หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

การกำหนดแนวทางการเลือกเครื่องมือการประเมินคุณภาพและความสามารถ

3.1 กำหนดวัตถุประสงค์การประเมินและระดับคุณภาพและความสามารถ

การประเมินความสามารถหรือประสิทธิภาพ หรือคุณภาพการทำงาน หรือการดำเนินงานเป็นกิจกรรมที่สำคัญในทุกสาขาวิชาการและการปฏิบัติงาน ซึ่งจำเป็นต้องพิจารณาว่าเป็นที่สัมฤทธิ์ผลตามที่ได้กำหนดหรือได้วางแผนไว้ล่วงหน้าอย่างไร หลักวิธีการของการออกแบบระบบงานและขั้นตอนการประเมินผลนั้น จะถูกออกแบบให้รวมอยู่ในระบบงานนั้นด้วย โดยปกติการกำหนดประสิทธิภาพหรือความสามารถในการทำงานของระบบงานใด ๆ นั้นจะเปลี่ยนแปลงและขึ้นกับประเภทของระบบงาน ตลอดจนแนวทางการประเมินผลของผู้ออกแบบและประมินผลของระบบงานนั้นเอง ให้มีกระบวนการดำเนินงานอย่างถูกต้องและเหมาะสมกับความต้องการเบื้องต้นของระบบที่ได้กำหนดไว้

การกำหนดความสามารถของระบบงานก็เป็นไปตามระดับการยอมรับได้ของผู้ใช้ระบบงานนั้น และจะกำหนดว่ามีความสามารถได้ดีเพียงใดปกติจะขึ้นอยู่กับ การกำหนดความถูกต้องของระบบงานไว้เพียงใด ซึ่งถือเสมือนว่าระบบงานนั้นได้ดำเนินไปอย่างมีค่าแก่ผู้ใช้ระบบงานเพียงใด อาจจะเป็นค่าทางเศรษฐกิจ เป็นอรรถประโยชน์ หรือผลตอบแทนของระบบงานนั้นได้รับจากการกำหนดความสามารถดังกล่าว

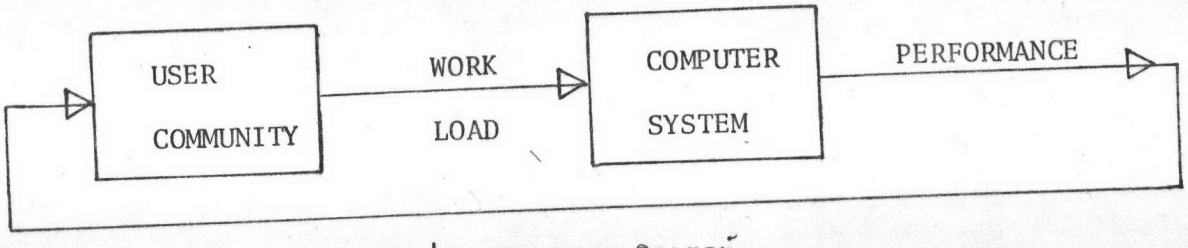
ประเภทของระบบในคอมพิวเตอร์มีหลายแบบด้วยกัน ตามชนิดหรือลักษณะธรรมชาติของระบบนั้น เช่น หน่วยงานคอมพิวเตอร์ (Computer Installation), ระบบงานคอมพิวเตอร์ (Computer System), ขอบข่ายงานคอมพิวเตอร์ (Computer Networks), ส่วนประกอบของระบบคอมพิวเตอร์ (System Components), โปรแกรมควบคุมระบบ (Operating System) โปรแกรมสำเร็จรูป (Package and Utilities) และโปรแกรมใช้งาน (Application Program) ความสามารถของระบบที่กล่าวมานั้นจะถูกประเมินโดย ผู้ออกแบบ, บริษัทผู้ผลิต, ผู้จัดการ ผู้ตรวจสอบบำรุงรักษา และผู้ใช้บริการ ดังนั้นในที่นี้จะขอกล่าวเฉพาะในส่วนระบบงานคอมพิวเตอร์ ซึ่งจะขอเน้นในส่วนประกอบฮาร์ดแวร์ (Hardware Components) และส่วนประกอบของซอฟท์แวร์ โดยจะขอเรียกส่วนประกอบทั้งสองลักษณะว่าเป็น ทรัพยากรของระบบ (Resources of the

System) ส่วนประกอบแต่ละหน่วยจะมีลักษณะเฉพาะของส่วนประกอบนั้น ซึ่งในที่นี้จะเรียกว่า พารามิเตอร์ของระบบงาน (System Parameter)

ต่อมาเราจะมาพิจารณาการกำหนดขอบเขตของระบบคอมพิวเตอร์ คำว่า "โปรแกรมควบคุมระบบการดำเนินงานของคอมพิวเตอร์" Computer Operating System เราอาจจะอธิบายได้ว่าเป็นกลุ่มของขั้นตอนหรือกระบวนการที่ทำขึ้นด้วยการกระทำของคนที่ทำขึ้นโดยอัตโนมัติที่จะจัดการ กลุ่มของผู้คนให้มาร่วมกันใช้หน่วยงานคอมพิวเตอร์ให้ได้อย่างมีประสิทธิภาพและประสิทธิผล ซึ่งหมายความว่าเราจะรวมโอเปอเรเตอร์ซึ่งปฏิบัติงานในห้องเครื่องคอมพิวเตอร์อยู่ในระบบงานของหน่วยงานคอมพิวเตอร์ด้วย

สำหรับหน่วยงานคอมพิวเตอร์ที่กล่าวถึงนั้น ประกอบไปด้วยระบบงานและสภาพแวดล้อมของระบบนั้น (Environment of a System) จะรวมถึงกลุ่มของผู้ใช้บริการ (User Community) เป็นบุคคลที่ใช้ระบบคอมพิวเตอร์ เพื่อที่จะได้ประโยชน์จากการประมวลผลข้อมูลและสารสนเทศ แต่ยังมีบุคคลอื่น เช่น ผู้บริหารหน่วยงาน สมาชิกในหน่วยงานประกอบเป็นส่วนของสภาพแวดล้อมของระบบด้วย

กลุ่มของผู้ใช้บริการ จะเป็นส่วนที่มีอิทธิพลต่อความสามารถของระบบในส่วนที่จะสร้างส่วนอินพุท (Input) ซึ่งจะเรียกอีกอย่างว่าภาวะงาน (Work Load) ซึ่งระบบคอมพิวเตอร์ต้องทำตัวเสมือนส่วนที่ผลิตผลลัพธ์ (Outputs) ของระบบงาน ดังรูปที่ 3.1



รูปที่ 3.1 รูปแบบของหน่วยงานระบบคอมพิวเตอร์

การกำหนดตัวชี้ความสามารถของระบบงาน จำนวนของตัวชี้ความสามารถของระบบนั้นจะกล่าวในรูปของคุณภาพที่ว่าง่ายต่อการใช้งาน มีลักษณะเป็นโครงสร้าง หรือคำสั่ง มีความสามารถในการทำงานสูงนั้น เป็นการวัดได้ยากและแทบจะเป็นไปไม่ได้ แต่ถ้าเรามากำหนดเป็นปริมาณของตัวชี้ความสามารถของระบบดังกล่าวเป็นตัวชี้ปริมาณผลผลิตคือ จำนวนหน่วยต่อเวลา, ตัวชี้ความเร็วของการตอบโต้ข้อมูล คือเวลาของการตอบโต้ข้อมูล และตัวชี้อรรถประโยชน์ของส่วนประกอบของ

ระบบ ซึ่งสิ่งเหล่านี้ยังไม่เป็นมาตรฐานโดยทีเดียว แต่ยังต้องพิจารณาควบคู่ไปกับระบบงานและภาระงานด้วย นั่นก็จะเป็นแนวทางให้เห็นว่าเราจะมาวัดเฉพาะภาระงาน หรือส่วนประกอบระบบงานอย่างใดอย่างหนึ่งไม่ได้ต้องพิจารณาร่วมกันทั้งสองสิ่ง

ตารางที่ 3.1 การแบ่งกลุ่มหลักในการกำหนดค่าดัชนีความสามารถและคุณภาพของระบบ (1)

กลุ่มดัชนี	ตัวอย่างค่าดัชนี	ข้อกำหนด
กลุ่มผลิตผล (Productivity)	. อัตราของทรูพุท (Throughput Rate) . อัตราผลิต (Production Rate) . อัตราความสามารถสูงสุด (Capacity or Maximum throughput Rate) . อัตราคำสั่งที่ทำงาน (Instruction Execution Rate) . อัตราการประมวลผล (Data Processing Rate)	การประมวลผลตามเวลา

ตารางที่ 3.1 (ต่อ)

กลุ่มดัชนี	ตัวอย่างค่าดัชนี	ข้อกำหนด
กลุ่มเวลาตอบสนอง (Responsiveness)	. เวลาการตอบสนอง (Response Time) . เวลาการทำงานจนได้งาน (Turnaround Time) . เวลาการตอบคำถาม (Reaction Time)	เวลาการสั่งอินพุทจนได้เอาท์พุท
กลุ่มอรรถประโยชน์ (Utilization)	. อรรถประโยชน์ฮาร์ดแวร์ (Hardware module utilization) . อรรถประโยชน์โปรแกรมควบคุมระบบ (Operating system module utilization) . อรรถประโยชน์ซอฟท์แวร์อื่น (Public software module utilization) . อรรถประโยชน์ฐานข้อมูล (Data base utilization)	อัตราส่วนระหว่างเวลากับระบบที่ใช้ในช่วงเวลา

3.2 การเลือกเครื่องมือการวัดและการประเมินผลความสามารถและคุณภาพ

การแบ่งการศึกษาการประเมินความสามารถของระบบคอมพิวเตอร์ จะแบ่งตามวัตถุประสงค์ของการศึกษาปกติแบ่งเป็น 3 แนวทางด้วยกัน ได้แก่

1. การศึกษาเพื่อคัดเลือก
2. การศึกษาเพื่อปรับปรุงงาน
3. การศึกษาเพื่อออกแบบงาน

การศึกษาเพื่อคัดเลือก จะถูกใช้กับกรณีที่มีปัญหาการออกแบบหน่วยงานระบบคอมพิวเตอร์และการจัดหาส่วนประกอบของคอมพิวเตอร์ ซึ่งอาจจะประกอบไปด้วย การคัดเลือกวิธีการประมวลผล, การคัดเลือกการติดตั้งเครื่องด้วยผู้ผลิตใด และการคัดเลือกส่วนประกอบซอฟท์แวร์ใด จะเห็นได้ว่าปัญหาเหล่านี้สามารถเริ่มพิจารณาได้จากวงการคอมพิวเตอร์ที่มีอยู่แล้ว หรือหน่วยงานที่คล้ายกันที่มีอยู่แล้วเป็นต้น

การศึกษาเพื่อปรับปรุงงาน จุดประสงค์เพื่อใช้กับหน่วยงานที่มีระบบคอมพิวเตอร์อยู่แล้วและต้องการที่จะเพิ่มความสามารถหรือลดค่าใช้จ่ายหรือทั้งสองอย่าง แนวทางการศึกษาแบบนี้มีวิธีการแก้ปัญหาได้โดยการปรับพารามิเตอร์ของส่วนประกอบที่จะให้เข้ากับภาระงานของหน่วยงานนั้น รวมไปถึงการเพิ่มความสามารถของระบบงาน โดยการเปลี่ยนหรือเพิ่มส่วนประกอบของฮาร์ด-แวร์เป็นต้น

การศึกษาเพื่อออกแบบ จุดประสงค์เพื่อออกแบบเครื่องคอมพิวเตอร์ ส่วนประกอบคอมพิวเตอร์, โปรแกรมควบคุม, โปรแกรม, และภาษาของโปรแกรม ซึ่งจะไม่เกี่ยวข้องกับผู้ใช้โดยตรง จึงจะไม่นำมากล่าวในที่นี้

ในการทำวิจัยครั้งนี้ จะทำการศึกษาเฉพาะเพื่อการปรับปรุงระบบงานคอมพิวเตอร์เท่านั้น เพื่อเป็นแนวทางในการใช้งานของผู้ใช้บริการระบบคอมพิวเตอร์ โดยแบ่งเทคนิคในการประเมินผลของความสามารถและคุณภาพเป็น 2 ส่วนดังนี้

3.2.1 เทคนิคการวัดระบบงานคอมพิวเตอร์ (Measurement Technique)

ในการวางแผนที่จะทำการศึกษาการประเมินความสามารถระบบคอมพิวเตอร์ต้องคำนึงถึงเทคนิคที่จะใช้ ซึ่งจุดประสงค์สำคัญประกอบด้วย

1. การตัดสินเลือกจะวัดความสามารถส่วนประกอบใด

2. การเลือกเครื่องมือในการวัดความสามารถ

3. ออกแบบการทดลองและประมาณค่าใช้จ่าย

ในการจัดตัวชี้ความสามารถและพารามิเตอร์ของระบบงานของหน่วยงานหนึ่ง ๆ จะทำโดยคำนวณค่านั้นจริง ๆ, ทำการวัดเป็นช่วงเวลาหรือเหตุการณ์ตามชนิดของข้อมูลที่เก็บรวบรวม ผลที่ได้จะเป็นการวัดโดยทางอ้อม คือการเก็บรวบรวมข้อมูลในช่วงเวลาหนึ่ง และมาทำการประมวลผลลดรูปแบบข้อมูลความสามารถ(Data Reduction Process) ภายหลัง การวัดตัวชี้ความสามารถเราจะแบ่งตามเหตุการณ์ของการเกิดได้เป็น 2 ลักษณะ คือ เหตุการณ์ภายในระบบ และเหตุการณ์ภายนอก ดังนั้นเราต้องมาพิจารณาถึงเครื่องมือในการวัดความสามารถ

<u>เครื่องมือการวัดความสามารถระบบคอมพิวเตอร์</u>

ระบบคอมพิวเตอร์มีเครื่องมือในการวัดความสามารถตามเหตุการณ์ ซึ่งมีความเร็วสูง และจะเก็บไว้ในสื่อข้อมูลอันหนึ่งอันใดแทนที่จะแสดงให้เห็นในขณะที่เกิดขึ้นจริง แต่ไม่ได้หมายความว่าไม่สามารถแสดงให้เห็นในขณะนั้นได้ เพียงแต่เครื่องมือที่ใช้นั้นเข้าไปใช้ทรัพยากรของระบบ (System Resources) ที่การปฏิบัติงานของระบบงานนี้ติดอยู่ เพราะการที่มีเครื่องมือเข้าไปติดอยู่นั้นอาจจะทำให้ความสามารถลดลง หรือการวัดได้ผลที่ไม่ตรงหรือคลาดเคลื่อนได้ ดังนั้น คุณสมบัติของเครื่องมือที่จะใช้วัดความสามารถควรจะมี ดังนี้

1. ความสามารถในการติดกับระบบที่จะทำให้เกิดการลดความสามารถในการทำงานของระบบงานและเก็บข้อมูลคลาดเคลื่อน

2. ความถูกต้องของข้อมูลที่วัด

3. ความเชื่อถือได้ในการใช้กับระบบงาน

4. ราคาของเครื่องมือดังกล่าว

5. ความง่ายต่อการใช้และการติดตั้ง

<u>ชนิดของเครื่องมือวัดความสามารถของระบบคอมพิวเตอร์</u> (Measurement Tool) (1)

การแบ่งชนิดของเครื่องมือวัดความสามารถระบบคอมพิวเตอร์แยกตามส่วนประกอบของระบบคอมพิวเตอร์ได้เป็น 3 ส่วนคือ ฮาร์ดแวร์ทูล (Hardware Tool), ซอฟท์แวร์ทูล

