หอสมุดกลาง จุฬาลงกรณมหาวิทยาลัย

บทที่ ๒

แบบจำลองข่ายแถวคอย (Queueing network model)

๒.๑ ปัญหาเกี่ยวกับแถวคอย

การเข้าแถวคอยเป็นปรากฏการณ์ที่พบเห็นได้เสมอในชีวิตประจำวัน เช่น การเข้าแถวคอยซื้อตั๋วชมภาพยนตร์ การเข้าแถวเพื่อคอยรับยาตามโรงพยาบาลต่างๆ เป็นต้น ในช่วงเวลาที่ต้องคอย เป็นเวลาที่เราสูญไปโดยเปล่าประโยชน์ ดังนั้นกิจการประเภทให้บริการทุกอย่าง ควรทำให้เวลาการคอยน้อยที่สุด นั่นคือ ความยาวของแถวคอย (Queue length) สั้นที่สุด

แถวคอย นอกจากจะเป็นคนแล้ว อาจเป็นสิ่งของที่รอรับการบริการก็ได้ เช่น เครื่องจักรที่รอการซ่อม โปรแกรมคอมพิวเตอร์ซึ่งรอการประมวลผล (process) จากเครื่องและอื่น ๆ ในที่นี้เราจะสนใจและพิจารณาเฉพาะแถวคอยที่เกิดขึ้นในระบบคอมพิวเตอร์ เท่านั้น

๒.๒ ระบบแถวคอย (queueing system)

แถวคอยจะเกิดขึ้นเมื่อความต้องการในการเข้ารับบริการมีมากเกินความสามารถในการให้บริการของผู้ให้บริการ โครงสร้างขั้นมูลฐานของระบบแถวคอย สามารถแบ่งออกได้ดังต่อไปนี้[๒๕]

๒.๒.๑ การเข้ามาของงาน (Arrival) อาจวัดเป็นอัตราการเข้ามา (Arrival rate) เช่น ๕ งานต่อนาที หรือวัดเป็น ช่วงห่างของเวลาระหว่าง ๒ งาน (Inter-arrival time) ซึ่งการเข้ามาของงานนี้อาจเข้ามาสม่ำเสมอคงที่ หรือเข้ามาไม่แน่นอน

๒.๒.๒ การให้บริการ อาจวัดเป็น

ก) อัตราในการให้บริการ (Service rate) เช่น ๑๐ งาน ต่อนาที

ข) เวลาที่ใช้ในการให้บริการ (Service time) เช่นใช้เวลา ๑ นาที ต่อ ๑ งาน

๒.๒.๓ เครื่องให้บริการ (Server)

อาจมีเครื่องเดียว (Single server) หรือหลายเครื่อง (Multiple servers)

๒.๒.๔ ระเบียบในการให้บริการ (Service discipline)

ที่สำคัญมีอยู่ด้วยกัน ๔ แบบ คือ

ก) แบบมาก่อนบริการก่อน (First Come First Serve, FCFS)

ข) แบบมาทีหลังบริการก่อน (Last Come First Serve, LCFS)

ค) แบบบริการโดยใช้วิธีสุ่มเอา (Service in Random Order, SIRO)

ง) แบบบริการตามลำดับความสำคัญ (Priority) ของงาน ซึ่งการจัดลำดับ ความสำคัญขึ้นอยู่กับ ความเหมาะสมในหลาย ๆ ด้าน แต่ทั้งนี้ย่อมต้องคำนึงถึง ความพอใจของลูกค้า ส่วนใหญ่ เป็นหลักด้วย

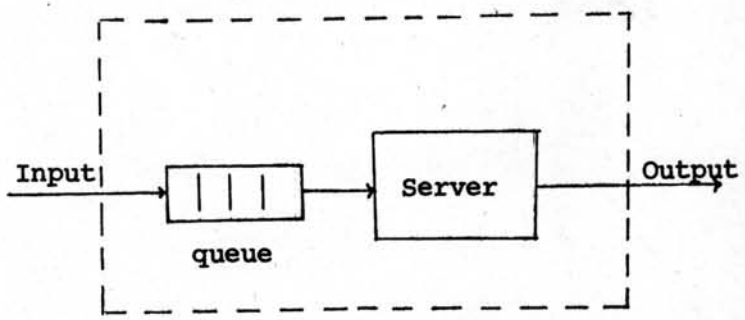
๒.๒.๕ จำนวนงานมากที่สุด ที่ยอมให้มีได้ในระบบ

ได้แก่จำนวนที่รอในแถวคอยรวมกับจำนวนที่กำลังรับบริการจากเครื่องให้บริการ

๒.๓ ลักษณะต่าง ๆ ของระบบแถวคอย

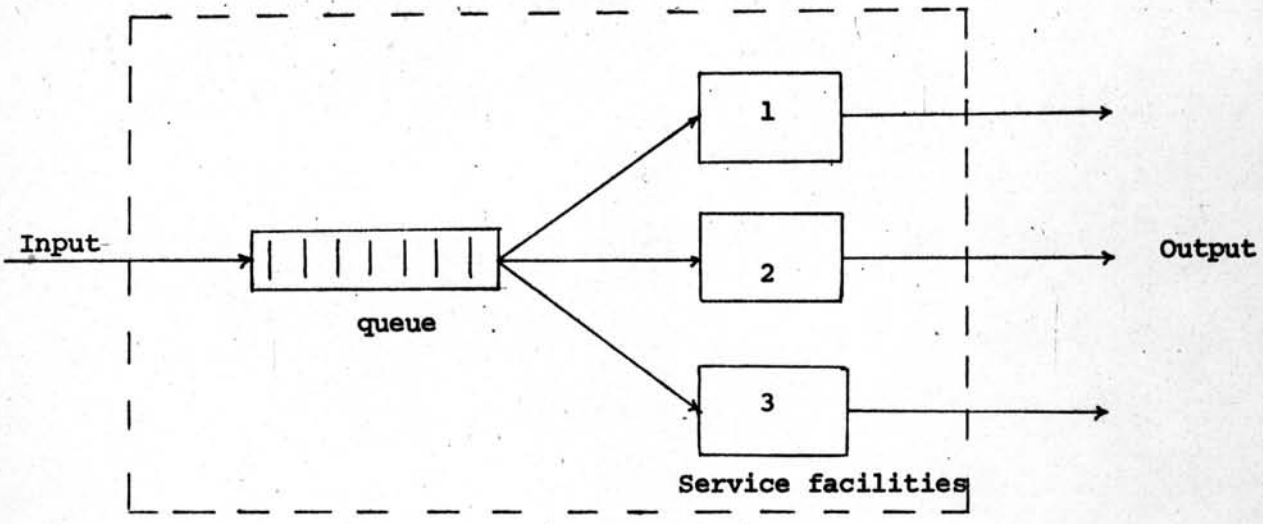
ระบบแถวคอยแบ่งเป็นลักษณะใหญ่ ๆ ได้ ๕ แบบ คือ

๒.๓.๑) แบบแถวคอยเดี่ยว และเครื่องให้บริการเดียว (Single queue, Single server) ระบบแถวคอยแบบนี้เป็นระบบขั้นพื้นฐานที่สุด ดังรูปที่ ๒.๑



