# บทที่ 2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ในการดำเนินงานวิจัยให้บรรลุวัตถุประสงค์ในการพัฒนาระบบต้นแบบการเฝ้าระวังข้อผิดพลาดเบื้องต้นอัตโนมัติ สำหรับระบบคลัสเตอร์ เว็บเซิร์ฟเวอร์ กรมตรวจบัญชีสหกรณ์ผู้วิจัยได้ศึกษาบทความ ทฤษฎีและงานวิจัยที่เกี่ยวข้องประกอบไปด้วย

2.1 ความหมายและหลักการทำงานของพีซีคลัสเตอร์

2.2 สถาปัตยกรรมของคอมพิวเตอร์คลัสเตอร์

2.3 แนวคิดเกี่ยวกับคลัสเตอร์ เว็บเซิร์ฟเวอร์

2.4 งานวิจัยที่เกี่ยวข้อง

## 2.1 ความหมายและหลักการทำงานของพีซีคลัสเตอร์

### 2.1.1 ระบบพีซีคลัสเตอร์ (PC Cluster)

ระบบพีซีคลัสเตอร์ (PC Cluster) หมายถึง การนำเครื่องคอมพิวเตอร์ส่วนบุคคลที่มีจำนวนมากกว่า 2 เครื่องขึ้นไปนำมาเชื่อมต่อกัน ผ่านเครือข่ายความเร็วสูงเพื่อให้เครื่องทุกเครื่องในระบบสามารถทำงานร่วมกันได้เสมือนเป็นเครื่องเดียวกัน ความแตกต่างระหว่างระบบคอมพิวเตอร์คลัสเตอร์กับระบบแลน (LAN) คือ ถึงแม้ว่าระบบ LAN จะเป็นการเชื่อมต่อเครื่องคอมพิวเตอร์จำนวนมากเข้าด้วยกันเหมือนระบบคอมพิวเตอร์คลัสเตอร์เพื่อใช้ทรัพยากรภายนอกร่วมกันแต่เครื่องแต่ละเครื่องจะมีระบบการทำงานที่อิสระแยกออกจากกัน ไม่สามารถออกคำสั่งให้เครื่องในระบบแลนทำการร่วมใช้ประโยชน์ทรัพยากรภายในเครื่อง เช่น CPU, Ram, Harddisk เป็นต้น ด้วยเหตุนี้จึงทำให้ระบบคอมพิวเตอร์คลัสเตอร์ได้รับความนิยมในการนำมาใช้ทำงานแทนระบบคอมพิวเตอร์สมรรถนะสูง

#### 2.1.1.1 องค์ประกอบสำคัญของเทคโนโลยีคลัสเตอร์มีอยู่ 3 ประการ คือ

1. เครื่องคอมพิวเตอร์และเครือข่ายความเร็วสูงซึ่งเป็นองค์ประกอบหลักของโครงสร้างระบบคอมพิวเตอร์คลัสเตอร์โดยเป็นการนำเครื่องคอมพิวเตอร์มากกว่า 2 ตัวขึ้นไปมาทำการเชื่อมต่อกันผ่านเครื่องข่ายความเร็วสูง

2. ซอฟต์แวร์ที่ใช้สนับสนุนในการบริหารการทำงานแบบคลัสเตอร์ เช่น Sun Grid Engine ซึ่งเป็นโปรแกรมจัดลำดับงานประเภทหนึ่ง (Job Scheduler)

3. โปรแกรมประยุกต์ที่มีความสามารถในการรองรับการทำงานกับแบบระบบคลัสเตอร์ได้

### 2.1.2 การทำงานของคอมพิวเตอร์คลัสเตอร์

หลักการทำงานของระบบคอมพิวเตอร์คลัสเตอร์ คือ ใช้โปรแกรมจัดลำดับงาน (Job Scheduler) ที่ติดตั้งไว้ยังเครื่องที่ทำหน้าที่กระจายงานซึ่งเราเรียกเครื่องคอมพิวเตอร์ที่ทำหน้าที่นี้ว่า เครื่องเฮดโหนด (Head Node) มาทำหน้าที่ควบคุมกระจายงานไปยังเครื่องคอมพิวเตอร์เครื่องอื่นๆ ที่อยู่ภายในระบบคอมพิวเตอร์คลัสเตอร์ให้ทำการประมวลผลซึ่งเราเรียกเครื่องคอมพิวเตอร์ที่ทำหน้าที่ประมวลผลงานนี้ว่า เครื่องคอมพิวส์โหนด (Compute Node) และหลังจากที่เครื่องคอมพิวส์โหนดดำเนินการประมวลผลงานเสร็จสิ้นเรียบร้อยแล้วเครื่องคอมพิวส์โหนดจะทำการส่งผลลัพธ์กลับไปยังเครื่องเฮดโหนด