(Software Tool) และ เฟิร์มแวร์ทูล (Firmware Tool) และมาพิจารณาดูแต่ละชนิด

1. ฮาร์ดแวร์ทูล (Hardware Tools) เป็นเครื่องมือที่วัดภายในระบบที่จะตรวจจับเหตุการณ์ที่เล็ก ๆ ที่จะแสดงให้เห็นเป็นทางด้านไฟฟ้า โดยมีหลายชนิดด้วยกัน เป็นต้นว่า

1.1 อาร์ดแวร์ทูลเฉพาะ (Fixed Hardware Tool) เป็นเครื่องมือที่ใช้อ่านเหตุการณ์ภายในของระบบคอมพิวเตอร์ แต่ไม่สามารถควบคุมโดยผู้ใช้ได้

1.2 โปรแกรมวายด์ทูล (Wired-Program Tool) เป็นเครื่องมือที่ใช้ตรวจวัดที่มีระดับโวลต์เตสูง ทำให้ไม่ไปรบกวนระบบคอมพิวเตอร์และวัดเหตุการณ์ภายในและแสดงผลเป็น Counter

2. ซอฟท์แวร์ทูล (Software Tool) เป็นเครื่องมือที่ประกอบด้วยคำสั่งที่รวมกันเก็บ Hardware และ Software System ที่เก็บรวบรวมข้อมูลที่ต้องการ ซึ่งจะกล่าวถึงรายละเอียดต่อไป

3. เฟิร์มแวร์ทูล (Firmware Tool) ลักษณะเหมือน Software tool แต่ช่วงเวลาเก็บรวบรวมข้อมูลในช่วงเวลาที่สั้นกว่า

สำหรับการคัดเลือกเครื่องมือในการทำการศึกษาเพื่อปรับปรุงการดำเนินงานนั้น เมื่อเปรียบเทียบกับความเป็นไปได้และค่าใช้จ่ายแล้ว ฮาร์ดแวร์ทูลเป็นเครื่องมือที่มีราคาสูงมาก และไม่เหมาะสม ดังนั้น การศึกษาครั้งนี้จะพิจารณาเฉพาะซอฟท์แวร์ทูล

วิธีการของการประเมินผลความสามารถของระบบคอมพิวเตอร์ ดังที่กล่าวมาแล้วว่าหลักการของการประเมินผลความสามารถคือวิธีการวัดและตรวจสอบความสามารถด้วยฮาร์ดแวร์ทูลหรือซอฟท์แวร์ทูลที่ประกอบในระบบโปรแกรมควบคุมนั้น ๆ ซึ่งจะต้องมีวิธีการที่มีความถูกต้องและความเชื่อถือ

แต่ในการทำการประเมินความสามารถ ยังมีปัญหาในการวัดและตรวจสอบอยู่สองลักษณะคือ ความยากในการที่จะออกแบบ หรือพัฒนาขั้นตอนทั้งระบบ หมายความถึงระบบจะไม่สามารถวัดได้ถ้าหากระบบไม่สามารถปฏิบัติงาน และอีกลักษณะคือการวัดเป็นกิจกรรม ซึ่งจะมีขั้นตอนซับซ้อนเพราะต้องพิจารณารวมถึงต้นทุนทั้งด้านคนและเครื่องคอมพิวเตอร์ไปด้วยกัน

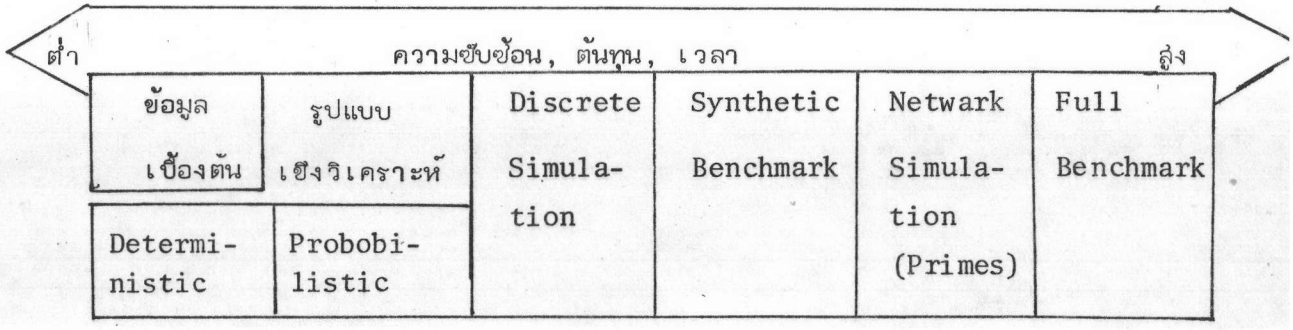
การกำหนดรูปแบบนั้น รูปแบบต้องสามารถจับส่วนสำคัญหรือพิจารณาในส่วนสำคัญความสามารถของระบบคอมพิวเตอร์ โดยรูปแบบของความสามารถจะมากำหนดลักษณะของระบบ

คอมพิวเตอร์ โดยรูปแบบของความสามารถจะมากำหนดลักษณะของระบบคอมพิวเตอร์ที่เป็นจริง ซึ่งขึ้นกับว่าเราวางแผนจะวัดความสามารถส่วนใดในรูปแบบนั้น ซึ่งรูปแบบจะเป็นการคาดคะเนเพื่อสัมพันธ์กับระบบงานจริง ถ้าได้รูปแบบที่แสดงถึงรายละเอียดมากเท่าใดยิ่งดี แต่ในด้านการจัดการให้ได้รูปแบบที่ละเอียดได้ จำเป็นต้องพิจารณาถึงค่าใช้จ่ายในด้านบุคคลากรและเครื่องมือ ดังนั้นจะพิจารณารูปแบบที่เป็นเบื้องต้น

ในขั้นของการออกแบบระบบคอมพิวเตอร์ในเรื่องการวัดและประเมินผลความสามารถ เราไม่จำเป็นต้องการความถูกต้องในการพยากรณ์หรือคาดคะเนส่วนบางส่วนอย่างสูงสุดมากไปกว่าให้ระบบปฏิเสธส่วนที่ผิดพลาดอันร้ายแรงจากการออกแบบที่จะได้วิธีการที่ดีที่สุด

3.2.2 เทคนิคการประเมินความสามารถและคุณภาพของระบบคอมพิวเตอร์

การประเมินผลความสามารถและคุณภาพเป็นการนำข้อมูลที่ได้จากการเก็บรวบรวมข้อมูล ความสามารถและคุณภาพมาทำการพัฒนาหารูปแบบสำหรับการพยากรณ์ ซึ่งข้อมูลจะได้มาในหลายรูปแบบและระดับของรายละเอียด รูปแบบอาจกำหนดแบบง่ายเพื่อเป็นพื้นฐานและแนวทางสำหรับการพัฒนาการตรวจสอบการดำเนินงานของระบบคอมพิวเตอร์ ดังนั้นสำหรับรูปแบบที่มีรายละเอียดมาตรฐานมากขึ้น และต้นทุนการผลิตสูงขึ้น ก็อาจจะได้รูปแบบที่เป็นจริงของระบบทั้งในส่วนฮาร์ดแวร์และซอฟท์แวร์ตามลักษณะของงานต่าง ๆ ที่เป็นจริง ความถูกต้องก็เกิดขึ้นจากการกำหนดลักษณะของงานที่ถูกต้องและสอดคล้องกับความถูกต้องของส่วนฮาร์ดแวร์และซอฟท์แวร์ที่ใช้ วิธีการกำหนดรูปแบบขึ้นอยู่กับความซับซ้อนของการเก็บรวบรวมข้อมูล ค่าใช้จ่ายด้านต่าง ๆ เวลาที่สูญเสียไป ดังรูปที่ 3.2 (10)



รูปที่ 3.2 แสดงความสัมพันธ์ระดับความซับซ้อนและวิธีการกำหนดรูปแบบ

ในการศึกษาครั้งนี้มีวิธีการกำหนดรูปแบบและตัวแปร 2 วิธีการ (1) คือ

1. รูปแบบจำลองเชิงวิเคราะห์ (Analytic Model)
2. การจำลองแบบ (Discrete Simulation)

3.2.2.1 รูปแบบจำลองเชิงวิเคราะห์ (Analytic Model)

รูปแบบจำลองเชิงวิเคราะห์เป็นการวิเคราะห์ที่ทำทฤษฎีของสถิติมาประยุกต์ ซึ่งประกอบด้วยระดับของรายละเอียดของข้อมูลที่มารวมกันในรูปแบบ วิธีการวิเคราะห์และความถูกต้องและเหมาะสมสำหรับการเลือกพารามิเตอร์ที่นำมาเข้าอยู่ในรูปแบบ

การวิเคราะห์จะแบ่งรูปแบบออกเป็น 2 ลักษณะ คือ

1 รูปแบบดีเทอร์มินิสติค (Deterministic Model) (1) เป็นรูปแบบที่แสดงถึงระบบคอมพิวเตอร์ต่าง ๆ ของภาระงานในแต่ละหน่วยงานคอมพิวเตอร์ พิจารณาและวิเคราะห์ถึงคุณสมบัติของตรรกวิทยาของการจัดการทรัพยากรที่มีการจัดปฏิทินที่เหมาะสมสำหรับการประมวลผลแบบมัลติโปรเซสเซอร์ (Optimal Schedule Multiprocessor system) และการจัดปฏิทินของการใช้หน่วยความจำที่สอง (Secondary Storage Device) โดยใช้รูปแบบของค่าเฉลี่ย (Mean Value Model) โดยพิจารณาค่าเฉลี่ยของภาระงานของงานซึ่งเป็นอินพุทของระบบที่เข้ามาด้วยค่าเฉลี่ยคงที่ (arrival rate)

2 รูปแบบความน่าจะเป็น (Probabilistic Model) (1) เป็นรูปแบบที่มีพื้นฐานทฤษฎีการรอคอย (Queuing Theory) ซึ่งนำการจัดกิจกรรมของการใช้และเกิดการรอคอยการใช้ทรัพยากรต่าง ๆ ของระบบคอมพิวเตอร์

3.2.2.1.1 รูปแบบดีเทอร์มินิสติค

หลักการของดีเทอร์มินิสติคใช้สำหรับพิจารณาและวิเคราะห์คุณสมบัติของระบบย่อย (Subsystem) หรือตรรกวิทยาของการจัดการทรัพยากรระบบซึ่งจะศึกษาในส่วนของการจัดปฏิทินที่เหมาะสมสำหรับระบบมัลติโปรเซสเซอร์และการวิเคราะห์การจัดขั้นตอนทำงานสำหรับหน่วยความจำที่สอง รูปแบบค่าเฉลี่ยจะนำมาใช้สำหรับภาระงานซึ่งประกอบด้วยงาน (job) ที่มีลักษณะเหมือนกันคือ ประกอบด้วยพารามิเตอร์ของภาระงานมาประกอบเป็นงานในรูปแบบเหมือนกัน ซึ่งจะสมมือนมีค่าเฉลี่ยคงที่เป็นค่าเฉลี่ยของการเข้ามาในระบบด้วยค่าอัตราคงที่ (mean arrival

rate) ซึ่งจะใช้เป็นค่าสำหรับการทำนายค่าดัชนีคุณภาพและความสามารถของระบบคอมพิวเตอร์

การพิจารณารูปแบบของซีพียูและหน่วยอินพุทและเอาท์พุท โดยมีการกำหนดพารามิเตอร์ดังตัวอย่างดังนี้

ตารางที่ 3.2 แสดงรายการพารามิเตอร์, สัญญลักษณ์และค่าของซีพียูและหน่วยอินพุทเอาท์พุท

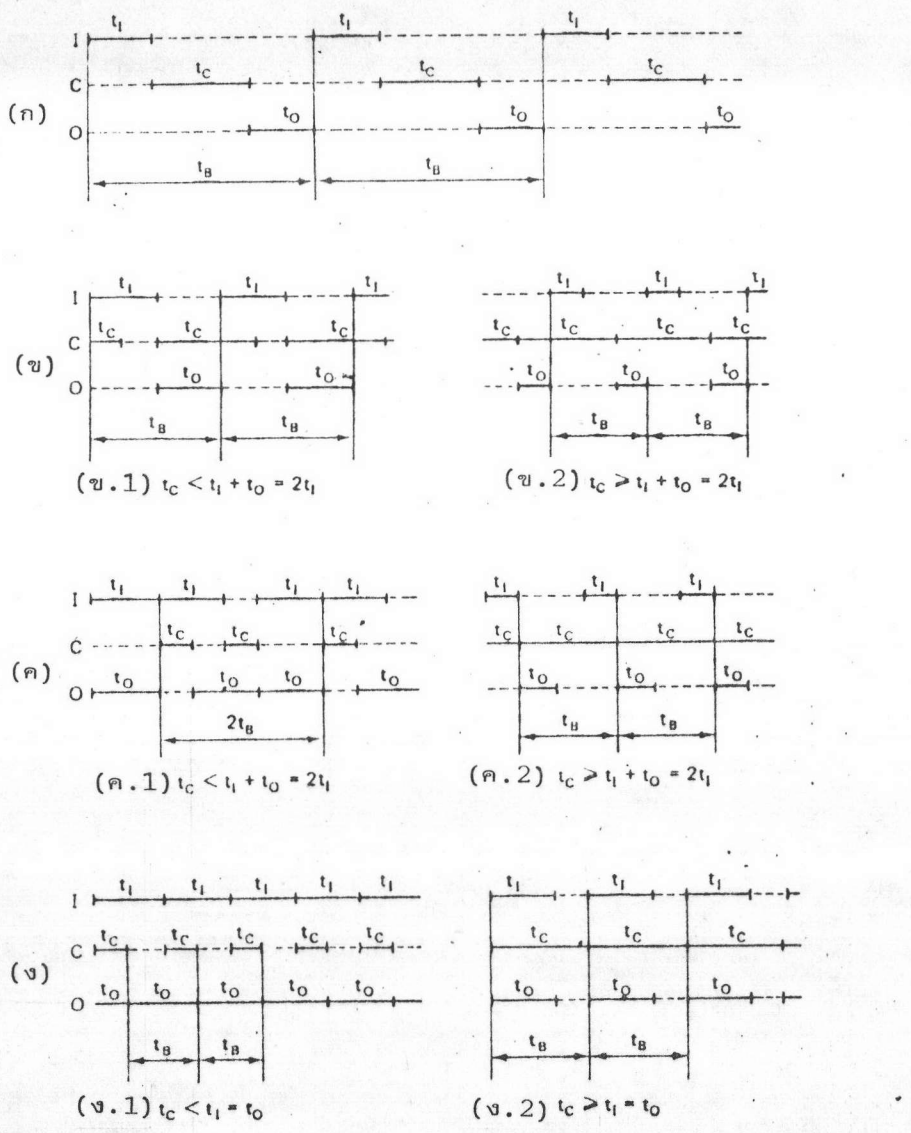
ชื่อพารามิเตอร์	สัญญลักษณ์	ค่า
ขนาดของหน่วยความจำ	M	2k, 4k, 6k,..... ไบท์
ขนาดของระเบียน	r	64 ไบท์
ขนาดของบล็อก	b	2k ไบท์
เวลาซีพียูประมวลผล 1 ระเบียน	t_{cpu}	1.6 ms (milli second)
เวลาที่ตั้งข้อมูลจากหน่วยความจำสำรอง	ta	15 ms
เวลาที่เปลี่ยนข้อมูลในหน่วยความจำสำรอง	t_w	10 (micro second)
จำนวนทาร์ติชั่นในหน่วยความจำหลัก	v	1, 2, 3,.....

