertahankan kebenaran logikan program, sehingga hasil yang diperoleh harus sama dengan hasil dari eksekusi secara serial.

**4. Contoh CPU yang menerapkan arsitektur superscalar :**

* 486, Pentium, Pentium Pro, keluarga Intel Pentium, Intel Pentium Pro, Intel Pentium II, Intel Pentium III, Intel Itanium, Intel Xeon, Intel Pentium 4, Intel Pentium M, Intel Core dari Intel Corporation; keluarga AMD K5, AMD K6, AMD Athlon, AMD Athlon 64, dan AMD Opteron

Superscalar Processor Design

• Use PowerPC 604 as case study

• Speculative Execution, Register Renaming, Branch Prediction

More Superscalar Examples

• MIPS R10000

• DEC Alpha 21264

**5. Implementasi Superscalar**

1. Proses fetch dari beberapa instruksi secara bersamaan.
2. Logika untuk menentukan ketergantungan sebenarnya yang meliputi nilai register.
3. Mekanisme untuk mengkomunikasikan nilai tersebut.
4. Mekanisme untuk menginisialisasi instruksi paralel.
5. Tersedianya sumber untuk eksekusi paralel dari beberapa instruksi.
6. Mekanisme processing instruksi dengan urutan yg sesuai.

Pada penjelasan diatas bias diterangkan bahwa untuk superscalar dapat digunakan untuk berbagai keperluan dan dapat diimplementasikan pada perangkat prosessor seperti :

**Pentium 4**

Pada Pentium, implementasi superscalar dapat dijabarkan sebagai berikut :

* 80486 – CISC.
* Pentium.
* ada beberapa komponen superscalar.
* 2 unit eksekusi integer yang terpisah.
* Pentium Pro – Full superscalar.
* Memperhalus models subsequent & Meningkatkan design superscalar.

**6. Permasalahan pada Superscalar**

Kemampuan dari computer superscalar dapat dilihat dari banyaknya instruksi yang dapat diproses secara parallel. Superscalar hingga derajat tertentu dapat terjadi jika computer mampu mengatasi permasalahan ketergantungan yang ada pada suatu instruksi, diantaranya adalah : ketergantungan data, ketergantungan procedural, ketergantungan sumber unit, ketergantungan output dan antiketergantungan. Dari lima ketergantungan tersebut, sampai sekarang belum ada computer yang mampu mengatasinya secara total. Bagaimanapun computer akan menemukan beberapa ketergantungan dalam menyelesaikan suatu program.

Dari rancangan computer yang ada, hanya mampu meminimalisasikan beberapa ketergantungan saja, walaupun demikian ternyata dengan usaha ini mampu diciptakan computer dengan kinerja yang jauh lebih baik dari generasi-generasi sebelumnya. Adapun usaha yang dilakukan dalam peningkatan kinerja superscalar tersebut diantaranya : desain pipeline (*out of* *order issue and out of order* *completion*), memperbanyak tahapan pada pipeline, penduplikasian sumber unit (contoh : unit fungsional /ALU, memori, bus, cache, dll), desain micro operation (pada system CICS), desain unit prediksi percabangan, dan renaming allocation. Kebanyakan computer lebih menekankan pada upaya penduplikasian unit fungsional baik integer maupun floating point. Untuk itu pada pembahasan ini akan kita bahas salah satu upaya peningkatan kinerja prosesor superscalar secara khusus dalam hal reconfigurasi *floating point unit* (FPU).