### 2.1.3 โครงสร้างของระบบพีซีคลัสเตอร์

ระบบพีซีคลัสเตอร์ คือ การนำเอาเครื่องพีซีคอมพิวเตอร์หลายๆตัวมาเชื่อมต่อกันด้วยเครือข่ายความเร็วสูงเพื่อให้ทำงานเหมือนเครื่องใหญ่เครื่องเดียว โครงสร้างของระบบพีซีคลัสเตอร์ประกอบด้วย เครื่องพีซีคอมพิวเตอร์ที่มีสมรรถนะสูงและเชื่อมต่อกันด้วยเครือข่ายความเร็วสูง เช่น ระบบ ATM Switch, Fast Ethernet Switch, Myrinet และเมื่อพิจารณาโครงสร้างของระบบพีซีคลัสเตอร์แล้ว จะพบว่า ระบบพีซีคลัสเตอร์ เป็นคอมพิวเตอร์แบบขนานที่มีหน่วยความจำแยกการเชื่อมต่อกันระหว่างเครื่องคอมพิวเตอร์ในระบบพีซีคลัสเตอร์ต่างจากการเชื่อมต่อกันของเครื่องคอมพิวเตอร์ในระบบแลน คือ เครื่องคอมพิวเตอร์ทุกเครื่องที่อยู่บนระบบแลนจะเป็นอิสระต่อกัน และไม่มีระบบซอฟต์แวร์ที่นำความสามารถของการประมวลผลแบบขนานและแบบกระจายมาใช้ และการทำงานเป็นลักษณะแยกกัน แต่ระบบพีซีคลัสเตอร์มีส่วนสำคัญสามส่วนคือ ระบบเครือข่ายความเร็วสูง ระบบซอฟต์แวร์ที่สนับสนุนการทำงานแบบคลัสเตอร์ และโปรแกรมประยุกต์ที่ใช้ขีดความสามารถของการประมวลผลแบบขนานหรือแบบกระจาย

โครงสร้างของระบบคลัสเตอร์ แบ่งออกเป็น 2 ชนิด คือ ระบบคลัสเตอร์แบบปิด และระบบคลัสเตอร์แบบเปิด ระบบคลัสเตอร์แบบปิด จะต่อคลัสเตอร์ผ่านเกตเวย์ที่ซ่อนทั้งระบบจาก โลกภายนอก ข้อดี คือ มีความปลอดภัยสูงและใช้อินเทอร์เน็ตแอดเดรสเพียงแอดเดรสเดียวเท่านั้น ข้อเสีย คือ แต่ละโหนดในระบบไม่สามารถช่วยกันบริหารข้อมูลจากภายนอกได้ ระบบคลัสเตอร์ แบบเปิด จะต่อกับเน็ตเวิร์คภายนอกโดยตรงทำให้ผู้ใช้เข้าถึงทุกโหนดในระบบคลัสเตอร์ได้โดยตรง ข้อเสีย คือ ความปลอดภัยจะต่ำลงมากเพราะต้องคอยดูแลทุกเครื่องในระบบ และยังต้องการหมายเลขอินเทอร์เน็ตแอดเดรสจำนวนมาก แต่ข้อดี คือ ระบบสามารถช่วยกันบริการข้อมูลได้ จึงเหมาะกับงานบริการข่าวสารเป็นจำนวนมาก เช่น ในระบบเซิร์ฟเวอร์สำหรับ www หรือ ftp ที่ขยายตัวได้