รูปที่ ๒.๑ ระบบแถวคอยที่มีแถวคอยเดียว เครื่องให้บริการเดียว

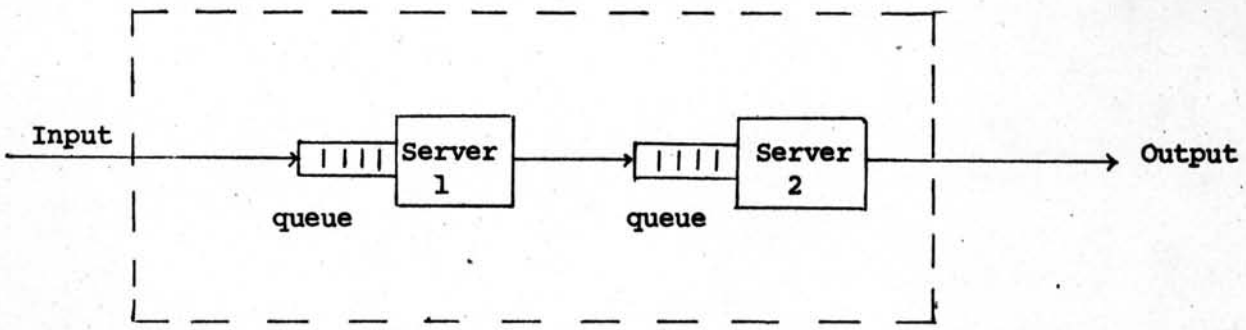
๒.๓.๒) แบบแถวคอยเดียวแต่มีเครื่องให้บริการมากกว่า ๑ เครื่องต่อขนานกัน (Single queue, Multiple servers in parallel) ดังรูปที่ ๒.๒



รูปที่ ๒.๒ ระบบแถวคอยที่มีแถวคอยเดียว เครื่องให้บริการหลายเครื่องต่อขนานกัน

จากรูปที่ ๒.๒ เมื่อเครื่องให้บริการเต็มหมด งานที่เข้ามาจะต้องรอในแถวคอย (queue) และเมื่อเครื่องใดว่าง งานที่อยู่ต้นแถว จะเข้ารับการบริการก่อน ถ้าหากเป็นการให้บริการแบบ FCFS

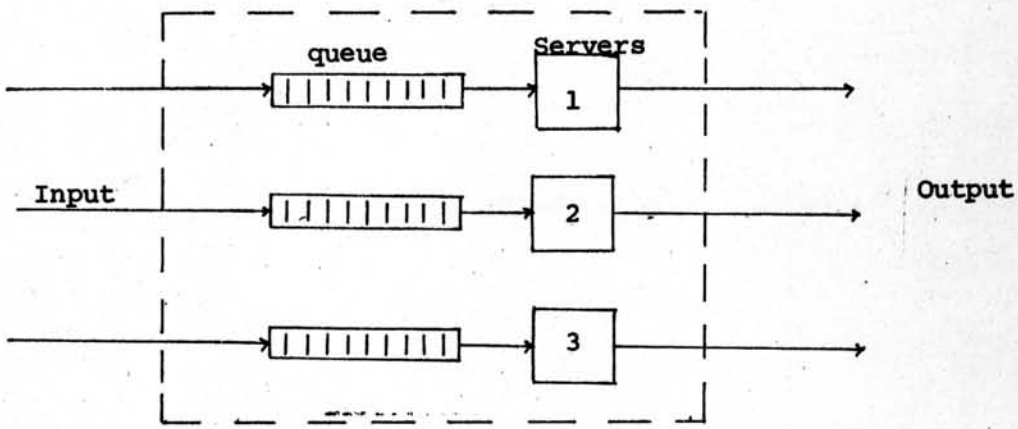
๒.๓.๓) มีแถวคอยแถวเดียวแต่มีเครื่องให้บริการหลายเครื่องต่ออนุกรมกัน (single queue, multiple servers in series) ดังรูปที่ ๒.๓



รูปที่ ๒.๓ ระบบแถวคอยที่มีแถวคอยเดียว เครื่องให้บริการหลายเครื่องต่ออนุกรมกัน

ในลักษณะนี้งานจะต้องผ่านทุก เครื่องให้บริการตั้งแต่เครื่องแรกจนเครื่องสุดท้ายจึงจะเสร็จการบริการ และออกจากระบบได้

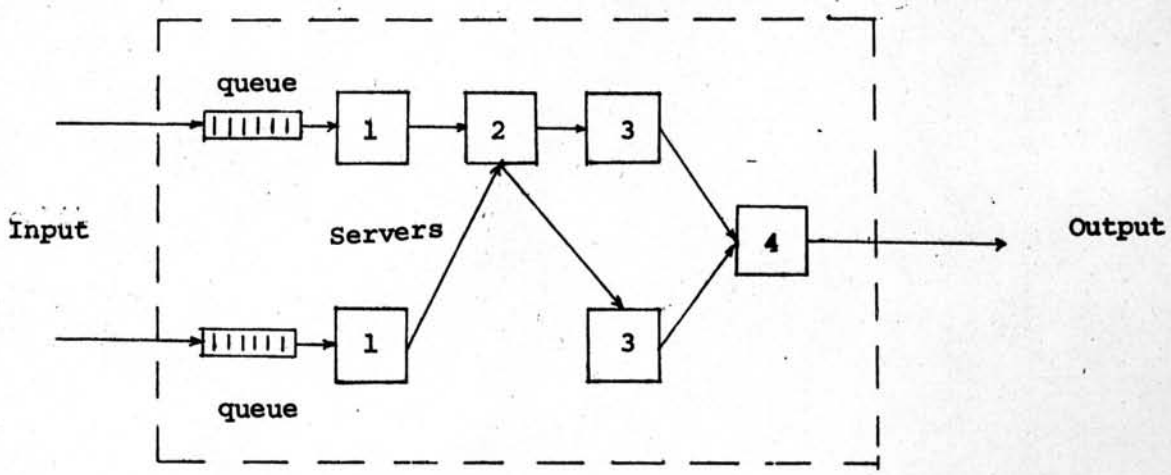
๒.๓.๔) มีหลายแถวคอย และหลายเครื่องให้บริการ (Multiple queues, Multiple servers) ดังรูปที่ ๒.๔



รูปที่ ๒.๔ ระบบแถวคอยที่มีหลายแถวคอย หลายเครื่องให้บริการ

จากรูปที่ ๒.๔ ภายในระบบ แบ่งเป็นระบบย่อย ๆ ซึ่งเป็นอิสระแก่กัน แต่งานที่เข้ามามาจากแหล่งเดียวกัน โดยทั่วไปลูกค้าจะไปเข้าแถวคอยแถวที่สั้นที่สุด เพราะคิดว่าจะได้รับการบริการเร็วกว่าไปรอในแถวคอยที่ยาว ซึ่งในบางครั้งก็ไม่แน่เสมอไป ถ้าหากความเร็วในการบริการของแต่ละเครื่องไม่เท่ากัน

๒.๓.๕) มีหลายแถวคอยและหลายเครื่องให้บริการ เช่นเดียวกับข้อ ๒.๓.๔ ต่างกันที่ระบบย่อยอาจมีการเกี่ยวข้องกันในการทำงาน เช่น งานในระบบย่อยระบบหนึ่งอาจต้องไปผ่านเครื่องให้บริการในอีกระบบย่อยหนึ่ง ดังรูปที่ ๒.๕



รูปที่ ๒.๕ ระบบแถวคอยที่มีหลายแถวคอย และหลายเครื่องให้บริการ

๒.๔ ทฤษฎีเกี่ยวกับแบบจำลองข่ายแถวคอย

สิ่งที่กระตุ้นให้มีการวิเคราะห์และพัฒนาแบบจำลองข่ายแถวคอยอย่างจริงจัง จนมีการใช้กันอย่างแพร่หลายในปัจจุบัน คือ ปัญหาเกี่ยวกับการวิจัยขั้นดำเนินงาน (operation Research, O.R.) ในปี ๑๙๕๗ Jackson[๖] ได้เขียนบทความที่เกี่ยวกับระบบมัลติเพิลดีไวซ์ (Multiple device) แต่ละดีไวซ์ (Device) มี ๑ เครื่องให้บริการหรือมากกว่า ๑ เครื่องให้บริการต่อขนานกัน โดยงานสามารถเข้าหรือออกจากระบบที่ใดก็ได้ ในบทความนี้ Jackson ได้กล่าวถึง ลูกค้า (Customer) ที่เข้ามารับบริการจากหน่วยงาน (department) ซึ่งถ้าเทียบกับในระบบคอมพิวเตอร์ ลูกค้าก็คืองานที่เข้ามารอรับบริการและหน่วยงาน ก็คือ ดีไวซ์ ซึ่งให้การบริการนั่นเอง