กำหนด g คือ blocking factor ขนาดของบล็อกหารด้วยขนาดของระเบียน

$$g = \frac{b}{r}$$

หน่วยความจำหลักถูกแบ่งเป็นพาร์ติชั่นจำนวน v พาร์ติชั่นซึ่งแต่ละพาร์ติชั่นมีขนาดเท่ากับขนาดของบล็อก

$$v = \frac{M}{b}$$



รูปที่ 3.3 แสดงเวลาในการทำงานเหลื่อมกันของเวลาคำนวณเวลาอินพุทและเวลาเอาท์พุท

ลักษณะของการประมวลผลเลื่อมกันสำหรับซีพียูและหน่วยอินพุทเอาท์พุท (I/O)

ลักษณะ	ลำดับการเหลื่อมกันคือขนาดพารัติขึ้น
1. ตามลำดับ (ไม่มีการเลื่อมกัน)	1
2. ล่องทางเฉพาะ (C-I หรือ O-C)	2
3. ล่องทางขนาน (C-I หรือ)-C หรือ I-O)	2
4. ล่ามทางขนาน (I-C-O)	3

ถ้า V = 1 ดังรูปที่ 3.3 (ก) จะได้

เวลาของอินพุท t_I = ta + btw

เวลาของการประมวลผล $t_c = g\ t_{cpu}$

เวลาของเอาท์พุท t_o = ta + btw = t_I

เมื่อให้เวลาของการประมวลผลระเบียนหนึ่ง เป็นเวลา 1/T

เวลาที่การประมวลผลบล็อกหนึ่ง เป็นเวลา t_B

ด้านนี้อัตราของทรูพุทในหนึ่งบล็อกต่อหนึ่งหน่วยเวลา เป็น

$$\frac{1}{T} = \frac{t_B}{g}$$

ดังนั้นถ้ามี g ระเบียนในแต่ละบล็อก

$$t_B = t_I + t_c + t_o$$

$$t_B = gt_{cpu} + 2\ t_a + 2\ btw$$

ถ้า v = 2 และเป็นแบบล่องทางเฉพาะคือ ให้เกิดการทำงานของ ไอ/โอ และซีพียู พร้อมกันแต่เฉพาะซีพียูกับอินพุทหรือเอาท์พุทกับซีพียู เท่านั้น โดยทำอินพุทและเอาท์พุทพร้อมกันไม่ได้

ถ้าเกิดเหตุการณ์ ไอ/โอ บราวด์ (I/O Bound) หมายถึงการใช้เวลาสำหรับหน่วย ไอ/โอ มากกว่าซีพียู ดังรูปที่ 3.3 (ข.1)

$$t_c < t_I + t_o = 2\, t_I$$

ถ้าเกิดเหตุการณ์ ซีพียู บราวด์ (CPU Bound) หมายถึงการใช้เวลาสำหรับซีพียูมากกว่าหน่วย ไอ/โอ ดังรูปที่ 3.3 (ข.2)

$$t_c > t_I + t_o = 2\, t_I \quad .$$

ส่วนถ้า v = 2 และล่องทางขนานกันจะได้ลักษณะที่คล้ายกัน ดังรูปที่ 3.3 (ค.1, ค.2)

ส่วนถ้า v = 3 จะเกิดได้ 2 ลักษณะคือ

$t_c < t_I = t_o$ รูปที่ 3.3 (ง.1)

$t_c > t_I = t_o$ รูปที่ 3.3 (ง.2)

ตารางที่ 3.3 จะแสดงความสัมพันธ์ในรูปสมการ ดังนี้

รูปที่ 3.3	สมการ
ก.	$1/T = tcpu + 2\ ta/g + 2\ tw = 1 + 2y + 2x$
ข. 1	$1/T = 2t_a/g + rtw = y + 2x$
ข. 2	$1/T = tcpu = 1$
ค. 1	$1/T = tcpu/2 + ta/g + rtw = \frac{1}{2} + 2y + x$
ค. 2	$1/T = tcpu = 1$
ง. 1	$1/T = ta/g + rtw = 3y + x$
ง. 2	$1/T = tcpu = 1$

เวลาที่ซีพียูใช้ประมวลผลหนึ่ง เรคคอร์ดได้อัตราส่วน

$$= \frac{1}{T_{tcpu}}$$

ผลที่ได้จากตัวอย่างดังนี้ ตารางที่ 3.4

ตารางที่ 3.4 แสดงการกำหนดค่าคงที่และตัวแปรตามรูปที่ 3.3 และตารางที่ 3.2

กรณี	กำหนด b คงที่ ค่า M ตัวแปร		กำหนดค่า M คงที่ (2k) ค่า b ตัวแปร	
	T (ระเบียน/วินาที)	ρ	T (ระเบียน/วินาที	ρ
รูปที่ 3.3 (ก.)	261.9	0.419	261.9	0.419
รูปที่ 3.3 (ข.1)	450.9	0.721	317.2	0.507
รูปที่ 3.3 (ค.1)	523.9	0.838	420.6	0.672
รูปที่ 3.3 (ง.2)	625	1.000	488.7	0.781

เมื่อกำหนดให้ T_M ค่าสูงสุด = 625 ระเบียนต่อวินาที

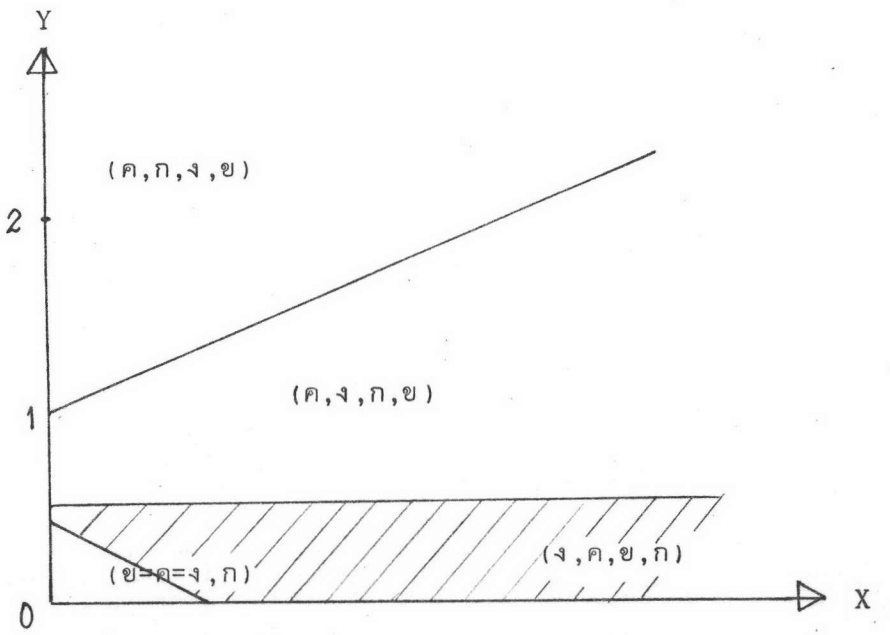
$$= T/T_M$$

ρ เป็นค่าดัชนีของความสามารถและคุณภาพโดยกำหนดค่าพารามิเตอร์ในรูปแบบเป็นค่าอรรถประโยชน์ของซีพียู ซึ่งสามารถกำหนดสมการได้เป็น

$$x = \frac{rtw}{t_{cpu}}$$

$$y = \frac{rta}{M t_{cpu}}$$

แทนค่า X, Y จากค่าดังกล่าวเรียงลำดับของการกำหนดวิธีการที่ ง. จะดีที่สุด เรียงตามมาด้วย ค, ข และ ก. ดังรูปที่ 3.4

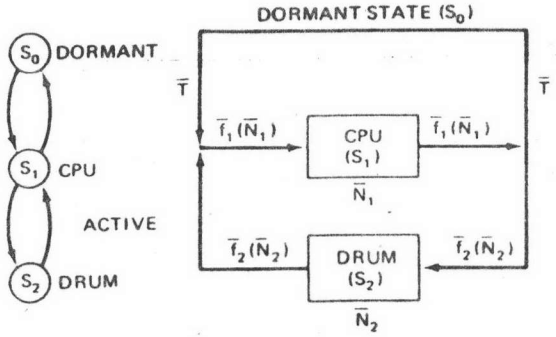


รูปที่ 3.4 แสดงการแบ่งเขตของการลำดับความสามารถและคุณภาพที่เหลื่อมกัน

การใช้รูปแบบค่าเฉลี่ยของระบบมัลติโปรแกรมมิง จากรูปที่ 3.5 แสดงถึงการทำงานแบบ Single Job และรูปแสดงถึงระบบมัลติโปรแกรมมิง ซึ่งกำหนดให้งาน (Job) ทุก ๆ งานเหมือนกัน การใช้เวลาซีพียูเป็น t_{cpu} มีการเรียกใช้ Drum N_{dr} และเมื่องานประมวลผลด้วย CPU เป็นเวลา $t_{cpu}/(N_{dr} + 1)$ จนจะจบงานให้ N_1 เป็นจำนวนงานโดยเฉลี่ยของสภาพ S_1 และให้ N_1 เป็นจำนวนงานโดยเฉลี่ยของสภาพ S_2

ดังนั้น $$N_1 + N_2 = N$$

ค่าเฉลี่ยของงานที่เข้าระบบงานที่ออกจากระบบ (ซีพียูและดรัม) เท่ากัน เมื่อทุกงานที่เข้าสภาพ S_1 เวลา $N_{dr}+1$ เวลา N_{dr} งานก็จะออกจาก S_1 และงานเข้าสภาพ S_2 และเวลาสุดท้ายก็ไปที่สภาพ S_o ดังนั้น



รูปที่ 3.5 แสดงรูปแบบของมิลติโปรแกรมมิง และสภาพของการทำงาน

จะได้อัตราการทำงานสูงสุด (Maximum througput rate)

$$\bar{f}_2(N_2) = \frac{N_{dr}}{N_{dr}+1}\ \bar{f}_1(N_1)$$

$$\bar{T} = \bar{f}_1(N_1) - \bar{f}_2(N_2) = \frac{\bar{f}_1(N_1)}{N_{dr}+1}$$

$$\bar{f}_1(\bar{N}_1) = 0 \quad \text{เมื่อ} \quad \bar{N}_1 = 0$$

$$= \frac{N_{dr}+1}{t_{cpu}} \quad \text{เมื่อ} \quad \bar{N}_1 \gtrsim 1$$

$$= \frac{\bar{N}_1\ N_{dr}+1}{t_{cpu}} \quad \text{เมื่อ} \quad 0 \lesssim N_1 \lesssim 1$$

จะได้อัตราการทำงานสูงสุด (Maxinum throughput rate)

$$T_{max} = \frac{1}{t_{cpu}}$$

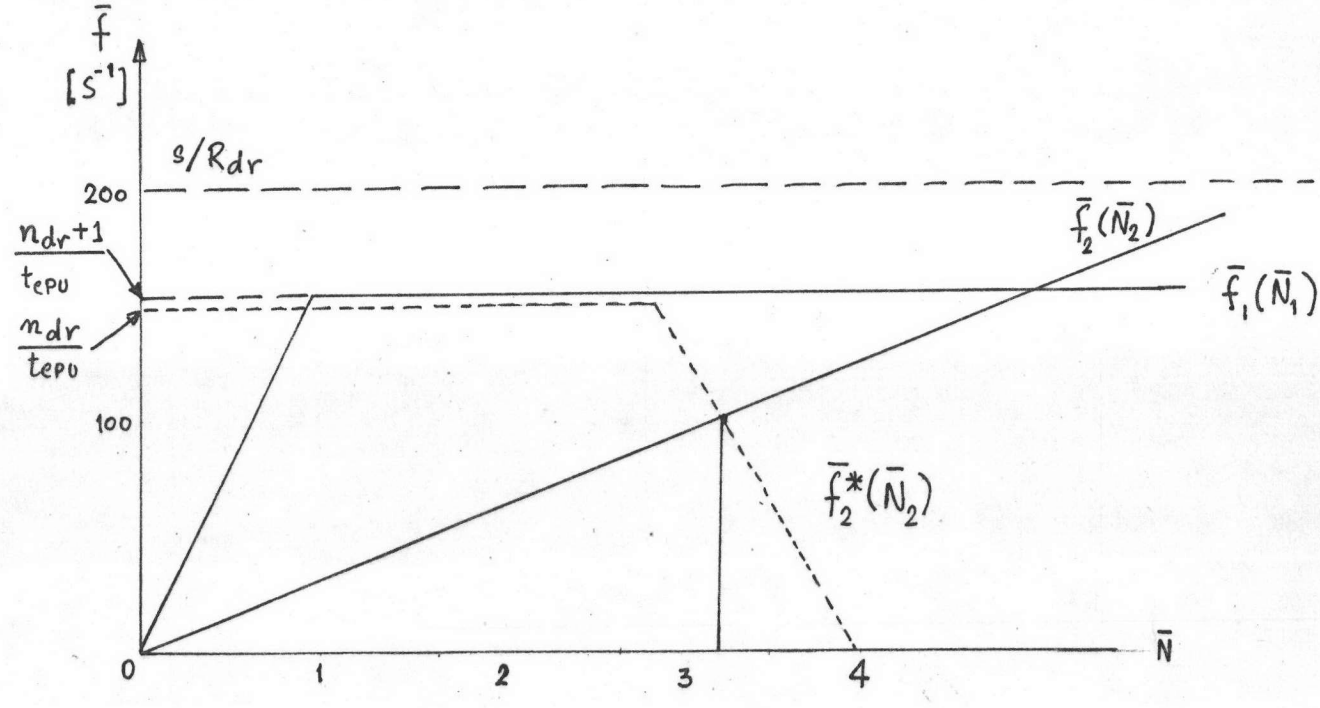
$$\bar{f}_{2max} = \bar{f}_2(\infty) = \frac{S}{R_{dr}}$$

$$\bar{f}_2^* = 0 \text{ for } N_2 = N$$

$$= \frac{n_{dr}}{t_{cpu}} \quad \text{for} \quad N_2 \leq N - 1$$

$$= \frac{n_{dr}}{t_{cpu}} (N - N_2) \quad \text{for} \quad N - 1 < N_2 < N$$

ดังนั้น ในการเพิ่มทรูพุทก็เป็นการลดเวลา t_{cpu} เพิ่มเวลา t_{cpu} ก็ลดทรูพุทและ สำหรับ N_{dr} จำนวนเรียกของดรัมต่อเวลา t_{cpu} ก็เป็นสัดส่วนกลับกับทรูพุท ดังรูปที่ 3.6



รูปที่ 3.6 แสดงอัตราค่าเฉลี่ยการเข้ารับและออกจากการบริการ จากรูปที่ 3.5

3.2.2.1.2 รูปแบบความน่าจะเป็น (Probabilistics) (3)

หลักการทางคณิตศาสตร์ที่จะพิจารณาการประเมินความสามารถของระบบคอมพิวเตอร์ ประกอบด้วย

1. รูปแบบหรือการแจกแจงของเวลาการใช้บริการ (Service Time Distributions) ใน Model (รูปแบบ) ของการใช้ทรัพยากรของระบบคอมพิวเตอร์มีส่วนหนึ่งที่สำคัญในการพิจารณาเพื่อศึกษา คือเวลาที่โปรแกรมใช้บริการจากทรัพยากรของระบบคอมพิวเตอร์ เช่น มีทรัพยากรอันหนึ่งคือตัวหน่วยประมวลผล เวลาที่โปรแกรมใช้บริการของหน่วยนี้จะประกอบไปด้วย เวลาของการทำงานตามคำสั่งของโปรแกรมและเวลาทั้งหมดที่เสียสำหรับส่วนไอโอ ต้องการใช้จากโปรแกรมตลอดเวลาการเรียกใช้หน่วยความจำ เพราะโปรแกรมเรียกใช้หน่วยความจำแบบเสมือน แต่ถ้าทรัพยากรหน่วยนั้นเป็นฐานแม่เหล็กที่เป็นแบบหัวเคลื่อนที่ เวลาที่โปรแกรมใช้บริการประกอบด้วยเวลาที่ไปหาตำแหน่งของข้อมูล และเวลาที่เปลี่ยนแปลงข้อมูล, ระยะทางของข้อมูลถูกเคลื่อนที่ ซึ่งขึ้นอยู่กับความเร็วของหน่วยฐานแม่เหล็กและปริมาณของข้อมูลที่ถูกเคลื่อนที่โดยการเรียกใช้จากโปรแกรม

