สถาปัตยกรรมคอมพิวเตอร์
Computer Architecture

การออกแบบและการวิเคราะห์
Design and Analysis

ดร. เกริก ภิรมย์โสภา
Krerk Piromsopa, Ph.D.

ฉบับ พ.ศ. 2560

สถาปัตยกรรมคอมพิวเตอร์
Computer Architecture

การออกแบบและการวิเคราะห์
Design and Analysis

ดร. เกริก ภิรมย์โสภา
Krerk Piromsopa, Ph. D.

ฉบับปรับปรุง พ.ศ. 2559

สำหรับ คุณผกาทิพย์ เด็กชายดีดี และ เด็กหญิงลดา
ขอบคุณสำหรับความหมายของชีวิต

คำนำ

สถาปัตยกรรมคอมพิวเตอร์ (ภาษาอังกฤษใช้คำว่า computer architecture) เป็นพื้นฐานสำคัญที่นักคอมพิวเตอร์ทุกคนต้องรู้ บ่อยครั้งมักจะมีผู้โต้แย้งว่า แม้จะไม่มีความเข้าใจอันดีในสถาปัตยกรรมคอมพิวเตอร์ เราก็คงยังเป็นนักคอมพิวเตอร์ได้ คำกล่าวนี้คงจะไม่จริงนัก เพราะทุกอย่างที่เป็นฮาร์ดแวร์ของคอมพิวเตอร์ ตั้งแต่คำสั่งที่ใช้ได้ การจัดเก็บข้อมูล การประมวลผล คือสิ่งที่กำหนดโดยสถาปัตยกรรมคอมพิวเตอร์ทั้งหมด ในฐานะวิศวกรคอมพิวเตอร์ และ นักวิทยาศาสตร์คอมพิวเตอร์ นั้น การออกแบบ และ การพัฒนา ระบบสถาปัตยกรรมคอมพิวเตอร์เป็นหน้าที่หลักโดยตรง ในฐานะนักพัฒนาซอฟต์แวร์ แม้จะมิได้ออกแบบหรือพัฒนาสถาปัตยกรรมคอมพิวเตอร์โดยตรง แต่ความเข้าใจอันดีในสถาปัตยกรรมคอมพิวเตอร์ ย่อมช่วยให้สามารถพัฒนาซอฟต์แวร์ได้ดีขึ้น

เพื่อให้เห็นภาพชัดเจนยิ่งขึ้น จึงขอยกตัวอย่างประกอบดังนี้ ในช่วงที่ผมกำลังศึกษาระดับปริญญาเอกในต่างประเทศนั้นมีการเรียนการสอนวิชาอัลกอริทึม (algorithm) ซึ่งอาจารย์ผู้สอนในตอนนั้น ได้ให้ผู้เรียนแข่งขันทำโปรแกรมเป็นโครงงานรายวิชาเพื่อแก้ปัญหาเอ็นพีบริบูรณ์ (NP-Complete) ให้ได้เร็วที่สุด หากใครเคยศึกษาปัญหาเอ็นพีบริบูรณ์มาบ้าง ย่อมทราบว่าปัญหาในลักษณะดังกล่าวไม่สามารถหาคำตอบได้โดยใช้เวลาเป็นพหุนาม กล่าวคือการหาคำตอบที่เหมาะสมนั้นต้องทดสอบวัดค่าจากทุกกรณีที่เป็นไปได้เพื่อให้ได้ทราบคำตอบที่ดีที่สุด เมื่อทราบเช่นนั้น ผมก็คาดเดาได้ไม่ยากว่า โปรแกรมของเพื่อนร่วมชั้นเรียนแต่ละคน คงจะมีลักษณะไม่แตกต่างกันมาก เพราะแนวทางคร่าว ๆ คือ ต้องทำการวนซ้ำเพื่อวิเคราะห์กรณีที่เป็นไปได้ทั้งหมด และ นำคำตอบที่ดีที่สุดออกมาแสดง ในการแข่งขันครั้งนี้ ผมได้ใช้ความรู้พื้นฐานด้านสถาปัตยกรรมคอมพิวเตอร์ของตนเอง เลือกใช้โครงสร้างข้อมูลแบบที่เหมาะสมและสามารถประมวลผลได้เร็วบนสถาปัตยกรรมแบบ Intel Pentium 3 และ Intel Pentium 4 (ซึ่งเป็นสถาปัตยกรรมที่ดีที่สุดในขณะนั้น) เอาชนะกลุ่มเพื่อนหลายคนในชั้นเรียน จนทำเวลาได้เป็นอันดับที่สองของชั้นเรียน ซึ่งเพื่อนร่วมชั้นเรียนต่างก็เป็นนักศึกษาปริญญาโทและปริญญาเอกที่มีความสามารถสูง ในครั้งนั้นเพื่อนร่วมชั้นที่ได้ผลการทำงานเป็นที่หนึ่งใช้เทคนิคที่แตกต่างกันไปกับวิธีการที่ผมใช้ ซึ่งซับซ้อนกว่า เข้าใจได้ยาก และ ในบางกรณีทำงานได้ช้ากว่ามาก แต่เมื่อนับเวลาโดยรวมแล้วสามารถทำได้เร็วกว่า จึงเป็นผู้ชนะไป