แบบจำลองข่ายแถวคอย ถูกนำมาใช้กับระบบคอมพิวเตอร์ได้เป็นผลสำเร็จครั้งแรกในปี ๒๙๖๕ โดย Scherr[๑๑] ได้ใช้แบบจำลองชนิดแคลสซิคอลมาชีนรีแพร์แมน (Classical machine repairman) ในการวิเคราะห์ ระบบไทม์แชริง (Time sharing) ของ MIT

และในปี ๑๙๗๑ Buzen ได้เสนอแบบจำลองแถวคอยรับบริการส่วนกลาง (Central server model) แอลกอริทึม (Algorithm) ที่ใช้ในการคำนวณค่าจากแบบจำลองได้อย่างรวดเร็ว และถูกต้องใกล้เคียงความจริง[๑๒,๑๓,๑๔]

แบบจำลองข่ายแถวคอย ได้รับการพัฒนาและใช้กันอย่างแพร่หลาย เนื่องมาจาก

(ก) สามารถใช้แทนระบบจริง (Actual System) ได้อย่างละเอียดและใกล้เคียงเกือบทุกส่วนของระบบ

(ข) ในการสมมติค่าต่าง ๆ เพื่อใช้ในการวิเคราะห์ ได้กระทำตามความเป็นจริง

(ค) ในการสมมติค่าต่าง ๆ จากแบบจำลอง สามารถทำได้อย่างมีประสิทธิภาพ

ในการใช้แบบจำลองข่ายแถวคอยเพื่อคำนวณค่าต่าง ๆ มีสิ่งที่ควรคำนึงถึงอยู่ ๓ ประการคือ

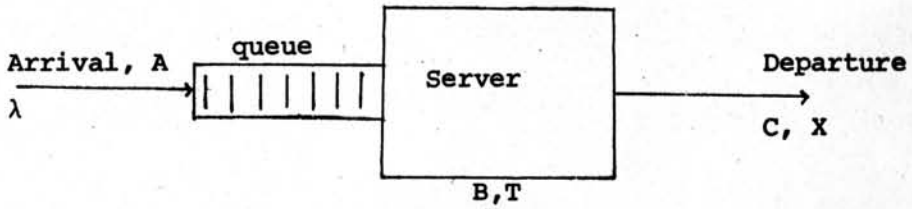
(ก) การกำหนดค่าใดค่าหนึ่ง ควรสามารถวัดค่านั้นได้อย่างถูกต้องและการสมมติค่า (assumption) ใดค่าหนึ่งก็ควรสามารถทดสอบค่านั้นได้โดยตรง (directly testable) ด้วย

(ข) ระบบต้องมีการถ่ายเทของงานที่สมดุลย์ (Job flow balanced) กล่าวคือ ที่ดีไวซ์ใด ๆ จำนวนงานที่เข้ามาในระบบจะต้องเท่า (หรือใกล้เคียงมากที่สุด) กับจำนวนงานที่ออกไปจากระบบในช่วงเวลาที่เราสังเกตการณ์ (observation period)

(ค) งานที่เสร็จการรับบริการจากเครื่องให้บริการหนึ่ง และจะไปยังเครื่องให้บริการอื่น โอกาสที่จะไปต้องไม่ขึ้นกับความยาวของแถวคอย

๒.๔.๑ แบบจำลองข่ายแถวคอยพื้นฐาน

ในระยะต้น ๆ ที่มีการวิเคราะห์เกี่ยวกับแบบจำลองข่ายแถวคอย ลักษณะของแบบจำลอง เป็นดังรูปที่ ๒.๖



รูปที่ ๒.๖ ระบบแถวคอยที่มี เครื่องให้บริการเดียว

จากรูปที่ ๒.๖ เราจะมองระบบเป็น แบลคบอกซ์ (Black box) คือจะไม่มีการแสดงให้เห็นลักษณะโครงสร้างภายใน เช่น การต่อวงจรหรือตำแหน่งของคิวซ์ต่าง ๆ ดังนั้นจึงไม่อาจจะแสดงความเกี่ยวข้องระหว่างพฤติกรรมที่เกิดขึ้นภายนอก (External behavior) กับส่วนประกอบต่าง ๆ ภายในระบบ (Internal Configuration) ได้ นอกจากนี้ในแบบจำลองยังได้สมมติว่าอัตราการเข้ามาของงานเป็นไปอย่างคงที่ และความยาวของแถวคอยจะยาวเท่าใดก็ได้ ไม่มีขีดจำกัด ซึ่งในระบบงานจริง ๆ มิได้เป็นเช่นนั้น

จากรูปที่ ๒.๖ สามารถคำนวณค่าพื้นฐาน (Basic quantities) ของระบบได้ดังนี้

กำหนดให้ T = ช่วงเวลาที่เราสังเกตการณ์ (observation period)

A = จำนวนงานที่เข้ามาในระบบในช่วงเวลาที่สังเกตการณ์

B = เวลาที่ระบบทำงานจริง ๆ (busy) ในช่วงเวลาที่สังเกตการณ์

C = จำนวนงานทั้งหมดที่ทำเสร็จและออกไปจากระบบในช่วงที่สังเกตการณ์

ค่าต่าง ๆ ที่จะคำนวณได้ คือ

λ = อัตราการเข้ามาของงาน

= A/T (Jobs/Sec) ----------2.1

X = อัตราขาออก (Output rate) ของงาน

= C/T (Jobs/Sec) ----------2.2

U = Utilization ของระบบ

= B/T ----------2.3

S = ค่าเฉลี่ยของเวลาที่ใช้ในการบริการ (Mean Service time) ต่อ ๑ งาน

= B/C, (Sec) ----------2.4

ค่า A,B,C เป็นค่าที่วัดได้ในช่วงเวลาสังเกตการณ์, T และค่า λ, X, U และ S เป็นค่าที่คำนวณได้ ซึ่งค่าต่าง ๆ เหล่านี้ อาจแตกต่างกันไป สำหรับช่วงเวลาสังเกตการณ์ที่ต่างกัน

จากสมการ 2.1-2.4 จะได้ว่า

U = XS ----------2.5

สมการ 2.5 เรียกว่า "Utilization law"