ทุก ๆ ส่วนประกอบที่เราจะมาพิจารณากัน เราได้จากการวัดตรวจข้อมูล และโดยการวิเคราะห์ตามรูปแบบของทรัพยากร โดยปกติในแง่ของการใช้งานและทางปฏิบัติแล้ว ลักษณะของเวลาการใช้บริการของทรัพยากรต่าง ๆ นั้น มีลักษณะเป็นแบบสุ่ม ซึ่งรูปแบบที่เราจะมาพิจารณาก็คือ การแจกแจงของเวลาที่ใช้ของทรัพยากรที่เหมาะสม ถ้าเวลาการใช้บริการของทรัพยากรต่าง ๆ ทำการจัดส่วนประกอบอื่น ๆ ตามเวลาที่เกิดขึ้นแบบสุ่ม คือ ไม่สามารถพยากรณ์ไว้ล่วงหน้า แต่เวลาที่ใช้บริการของทรัพยากรนั้น ไม่สามารถกำหนดไว้ล่วงหน้าได้ หรือเป็นแบบสุ่ม แต่ก็ไม่ใช่ว่าเราไม่สามารถกำหนดลักษณะของทรัพยากรได้ เช่น เราไม่ทราบล่วงหน้าว่าจะใช้เวลากับทรัพยากรหนึ่ง ๆ เท่าใดได้ แต่เราพอจะทราบได้เป็นโดยเฉลี่ยของเวลาที่ใช้บริการ ลักษณะของค่าเฉลี่ยของเวลาที่ใช้กับหน่วยหนึ่ง ๆ ของทรัพยากรนั้น เราก็นำมาเป็นแนวทางกำหนดรูปแบบได้

การกำหนดรูปแบบในรายละเอียด เราต้องทราบการแจกแจงความน่าจะเป็นของการใช้บริการ โดยตัวแปรต่าง ๆ เป็นอิสระต่อกัน คือ เวลาการใช้บริการในหน่วยทรัพยากรต่าง ๆ เป็นอิสระต่อกัน คือการใช้บริการของหน่วยอินพุท และเอาท์พุทไม่ได้ขึ้นกับเวลาที่ใช้บริการของ

หน่วยความจำประมวลผลกลาง เป็นต้น

เมื่อเรากำหนด $P(x)$ เป็น ความน่าจะเป็นของ X

จะได้ว่า $0 \leq P(x) \leq 1$ ทุก ๆ ค่าของ X

และ $\sum_x P(x) = 1$

แต่โดยปกติแล้ว ไม่เป็นงานที่จะกำหนดงานด้วยการแจกแจงของรูปแบบนั้นได้โดยตรง แต่เราก็สามารถกำหนดค่าง่าย ๆ ที่ใช้พิจารณาด้วยค่าเฉลี่ยของการแจกแจงแบบนี้ได้ โดยให้ $E(x)$ เป็นค่าเฉลี่ย

$$E(x) = \sum_x xP(x)$$

และบางครั้งเราอาจจะเลือก moments และค่าเฉลี่ย moments เป็นค่าเฉลี่ยหรือค่าที่คาดหมายของการใช้เวลาในกำลังที่ n^{th} เช่น

$$E(x^n) = \sum_x x^n P(x)$$

เมื่อค่าเฉลี่ย moment ที่ n ก็คือ ค่าคาดหมายของผลต่างของเวลาที่ใช้บริการกับค่าเฉลี่ย เช่น

$$E\ (x-E\ x)^n = \sum_x (x-E\ x)^n P(x)$$

ค่าเฉลี่ย moment ที่ 2 ก็คือค่าแปรผันของการแจกแจงของเวลาที่ใช้บริการ เมื่อถอดรากที่สองของค่าแปรผันก็คือ ค่าความเบี่ยงเบนมาตรฐาน (σ_x)

$$\sigma_x^2 = E(x^2) - (E(x))^2$$

ค่าเฉลี่ยของเวลาการใช้บริการเป็นตัวกำหนดค่าเฉลี่ยของการใช้เวลาต่อทรัพยากรหนึ่ง ๆ ค่าแปรผัน คือ ค่ากำหนดค่าการแปรปรวนของค่าเฉลี่ย จะมีค่าที่วัดความแปรปรวนกับค่าเฉลี่ยของการใช้บริการเป็นค่าสัมประสิทธิ์ของความแปรปรวน เช่น

$$Cx = \frac{\sigma_x}{E[x]}$$

เวลาการใช้บริการของหน่วยประมวลผลโดยปกติจะมีค่าความแปรปรวนสูง ค่าของสัมประสิทธิ์ของความแปรปรวน จะมีค่าเท่ากับ 10 หรือมากกว่าก็ไม่แปลก แต่หน่วยอินพุทจะมีค่าแปรปรวนต่ำ ค่าสัมประสิทธิ์ก็ต่ำด้วย ปกติจะต่ำกว่า 1 แต่ถ้าค่าสัมประสิทธิ์ความแปรปรวนเป็นศูนย์ก็หมายถึงทุก ๆ ค่าเวลาการใช้บริการเท่ากันทุก ๆ ครั้ง

ถ้าตัวแปรของเวลาการใช้บริการมีการกระจายเป็นแบบค่าคงที่ (Discrete) โดยมีการแจกแจงความน่าจะเป็นแบบเรขาคณิต (Geometric Distribution) เช่น กำหนด $0 < P < 1$ เป็น

$$P(x) = (1-P)^{x-1} P \; ; \; x = 1, 2, 3, \ldots$$

จะมีค่าเฉลี่ยเป็น

$$E(x) = \frac{1}{P}$$

และ $$E(x)^2 = \frac{2-P}{P^2}$$

ได้ค่าความเบี่ยงเบนมาตรฐาน $$\sigma_x = \frac{\sqrt{1-P}}{P}$$

สัมประสิทธิ์ความแปรปรวน $Cx = 1-p$

ถ้าตัวแปรของข้อมูลที่เรานำมาพิจารณาลักษณะของเวลาใช้บริการเป็นแบบต่อเนื่อง (Continuous) จะมีตัวกำหนด 2 ส่วน คือ Probability Density function ($f_x(x_o)$) และ probability distribution function $F_x(X_o)$

Probability density function กำหนดเป็น

$$fx(Xo) = \frac{dF_x(x_o)}{dx_o}$$

$$F_x(x_o) = \int_{-\infty}^{x_o} f_x(x_o)\,dx_o$$

และ $$f_x(x_o)\,dx_o = 1$$

จากข้อกำหนดการหาค่าเฉลี่ย โมเมนต์ (moment) และค่าแปรปรวน ดังนี้

$$E(x) = \int_{-\infty}^{\infty} x f_x(x_o)\,dx_o$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f_x(x_o)\,dx_o$$

และ $$\sigma_{x^2} = \int_{-\infty}^{\infty} (x-E(x))^2 f_x(x_o)\,dx_o$$

ถ้าการกระจายของตัวแปรเป็นการกระจายความน่าจะเป็นแบบยูนิฟอร์ม (uniform)

$$f_x(x_o) = \begin{cases} \frac{1}{b-a}, & a \leq x_o \leq b \\ 0, & \end{cases}$$

$$F_x(x_o) = \begin{cases} 0, & x_o < a \\ \frac{x_o-a}{b-a}, & a \leq x_o \leq b \\ 1, & x_o > b \end{cases}$$

$$E(x) = \frac{a+b}{2}$$

$$E(x^2) = \frac{a^2+ab+b^2}{3}$$

$$\sigma_{x^2} = \frac{(b-a)^2}{12}$$

และ $$C_x = \frac{b-a}{(b+a)\sqrt{3}}$$

ดังนั้น เราจะได้สมการของความน่าจะเป็นในช่วง $a < x \leq b$ เช่น

$$P[a<x\leq b] = \int_a^b f_x(x_o)\,dx_o$$

หรือ $$P\left[a<x\leq b\right] = F_x(b) - F_x(a)$$

การแจกแจงความน่าจะเป็นอีกแบบหนึ่งที่นิยมใช้คือ การแจกแจงความน่าจะเป็นแบบเอกซ์โพเนนเชียล (Exponential Distribution) โดยมีอัตราการให้บริการ = a

$$f_x(x_o) = \begin{cases} 0 & , x_o\ 0 \\ ae^{-ax_o} & , x_o\ 0 \end{cases}$$

$$F_x(x_o) = \begin{cases} 0 & , x_o\ 0 \\ 1-e^{-ax_o} & , x_o\ 0 \end{cases}$$

$$E\ x = \frac{1}{a},$$

$$E\ x^2 = \frac{2}{a^2}$$

$$\sigma x^2 = \frac{1}{a^2}$$

และ $$Cx = 1$$

การนำการแจกแจงความน่าจะเป็นแบบเอกซ์โพเนนเชียล (Exponential Distribution) มาใช้กับรูปแบบความสัมพันธ์ทางคณิตศาสตร์ของระบบคอมพิวเตอร์ โดยแบ่งลักษณะของการแจกแจงน่าจะเป็นสองแบบตามความแปรปรวน ถ้าการแจกแจงความน่าจะเป็นของตัวแปร ที่มีการแปรปรวนมาก ที่จะใช้การแจกแจงความน่าจะเป็นแบบไฮเปอร์เอกซ์โพเนนเชียล (Hyper exponential) ส่วนการแจกแจงความน่าจะเป็นของตัวแปรที่มีค่าการแปรปรวนน้อย ก็จะเลือกใช้แบบไฮโปเอกซ์โพเนนเชียล (Hypo exponential)

2. <u>หลักการจากจัดลำดับงานเพื่อเข้ารับบริการ (Scheduling Algorithm)</u>

นอกเหนือจากค่าเฉลี่ยของเวลาที่ใช้บริการของทรัพยากรหน่วยต่าง ๆ ของระบบแล้ว

ลักษณะที่สำคัญเกี่ยวกับทรัพยากรของระบบยังมีอีกสิ่งหนึ่ง คือ การจัดลำดับงานหรือโปรแกรมเพื่อเข้ารับบริการจากทรัพยากรของระบบในหน่วยต่าง ๆ เป็นหลักการที่จะตัดสินใจว่า ทรัพยากรจะถูกจัดให้ทำงานโปรแกรมส่วนใดและให้สำเร็จได้หรือเรียกอีกอย่างหนึ่งว่าการจัดคิวของโปรแกรมให้ไปให้ทรัพยากรส่วนต่าง ๆ ของระบบ

หลักในการจัดงานเพื่อเข้ารับบริการ มีหลายวิธีการดังนี้คือ

1. การจัดแบบงานใดเข้ามาก่อนก็รับบริการก่อน (First-Come First-Served FCFS) สำหรับแบบ FCFS นี้ไม่เหมาะสำหรับใช้กับทรัพยากรที่มีความแปรปรวนของเวลาให้บริการสูง สำหรับการจัดแบบนี้งานที่รับบริการจะถูกจัดตามเวลาที่งานนั้นเข้ามารับบริการ งานใดที่เข้ามาก่อนก็รับบริการก่อนตามเวลาที่เข้ามาไม่มีการแซงคิว

2. การจัดแบบงานใดเข้ามาทีหลังก็รับบริการก่อน (Last Come First Served LCFS) แบบที่ตรงข้ามกับแบบแรก ทุก ๆ งานที่เข้ามาทีหลังก็จะถูกจัดให้รับบริการก่อนเป็นการจัดคิวหรือแซงคิว งานที่เข้ามาก่อน ก็เกิดปัญหาว่าจะแซงคิว งานที่รับบริการอยู่แล้วให้หยุดก่อนเอางานที่เพิ่งเข้ามาใหม่ไปใช้บริการหรือจะแซงคิวงานที่จะถูกเลือกให้รับบริการงานถัดไป

3. การจัดแบบงานใดเข้ามาทีหลังก็รับบริการก่อนโดยการแซงคิว (Last Come First Served Preemptive Resume LCFSFR) คือการจัดที่ต้องแซงคิวทุกครั้งที่เขามารับบริการ ซึ่งเป็นลักษณะของการให้บริการของหน่วยประมวลผลในระบบอินเทอร์แอกทีฟ (Interactive System)

โดยทั่ว ๆ ไป การจัดงานเพื่อรับบริการของหน่วยประมวลผลจะจัดเป็นแบบราวด์โรบิน (Round-Robin LRR) หรือบางทีเรียกว่า ไทม์สไลซิงค์ (Time-Slic) ซึ่งเป็นวิธีการที่กำหนดว่าในช่วงของเวลาซึ่งเรียกว่า ควันตัม (Quantum) หรือ ไทยม์สไลซ์ (Time-Slic) งานที่รับบริการจะถูกจัดแบบงานใดเข้ามาก่อนรับบริการก่อน (FCFS) ตลอดเวลาภายในเวลาที่กำหนด แต่ถ้างานใดทำงานในช่วงเวลาควันตัม นั้น ยังไม่เสร็จงานนั้นจะถูกแซงคิวด้วยงานอื่นแล้วงานนั้นจะต้องไปต่อท้ายของคิวงานที่จะเข้ารับบริการใหม่ เฉพาะฉะนั้น เวลาที่งานหนึ่ง ๆ จะเข้ารับบริการจากทรัยากรนี้ต้องแบ่งเวลาเป็นหลาย ๆ ควันตัม ซึ่งจะทำให้หน่วยประมวลผลได้ถูกใช้ร่วมกันจากหลาย ๆ งาน อย่างมีประสิทธิภาพมากขึ้น งานที่ใช้เวลาในการบริการสั้นก็ไม่จำเป็นต้องถูกบังคับให้รองานที่ใช้เวลาการบริการก่อนหน้านี้เสร็จไปก่อน แต่การที่จะเปลี่ยนจาก

งานที่ให้บริการอยู่ไปให้งานใหม่นั้นก็เสียเวลาไปบ้าง ซึ่งปกติจะน้อยมากเป็น 100 microsecond (100 ของเศษหนึ่งล่วนพันวินาที) ต่อการแซงต่อหนึ่งครั้ง ดังนั้นการเลือกช่วงของควันตัมให้มากกว่าเวลาที่เสียไปต่อการเปลี่ยนงานอันเกิดจากการแซงคิวของงานแต่ถ้าควันตัมเป็นเวลามาก ๆ ก็เสมือนเป็นแบบ FCFS ซึ่งทำให้การใช้ทรัพยากรนั้นมีประสิทธิภาพน้อยลง

แต่ถ้าเรามาคำนึงถึงให้มีค่าเฉลี่ยของเวลาการถามตอบข้อมูลน้อยที่สุด (Mean Response Time) การจัดแบบใช้เวลาสั้นเข้าทำงานก่อน (Short-Remaining-Time-First) (SRTF) งานที่เหลือเวลาในการรับบริการน้อยที่สุดทำก่อน คืองานที่เข้ามาใหม่มีเวลาการใช้บริการทรัพยากรนั้นน้อยกว่างานที่รับบริการอยู่ งานที่เข้ามาใหม่จะถูกแซงคิวให้ได้ทำงานทันที ลักษณะนี้เหมาะสมสำหรับล่วนอินพุทและเอาท์พุท เช่น ของดรัมแม่เหล็กแซงคิวด้วยเวลาที่ลาเทนซีสั้นที่สุดทำงานก่อน (Shortest-Latency-Time-First) ล่วนของฐานแม่เหล็กก็เป็นเวลาที่ซีคสั้นที่สุดทำงานก่อน (Shortest-seek-time-first)

ที่กล่าวมาแล้วยังไม่คำนึงถึงกรณีที่มีการให้โอกาสเปลี่ยนแปลงการเข้ารับบริการจากภายนอก (External Priority)

<u>ความสามารถของระบบคอมพิวเตอร์กับความสัมพันธ์ของการแจกแจงความน่าจะเป็นของเวลาการให้บริการกับการจัดงานเข้ารับบริการ</u>

