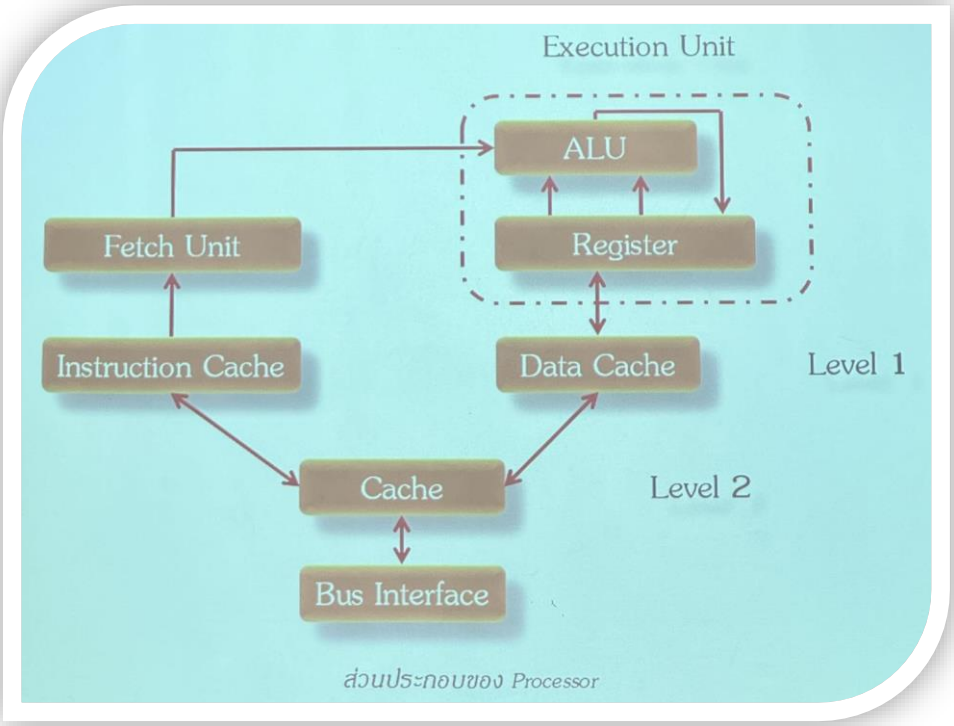

ภาพที่เห็นนี้เป็นแผนภาพที่แสดงส่วนประกอบของตัวประมวลผล (Processor) โดยแบ่งออกเป็นหน่วยย่อยต่าง ๆ ดังนี้

1. Fetch Unit : หน่วยนี้ทำหน้าที่ดึงคำสั่งจากหน่วยความจำ (Memory) มายังตัวประมวลผล
2. Instruction Cache: เป็นหน่วยความจำชั่วคราวที่ใช้เก็บคำสั่งที่ดึงมาจากหน่วยความจำหลักเพื่อรอการประมวลผล
3. Cache: เป็นหน่วยความจำความเร็วสูงที่อยู่ใกล้กับ CPU เพื่อเก็บข้อมูลที่ใช้บ่อย ช่วยลดเวลาการเข้าถึงข้อมูล
4. Bus Interface : เป็นส่วนที่เชื่อมต่อระหว่างตัวประมวลผลกับส่วนต่าง ๆ ของคอมพิวเตอร์ เช่น หน่วยความจำหลัก (RAM), อุปกรณ์ I/O และอื่น ๆ
5. Execution Unit : หน่วยที่ทำหน้าที่ประมวลผลคำสั่งต่าง ๆ ภายใน Execution Unit จะแบ่งเป็น
   - ALU (Arithmetic Logic Unit)*: หน่วยที่ทำหน้าที่ประมวลผลทางคณิตศาสตร์และตรรกศาสตร์
   - Register : หน่วยความจำภายใน CPU ที่มีความเร็วสูง ใช้เก็บข้อมูลชั่วคราวระหว่างการประมวลผล
6. Data Cache เป็นหน่วยความจำชั่วคราวที่ใช้เก็บข้อมูลที่ถูกดึงมาจากหน่วยความจำหลักเพื่อรอการประมวลผล
7. Level 1 & Level 2 : แสดงระดับของหน่วยความจำแคช โดยทั่วไป L1 Cache จะอยู่ใกล้กับตัวประมวลผลมากที่สุดและมีความเร็วสูงที่สุด ในขณะที่ L2 Cache จะมีความจุมากกว่าแต่ความเร็วต่ำกว่า L1 Cache