ประสบการณ์ในครั้งนั้นชี้ให้เห็นว่า หากมีความเข้าใจในสถาปัตยกรรมคอมพิวเตอร์ดีพอ แม้การเขียนโปรแกรมง่าย ๆ ในภาษาระดับสูง ก็สามารถจะดึงสมรรถนะของหน่วยประมวลผลกลางออกมาใช้ได้ทุกหยด ต่างจากเพื่อนหลายคนในห้องเรียนครั้งนั้น ที่เขียนโปรแกรมลักษณะเดียวกันหรือคล้ายคลึงกับของผม แต่เนื่องจากความเข้าใจในสถาปัตยกรรม

คอมพิวเตอร์อาจไม่ลึกซึ้งมากพอ จึงไม่สามารถที่จะดึงความสามารถบางอย่างของหน่วยประมวลผลออกมาใช้ได้ ทำให้ได้สมรรถนะที่ได้แย่กว่าของผมมาก แน่นอนว่าผมมีประสบการณ์ที่คล้ายกันนี้ในหลายงานที่ผมเคยมีส่วนเกี่ยวข้อง สิ่งที่ยกขึ้นมานี้เป็นเพียงกรณีศึกษาอันหนึ่งเท่านั้น (เคยมีนิสิตมาถามผมว่าอัลกอริทึมที่เค้าใช้ก็เหมือนกับของผม แต่ทำไมของเค้าถึงช้ากว่าของผมมาก ผมอมยิ้มในใจแล้วตอบนิสิตว่า อีกครึ่งหนึ่งอยู่ที่โครงสร้างข้อมูลที่เลือกใช้ และการเลือกใช้คำสั่งที่เหมาะสมด้วยครับ)

ดังนั้นผมจึงอยากใช้หนังสือเล่มนี้ส่งข้อความถึง นิสิต นักศึกษา และ นักเรียนทั้งหลาย ว่าความเข้าใจอันดีในสถาปัตยกรรมคอมพิวเตอร์ นอกจากจะเป็นพื้นฐานสำคัญสำหรับการศึกษาด้านคอมพิวเตอร์ ยังเป็นสิ่งจำเป็นสำหรับการทำงานในสาขาวิชาชีพทางด้านคอมพิวเตอร์ทุกสาขา ผมหวังว่า การได้อ่านหนังสือที่ดี การได้ทดลอง(ออกแบบ) รวมถึงการวิเคราะห์ที่ดี จะช่วยให้สามารถศึกษาได้อย่างถ่องแท้มากขึ้น