ถ้าหากการถ่ายเทของงานสมดุลย์

นั่นคือ A = C

∴ λ = X

และ U = λS ----------2.6

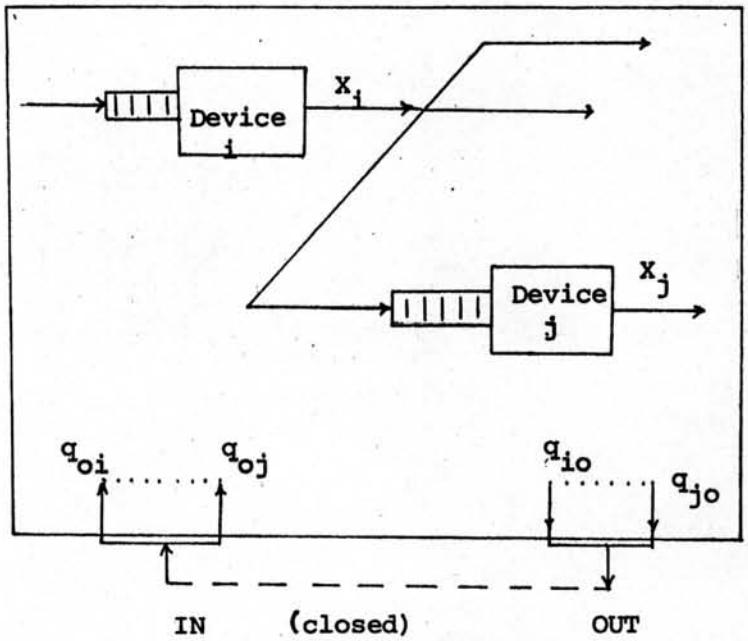
ในช่วงเวลาของการสังเกตการณ์ การถ่ายเทของงาน อาจไม่สมดุลย์ แต่เราสามารถประมาณได้ว่าสมดุลย์ ถ้าหากช่วงเวลาในการสังเกตการณ์นานพอ ที่จะทำให้ค่า $\frac{A-C}{C}$ มีค่าน้อยมาก

การสมมติว่าการถ่ายเทของงานสมดุลย์นี้เป็นตัวอย่างของการสมมติค่า ซึ่งสามารถทดสอบได้ กล่าวคือ ในบางช่วงของการสังเกตการณ์ การถ่ายเทของงานอาจไม่สมดุลย์ตามที่เราสมมติ แต่เราสามารถทดสอบได้ว่ามันสมดุลย์หรือไม่ และถ้าไม่สมดุลย์ก็สามารถวัดได้ว่า มี

ความผิดพลาดไปมากน้อยเท่าใด

๒.๔.๒ ชนิดของแบบจำลองข่ายแถวคอย

แบบจำลองข่ายแถวคอยในยุคต้น ๆ เป็นลักษณะซิงเกิลรีซ็อส (Single resource) (ดังได้กล่าวมาแล้วในหัวข้อ ๒.๔.๑) ซึ่งเหมาะที่จะแทนระบบคอมพิวเตอร์ที่ไม่ค่อยซับซ้อน ไม่สามารถนำมาใช้อธิบายการทำงานของระบบใหม่ ๆ ที่สลับซับซ้อนมาก ๆ ได้ จึงมีการปรับปรุงแบบจำลองให้เป็นแบบมัลติเพิลรีซ็อส (Multiple resources) ดังรูปที่ ๒.๗



รูปที่ ๒.๗ ระบบแถวคอย แบบมัลติเพิลรีซ็อส

จากรูปที่ ๒.๗ งานจะเข้าสู่ระบบที่จุด IN และสามารถเข้ารับบริการที่ดีไวซ์ใดก็ได้ เมื่อรับบริการเสร็จแล้วจะออกจากระบบที่จุด OUT

ลักษณะของแบบจำลองข่ายแถวคอยที่ใช้กันอยู่แบ่งเป็น ๓ แบบคือ

ก) แบบจำลองข่ายแถวคอยปิด (Closed queueing network model)

ข) แบบจำลองข่ายแถวคอยเปิด (Open queueing network model)

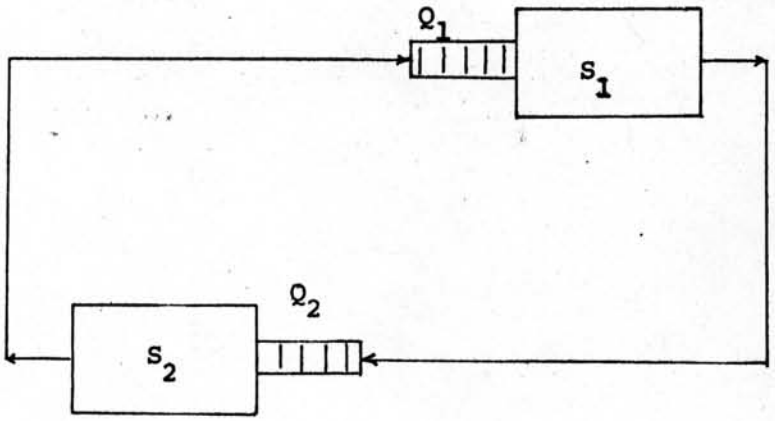
ค) แบบจำลองข่ายแถวคอยชนิดผสม (Mixed queueing network model)

ก) แบบจำลองข่ายแถวคอยปิด คือข่ายแถวคอยที่มีฟีดแบค (feedback) จาก OUT มาเข้าที่ IN และมีคุณลักษณะดังต่อไปนี้

(๑) จำนวนงานในระบบคงที่

(๒) อัตราขาออกของงาน สามารถหาค่าได้

(๓) ไม่มีเอ็กซ์เทอร์นอล ซ๊อส (External source)



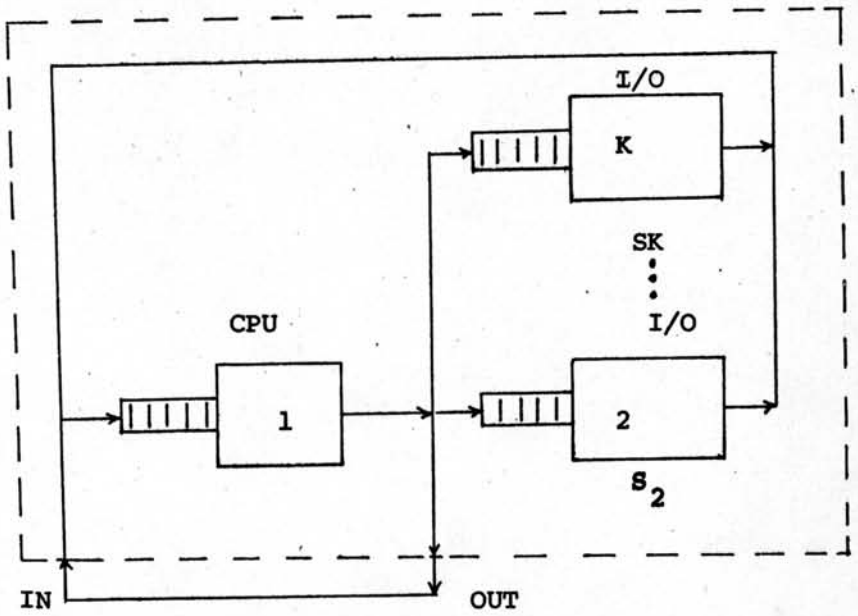
รูปที่ ๒.๔ แบบจำลองข่ายแถวคอยปิด

แบบจำลองข่ายแถวคอยปิดมักใช้กับระบบที่ยอมให้หลายงานทำพร้อมกัน (Multiprogramming) ซึ่งระบบคอมพิวเตอร์ที่ทันสมัยส่วนใหญ่สามารถทำงานแบบนี้ได้ แบบจำลองชนิดนี้จึงเหมาะมาก เพราะสามารถแทนค่าต่าง ๆ ในระบบได้อย่างละเอียดและใกล้เคียงกับสภาพของจริงแบบจำลองแบบนี้ ยังแบ่งได้เป็นแบบใหญ่ ๆ ได้ ๒ แบบคือ

(๑) แบบจำลองแถวคอยรับบริการส่วนกลาง (Central server model, CSM)

(๒) แบบจำลองแถวคอยวนเวียน (Cyclic queue model, CQM)