คำศัพท์เพิ่มเติมที่เกี่ยวข้องกับภาพนี้

- Processor : ตัวประมวลผลกลางหรือซีพียู เป็นหน่วยที่ทำหน้าที่ประมวลผลข้อมูลและคำสั่งต่าง
- CPU (Central Processing Unit) : ตัวประมวลผลกลางของคอมพิวเตอร์

- Memory : หน่วยความจำที่ใช้เก็บข้อมูลและคำสั่ง
- RAM (Random Access Memory) : หน่วยความจำหลักที่ใช้เก็บข้อมูลและคำสั่งชั่วคราว
- I/O (Input/Output) : อุปกรณ์หรือหน่วยที่ใช้ในการรับข้อมูลเข้าหรือส่งข้อมูลออก
- Instruction: คำสั่งที่ใช้ในการควบคุมการทำงานของตัวประมวลผล
- Data: ข้อมูลที่ใช้ในการประมวลผล
- Register File: กลุ่มของรีจิสเตอร์ภายใน CPU
- Microarchitecture : สถาปัตยกรรมภายในของ CPU ที่กำหนดวิธีการทำงานของหน่วยต่าง ๆ
- Pipeline : โครงสร้างการประมวลผลที่แบ่งการทำงานของคำสั่งออกเป็นขั้นตอนต่าง ๆ ซึ่งสามารถประมวลผลคำสั่งหลาย ๆ คำสั่งในเวลาเดียวกันได้
- Clock Speed : ความเร็วในการทำงานของ CPU วัดเป็นหน่วย Hz (Hertz) ซึ่งหมายถึงจำนวนรอบการทำงานต่อวินาที
- Core : หน่วยประมวลผลย่อยภายใน CPU ซึ่งสามารถทำงานแยกกันได้ในหลาย ๆ คำสั่งพร้อมกัน
- Multithreading : การประมวลผลหลาย ๆ เธรดพร้อมกันภายใน CPU
- Instruction Set Architecture (ISA) : ชุดของคำสั่งที่ CPU สามารถประมวลผลได้
- Branch Prediction : เทคนิคที่ใช้ในการคาดการณ์เส้นทางที่โปรแกรมจะดำเนินการต่อไป เพื่อลดความล่าช้าในการประมวลผล
- Out-of-Order Execution : การประมวลผลคำสั่งในลำดับที่แตกต่างจากที่ปรากฏในโปรแกรม เพื่อลดการรอคอยข้อมูลหรือทรัพยากร
- Superscalar : สถาปัตยกรรมที่สามารถประมวลผลหลายคำสั่งในเวลาเดียวกันภายในรอบนาฬิกาเดียวกัน
- Cache Coherency : การรักษาความสอดคล้องของข้อมูลในหน่วยความจำแคชที่ถูกแบ่งปันระหว่างหลาย ๆ คอร์หรือหลาย ๆ หน่วยประมวลผล
- Virtual Memory : เทคนิคการจัดการหน่วยความจำที่ทำให้โปรแกรมสามารถใช้หน่วยความจำได้มากกว่าที่มีอยู่จริง
- Memory Management Unit (MMU) : หน่วยที่จัดการการแปลงที่อยู่หน่วยความจำเสมือนเป็นที่อยู่หน่วยความจำจริง
- Prefetching : การดึงข้อมูลหรือคำสั่งล่วงหน้ามาเก็บในแคชเพื่อเตรียมพร้อมสำหรับการประมวลผล
- TDP (Thermal Design Power) : ค่าการใช้พลังงานสูงสุดที่ระบบระบายความร้อนต้องสามารถจัดการได้ในสภาวะการทำงานปกติ
- Micro-ops (Micro-operations) : การแบ่งคำสั่งระดับสูงเป็นคำสั่งระดับล่างเพื่อการประมวลผลภายใน CPU
- Scalar Processing : การประมวลผลคำสั่งเดียวในเวลาเดียวกัน
- Vector Processing : การประมวลผลคำสั่งเดียวที่สามารถดำเนินการกับชุดข้อมูลหลาย ๆ ชุดพร้อมกัน