หนังสือเล่มนี้พยายามวางพื้นฐานที่สำคัญ เริ่มจากความเข้าใจในสถาปัตยกรรมคอมพิวเตอร์ไปถึงการวิเคราะห์สถาปัตยกรรมคอมพิวเตอร์ที่หลากหลายแบบต่าง ๆ อย่างไรก็ตามศาสตร์ด้านนี้มีการเปลี่ยนแปลงอยู่ตลอดเวลา และ สถาปัตยกรรมที่ถูกมองว่าดีในเวลาหนึ่งมักจะถูกหักล้างในเวลาถัดมาด้วยเทคโนโลยีและความต้องการที่เปลี่ยนไป ผมจึงพยายามจัดวางเนื้อหา ให้เหมาะกับการสอนในหนึ่งภาคการศึกษา ซึ่งสอดคล้องกับหลายวิชา เช่น Computer System Architecture ซึ่งเป็นวิชาบังคับในหลักสูตรวิศวกรรมศาสตร์บัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ ของ ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย (ในบางหลักสูตรอาจจะใช้ชื่อวิชาว่า Computer Organization and Architecture) ซึ่งอาจจะไม่ครอบคลุมถึงสถาปัตยกรรมทุกรูปแบบ รวมถึงเนื้อหาในบางด้าน อาจจะไม่ลงรายละเอียดในเชิงลึกมากนัก แต่จะเน้นในเชิงวิเคราะห์และการนำไปใช้เป็นหลัก ในบางช่วงบางตอนที่ต้องการให้เห็นโครงสร้าง ก็จะทำการอธิบายด้วยแผนภาพ (ในการปรับปรุงหนังสือเล่มนี้ครั้งถัดไป ผู้เขียนวางแผนจะใช้ภาษา Verilog HDL และใช้การ Simulation เพื่อประกอบความเข้าใจ)

ทั้งนี้ขอออกตัวว่า ผมมิได้มองว่าตัวเองเป็นผู้เชี่ยวชาญด้านสถาปัตยกรรมคอมพิวเตอร์แต่อย่างใด (แน่นอนว่าคนที่ทำงานในสายวิชาชีพนี้ หากบอกว่าตัวเองเชี่ยวชาญในด้านใด คงจะเป็นความคิดที่ไม่ค่อยดีนัก เพราะศาสตร์ด้านนี้เปลี่ยนแปลงรวดเร็วทุกวันจนสิ่งที่เราเรียนรู้พร้อมจะตกรุ่นได้ทุกเวลา) แม้ผมจะมีงานวิจัยและได้เคยยื่นขอจดสิทธิบัตร (pending) เกี่ยวกับสถาปัตยกรรมคอมพิวเตอร์ที่มีความสามารถด้านความมั่นคงปลอดภัยอยู่บ้างก็ตาม หากแต่ผมเพียงอยากเห็นนักคอมพิวเตอร์รุ่นใหม่ (โดยเฉพาะกลุ่มที่เป็นคนไทย) เข้าใจสถาปัตยกรรมคอมพิวเตอร์และนำไปประยุกต์ใช้ประโยชน์ เพื่อให้สามารถพัฒนาโปรแกรมที่มีสมรรถนะมากขึ้น หรืออย่างน้อย ก็สามารถที่จะเลือกใช้สถาปัตยกรรมคอมพิวเตอร์ที่เหมาะสมกับงานที่ต้องการใช้ได้

จากความชอบส่วนตัวของผม และ ประสบการณ์ที่เกี่ยวข้องกับการเรียนการสอนด้านฮาร์ดแวร์และสถาปัตยกรรมคอมพิวเตอร์กว่าสิบปี (ผมเริ่มสอนวิชาสถาปัตยกรรมคอมพิวเตอร์ครั้งแรกในฐานะอาจารย์พิเศษที่มหาวิทยาลัยเอกชน ตั้งแต่ยังเรียนปริญญาโท) ผมจึงอยากถ่ายทอดประสบการณ์ และ ความรู้ ที่ส่วนหนึ่งได้จากค้นคว้าและการลองถูกลองผิด ให้กับ นิสิต นักศึกษาและคนรุ่นต่อไปได้ทราบ โดยหวังว่านอกจากจะเป็นการช่วยให้นิสิต นักศึกษา สามารถเรียนลัดได้เร็วขึ้นแล้ว ยังจะเป็นการสืบต่อองค์ความรู้ด้านสถาปัตยกรรมคอมพิวเตอร์ (ซึ่งปัจจุบัน นักพัฒนาซอฟต์แวร์รุ่นใหม่มักไม่ค่อยให้ความสนใจหรือไม่เห็นความสำคัญ) ให้เป็นพื้นฐานสำคัญ และยังคงอยู่ต่อไป