แบบจำลองแถวคอยรับบริการส่วนกลาง



รูปที่ ๒.๔ แบบจำลองแถวคอยแบบรับบริการส่วนกลาง

จากรูปที่ ๒.๔ บอกซ์ (box) ที่ ๑ ถึง K เป็นเครื่องให้บริการบอกซ์ที่ ๑ อาจเป็น CPU และบอกซ์ที่ ๒ ถึง K เป็นอุปกรณ์ขาเข้าขาออก (I/O device) แต่ละดีไวซ์ จะมีแถวคอยของตัวเอง จำนวนงานทั้งหมดในระบบจะเท่ากับผลรวมของจำนวนงานทั้งหมดในแต่ละดีไวซ์ เช่น ถ้าที่ดีไวซ์ ๑,๒, และ ๓ มีงานอยู่ ๓, ๐ และ ๑ ตามลำดับ โดยที่ดีไวซ์ ๑ จะมีงานอยู่ในแถวคอย ๒ งาน และกำลังรับบริการอยู่อีก ๑ งาน ดีไวซ์ที่ ๒ ไม่มีงาน จึงว่างอยู่เฉย ๆ (idle), ดีไวซ์ที่ ๓ มีงานกำลังรับบริการอยู่ ๑ งาน โดยไม่มีงานอยู่ในแถวคอย สรุปแล้วทั้งระบบจะมีจำนวนงานคงที่ ๔ งาน ถ้าหากมีงานใดงานหนึ่งทำเสร็จและออกจากระบบไปจะต้องมีงานอื่นเข้ามาแทนที่ทันที เพื่อให้งานในระบบมีจำนวนคงที่เสมอ

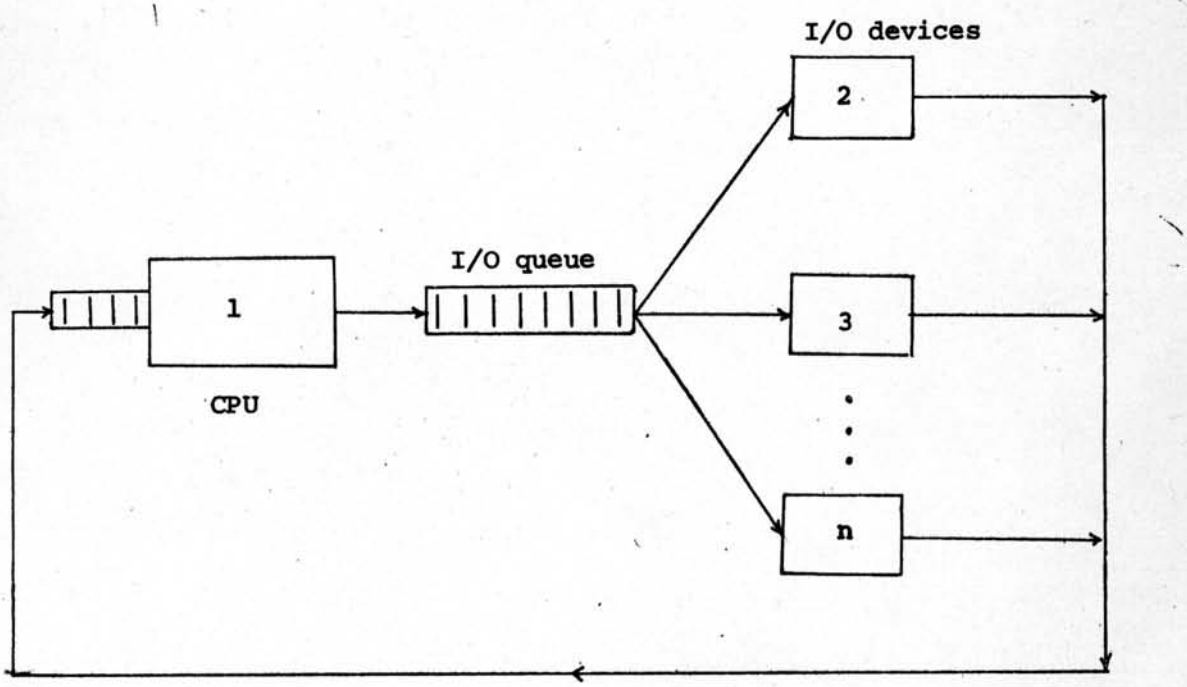
ในการรับบริการเมื่องานเสร็จจากดีไวซ์ที่ ๑ แล้ว อาจยังต้องไปรับการบริการต่อที่ดีไวซ์อื่น ๆ อีก โดยมันจะเลือกไปยังดีไวซ์ที่มันต้องการรับบริการ เช่น ต้องการพิมพ์ผล ก็ไปยังเครื่องพิมพ์ (printer) เป็นต้น ในกรณีที่มีดีไวซ์ที่เหมือน ๆ กันหลายเครื่อง การที่มันจะไปยังเครื่องใดย่อมขึ้นอยู่กับ โอกาส (probability) ที่จะไปยังเครื่องนั้น ว่ามีมากน้อยเพียงใด ซึ่งแทนด้วย q_{ij}

ในที่นี้ q_{ij} คือโอกาสที่งานจะไปยังดีไวซ์ j หลังจากเสร็จการรับบริการจากดีไวซ์ i

ถ้าหากไม่มีการเชื่อมโยงติดต่อกันระหว่าง ดีไวซ์ i และ j ค่า $q_{ij} = 0$ และถ้ามีเส้นทางเดียว จาก i ไป j ค่า $q_{ij} = 1$

<u>แบบจำลองแถวคอยวนเวียน</u>

ข้อแตกต่างของแบบจำลองแถวคอยวนเวียนกับแบบจำลองแถวคอยรับบริการส่วนกลาง คือแบบ จำลองแถวคอยวนเวียนมีแถวคอยที่ I/O เพียงแถวเดียว (ดังรูปที่ ๒.๑๐) ขณะที่แบบจำลองแถวคอยรับบริการส่วนกลางมีแถวคอยของตัวเองทุกดีไวซ์



รูปที่ ๒.๑๐ แบบจำลองแถวคอยวนเวียน

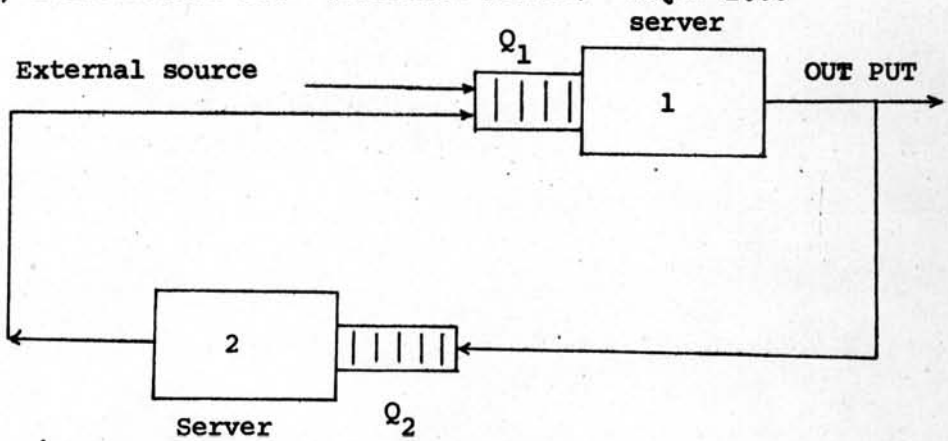
เราอาจเลือกแบบใดแบบหนึ่งใน ๒ แบบนี้ ก็เพียงพอที่จะใช้ทำนายคุณลักษณะของระบบ (system performance) ได้ ซึ่งทั้ง ๒ แบบให้ผลใกล้เคียงกัน(๑๕)