จากโครงสร้างของพีซีคลัสเตอร์จะเห็นได้ว่าระบบคลัสเตอร์และระบบแลนนั้นมีความแตกต่างกันอย่างเห็นได้ชัดการอิสระต่อกันของระบบแลนและการช่วยกันประมวลผลของระบบคลัสเตอร์ที่ทำให้หน่วยประมวลผลนั้นมากขึ้นตามจำนวนเครื่องที่ทำคลัสเตอร์ ดังนั้นผู้วิจัยได้เลือกระบบคลัสเตอร์แบบปิดมาประยุกต์ใช้ในงานวิจัยเนื่องจากมีความปลอดภัยสูงโดยใช้อินเทอร์เน็ตแอดเดรสเพียง 1 แอดเดรสเป็นเกตเวย์ที่ให้ผู้ใช้บริการเว็บไซต์เรียกใช้งานทั้งยังเป็นแอดเดรสที่ซ่อนอินเทอร์เน็ตแอดเดรสจริงๆของเครื่องคอมพิวเตอร์คลัสเตอร์ภายในจึงเหมาะที่จะนำไปใช้กับเว็บเซิร์ฟเวอร์ที่ให้บริการทางด้านเว็บไซต์อย่างยิ่ง

### 2.1.4 แนวโน้มของพีซีคลัสเตอร์

ในปัจจุบันมีแนวโน้มทางเทคโนโลยีหลายประการที่สนับสนุนการก้าวไปสู่ระบบคลัสเตอร์ เช่น

1. โปรเซสเซอร์ของเครื่องคอมพิวเตอร์ในปัจจุบันมีสมรรถนะแทบจะไม่แตกต่างจากเครื่องซูเปอร์คอมพิวเตอร์ในอดีต ดังนั้นการเชื่อมระบบพีซีเข้าด้วยกันเพื่อทำงานเป็นซูเปอร์คอมพิวเตอร์จึงมีสมรรถนะใกล้เคียงกับเครื่องซูเปอร์คอมพิวเตอร์

2. เนื่องจากฮาร์ดแวร์คอมพิวเตอร์ในปัจจุบันมีการแข่งขันในเรื่องการตลาดเป็นจำนวนมาก ซึ่งผลที่เกิดขึ้น คือ ราคาเครื่องพีซีคอมพิวเตอร์ที่ถูกลงมากเมื่อเทียบกับประสิทธิภาพที่สูงขึ้น นอกจากนี้การแข่งขันของการตลาดที่สูงขึ้นในทุกวันยังทำให้เงินทุนที่ลงทุนให้กับการวิจัยพัฒนาเป็นไปได้อย่างรวดเร็ว ทำให้ได้ทรัพยากรที่มีประสิทธิภาพสูงขึ้นมากในราคาที่เท่าเดิม

3. สิ่งที่เกิดขึ้นตามมาสำหรับการใช้เครื่องซูเปอร์คอมพิวเตอร์ คือ ค่าบำรุงรักษาที่สูงมาก นอกจากนี้ เมื่อเทคโนโลยีใหม่มาแทนที่เทคโนโลยีตัวเก่า การเพิ่มความสามารถของระบบต้องใช้เงินทุนจำนวนมากและเป็นไปได้ยากในขณะที่ การเพิ่มความสามารถในระบบพีซีคลัสเตอร์ทำได้ทีละน้อยในราคาถูกกว่า นอกจากนั้นเครื่องที่นำออกจากระบบยังสามารถนำไปประยุกต์ใช้เพื่อประโยชน์อื่นๆ ต่อได้อีก

4. ความก้าวหน้าของซอฟต์แวร์ เช่น ระบบปฏิบัติการลีนุกซ์ ซึ่งเป็นระบบปฏิบัติการที่มีประสิทธิภาพสูงแต่ไม่ต้องเสียค่าใช้จ่ายในการใช้งาน ระบบพัฒนาโปรแกรมแบบขนานโดยใช้ PVM (Parallel Virtual Machine) และ MPI (Message Passing Interface) ทำให้สามารถสร้างและใช้ความสามารถของระบบคลัสเตอร์ได้

5. ระบบพีซีเป็นเทคโนโลยีที่คนส่วนใหญ่คุ้นเคยทำให้สามารถบำรุงรักษาระบบได้ง่ายกว่า [1]

## 2.2 สถาปัตยกรรมของคอมพิวเตอร์คลัสเตอร์

### 2.2.1 ระบบคลัสเตอร์ในปัจจุบัน

ระบบคลัสเตอร์ในปัจจุบันสามารถจำแนกได้ 3 รูปแบบคือ

#### 2.2.1.1 คลัสเตอร์ประสิทธิภาพสูง (High Performance Cluster: HPC)

ใช้สำหรับประมวลผลกับข้อมูลขนาดใหญ่ ที่ต้องการกำลังประมวลผลสูงโดยให้ได้ผลลัพธ์ในเวลาอันรวดเร็วซึ่งส่วนมากเป็นงานทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์เช่น การประมวลผลภาพ (Image Processing), การสร้างแบบจำลอง (Simulation Modeling), การประมวลผลฟูลอิตไดนามิกส์ (Computational Fluid Dynamic) เป็นต้น