เมื่อพิจารณาถึงความสามารถของระบบจะทำการวัดค่าเฉลี่ยของจำนวนงานที่งานเข้ามารับบริการของระบบต่อหน่วยของเวลา เรียกว่า ทรูพุท (Throughput) ถ้าค่าทรูพุทมากที่สุด (Maximum throughput) ในรูปแบบที่เรากำหนดก็คือค่าสูงสุดของค่าอรรถประโยชน์ของทรัพยากรนั้น ตลอดจนถ้าค่าเฉลี่ยการถามตอบข้อมูลน้อยที่สุด (Minimum Mean Response Time) ก็คือภาระงานยังคงสามารถดำเนินการต่อไปได้

ถ้าเวลาให้บริการมีการแปรปรวนมากคือ $Cx > 1$ งานล่วนใหญ่ใช้เวลาในการบริการมากกว่าค่าเฉลี่ยของเวลาใช้บริการ ซึ่งจะทำให้ความสามารถลดลงเมื่อเราเลือกให้งานเข้ารับบริการแบบ FCFS เพราะงานล่วนใหญ่จะไปรอที่หน่วยประมวลผล ทำให้ล่วนอินพุทและเอาท์พุทมีเวลาการรอคอยให้บริการมาก ทำให้ความสามารถของการทำงานแบบมัลติโปรแกรมมิ่งลดลง แต่ถ้าเป็นแบบราวด์โรบิน งานที่ใช้เวลานั้นก็ไม่ต้องเสียเวลาให้งานที่ใช้เวลามากทำงานเสร็จก็สามารถแซงคิวเข้าไปรับบริการได้

ความสัมพันธ์ระหว่างการวัดความสามารถของระบบคอมพิวเตอร์

1. ทรูพุท, ค่าอรรถประโยชน์และเวลาเฉลี่ยของการให้บริการของระบบคอมพิวเตอร์ (Throughput, Utilization and Mean Service Time) ถ้ามีทรัพยากรหน่วยหนึ่ง และกำหนดค่าเฉลี่ยของเวลาที่ให้บริการหนึ่งงานเป็นค่า $E(X)$ และอัตราเฉลี่ยของการให้บริการเป็น $1/E(x)$ เราจะเรียกค่าอรรถประโยชน์เมื่ออัตราของเวลาที่ทรัพยากรนั้นทำงานว่า U เพราะฉะนั้น

$$\text{ค่าทรูพุท} = \frac{U}{E(x)}$$

ในกรณีที่มีทรัพยากรเหมือนกัน K หน่วย จะได้ค่า

$$\text{ทรูพุท} = \frac{KU}{E(x)}$$

2. ทรูพุท, ค่าเฉลี่ยความยาวของการรอคอย, ค่าเฉลี่ยของเวลาการถามตอบข้อมูล (Throughput, Mean Queue lenght, Mean Response time)

เมื่อกำหนดค่าเฉลี่ยความยาวของการรอคอย คือจำนวนงานที่รอที่จะใช้บริการของทรัพยากรหนึ่ง ๆ และเวลารอคอยหรือเวลาที่ถามตอบข้อมูลคือเวลาระหว่างงานเข้ามายังทรัพยากรจนสิ้นสุดการใช้ทรัพยากรนั้น

$$L = E(1) = \sum_{L=1}^{\infty} 1P(1)$$

$P(1)$: ความน่าจะเป็นของความยาวของการรอคอย 1 เพราะฉะนั้นเวลาเฉลี่ยของการรอคอย (Q)

$$Q = E(g) = \int_{o}^{\infty} g_o f_g(g_o)\, dg_o$$

ซึ่งจะได้ค่าเป็น

$$L = \lambda Q$$

เป็นทรูพุท

สมมติเมื่อกำหนดเวลาเริ่มต้น ตั้งแต่ 0 ถึง T ให้ n(T) เป็นลำดับของงานที่เสร็จสิ้นเมื่อเวลา T

$$\lambda = \frac{n(T)}{T}$$

A(T) คือพื้นที่ของ 1(f) ตั้งแต่ 0 ถึงเวลา T

$$A(T) = \int_0^T 1(t)\,dt$$

$$L = \frac{A(T)}{T}$$

$$Q = \frac{A(T)}{nT}$$

$$L = \frac{A(T)}{T} = \frac{n(T)}{T} \cdot \frac{A(T)}{n(T)} = Q$$

L : ความยาวเฉลี่ยของการรอคอยเพื่อรับการบริการจากทรัพยากรหนึ่ง รวมงานที่กำลังรับบริการด้วย

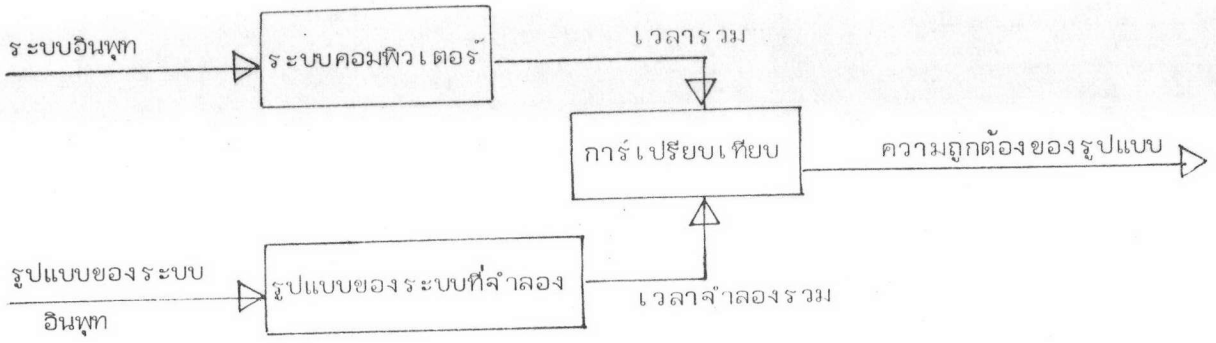
Q : ค่าเฉลี่ยของเวลาที่รอคอยเพื่อรับบริการรวมเวลาที่รับบริการด้วย

ปัญหาที่สำคัญของการจัดงานเพื่อเข้ารับบริการมีอยู่ว่า งานใดควรรับบริการอยู่เดี่ยวนี้ และถ้างานใดที่รับบริการอยู่จะให้รับบริการต่อไปจนเสร็จหรือควรจะเอางานอื่นมาแซงคิวหรือลัดคิวทำงานก่อน

3.2.2.2 การจำลองแบบ (Simulation)

การจำลองแบบเป็นรูปแบบของพฤติกรรมของระบบโดยมีเวลาเข้ามาเกี่ยวข้องการลังเกตพฤติกรรมตามเวลาของระบบภายใต้จำลองแบบระบบอินพุทให้ได้ผลออกมา ความสำคัญของการจำลองแบบคือ ความเชื่อถือได้แบบที่จำลองโดยการวางแผนแบบจำลองและทดสอบ แต่ก็ไม่ได้รับประกันว่ารูปแบบจะถูกต้องแน่นอน ความถูกต้องจะได้จากค่าดัชนีความสามารถและคุณภาพที่เลือกเข้ามาทำการศึกษาและนำไปใช้กับระบบที่มีเงื่อนไขตามลักษณะของระบบอินพุทที่นำมาทดลองแบบดังรูปที่ 3.7

รูปที่ 3.7 วิธีการการเปรียบเทียบรูปแบบ



การจำลองแบบขึ้นอยู่กับส่วนควรทางคณิตศาสตร์ที่แทนค่าพฤติกรรมเป็นตัวเลข และวิธีการสำหรับการแก้ไขปัญหาในการจำลองแบบจะมีพารามิเตอร์ที่เป็นอินพุทคือภาระงานและพารามิเตอร์ที่เป็นเอาท์พุท คือค่าดัชนีความสามารถและคุณภาพ โดยมีรูปแบบเป็นเหตุการณ์ตามเวลา เป็นการจำลองแบบดีสครีท (Discrete Simulation) ซึ่งประกอบด้วยวิธีการดังนี้

1. การกำหนดและการสร้างตัวจำลองแบบ (Formation and Construction)
2. ความสมบูรณ์และความถูกต้องของตัวจำลองแบบ (Calibration and Validation)
3. การออกแบบการทดสอบจำลองแบบ (Design of Simulation)
4. การแปลผลของการจำลองแบบ (Interpretation of Simulation Result)

3.2.2.2.1 การกำหนดและการสร้างตัวจำลองแบบ จำเป็นต้องทราบถึงจำนวนการตัดสินใจที่จะทำการจำลองสิ่งแรกคือการกำหนดวัตถุประสงค์ว่าจะเอารายละเอียด และความถูกต้องของรายละเอียดเพียงใด สิ่งที่สองคือ พฤติกรรมของระบบคอมพิวเตอร์ ดังนั้นการกำหนดและการสร้างตัวจำลองแบบควรมีการกำหนด

1 ระดับของรายละเอียดเป็นการหาระดับอัตราการจำลองกับเวลาจริงของการกำหนดอินพุท การประมวลผลตามรูปแบบของภาระงานซึ่งขึ้นอยู่กับการศึกษาลักษณะรูปแบบคอมพิวเตอร์ในรายละเอียดได้มากเพียงใด และวัตถุประสงค์ความถูกต้องของรูปแบบเพียงใด

2 ความยืดหยุ่นของตัวจำลองแบบในการสามารถปรับให้เข้ากับระบบอินพุทหรือรูปแบบที่กำหนดเปลี่ยนแปลงได้มากน้อยอย่างไร

3 ภาษาของการโปรแกรมจำลองแบบ ซึ่งจะมีผลต่อการสร้าง, ทดสอบ, การวิ่งและการปรับปรุงรักษา

4 รูปแบบของโครงสร้างของตัวจำลองแบบ เช่น โครงสร้างของระบบฮาร์ดแวร์ โครงสร้างของภาระงาน

5 ตัวแปรที่จะเป็นผลลัพธ์เพื่อให้ได้ค่าดัชนีความสามารถและคุณภาพ และตัวแปรที่เป็นภาระงานของระบบอินพุท

6 การทดสอบแบบจำลองโดยการวิ่งตามลักษณะเวลาที่กำหนด และรูปแบบของโครงสร้าง

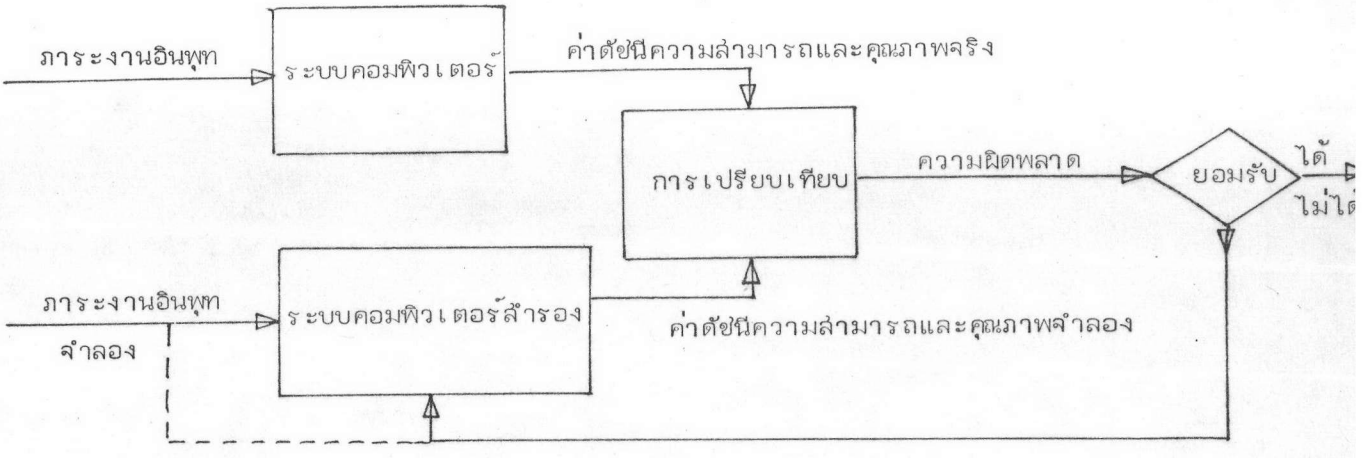
โครงสร้างของตัวแปรมีการกำหนดด้วย 2 คำสั่งคือ

1. ลำดับของเหตุการณ์จะต้องมีการสร้างและการอัพเดทอย่างต่อเนื่อง
2. การทำงานของเหตุการณ์ ว่าแต่ละเหตุการณ์ที่เกิดขึ้นจะมีการทำงานอะไรบ้างอย่างไร

การจำลองรูปแบบภาระงานมี 3 วิธีการคือ

1. จำลองการตรวจสอบ
2. จำลองโปรแกรม (Program-driven Simulators)
3. จำลองการแจกแจงของการงาน (Distribution Simulators)

3.2.2.2 ความสมบูรณ์และความถูกต้องของตัวจำลองแบบ ในความถูกต้องของรูปแบบที่จำลองจำทำการเปรียบเทียบค่าผลลัพธ์ คือค่าดัชนีความสามารถและคุณภาพที่ได้จากการจำลองแบบและการทำงานจริง จะให้ค่าความแตกต่างที่ยอมรับได้หรือไม่ ดังรูปที่ 3.8



รูปที่ 3.8 แสดงขั้นตอนของการเปรียบเทียบความถูกต้องของรูปแบบ ของรูปแบบ

ความสมบูรณ์และความถูกต้องของรูปแบบ ถ้ายังไม่เป็นที่พอใจจะต้องมีการปรับปรุง และตรวจสอบกระบวนการใหม่ ซึ่งจะมีสาเหตุจาก

1. การกำหนดที่ไม่ถูกต้องในรูปแบบของรายละเอียด

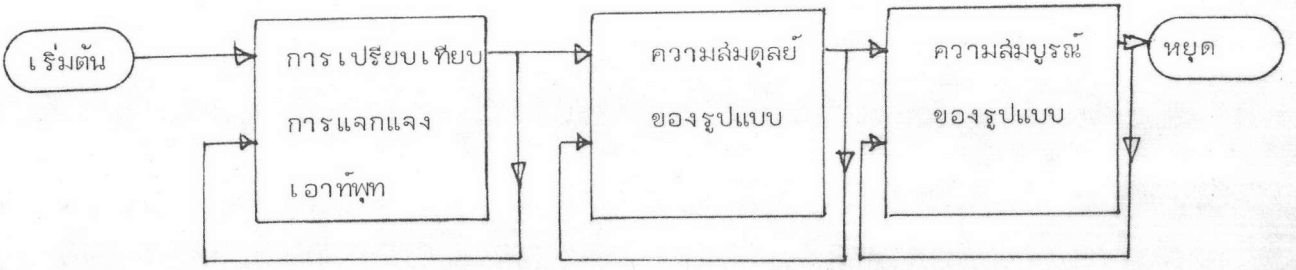
2. การแก้ปัญหาไม่ถูกต้องในรูปแบบ

การปรับปรุงของแบบจำลองกระทำได้ดังนี้คือ

1. การเปลี่ยนแปลงการทำงานของเหตุการณ์ต่าง ๆ หรือการเพิ่มเติมการทำงานต่าง ๆ

2. การเปลี่ยนแปลงในส่วนประกอบของรูปแบบ

3. การเปลี่ยนแปลงค่าพารามิเตอร์



รูปที่ 3.9 แสดงขั้นตอนการปรับปรุงการจำลองแบบ

3.2.2.2 การออกแบบการทดสอบแบบจำลอง แบ่งออกเป็น 2 ส่วนคือ

1 การออกแบบการทดสอบ ตามลักษณะของภาระงานในการตรวจสอบ ตรวจวัดข้อมูลโดยแบ่งเป็น ระดับ และปัจจัยของภาระงาน และเลือกปัจจัยและระดับที่เป็นไปได้ มาทำการทดสอบจำลองแบบ