ข) แบบจำลองข่ายแถวคอยเปิด มีลักษณะที่สำคัญดังนี้

004946

(๑) รู้อัตราขาออกของงาน แต่จำนวนงานในระบบเปลี่ยนแปลงได้

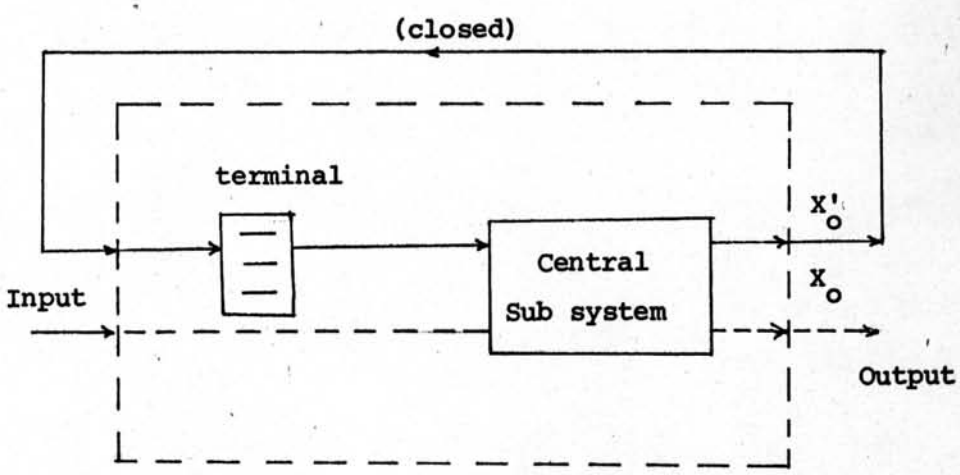
(๒) มีเอกซ์เทอร์นอล ซ็อส (External Source) ดังรูปที่ ๒.๑๑



รูปที่ ๒.๑๑ แบบจำลองแถวคอยเปิด

ค) แบบจำลองแถวคอยชนิดผสม

เป็นแบบที่มีทั้งแบบเปิดและแบบปิดผสมกัน กล่าวคือใช้แบบปิดสำหรับงานชนิดหนึ่ง และใช้แบบเปิดกับงานอีกชนิดหนึ่ง ดังตัวอย่างในรูปที่ ๒.๑๒



รูปที่ ๒.๑๒ แบบจำลองแถวคอยชนิดผสม

i 17420039

จากรูป ๒.๑๒ ใช้แบบปิดในกรณี อินเตอร์แอ็กทีฟเวิร์คโลด (Interactive workload) ซึ่งงานในระบบได้มาจากเทอร์มินอล (Terminal) ซึ่งมีจำนวนคงที่ = N ดังนั้น จำนวนงานในระบบ ย่อมคงที่ = N ด้วย สำหรับแบบเปิดใช้กับแบชท์เวิร์คโลด (Batch workload) งานที่เข้าสู่ระบบของแบบเปิดอาจได้มาจาก สถานีป้อนข้อมูลที่อยู่ห่างไกล (remote job entry station) โดยจำนวนงานในระบบไม่จำเป็นต้องคงที่

๒.๔.๓ <u>การถ่ายเทของงานในระบบ</u> (Job flow in the system)

กำหนดให้ A_o = จำนวนงานที่เข้ามาในระบบ

C_o = จำนวนงานที่ออกไปจากระบบ

C_{ij} = จำนวนครั้งที่งานขอรับบริการ (request service) ที่ดีไวซ์ j หลังจากเสร็จจากดีไวซ์ i

และ o เป็นส่วนที่อยู่ภายนอกระบบ ดังนั้นเราจะกำหนดว่า

A_{oj} = จำนวนงานที่เข้ามาในระบบเป็นครั้งแรกที่ดีไวซ์ j

C_{io} = จำนวนงานที่รับการบริการครั้งสุดท้ายที่ดีไวซ์ i แล้วออกจากระบบไป

$\therefore C_i$ = จำนวนงานที่เสร็จจากดีไวซ์ i

$$= \sum_{j=o}^{K} C_{ij} \quad \text{for } i = 1,\ldots, K \text{------}\rightarrow 2.8$$

$$A_o = \sum_{j=1}^{K} A_{oj} \text{------}\rightarrow 2.9$$

$$C_o = \sum_{i=1}^{K} C_{io} \text{------}\rightarrow 2.10$$

ในระบบที่เป็นแบบแถวคอยปิด ค่า $A_o = C_o$

สมดุลย์การถ่ายเทของงาน (Job flow balance)

กำหนดให้ S_i = ค่าเฉลี่ยของเวลาที่ใช้ในการบริการต่อ ๑ งาน

$= B_i/C_i$

q_{ij} = โอกาสที่งานจะไปยังดีไวซ์ j หลังจากเสร็จการรับบริการที่ดีไวซ์ i แล้ว

X_i = อัตราการเสร็จของงาน (completionrate) จากดีไวซ์ i

$= C_i/T$

X_o = อัตราของจำนวนงานที่ออกจากระบบ

R_i = เวลาตอบสนอง (response time)

A_i = จำนวนครั้งของงานที่เข้ามายังดีไวซ์ i

C_i = จำนวนครั้งของงานที่ออกไปจากดีไวซ์ i

จากหลัก การสมดุลย์ของการถ่ายเทของงาน (principle of job flow balance) ซึ่งกล่าวว่า

"งานที่เข้ามายังระบบจะเท่ากับงานที่ออกไปจากระบบ"

นั่นคือไม่มีความแตกต่างกันระหว่าง A_i และ C_i หรือต่างกันน้อยมาก ซึ่งดูได้จากความยาวของแถวคอย ตอนเริ่มสังเกตการณ์ (initial queue length) ควรจะเท่ากับความยาวของแถวคอย เมื่อเลิกสังเกตการณ์ (Final queue length)

จะได้สมการ $C_j = A_j = \sum_{i=o}^{K} C_{ij} \rightarrow j = o,1\text{------}K$

แต่ $q_{ij} = C_{ij}/C_i \rightarrow i = 1, \text{------}K$

$$\therefore \quad C_j = \sum_{i=o}^{K} C_i \, q_{ij} \longrightarrow \quad 2.11$$

$$\text{และ } \therefore \quad X_i = C_i/T \longrightarrow \quad 2.12$$

$$\therefore \quad X_j = C_j/T \longrightarrow \quad 2.13$$

แทนค่าสมการ 2.11 และ 2.12 ลงใน 2.13 จะได้

$$X_j = \sum_{i=o}^{K} X_i \, q_{ij} \rightarrow j = o\text{-----}K \rightarrow \quad 2.14$$

สมการ 2.14 เรียกว่า "Job flow balance equation"

<u>อัตราส่วนการขอรับบริการ</u> (Visit ratio, V)

คือ ค่าเฉลี่ยของจำนวนครั้งที่งานหนึ่ง ๆ ขอรับบริการจากดีไวซ์

ให้ V_i = จำนวนครั้งที่งานหนึ่ง ๆ มายัง ดีไวซ์ i

$$= X_i/X_o \longrightarrow \quad 2.15$$

หรือ $V_i = C_i/C_o \longrightarrow \quad 2.16$

จาก 2.15 จะได้ $X_i = V_i \, X_o \longrightarrow \quad 2.17$

สมการ 2.17 คือ "forced flow law" ซึ่งกล่าวว่า "การถ่ายเทของงานที่ส่วนหนึ่งส่วนใดของระบบสามารถแทนการถ่ายเทของงานในทุก ๆ ส่วนของระบบได้"

แทนค่า X_i ใน สมการ 2.14 จะได้

$$V_j = q_{oj} + \sum_{i=1}^{K} V_i q_{ij} \quad j = 1, \text{-----}K \qquad \text{..2.18}$$