#### 2.2.1.2 คลัสเตอร์ประมวลผลปริมาณงานสูง (High Throughput Cluster : HTC)

ใช้สำหรับประมวลผลกับงานที่มีจำนวนมากโดยที่แต่ละงานเป็นอิสระต่อกันใช้เวลาประมวลผลนานเป็นเดือนหรือปีโดยเป็นงานทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์ แต่ส่วนใหญ่จะเป็นงานวิจัยทางด้านวิเคราะห์ทางเคมี, ชีววิทยาและฟิสิกส์ เช่น ค้นหาตัวยา (Drug Discovery), การวิเคราะห์โครงสร้าง (Structural Analysis) และเว็บเซอร์วิส (Web Service) เป็นต้น

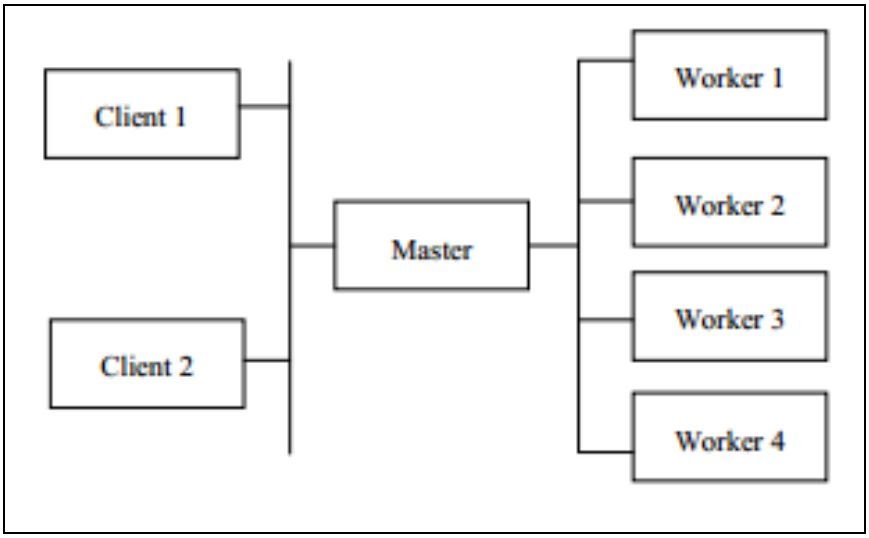
#### 2.2.1.3 คลัสเตอร์ที่มีสภาพพร้อมใช้งานสูง (High Availability Clustering : HAC )

เพื่อใช้เพิ่มความเสถียรภาพของการประมวลผล ใช้สำหรับงานที่สำคัญและหยุดการประมวลผลไม่ได้ โดยต้องการประมวลผลคู่ขนาน จำนวน 2 หน่วยประมวลผลอย่างน้อย ซึ่งส่วนใหญ่เป็นงานทางด้านธุรกิจการเงินและธนาคารเช่น ระบบบัญชี (Accounting System), ระบบการเงิน (Financial System) เป็นต้น