สุดท้ายนี้ผู้เขียนขอขอบคุณท่าน รองศาสตราจารย์ ดร. สมชาย ประสิทธิ์จูตระกูล ที่ได้ให้คำแนะนำด้านการเขียนและการจัดพิมพ์ หนังสือเล่มนี้ประสบความสำเร็จได้ส่วนหนึ่งเพราะคำแนะนำของท่านครับ

ดร. เกริก ภิรมย์โสภา

ผู้ช่วยศาสตราจารย์
ภาควิชาวิศวกรรมคอมพิวเตอร์
จุฬาลงกรณ์มหาวิทยาลัย
Krerk.P@chula.ac.th
1 ธันวาคม 2559

สารบัญ

ดรรชนีรูปภาพ

ดรรชนีตาราง

ดรรชนีสมการ

1 บทนำ

1.1 รู้ก่อนเรียน

เพื่อให้การศึกษาวิชาสถาปัตยกรรมคอมพิวเตอร์ง่ายขึ้น ผู้เรียนควรมีทักษะในการออกแบบวงจรตรรกะ หรือวงจรดิจิตอล โดยองค์ความรู้ทางวงจรดิจิตอลที่มีนั้น ควรจะเพียงพอให้สามารถออกแบบได้ทั้งวงจรแบบผสม (combinational logic circuit) และวงจรเชิงลำดับ (sequential circuit) นอกจากนี้ผู้เรียนควรจะมีความรู้พื้นฐานทางซอฟต์แวร์พอสมควรเพื่อประกอบความเข้าใจ กล่าวคือ ผู้เรียนควรมีประสบการณ์ในการเขียนโปรแกรม (ไม่จำกัดภาษา) ทั้งนี้หากผู้เรียนมีประสบการณ์กับภาษาแอสเซมบลี (assembly) ไม่ว่าจะเป็นสถาปัตยกรรมใด จะช่วยให้สามารถเข้าใจเนื้อหาบางส่วนได้ง่ายและรวดเร็วยิ่งขึ้นเป็นพิเศษ แต่หากไม่มีความคุ้นเคยกับภาษาแอสเซมบลี อย่างน้อยผู้เรียนควรจะเข้าใจหลักการเบื้องต้น เช่น ตัวแปร การจองหน่วยความจำ ประโยคเงื่อนไขและการเรียกใช้งานโปรแกรมย่อย ในกรณีของพื้นฐานความรู้ด้านการเขียนโปรแกรม ปัจจุบันไม่น่าจะเป็นประเด็นแต่อย่างไร เนื่องการจากเรียนการสอนทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์สมัยใหม่ มักสอดแทรกองค์ความรู้เกี่ยวกับการเขียนโปรแกรมเข้าไปอยู่แล้ว เช่น นิสิตนักศึกษาชั้นปีที่หนึ่งในหลายหลักสูตรทางด้านคอมพิวเตอร์ มักจะได้เรียนวิชาการทำโปรแกรมด้วยภาษาจาวา (Java programming) หรือ การทำโปรแกรมด้วยภาษาไพธอน (Python programming) และนิสิตนักศึกษาชั้นปีที่สองทางสายวิศวกรรมคอมพิวเตอร์ มักได้เรียนภาษาซี (C programming) และพื้นฐานการออกแบบวงจรตรรกะหรือวงจรดิจิตอล

เนื้อหาในหนังสือเล่มนี้เรียบเรียงมาจากตำราต่าง ๆ ทั้งฉบับคลาสสิก (ตำนาน) ฉบับอ้างอิง ฉบับอ่านเล่น รวมถึงงานตีพิมพ์และงานวิจัยด้านต่าง ๆ เนื่องจากสถาปัตยกรรมคอมพิวเตอร์เป็นศาสตร์ที่มีการพัฒนา ปรับปรุงและเปลี่ยนแปลงอยู่สม่ำเสมอ ผู้เรียนควรค้นคว้าเพิ่มเติมจากแหล่งความรู้อื่นประกอบด้วย