และ $V_o = 1$

ในที่นี้ q_{oj} ก็คือโอกาสที่งานจากภายนอกจะเข้าสู่ระบบเป็นครั้งแรกที่ device j สมการ 2.18 เรียกว่า "Visit ratio equation"

<u>เวลาตอบสนองของระบบ</u> (System response time)

ให้ $n_i(t)$ = จำนวนงานทั้งหมดในระบบของ ดีไวซ์ i

= จำนวนงานในแถวคอย + จำนวนงานที่กำลังรับบริการ

W_i = พื้นที่ใต้กราฟของ $n_i(t)$ ในช่วงที่สังเกตการณ์

$\bar{n}_i$ = ค่าเฉลี่ยของ $n_i(t)$

$= W_i/T$ ⟶ 2.19

ให้ R_i = ค่าเฉลี่ยของเวลาตอบสนองที่ ดีไวซ์ i

$$R_i = \frac{W_i}{C_i} \qquad \longrightarrow \text{2.20}$$

แทนค่า $W_i = R_i C_i$ ใน 2.19

$$\therefore \bar{n}_i = \frac{R_i C_i}{T}$$

$$\bar{n}_i = R_i X_i \qquad \longrightarrow \text{2.21}$$

สมการ 2.21 เรียกว่า "little's law"

ถ้าต้องการหา เวลาตอบสนอง ของทั้งระบบ หาได้จากสมการ 2.22

$$R = \bar{N}/X_o \longrightarrow 2.22$$

โดย $\bar{N} = \bar{n}_1 + \bar{n}_2 + \cdots + \bar{n}_K$

ถ้าไม่รู้ค่า $\bar{N}$ และ X_o ใช้วิธี ทำโดย

$$\therefore \bar{n}_i = X_i R_i$$

แต่ $X_i = V_i X_o$

$$\therefore \bar{n}_i = V_i X_o R_i,$$

$$\bar{n}_i/X_o = V_i R_i$$

$$\bar{N}/X_o = R = \sum_{i=1}^{K} V_i R_i \longrightarrow 2.23$$

สมการ 2.23 เรียกว่า "General response time law"

สมการนี้ใช้ได้ทั้งระบบที่เป็นแบบ ปิด หรือ เปิด และในกรณีที่การถ่ายเทของงานไม่สมดุลย์

แต่ถ้าระบบเป็นแบบปิด และเวิร์คโลด (Workload) เป็นแบบอินเตอร์แอ็กทีฟ (Interactive) จะหาค่า เวลาตอบสนอง ได้จากสมการ

$$R = M/X_o - Z \longrightarrow 2.24$$

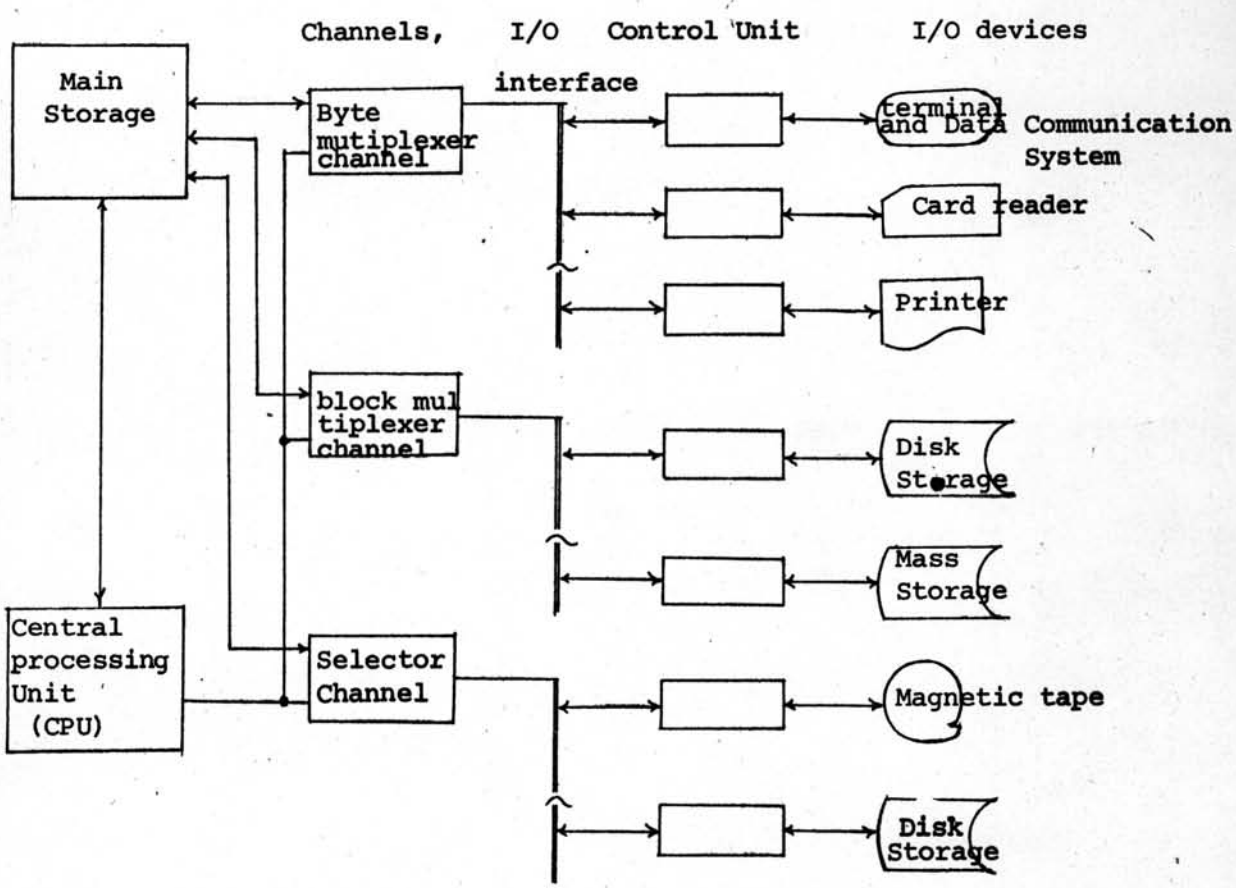
โดย M = จำนวนเทอร์มินอล

Z = เวลาที่ใช้ในการคิด (Think time) ว่างานต่อไปซึ่งจะป้อนเข้าสู่ระบบที่เทอร์มินอลเป็นงานใด

สมการ 2.24 เรียกว่า "Interactive response time formula"

๒.๕ โครงสร้างของคอมพิวเตอร์ระบบ IBM 370/138[๑๖,๑๗,๑๘]

เครื่องคอมพิวเตอร์ที่จะ ศึกษา เพื่อวัดคุณลักษณะของระบบคือ คอมพิวเตอร์ระบบ IBM 370/138 ซึ่งมีโครงสร้างดังแสดงในรูปที่ ๒.๑๗



รูปที่ ๒.๑๓ "โครงสร้างทั่วไปของระบบ IBM 370/138

๒.๕.๑ หน่วยความจำหลัก (Main memory)

คอมพิวเตอร์ระบบ IBM 370 จะมีความจุ (capacity) ตั้งแต่ 64 KB (65,536 bytes) ถึง 8 MB (8,388,608 bytes) ทั้งนี้ขึ้นอยู่กับรุ่นของเครื่องคอมพิวเตอร์สำหรับ รุ่น 138 มีหน่วยความจำหลักตั้งแต่ 512 KB (524,288 bytes) ถึง 1 MB (1,048,576 bytes)

๒.๕.๒ หน่วยบังคับการกลาง (Central processing Unit, CPU)