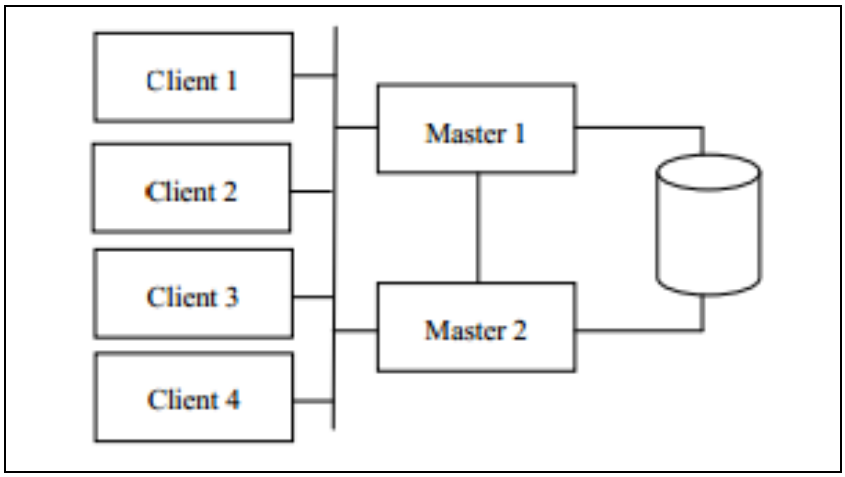
### 2.2.2 สถาปัตยกรรมทางด้านฮาร์ดแวร์

สถาปัตยกรรมทางด้านฮาร์ดแวร์ในการเชื่อมต่อของหน่วยประมวลผลผ่านเครือข่ายเน็ตเวิร์คของระบบคลัสเตอร์มีความแตกต่างกันโดย การเชื่อมต่อของคลัสเตอร์ประสิทธิภาพสูง (HPC) และคลัสเตอร์ประมวลผลปริมาณงานสูง (HTC) มีลักษณะของการเชื่อมต่อเหมือนกัน ส่วนลักษณะโครงสร้างการเชื่อมต่อคลัสเตอร์มีสภาพพร้อมใช้งานสูง (HAC) จะมีความแตกต่างกันที่ใช้ฮาร์ดแวร์และหน่วยประมวลผลน้อยกว่าแต่ยังคงความเสถียรของระบบที่ไม่สามารถให้ระบบนั้นหยุดทำงานได้ สถาปัตยกรรมของการเชื่อมต่อหน่วยประมวลผลแบบ HPC และ HTC สามารถใช้งานได้อย่าง

กว้างขวางในสถาบันอุดมศึกษาและหน่วยงานของรัฐมากกว่าคลัสเตอร์มีสภาพพร้อมใช้งานสูงแบบ HAC ที่เหมาะสำหรับงานทางด้านธุรกิจ [2]




**รูปที่ 2.1** การเชื่อมต่อหน่วยประมวลผลของคลัสเตอร์แบบ HPC และ HTC [2]




**รูปที่ 2.2** การเชื่อมต่อหน่วยประมวลผลของคลัสเตอร์แบบ HAC [2]

จากรูปที่ 2.1 และ 2.2 จะเห็นได้ว่าการเชื่อมต่อของของคลัสเตอร์แบบ HPC นั้นจะมีหน่วยประมวลผลที่มากกว่าคลัสเตอร์แบบ HAC ทั้งยังสามารถเพิ่มฮาร์ดแวร์ให้มีหน่วยประมวลผลได้มากขึ้นได้เรื่อยๆ แต่ในคลัสเตอร์แบบ HAC นั้นยังคงมีการเชื่อมต่อการประมวลผลแบบคู่ขนาน คือ คงสภาพให้ทำงานได้ต่อเนื่องไม่ให้ระบบหยุดทำงานแต่ไม่สามารถหน่วยประมวลผลได้ [2]

การทำงานทั้งสองรูปแบบนั้นมีการทำงานที่แตกต่างกันแต่ยังคงเป็นโครงสร้างการทำงานในรูปแบบคลัสเตอร์ ในรูปแบบ HAC นั้นมีการทำงานคอยแทนกันและกันอยู่แต่โดยปกตินั้นจะทำงานอยู่เพียงเครื่องเดียวถ้าเครื่องที่ทำงานอยู่เสียอีกเครื่องที่คอยจะทำหน้าที่แทน ส่วนคลัสเตอร์แบบ HPC นั้นเป็นการทำงานในรูปแบบการช่วยกันประมวลผล และมีความสามารถลดหรือเพิ่มเครื่องประมวลผลได้ ทั้งสองรูปแบบนี้สามารถนำมาทำงานร่วมกันได้โดยงานวิจัยในครั้งนี้จะออกแบบการทำงานด้วยการประยุกต์ทั้งสองรูปแบบทั้ง HPC และ HAC เข้าด้วยกันเพียงแต่หน้าที่ของแต่ละรูปแบบจะทำหน้าที่ต่างกัน