2 เวลาในการวิ่งแบบจำลอง โดยปกติจะทำการวิ่งเพื่อให้ได้ค่าดัชนีความสามารถและคุณภาพตามที่ประมาณไว้ให้ได้ถึงระดับความถูกต้องที่ยอมรับได้

3.2.2.2 การแปลผลการจำลองแบบ ในการจำลองแบบจะให้ค่าดัชนีความสามารถและคุณภาพซึ่งมีรูปแบบตามที่จำลอง ซึ่งจะให้ผลได้ใกล้เคียงเมื่อได้ข้อมูลและสมการของรูปแบบ ซึ่งจะแสดงในรูปของกราฟ หรือตาราง ซึ่งการเปรียบเทียบผลการประเมินกับการวิเคราะห์ทางสถิติแล้วจะให้ความเชื่อมั่นในการตัดสินใจแก่ผู้ประเมินผลมากกว่า

การจำลองแบบเป็นการกำหนดรูปแบบและตัวแปรของระบบคอมพิวเตอร์ ที่มีต้นทุนสูงในส่วนของเวลาการทำโปรแกรม การตรวจสอบโปรแกรม และเวลาของการวัดโปรแกรมจำลองนี้ ซึ่งเป็นผลให้ไม่สามารถทำการจำลองแบบนี้ได้ เพราะเป็นค่าใช้จ่ายที่สูงและไม่คุ้มค่ากับการนำไปบำรุงการดำเนินการในหน่วยงานที่มีการให้บริการ แต่สำหรับทางด้านการทดลองโดยทำแบบสุ่มตัวอย่าง เราอาจจะจัดหาโปรแกรมจำลองแบบสำเร็จรูปมาใช้ได้หลายโปรแกรม เช่น โปรแกรมสเคอร์ท (System and Computer Evaluation and Review Technique) โปรแกรมเคส (Computer Aided System Evaluation, CASE) เป็นต้น

3.3 วิธีการเลือกและกำหนดตัวแปร

ในการเลือกและกำหนดตัวแปรสำหรับการประเมินความสามารถของระบบคอมพิวเตอร์สามารถกำหนดได้จาก

3.3.1 ชนิดของภาระงาน อรรถประโยชน์และระดับการบริการ การพิจารณาถึงลักษณะการทำงานของระบบ สามารถแบ่งออกได้เป็น 3 ส่วนดังนี้

ก. ภาระงาน (Work Load) โดยมีหน่วยในการวัดและตรวจสอบเป็นจำนวนงานต่อชั่วโมง สำหรับงานแบบแบทช์ หรือจำนวนทรานแซกชั่นต่อวินาที สำหรับงานแบบการดึงข้อมูลระบบแฟ้มข้อมูลหรือฐานข้อมูล หรือจำนวนอินเตอร์แอกชั่นต่อวินาที สำหรับงานแบบการติดต่อสอบถามกับเครื่องโดยตรง

ข. อรรถประโยชน์ของทรัพยากรในส่วนของฮาร์ดแวร์ อันได้แก่ อรรถประโยชน์ซีพียู ช่องสถานี จานแม่เหล็ก เทป และเครื่องพิมพ์

ค. ระดับการบริการแก่ผู้ใช้บริการ ถ้าเป็นแบบออนไลน์เป็นจำนวนเวลาที่ตอบสนองคำถามกลับมา หรือถ้าเป็นแบบ เป็นจำนวนเวลาตั้งแต่ส่งงานจนได้รับงานจากระบบคอมพิวเตอร์

3.3.2 การกำหนดตัวแปรตามโปรแกรมควบคุมระบบ การพิจารณาโปรแกรมควบคุมระบบตามหน้าที่ของการทำงานสามารถแบ่งได้ดังนี้

ก. โปรแกรมควบคุมระบบโอเอส/วีเอสวัน สำหรับโปรแกรมควบคุมระบบโอเอส/วีเอสวัน กำหนดได้ดังนี้

1. ชนิดของงาน (Jobs) ตามเวลา

1.1 ชนิดใช้เวลาน้อย (Short Run)

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

1.2 ชนิดใช้เวลาปานกลาง (Medium Run)

1.3 ชนิดใช้เวลามาก (Long Run)

2. ชนิดของงาน (Job Steps) ตามการใช้ทรัพยากร

2.1 ชนิดใช้หน่วยความจำกลางมาก (CPU Bound)

2.2 ชนิดใช้หน่วยอุปกรณ์อินพุทและเอาท์พุทมาก (I/O Bound)

3. จำนวนการเรียกใช้ข้อมูลจากอุปกรณ์อินพุทและเอาท์พุทต่องาน (EXCP'S PER JOB)

4. จำนวนการเรียกใช้ข้อมูลจากอุปกรณ์อินพุทและเอาท์พุทต่องานย่อย (EXCP'S PER JOB STEP)

5. จำนวนค่าเฉลี่ยของการใส่เทปต่องาน (Average Tape Mounts Per Job)

6. จำนวนค่าเฉลี่ยเวลาในการใช้เครื่องพิมพ์รายงานต่องาน (Printer Set Up Time Per Job)

7. จำนวนทรัพยากรที่งานจำเป็นต้องเรียกใช้ต่องาน

7.1 ช่องสถานี

7.2 จานแม่เหล็ก

7.3 เทป

7.4 เครื่องพิมพ์รายงาน

8. จำนวนอุปกรณ์อินพุทและเอาท์พุทที่ใช้โดยงานหนึ่ง

9. จำนวนเวลาที่ใช้ซีพียู (Total Batch CPU Hours Per Day)

10. จำนวนเวลาที่ใช้หน่วยประมวลผลกลางต่องานและงานย่อย (CPU Time Per Job and Per Job Step)

11. จำนวนเวลาที่งานหนึ่งหรืองานย่อยหนึ่งใช้ (Elapse Time Per Job and Elapse Time Per Job Steps)

ข. โปรแกรมควบคุมระบบซีไอซีเอล/วีเอล สำหรับโปรแกรมควบคุม

1. ทรานแซกชั่นตามสถานที่และชนิด (Transaction by Location and Type)

2. จำนวนการเรียกใช้ข้อมูลตามชนิดของทรานแซกชั่น (EXCP'S by Transaction Type)

3. จำนวนการเรียกใช้หน่วยอินพุทและเอาท์พุทของงานที่สำคัญ (I/O Device Required by Critical Application)

4. จำนวนการใช้ทรัพยากรหน่วยอินพุทและเอาท์พุทตามชนิดของทรานแซกชั่น (I/O Resource Consumption by Transaction Type)

5. จำนวนเวลาที่ใช้สำหรับหน่วยประมวลผลกลางต่อวัน (Total CPU Required Hours Per Day)

6. จำนวนเวลาต่อชนิดของทรานแซกชั่น (CPU Required Time Per Transaction Type)

7. เวลาการตอบสนองการสอบถามตามชนิดของทรานแซกชั่นของหน่วยประมวลผลกลางและเทอร์มินัล (Response Time by transaction type of CPU and terminal)

3.3.3 การกำหนดตัวแปรตามข้อมูลตามความต้องการของบุคคลากรในหน่วยงานคอมพิวเตอร์ ซึ่งสามารถแบ่งได้เป็น 5 กลุ่ม (9)

1. ฝ่ายปฏิบัติการคอมพิวเตอร์
2. ฝ่ายโปรแกรมเมอร์ระบบคอมพิวเตอร์
3. ฝ่ายผู้ใช้บริการ
4. ฝ่ายพัฒนาโปรแกรมงานต่าง ๆ
5. ผู้บริหารหน่วยงาน

ตัวอย่างเช่น ข้อมูลที่จำเป็นสำหรับฝ่ายปฏิบัติการคอมพิวเตอร์จะมีข้อมูลเป็นดังนี้

1. จำนวนเวลาของหน่วยประมวลผลการที่สามารถใช้ได้ต่อวัน
2. จำนวนเวลาของหน่วยประมวลผลการที่ถูกใช้ไปจริงต่อวัน
3. จำนวนเวลาของหน่วยประมวลผลการที่ถูกใช้โดยงาน วีเอ็ม/เอสพี

4. จำนวนเวลาของหน่วยประมวลผลการที่ถูกใช้โดยงาน

โอเอส/วีเอสวัน

5. จำนวนเวลาของหน่วยประมวลผลการที่ถูกใช้โดยงาน

ซีไอซีเอส/วีเอส

6. อรรถประโยชน์ของทรัพยากรเป็นช่วงของเวลาต่อวัน

7. จำนวนเวลาทั้งหมดที่งานแบบแบทช์ที่สำคัญใช้

8. เวลาการตอบสนองการถามตอบข้อมูลทีเทอร์มินัลและหน่วย

ประมวลผลกลาง

9. ปริมาณทรัพยากรที่จำเป็นต้องใช้สำหรับงานที่สำคัญ

10. ปริมาณทรัพยากรที่จำเป็นต้องใช้โดยเฉลี่ยต่องาน

3.4 การกำหนดวิธีการตัดสินใจและมาตรฐานการประเมินความสามารถ

ในการประเมินความสามารถและคุณภาพคอมพิวเตอร์จะมีวิธีการตัดสินใจและมีมาตรฐานการประเมินความสามารถและคุณภาพจากสิ่งต่าง ๆ ดังนี้

3.4.1 ปัจจัยสำคัญสำหรับการประเมินความสามารถที่เกิดจากการดำเนินงานได้แก่

3.4.1.1 ความต้องการของการบริการของระบบคอมพิวเตอร์ ทั้งที่เป็นระบบแบทช์หรือออนไลน์ ภายใต้โปรแกรมควบคุมระบบต่าง ๆ เป็นอย่างไร

3.4.1.2 ความสามารถและคุณภาพของทรัพยากรระบบนั้น มีความสามารถเป็นเช่นไร

3.4.1.3 ความสามารถและคุณภาพที่กำหนดตามข้อกำหนดของเครื่องหรือบริษัทผู้ผลิต

3.4.1.4 ความสามารถและคุณภาพที่ระบบจะอำนวยความสะดวกแก่ผู้ใช้สามารถจะใช้ความสามารถและคุณภาพอีกเป็นจำนวนเท่าไรบ้าง

3.4.1.5 ภาระงานของระบบ ความเปลี่ยนแปลงเกี่ยวกับภาระงาน ลักษณะและปริมาณของภาระงานในปัจจุบันและอนาคตเป็นอย่างไร

3.4.2 กระบวนการสำหรับการประเมินความสามารถและคุณภาพ ตามปัจจัยของการประเมินความสามารถและคุณภาพ

3.4.2.1 การกำหนดระดับของวัตถุประสงค์ของการให้บริการ เพื่อที่จะกำหนดความสามารถและคุณภาพของระบบที่เหลือและเพียงพอต่อการให้บริการ

3.4.2.2 การพยากรณ์ภาระงาน เพื่อที่จะทำการเปลี่ยนการพยากรณ์ในความต้องการของภาระงานให้เป็นความต้องการเพื่อการประเมินผลความสามารถคุณภาพและการวางแผน

3.4.2.3 การกำหนดสิ่งไม่สมดุลย์ เพื่อที่กำหนดการพัฒนางานและกำหนดงานที่ไม่สมดุลย์และหาทางเลือกเพื่อปรับให้อยู่ในภาพสมดุลย์

3.4.2.4 การเสนอผู้บริหาร เพื่อที่จะกำหนดแนวทางการนำส่วนฮาร์ดแวร์ หรือซอฟท์แวร์มาสู่ระบบเพิ่มขึ้น หรือการวางแผนการใช้ทรัพยากรในระบบใหม่

3.4.3 มาตรฐานสำหรับการประเมินความสามารถและคุณภาพ มีวิธีการที่ประกอบกัน ดังนี้

3.4.3.1 การกำหนดภาระงานตามเวลา เพื่อที่จะแบ่งลักษณะของภาระงานต่าง ๆ

ให้ทราบถึงภาระงานลักษณะนั้นตามช่วงเวลา

3.4.3.2 การกำหนดหน่วยงานผู้ใช้บริการ แบ่งตามสายงานหรือหน่วยงานที่เข้ามาใช้บริการ เพื่อทราบถึงลักษณะงานที่ใช้ และจำนวนปริมาณที่ประมวลผล

3.4.3.3 การเก็บรวบรวมข้อมูลความสามารถและคุณภาพตามภาระงานและเวลาและลักษณะของโปรแกรมควบคุมระบบ

3.4.3.4 กำหนดคุณลักษณะของภาระงานปัจจุบัน ตามความสามารถและคุณภาพของระบบนั้น

3.4.3.5 การพยากรณ์เพื่อการกำหนดโครงการของงานต่าง ๆ ในอนาคตที่จะเข้ามาประมวลผลร่วมกับงานในปัจจุบัน

3.4.3.6 การกำหนดคุณลักษณะของภาระงานและความสามารถและคุณภาพระบบในอนาคตนั้น จะต้องทราบถึงข้อมูลต่าง ๆ ดังนี้

ก. การเพิ่มขึ้นของการใช้งานในปัจจุบัน โดยแบ่งเป็น

ก.1 แบบออนไลน์ โดยคำนวณเป็นจำนวนทรานแซกชั่นต่อวินาที

ก.2 แบบแบทช์ โดยกำหนดเป็นจำนวนงานหรือความซับซ้อนของงาน

ข. การเพิ่มขึ้นของปริมาณงาน

ข.1 โดยการเปรียบเทียบกับงานใกล้เคียง หรือลักษณะคล้ายกัน

ข.2 โดยการใช้เวลาซีพียูขณะการพัฒนาระบบงานนั้น ๆ

ค. อุปสรรคของระบบที่ช่อนเร้น ซึ่งเกิดขึ้นจากการจัดการ

ค.1 จำนวนเทอร์มินัลใช้งานไม่เพียงพอ

ค.2 การกำหนดเวลาของภาระงานที่เข้าทำงานไม่ถูกต้อง

ค.3 การเกิดจุดอ่อนในระบบ

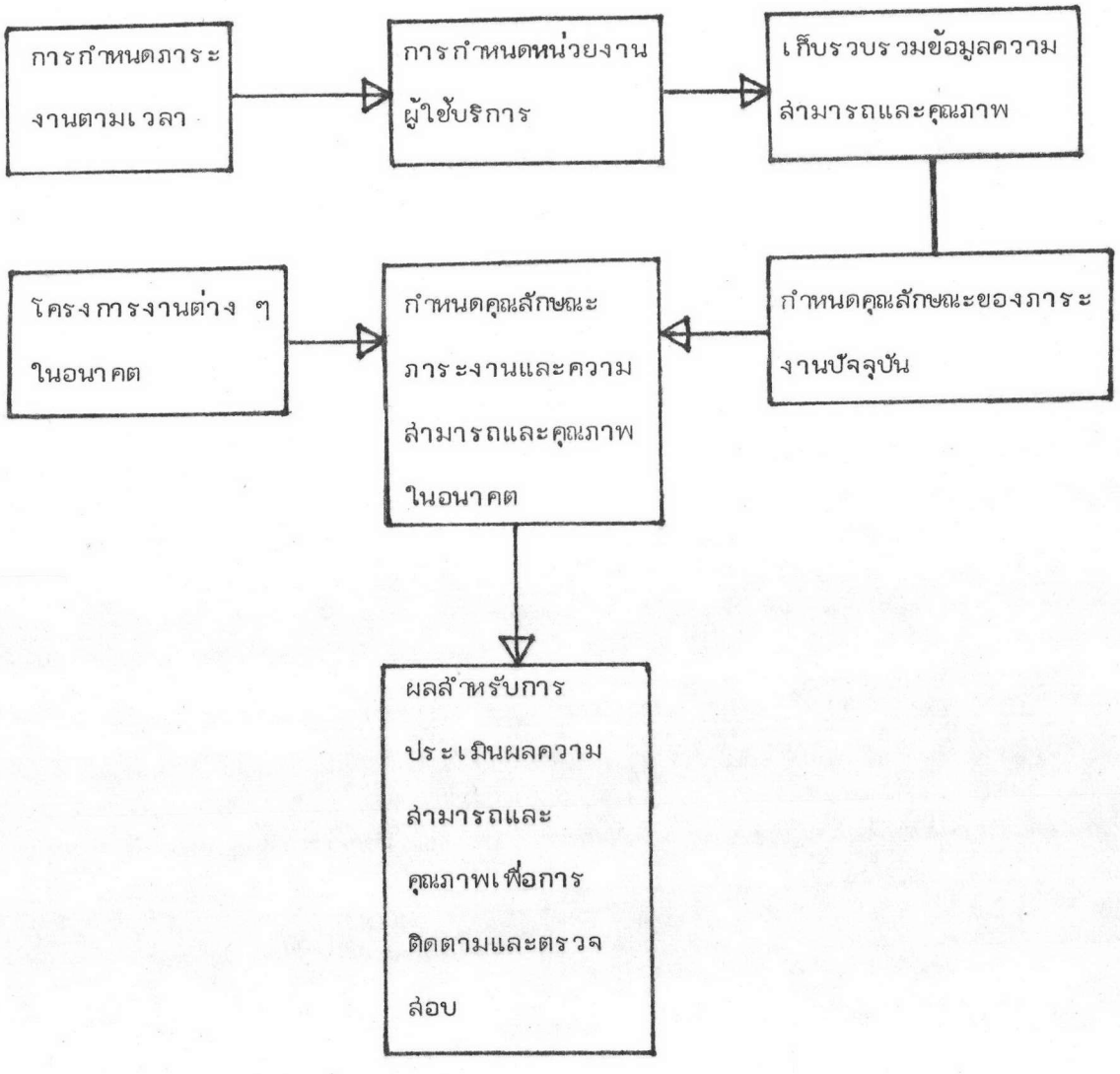
ง. บุคลากรในการใช้ระบบคอมพิวเตอร์ในส่วนของ