โครงสร้างและสถาปัตยกรรมคอมพิวเตอร์ที่อธิบายในที่นี้ จะอธิบายโดยอาศัยพื้นฐานจากสถาปัตยกรรมชุดคำสั่งสมมุติ ชื่อ nanoLADA ซึ่งเป็นฉบับย่อของ LADA (Light-weighted Architecture for Design and Analysis) เป็นหลัก โดย LADA เป็นสถาปัตยกรรมคอมพิวเตอร์ประเภท RISC แบบ 32 บิต ที่ผู้เขียนพัฒนาขึ้นเพื่อประกอบการเรียนการสอนโดยเฉพาะ (ผู้เรียนยังไม่ต้องกังวลว่า สถาปัตยกรรมคอมพิวเตอร์ประเภท RISC คืออะไร เมื่อถึงเนื้อหาส่วนที่เกี่ยวข้องก็จะเข้าใจเอง) สถาปัตยกรรม LADA มุ่งเน้นให้มีโครงสร้างเข้าใจได้ง่าย ไม่ซับซ้อน เหมาะกับการศึกษาในระดับเบื้องต้น ในส่วนที่

เป็นการกล่าวถึงเรื่องต่อยอดต่าง ๆ อาจมีการแทรกเนื้อหาของสถาปัตยกรรมอื่นเพื่อประกอบการอธิบายเพิ่มเติมบ้าง การแทรกรายละเอียดของสถาปัตยกรรมเหล่านี้ เน้นเพื่อให้เกิดการเปรียบเทียบและเห็นข้อแตกต่าง ทั้งนี้ผู้เขียนจะไม่แทรกเนื้อหาให้มากเกินไป โดยการสอดแทรกจะขึ้นกับความต่อเนื่องและความเหมาะสมของเนื้อหาเป็นหลัก ในตัวอย่างประกอบคำอธิบาย จะพยายามใช้สถาปัตยกรรม LADA เป็นพื้นฐาน (ยกเว้นกรณีที่ต้องการเปรียบเทียบกับสถาปัตยกรรมอื่น) เพื่อให้เกิดความต่อเนื่องของเนื้อหา

1.2 ส่วนประกอบของระบบคอมพิวเตอร์

ระบบคอมพิวเตอร์โดยทั่วไป ประกอบด้วยหน่วยประมวลผลกลาง แป้นพิมพ์ เม้าส์ จอภาพ เครื่องพิมพ์ ระบบสื่อประสม (multimedia) และอุปกรณ์ต่อพ่วงอีกมากมาย โครงสร้างดังกล่าวสามารถจำแนกตามการทำงานได้เป็น

- หน่วยประมวลผลกลาง (central processing unit)
- อินพุต (keyboard, mouse, etc…)
- เอาท์พุต (display, printer, etc…)
- หน่วยความจำ (memory)
- หน่วยจัดเก็บข้อมูล (storage, disk, cd, tape, etc)

การแบ่งในลักษณะนี้ เป็นการแบ่งส่วนประกอบของคอมพิวเตอร์ตามมุมมองของผู้ใช้งานทั่วไป ในทางสถาปัตยกรรมนิยมแบ่งส่วนประกอบของคอมพิวเตอร์เป็นสามส่วนหลักตามลักษณะของหน้าที่ดังแสดงในรูปที่ 1.1 ได้แก่ หน่วยประมวลผลกลาง หน่วยความจำ และระบบอินพุตเอาท์พุต จะสังเกตเห็นว่า อินพุต เอาท์พุต และ หน่วยจัดเก็บข้อมูล ถูกมองรวมเป็นระบบเดียวกันทั้งหมด ส่วนหนึ่งเนื่องจากในทางสถาปัตยกรรม อุปกรณ์เหล่านี้เชื่อมต่อกับหน่วยประมวลผลกลางผ่านระบบสายสัญญาณย่อยอินพุต/เอาต์พุต (I/O subsystem) เหมือนกัน ต่างจากหน่วยความจำที่มักจะมีสายสัญญาณเชื่อมต่อเฉพาะสำหรับหน่วยความจำ (memory bus)

หนังสือเล่มนี้มุ่งความสนใจไปที่หน่วยประมวลผลกลางเป็นหลักก่อน จากนั้นจึงกล่าวถึงองค์ประกอบรอบข้าง เช่น หน่วยความจำและส่วนต่อขยายต่าง ๆ เมื่อส่วนดังกล่าวมีผลกับสมรรถนะของระบบ จนส่งผลให้การออกแบบและการวิเคราะห์แตกต่างไป