## 2.3 แนวคิดเกี่ยวกับคลัสเตอร์เว็บเซิร์ฟเวอร์

Cluster Server เป็นระบบคอมพิวเตอร์คลัสเตอร์ที่ประยุกต์ใช้หลักการกระจายภาระงาน (Load Balancing) บริการเว็บให้กับเครื่องเว็บเซิร์ฟเวอร์หลายเครื่องช่วยกันทำงาน เพื่อให้ระบบสามารถรองรับกับปริมาณผู้ใช้งานเว็บไซต์ที่เข้ามาใช้งานพร้อมกันจำนวนมากได้ ระบบคลัสเตอร์ มุ่งแก้ไขปัญหา 2 เรื่องคือระบบไม่สามารถให้บริการได้และระบบตอบสนองช้า โดยเจาะจงลงไปในส่วนของเครื่องคอมพิวเตอร์ที่ให้สามารถให้บริการได้ตลอดเวลา โดยมีอัตราการหยุดให้บริการต่ำที่สุดการทำงานในลักษณะนี้เรียกว่า High Availability หรือ HA ในภาษาไทยจะใช้คำว่า ภาวะทนต่อความผิดพร่องสูง ซึ่งจะแบ่งวิธีการทำงานออกได้เป็น 2 รูปแบบดังต่อไปนี้

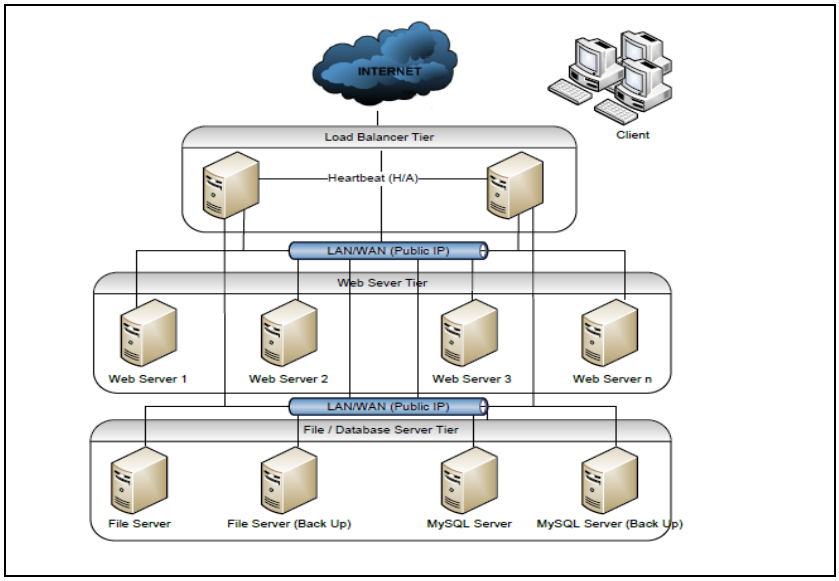
1. Active/Standby หรือ Fail Over ในรูปแบบนี้จะพบในลักษณะของการทำงานของเครื่องคอมพิวเตอร์ 2 เครื่องร่วมกัน โดยที่เครื่องให้บริการหลัก (Active) ค่อยให้บริการอยู่ และเครื่องสำรอง (Standby) ทำหน้าที่ในการปรับปรุงข้อมูลต่างๆ ให้เหมือนกับเครื่องหลักอยู่เสมอและค่อยตรวจสอบว่าเครื่องหลักยังให้บริการได้เป็นปรกติอยู่หรือไม่ ถ้าพบว่าเครื่องหลักไม่สามารถให้บริการได้เครื่องสำรองจะให้บริการแทน

2. Active/Active รูปแบบนี้เครื่องคอมพิวเตอร์ทุกเครื่องในระบบจะให้บริการในงานเดียวทั้งหมดซึ่งในระบบอาจจะมีหลายเครื่อง ถ้ามีเครื่องใดเครื่องหนึ่งให้บริการไม่ได้ เครื่องที่เหลือก็ยังคงให้บริการอยู่เพียงแต่ผู้ใช้งานอาจจะรู้สึกว่าระบบให้บริการช้าลงเหตุเพราะมีเครื่องให้บริการน้อยลงนั้นเอง [3]

### 2.3.1 หลักการทำงานคลัสเตอร์เว็บเซิร์ฟเวอร์

หลักการทำงานคลัสเตอร์ เมื่อเครื่อง Load Balance ได้รับ Request มาตัว Load Balance จะนำ Request นั้นไปกระจายโหลดไปยังเครื่องเซิร์ฟเวอร์ต่างๆที่อยู่ในระบบเว็บคลัสเตอร์โดยมีอัลกอริทึมหลายๆ รูปแบบ เช่น Round-Robin, Weighted Round-Robin Scheduling, Lease Connection, Weight Lease