ง.1 จำนวนผู้ใช้บริการ (End User)

ง.2 จำนวนผู้ใช้บริการที่มีประสบการณ์ (Experience End User)

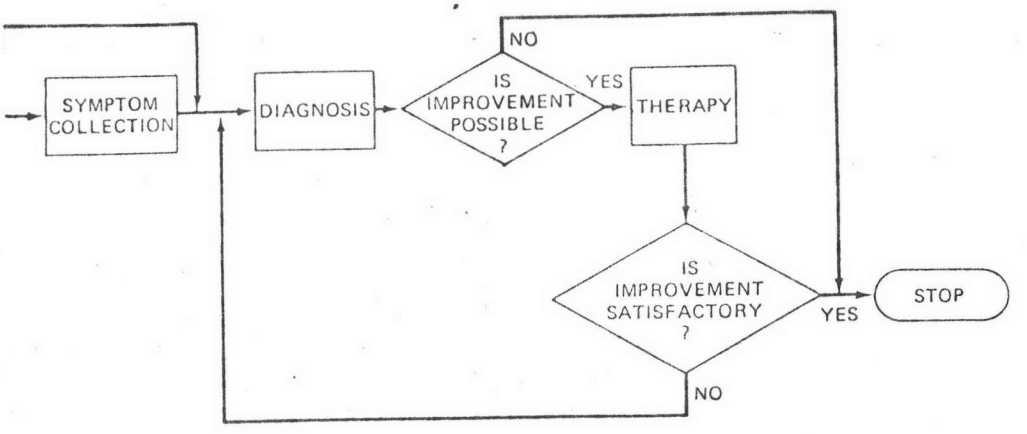
3.4.3.7 ผลลำหรับการประเมินความสามารถและคุณภาพเพื่อการติดตามและตรวจสอบระบบ

รูปที่ 3.10 แสดงการไหลของการประเมินความสามารถและคุณภาพเพื่อการติดตามและตรวจ-
ล่อบระบบ



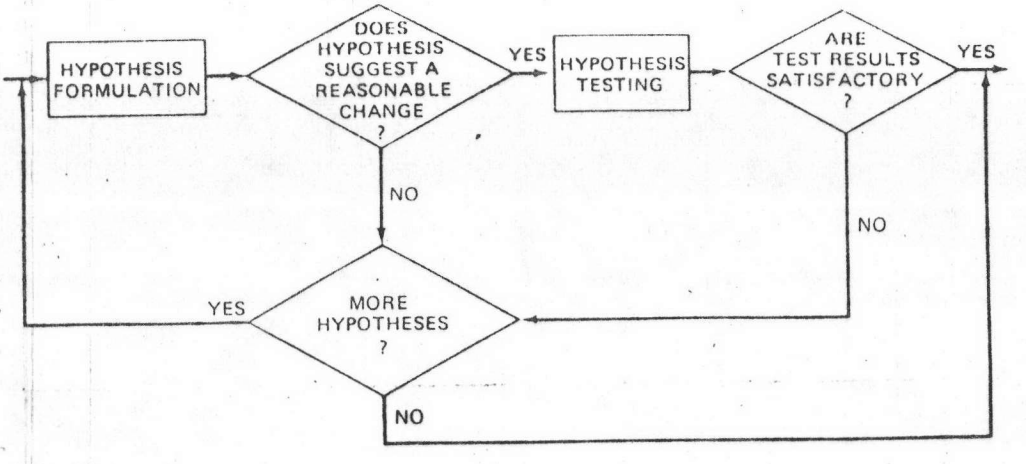
3.5 วิธีการวางแผนปรับปรุงการดำเนินงาน

สิ่งสำคัญในการทำการประเมินผลความสามารถและคุณภาพของระบบนั้น เพื่อที่จะนำผลการประเมินไปปรับปรุงการดำเนินงาน ความได้เปรียบทางด้านการลงทุนของหน่วยงานคอมพิวเตอร์ที่มีการจัดและตรวจล่อบความสามารถ เพราะภาระงานชนิดเดียวกันจะสามารถประมวลผลในเวลาที่น้อยกว่า การทำงานของบุคลากรลดน้อยลง ให้บริการแก่ผู้ใช้บริการได้มากขึ้น การที่จะต้องเพิ่มอุปกรณ์เพื่อการดำเนินงานก็จะจัดสรรได้เหมาะสมยิ่งขึ้น ดังนั้น จึงต้องมีวิธีการวางแผนปรับปรุงการดำเนินงานสำหรับการนำผลการประเมินความสามารถของระบบคอมพิวเตอร์ดังรูปที่ 3.11



รูปที่ 3.11 แสดงการไหลของการปรับปรุงความสามารถและคุณภาพ

จากอาการของสภาพระบบ (Symptoms) เราสามารถทำการเก็บรวบรวมได้จากการทำประเมินความสามารถแล้วพบจุดบอด (Bottleneck) เกิดขึ้นในระบบหรือจากรายงานสถิติต่าง ๆ เป็นต้น จากนั้นเราก็จะทำการวินิจฉัยอาการของสภาพระบบ (Diagnosis) โดยปกติการวินิจฉัยจะต้องศึกษาและรู้ถึงสภาพของอาการ และพบว่าเกิดมีจุดบกพร่องของความสามารถ (Performance Bug) ณ จุดใดบ้าง และสามารถลดจุดบกพร่องหรือปรับเปลี่ยนให้มีสภาพที่ดีขึ้นได้หรือไม่ ดังนั้นต้องทำการตั้งสมมติฐาน ดังรูปที่ 3.12



รูปที่ 3.12 แสดงการไหลการกำหนดสมมติฐานของการปรับปรุงดำเนินการ

สมมติฐานดังกล่าวมีข้อกำหนดเบื้องต้นที่จะทำการศึกษา คือ

1. การกำหนดสมมติฐานนั้น ได้ข้อมูลจากการเก็บรวบรวมข้อมูลจากส่วนต่าง ๆ ของระบบ เพื่อผู้เก็บรวบรวมข้อมูลเพียงบางส่วนของหน่วยงานอาจจะได้ข้อมูลของระบบไม่ครบถ้วน

2. ความรู้พื้นฐานของการกำหนดสมมติฐานขึ้นอยู่กับการทราบถึงค่าดัชนีความสามารถภายใต้ข้อจำกัดของภาระงานและระบบงานของหน่วยงานนั้น ประสบการณ์จะเป็นสิ่งที่ทำให้กำหนดสมมติฐานได้ตรงยิ่งขึ้น

3. การตั้งสมมติฐานต้องคำนึงถึงต้นทุนในการศึกษาเพื่อให้ได้ข้อมูลเหล่านั้นมา และเมื่อใช้สมมติฐานดังกล่าวได้แต่เพียงระยะสั้นนั้น อาจจะไม่เหมาะสมกับการลงทุน

4. สมมติฐานดังกล่าวผู้วิเคราะห์ต้องสามารถจะยอมรับหรือปฏิเสธได้ ซึ่งจะมีการกำหนดเพื่อการเปลี่ยนแปลงอุปกรณ์ เปลี่ยนแปลงระบบคอมพิวเตอร์ หรือการจัดภาระงาน

วิธีการสำหรับการกำหนดสมมติฐาน

1. กำหนดให้มีสภาพการทำงานที่เหมือนกัน กำหนดสมมติฐานให้เหมือนกับการทำปรับปรุงการดำเนินงานครั้งก่อน ๆ หรือ จากหน่วยงานที่มีระบบใกล้เคียง

2. ค้นหาค่าดัชนีบางอย่างที่สำคัญ ค่าดัชนีความสามารถ (Performance Indices)

3. กำหนดโดยรูปแบบตามเวลา การกำหนดค่าตั้งในเชิงปริมาณตามเวลา

4. กำหนดโดยความสัมพันธ์ของพารามิเตอร์ของระบบคอมพิวเตอร์และภาระงานภายใต้ระบบคอมพิวเตอร์นั้น

5. กำหนดตามความไม่สอดคล้อง ตามงาน (Job Relate) และตามระบบ (System Relate)

การทำการรักษา (Therapy) เป็นวิธีการแก้ไขและปรับปรุงระบบคอมพิวเตอร์นั้นมีหลายวิธีการ เช่น

1. การเปลี่ยนแปลงระบบ System Change

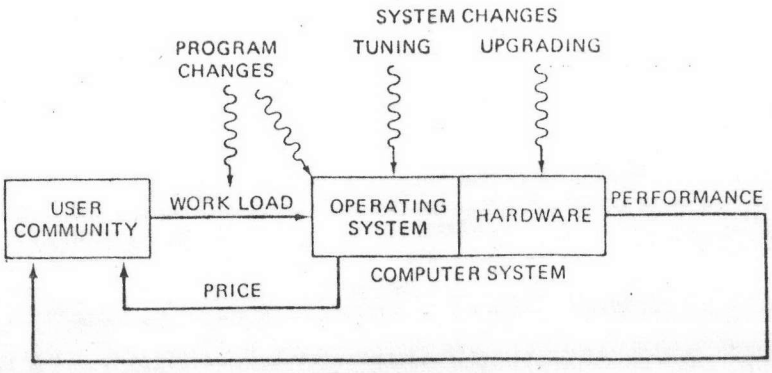
2. การเปลี่ยนแปลงโปรแกรม Program Change

ในการเปลี่ยนแปลงโปรแกรมนั้น บุคคลากรในส่วนงานพัฒนาโปรแกรมอาจจะมีส่วนปรับปรุงโปรแกรมงานให้มีประสิทธิภาพขึ้น ทำให้ความสามารถของระบบเปลี่ยนแปลงในทางที่ดีได้ ส่วนการเปลี่ยนแปลงระบบนั้น จะพิจารณาได้ลักษณะ ดังรูปที่ 3.13 คือ

1. การเปลี่ยนแปลง คือส่วนฮารัดแวร์ วิธีการนี้จะเรียกว่าการเพิ่มความสามารถเครื่อง (Upgrading)

2. การเปลี่ยนแปลงโปรแกรมควบคุมระบบ วิธีการนี้จะเรียกว่า การปรับสภาพความสามารถ (Tuning)

การเปลี่ยนแปลงส่วนของฮารัดแวร์ (Upgrading)การนำผลจากการประเมินผลความสามารถมาพิจารณายังต้องพิจารณาในส่วนการวางแผนการลงทุน การขยายงานของหน่วยงานถือว่ามีปัจจัยมาร่วมพิจารณา ดังนั้น จึงขอกล่าวในส่วนของการปรับสภาพความสามารถ (Tuning)



รูปที่ 3.13 แสดงความสัมพันธ์การปรับความสามารถของระบบ

การปรับสภาพของระบบคอมพิวเตอร์เป็นวิธีการเปลี่ยนแปลงความสามารถและคุณภาพของปัจจัยต่าง ๆ ภายในระบบคอมพิวเตอร์ เพื่อให้ได้ระดับกับความพอใจในการบริการด้านประมวลผลของหน่วยงาน การปรับสภาพระบบคอมพิวเตอร์แบ่งออกเป็นหลายประเภทดังนี้

3.5.1.1 พารามิเตอร์ของโปรแกรมควบคุมระบบ โดยมีวิธีการปรับพารามิเตอร์ของโปรแกรมควบคุมระบบดังนี้

3.5.1.1.1 ระดับมัลติโปรแกรมมิ่ง (Degree of Multiprogramming) เป็นการกำหนดพาร์ติชั่นสำหรับการทำงานแบบมัลติโปรแกรมมิ่งว่ามีกี่พาร์ติชั่นจะปรับให้มากขึ้นหรือลดน้อยลง และขนาดหน่วยความจำของแต่ละพาร์ติชั่น

3.5.1.1.2 ขนาดของบัฟเฟอร์ (Buffer Size) สำหรับการอ่านและ

บันทึกข้อมูลที่เป็นบล็อก การมีขนาดขึ้นอยู่กับบัฟเฟอร์ ดังนั้นจึงต้องปรับขนาดของบัฟเฟอร์เพื่อลดการเรียกหน่วยอินพุท หน่วยเอาท์พุท

3.5.1.1.3 เวลาของการทำงาน (Time Quantum) โดยการปรับเวลาของการทำงานของแต่ละกลุ่มของงาน

3.5.1.1.4 การกำหนดลำดับความสำคัญของงานแต่ละกลุ่ม (Priority Classes) ที่จะให้งานในกลุ่มงานใดมีสิทธิ์เข้าทำงานในหน่วยประมวลผลตามลำดับความสำคัญ

3.5.1.1.5 การกำหนดขนาดของเพจ (Page Size) จะกำหนดให้แต่ละเพจของการทำเพจจิง

3.5.1.2 การจัดการทรัพยากรระบบ (Resource Management) โดยมีวิธีการปรับการจัดการทรัพยากรระบบ ดังนี้

3.5.1.2.1 การจัดขั้นตอนเข้าทำงานของซีพียู, ของงาน และของงานแม่เหล็ก

3.5.1.2.2 การจัดแฟ้มข้อมูล (Access Method)

3.5.1.2.3 การจัดการหน่วยความจำเกี่ยวกับการกำหนดที่ (Allocation) การแทนที่ (Replacement) การโหลด (Loading) และการสลับ (Swapping)

3.5.1.2.4 การจัดการหน่วยอินพุท หน่วยเอาท์พุท เป็นการจัดหน่วยอินพุท หน่วยเอาท์พุทตามช่องสถานี หน่วยควบคุมชนิดของหน่วยอินพุท หน่วยเอาท์พุทให้มีการใช้หน่วยช่องสถานีให้เหมาะสมไม่ไปหนักที่จุดใดจุดหนึ่งมากเกินไป