เป็นศูนย์กลางควบคุมการทำงานทั้งหมดของระบบ หน้าที่การทำงานของหน่วยบังคับการกลาง ควบคุมโดยไมโครโปรแกรม (Microprogram)

ในการ เอ็กซ์ซีคิว (Execute) คำสั่งต่าง ๆ หน่วยบังคับการกลางสามารถประมวลผลได้ในลักษณะเลขจำนวนเต็ม (Integer) และเลขทศนิยมที่มีความยาวคงที่ (Fixed length) หรือเลขฐานสิบ (decimal) ที่มีความยาวไม่คงที่ (Variable length) แต่ไม่เกิน ๑๖ หลัก หรือข้อมูลอื่น ๆ ที่มีความยาวคงที่หรือไม่ก็ได้ แต่ต้องไม่เกิน ๒๕๖ ตัวอักษร

คอมพิวเตอร์ IBM 370/138 มี ซี.พี.ยู ไซเกิลไทม์ (CPU cycle time) ต่าง ๆ กันคือ

(๑) เมื่อทำงานใน BC mode (Basic control mode) จะมีไซเกิลไทม์ (Cycle time) ต่ำสุด 275 ns. และสูงสุด 1430 ns. ขึ้นอยู่กับงานที่ทำ

(๒) เมื่อทำงานใน EC mode (Extended control mode) จะมีไซเกิลไทม์ สูงสุดเพิ่มเป็น 1485 ns.

๒.๕.๓ แชนแนล (Channels)

เป็นตัวควบคุม I/O devices และหน่วยควบคุม (control units) โดยตรง แชนแนล อาจแยกเป็นส่วนอิสระ หรือติดอยู่กับ CPU ก็ได้ แต่ละแชนแนลแบ่งออกเป็นสับแชนแนล (Subchannel) และแต่ละสับแชนแนลอาจควบคุม อุปกรณ์ขาเข้าขาออกได้มากกว่าหนึ่งหน่วย เรียก สับแชนแนล แบบนี้ว่า แชร์สับแชนแนล (Shared subchannel) บางสับแชนแนล ก็ควบคุมอุปกรณ์ขาเข้าขาออกได้เพียงหน่วยเดียว เรียกว่า นอนแชร์สับแชนแนล

(Non shared subchannel)

แชนแนล สามารถแบ่งออกได้เป็น ๓ แบบคือ

๑) ไบท์มัลติเพลกเซอร์แชนแนล (Byte-multiplexer channel) สามารถทำงาน (Operate) ได้ ๒ โมด (mode) คือ

(ก) ไบท์โมด (Byte mode) สามารถทำงานพร้อมกันได้ คราวละมากกว่า ๑ ดีไวซ์ แต่ความเร็วในการส่งข้อมูลจะต่ำ (low data rate) อุปกรณ์ที่ทำงานในโมด นี้ได้แก่ เครื่องอ่านบัตร, เครื่องเจาะบัตร (card punch) และเครื่องพิมพ์ (Printer) เป็นต้น

(ข) เบิสร์ทโมด (Burst mode) ยอมให้มีการทำงานได้ครั้งละ ๑ ดีไวซ์เท่านั้น โดยความเร็วในการส่งข้อมูลสูง (high data rate) อุปกรณ์ที่ทำงานใน mode นี้ได้แก่ เทปแม่เหล็ก (magnetic tape) และจานแม่เหล็ก (disk storage)

ไบท์มัลติเพลกเซอร์แชนแนล อาจมีจำนวน สับแชนแนล เป็น 64, 128 หรือ 256 สับแชนแนล

อัตราความเร็วสูงสุดในการส่งข้อมูลของ ไบท์มัลติเพลกเซอร์แชนแนลคือ 41 KB/S เมื่อทำงานในไบท์โมด และ 149 KB/S เมื่อทำงานใน เบิสร์ทโมด

๒) บลอคมัลติเพลกเซอร์แชนแนล (Block multiplexer channel) สามารถทำงานแบบใดแบบหนึ่ง ใน ๒ แบบ คือ

(ก) ซีเลคเตอร์โมด (Selector mode) ลักษณะการทำงานจะเหมือนกับการทำงาน ของ ซีเลคเตอร์แชนแนล (Selector channel) ซึ่งจะกล่าวต่อไป

(ข) บลอคมัลติเพลคโมด (Block multiplex mode) ยอมให้มีการส่งข้อมูลตั้งแต่ ๒ ชุดขึ้นไป (two or more message) โดยผ่านแชนแนลเดียวกันได้

ไบท์มัลติเพลกเซอร์แชนแนล ต่างจาก บลอคมัลติเพลกเซอร์แชนแนล คือ บลอคมัลติเพลกเซอร์ สามารถทำงานกับอุปกรณ์ขาเข้าขาออกที่มีความเร็วสูงมากได้ ขณะที่ ไบท์มัลติเพลกเซอร์แชนแนล ทำไม่ได้ นอกจากนี้ การส่งข้อมูล แต่ละครั้งก็ส่งได้จำนวนมากกว่าด้วย อัตราความเร็วสูงสุดในการส่งข้อมูล อยู่ในช่วงตั้งแต่ 1 , 2 ถึง 3 MB/S

ปกติ คอมพิวเตอร์ ระบบ 370 รุ่น 138 จะมี บลอคมัลติเพลกเซอร์แชนแนล 2 แชนแนล

๓) ซีเลขเตอร์แชนแนล สามารถส่งผ่านข้อมูลได้ครั้งละ 1 ดีไวซ์ โดยอาจเป็น ดีไวซ์ ความเร็วสูงหรือต่ำก็ได้ แต่เหมาะกับ ดีไวซ์ ที่มีความเร็วสูงมากกว่า เพราะทำงานใน เบิสร์ทโมด

แต่ละซีเลคเตอร์แชนแนล สามารถต่อกับ หน่วยควบคุมอุปกรณ์ขาเข้าขาออก (I/O Control unit) ได้ถึง 8 หน่วย และต่อกับดีไวซ์ได้ถึง 256 หน่วย อัตราการส่งผ่านข้อมูลสูงสุด ของซีเลคเตอร์แชนแนลจะอยู่ในช่วง 1.3 ถึง 1.85 MB/S

๒.๕.๔ <u>ส่วนเชื่อมโยงอุปกรณ์ขาเข้าขาออก</u> (I/O interface)

เป็นตัวเชื่อมโยงระหว่างอุปกรณ์ขาเข้าขาออกกับแชนแนล โดยผ่านทางหน่วยควบคุม

๒.๕.๕ <u>หน่วยควบคุม</u> (Control Unit)

เป็นตัวช่วยในการเชื่อมโยงระหว่างอุปกรณ์ขาเข้าขาออกกับแชนแนล หน่วยควบคุมอาจอยู่ในส่วนหนึ่งของอุปกรณ์ขาเข้าขาออก หรืออยู่กับหน่วยบังคับการกลางหรืออยู่แยกเป็นอิสระต่างหากก็ได้

๒.๕.๖ <u>อุปกรณ์ขาเข้าขาออก</u>

เป็นอุปกรณ์ที่อาจใช้งานในด้านต่าง ๆ คือ

๑) ใช้เป็น ออกซิลลารีสตอเรจ (Auxilary storage) เช่น เทป แม่เหล็ก ไดเรคแอ็กเสสสตอเรจดีไวซ์ (Direct access storage device (DASD)) เช่น แผ่นจานแม่เหล็ก และดรัม (Drum)

๒) เป็นตัวรับข้อมูลเข้า เช่น เครื่องอ่านบัตร เป็นต้น

๓) เป็นตัวแสดงผล เช่น จอภาพ (CRT) เครื่องพิมพ์