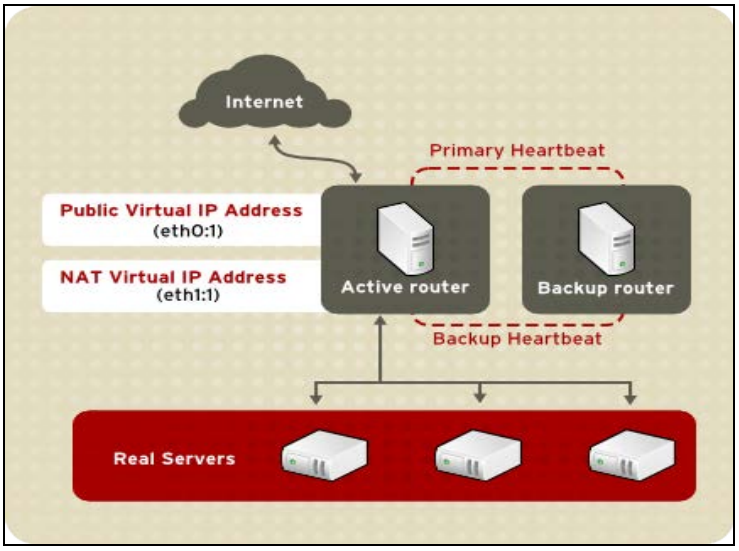
Connection เป็นต้น และเมื่อเครื่องเซิร์ฟเวอร์ได้รับ Request แล้วก็ตอบกลับไปยังเครื่องไคลเอ็นต์ ซึ่งจะมีวิธีตอบกลับที่เป็นที่นิยม 2 วิธีคือ วิธีตอบกลับแบบ NAT และ แบบ Direct Route




**รูปที่ 2.3** หลักการทำงานของ Web Cluster Diagram [3]

### 2.3.2 กระบวนการตอบกลับแบบ NAT

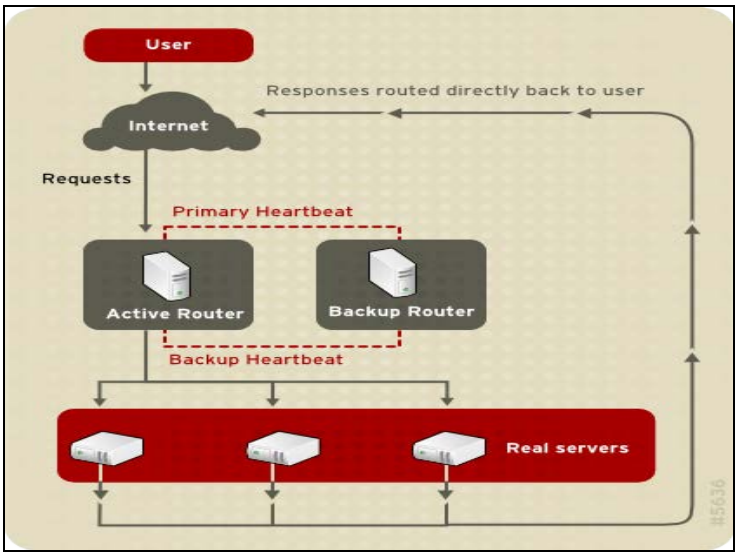
วิธีการตอบกลับแบบ NAT(Network Address Translation) จะมีรูปแบบการตอบกลับโดยที่เมื่อเครื่องเซิร์ฟเวอร์ Load Balance (router) ได้รับคำร้อง (Request) มาจากเครื่องที่ขอรับบริการ (Client) แล้วตัวเครื่อง Load Balance จะทำการส่งต่อคำร้องนั้นไปยังเซิร์ฟเวอร์ในกลุ่มคลัสเตอร์ (Real Server) เพื่อทำการประมวล Request ที่ได้รับมา เมื่อเครื่อง Real Server ประมวลผลเสร็จ จะส่งข้อมูลตอบกลับไปยัง Load Balance เพื่อให้ Load Balance ส่งต่อข้อมูลกลับไปยังเครื่องที่ร้องขอข้อมูล (Client) ก็จะจบกระบวนการตอบกลับแบบ NAT ซึ่งวิธีการตอบกลับนี้จะพบได้ใน Hardware Load Balance ส่วนใหญ่ในท้องตลาดนั้นเอง