3.5.1.2.5 การจัดการทำงานของโอเปอเรเตอร์ เป็นการจัดลำดับงานเข้าทำงานด้วยโอเปอเรเตอร์ การจัดการเกี่ยวกับการใส่ตู้เทป การใส่ตู้จานแม่เหล็ก การจัดงานเข้าทำงานภายใต้การควบคุมของโอเปอเรเตอร์ ตลอดจนการจัดวางหน่วยอินพุทหน่วยเอาท์พุทให้สอดคล้องเกี่ยวกับการทำงาน

3.5.1.3 การติดตั้งการติดต่อ (Interconnections) สำหรับการติดต่อระหว่างของหน่วยต่าง ๆ ดังนี้

3.5.1.3.1 การติดต่อระหว่างช่องสถานีกับหน่วยอินพุท หน่วยเอาท์พุท เอาท์พุทและหน่วยควบคุม (Channel-Device Connection)

3.5.1.3.2 การติดต่อระหว่าง หน่วยอินพุท และหน่วยเอาท์พุท (Device-Bus Connection)

3.5.1.4 การจัดที่สำหรับข้อมูล สำหรับข้อมูลที่จะจัดการที่สำคัญ แบ่งเป็น

3.5.1.4.1 การจัดข้อมูลในส่วนแฟ้มข้อมูล ที่จะเก็บไว้ในหน่วยความจำแบบลำดับขั้น (Hierarchy) อย่างไร

3.5.1.4.2 การจัดข้อมูลที่เป็นโปรแกรมควบคุมระบบจะอยู่ในหน่วยจานแม่เหล็กใดบ้าง

3.5.1.5 การปรับความสามารถและคุณภาพของโปรแกรมต่าง ๆ ได้แก่

3.5.1.5.1 ส่วนการใช้งาน (Overhead) ของโปรแกรมควบคุมระบบ

3.5.1.5.2 เวลาของการประมวลผล และขนาดของหน่วยความจำที่ใช้หรือต้องการ สำหรับการใช้โปรแกรมที่ใช้บ่อย ๆ เช่น โปรแกรมการแปล (Compiler) โปรแกรมตรวจล่อบ (Editer) โปรแกรมระบบฐานข้อมูล (Data Base System) โปรแกรมการเรียงลำดับ (Sort)

3.5.1.5.3 รูปแบบการอ้างอิงของหน่วยความจำ

3.5.1.6 การปรับนโยบายเกี่ยวกับราคาของการจัดงานทำงานในส่วนต่าง ๆ ดังนี้

3.5.1.6.1 การกำหนดราคาหรือสิ่งจูงใจบางอย่าง สำหรับช่วงเวลาที่มีงานน้อย (Light-Load Period) เพื่อให้ผู้ใช้สนใจมาทำงานในช่วงเวลาดังกล่าว

3.5.1.6.2 การกำหนดข้อลงโทษ หรือการงดทำงานของโปรแกรมที่ไม่มีประสิทธิภาพ

3.5.1.6.3 การกำหนดข้อลงโทษ หรือกำหนดราคาสูงกว่าสำหรับการใช้ทรัพยากรระบบจำนวนมากพร้อม ๆ กัน (Heavily Loaded Resource)

3.5.1.7 การปรับนโยบายการยอมรับงานเพื่อทำการประมวลผลโดยพิจารณาถึงส่วนต่าง ๆ เช่น

3.5.1.7.1 กำหนดประเภทงานของผู้ใช้ที่จะมีผลต่อการจัดลำดับงานเข้าทำงาน

3.5.1.7.2 กำหนดข้อจำกัดของการใช้ทรัพยากรต่าง ๆ (Load Limitation) เวลาและช่วงเวลา เป็นต้น

3.5.2 การปรับสภาพระบบคอมพิวเตอร์ในโปรแกรมควบคุมระบบ วีเอ็ม/เอสพี เพื่อเพิ่มประสิทธิภาพความสะดวกและคุณภาพของระบบคอมพิวเตอร์ได้หลายส่วน เป็นต้นว่า

3.5.2.1 การปรับการใช้งานของโปรแกรมควบคุม โดยการกำหนดหรือเลือกใช้พารามิเตอร์ให้เหมาะสม เช่น

3.5.2.1.1 การกำหนด V= R เพื่อให้โปรแกรมควบคุมจัดการหน่วยความจำและการเพจจิง ก่อนที่จะทำหน่วยอินพุท หน่วยเอาท์พุทให้เครื่องเสมือนเพื่อไม่ให้เกิดการทำเพจจิงสองครั้ง (Double Paging)และจัดให้หน่วยความจำหลักสามารถบรรจุได้เท่ากับหน่วยความจำเสมือน

3.5.2.1.2 การเลือกใช้ วีเอ็มเอสเซสซิสเจอร์ (VM Assist Feature) เพื่อลดเวลาการทำงานของหน่วยความจำหลักจริง

3.5.2.1.3 การเลือกใช้วีเอ็ม/เอสพี โปรแกรมควบคุมสนับสนุนเพิ่มเติม (VM/SP Extended Control Program Support) เพื่อลดการทำงานของซูเปอร์ไวเซอร์ (Supervisor State)

3.5.2.2 การปรับการสตาร์ทหน่วยอินพุทหน่วยเอาท์พุท (Start I/O) ของเครื่องเสมือนโดยมีวิธีการต่าง ๆ ดังนี้

3.5.2.2.1 การกำหนดใช้ขนาดของบล็อกของข้อมูลเพิ่มขึ้น (Large Blocking) สำหรับแฟ้มข้อมูลของผู้ใช้เพื่อลดจำนวนการสตาร์ทหน่วยอินพุทหน่วยเอาท์พุท

3.5.2.2.2 การกำหนดที่ของแฟ้มข้อมูลไว้ล่วงหน้า (Preallocate)

3.5.2.2.3 การกำหนดทางเลือกของโปรแกรมควบคุมระบบ เพื่อลดการสตาร์ทหน่วยอินพุท หน่วยเอาท์พุท

3.5.2.2.4 การเลือกใช้อุปกรณ์ที่เร็วกว่ามาแทน เช่น การเลือกใช้แฟ้มข้อมูลชั่วคราว (Temporary File) แทนการเรียกใช้หน่วยอินพุท หน่วยเอาท์พุทจริง

3.5.2.3 การปรับสภาพด้วยการกำหนดทางเลือกต่าง ๆ ของโปรแกรมควบคุมระบบวีเอ็ม/เอสพี เช่น

3.5.2.3.1 การล็อกเพจ (Lock Page) เป็นการกำหนดให้ฟิตเพจไว้ในหน่วยความจำจริง สำหรับงานที่ใช้จำนวนมาก

3.5.2.3.2 การกำหนดสำรองเพจ (Reserved Page Frames) เพื่อกำหนดให้เพจที่มีการเรียกใช้ร่วมกันสำรองไว้ สำหรับงานที่สำคัญและงานที่วิ่งในระบบออนไลน์

3.5.2.3.3 การกำหนดลำดับความสำคัญ เพื่อให้งานที่สำคัญมีโอกาสเข้าทำงานเร็วขึ้น

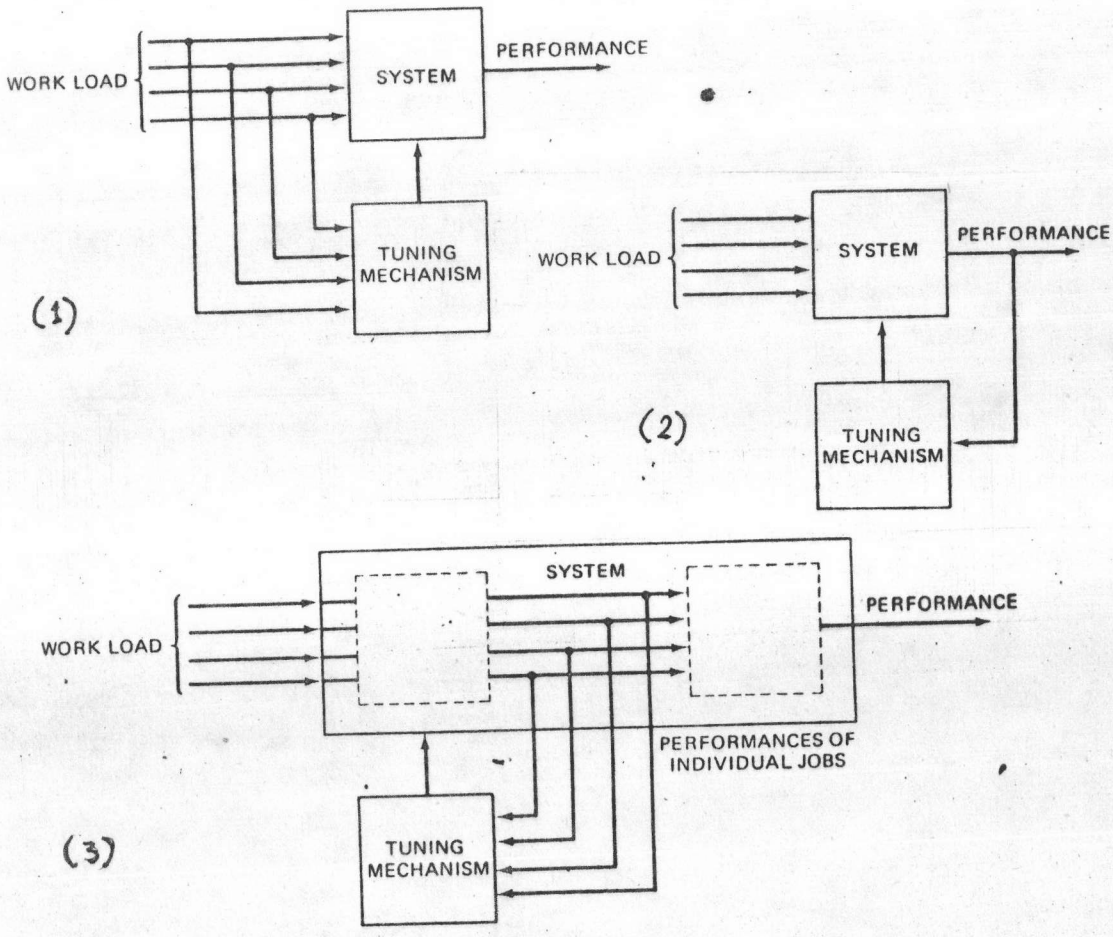
3.5.2.3.4 การกำหนดความชอบของการทำงานให้มีเวลาในการประมวลเพิ่มขึ้นจากที่กำหนด

3.5.2.3.5 การเลือกใช้ วีเอ็ม/เอล่เซล่ ซึ่งเป็นลักษณะพิเศษของฮาร์ดแวร์ของหน่วยประมวลผล เพื่อเพิ่มความล่ามารถและคุณภาพของวีเอ็ม/เอล่พี

3.5.3 การปรับล่ภาพของโปรแกรมควบคุมระบบและภาระงานอีกอย่างหนึ่ง คือการกำหนดเป็นการปรับล่ภาพโดยอัตโนมัติ (Automated Tuning) โดยการตรวจล่อบค่าปริมาณต่าง ๆ ที่เหมาะล่มแล้วกำหนดเป็นเครื่องมือการตรวจล่อบต่อเนื่องโดยมีรูปแบบ 3 แบบ คือ

1. การตรวจล่อบพารามิเตอร์ภาระงาน (Feed Forward Control)
2. การกำหนดค่าดัชนีความล่ามารถและคุณภาพ (Feed Back Control)
3. การกำหนดแบบผล่ม (Hybrid Mechanism) รวมการทำงานทั้งล่องร่วมกัน

ดังรูปที่ 3.14



รูปที่ 3.14 แสดงการใช้การปรับสภาพอัตโนมัติของความสามารถของระบบ

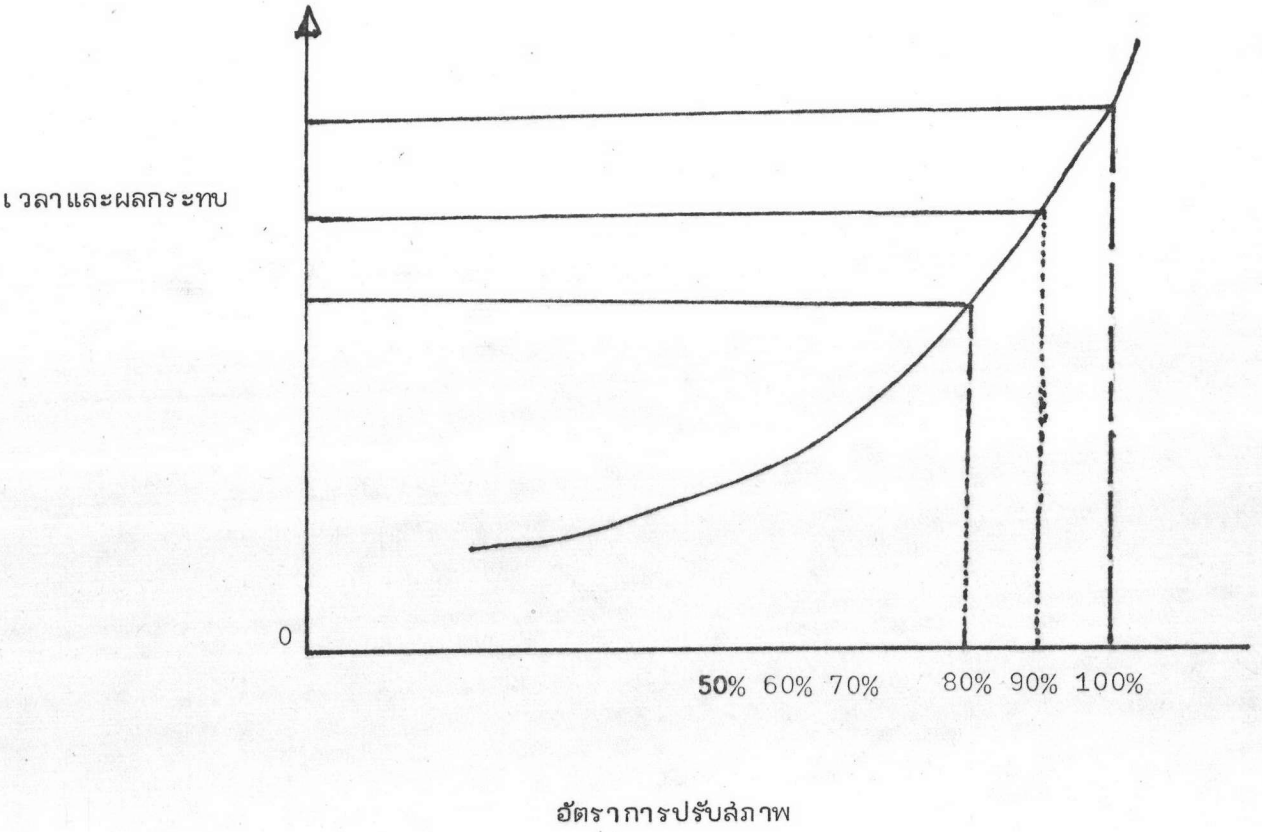
3.5.4 การปรับสภาพระบบคอมพิวเตอร์ต้องคำนึงถึงสิ่งต่าง ๆ เช่น

1. หลีกเลี่ยงที่จะไม่เปลี่ยนแปลงในสิ่งที่ไม่จำเป็นของภาระงาน

2. หลีกเลี่ยงการรอคอยที่ไม่จำเป็นเพื่อปรับสภาพความสามารถและคุณภาพ

3. การแลกเปลี่ยนเมื่อเปลี่ยนแปลงความสามารถและคุณภาพด้านหนึ่งอาจจะต้องกระทบกับงาน หรือภาระงานด้านใดด้านหนึ่ง หรือแลกเปลี่ยนกับสิ่งอื่น

4. การปรับสภาพควรหยุดเมื่อใด และคุ้มค่ากับการลงทุนปรับสภาพ หรือไม่



รูปที่ 3.15 แสดงการเปรียบเทียบเวลาและผลกระทบกับอัตราการปรับสภาพ