**รูปที่ 2.4** กระบวนการตอบกลับแบบ NAT [3]

### 2.3.3 วิธีการตอบกลับแบบ Direct Routing

วิธีการตอบกลับแบบ Direct Routing จะเป็นการตอบกลับข้อมูลจากเครื่องในกลุ่มคลัสเตอร์ (Real Server) กลับไปยังเครื่องที่ขอรับบริการ (Client) โดยตรง โดยไม่ต้องผ่านตัว Load Balance โดยเมื่อเครื่อง Load Balance ได้รับคำร้องจากเครื่องที่ขอรับบริการ (Client) เครื่อง Load Balance จะส่งคำร้องต่อไปยังกลุ่มคลัสเตอร์ (real server) และเมื่อกลุ่มเซิร์ฟเวอร์ได้รับคำร้องแล้ว จะตอบกลับไปยังเครื่องที่ขอรับบริการโดยตรง โดยที่ไม่ผ่าน Load Balance ซึ่งวิธีการตอบกลับแบบ Direct Route จะมีประสิทธิภาพในการทำงานมากกว่าการตอบกลับแบบ NAT เนื่องจากระบบจะไม่เป็นคอขวดที่เครื่อง Load Balance ในกรณีที่มี Traffic ของการรับส่งข้อมูลมากๆ




**รูปที่ 2.5** กระบวนการตอบกลับแบบ Direct Routing [3]

คอมพิวเตอร์แต่ละเครื่องในระบบคลัสเตอร์จะถูกเรียกว่า “โหนด (Node)” อาจจะมีโหนดที่ทำหน้าที่ควบคุมการทำงานของโหนดอื่น ๆ ในระบบอีกชั้น เรียกว่า “Front-end Node” ส่วนโหนดอื่นจะทำหน้าที่ประมวลผลเป็นหลัก เรียกว่า “Compute Node” แต่ละโหนดจะสร้างระบบที่เสมือนเป็นเครื่องเดียว โดยใช้วิธีการต่าง ๆ เช่น การใช้งานระบบ Network Information System (NIS) เพื่อให้ผู้ใช้ (User) สามารถใช้งานร่วมกันได้ทุกโหนด ทำให้ผู้ใช้สามารถล็อกอิน (Login) เพื่อใช้งานในโหนดใดๆ ภายใต้ระบบคลัสเตอร์เดียวกัน นอกจากนั้นภายในระบบคลัสเตอร์อาจจะมีการใช้งานซอฟต์แวร์ต่างๆ เพื่อการติดตั้งใช้งาน, การจัดลำดับงานที่ทำในระบบ, การดูแลบริหารระบบ และซอฟต์แวร์เพื่อการประมวลผลแบบขนาน (Parallel Computing) [3]

### 2.3.4 การติดตั้งระบบคลัสเตอร์ (Installation Clustering System)

เมื่อเตรียมอุปกรณ์ที่จะนำมาทำคลัสเตอร์ (เครื่องคอมพิวเตอร์สมรรถนะสูง, ระบบเครือข่ายความเร็วสูง) จากนั้นเตรียมซอฟต์แวร์ในระบบคลัสเตอร์ เช่น ระบบปฏิบัติการ (Linux,Solaris, BSD) ส่วนใหญ่มักจะเลือกใช้ลีนุกซ์ (Linux) โดยลีนุกซ์นั้นมีหลาย Distribution เช่น RedHat, Debian, Turbo Linux, Slackware เป็นต้น ซอฟต์แวร์ที่ใช้ในระบบคลัสเตอร์ควรเป็นโปรแกรมแบบขนาน โปรแกรมแบบขนานบนระบบคลัสเตอร์นั้นจะใช้วิธีการที่เรียกว่าการโปรแกรมแบบส่งผ่านข้อความ (Message Passing) การโปรแกรมในลักษณะนี้ทำได้โดยการกระจายงานขนาดใหญ่ไปยังหลายๆ เครื่องให้ทำงานพร้อมกัน และใช้การแลกเปลี่ยนข่าวสารผ่านเครือข่ายในการติดต่อระหว่างกลุ่มของโปรแกรมที่ช่วยกันทำงาน ระบบโปรแกรมแบบขนานที่ใช้งานเป็นมาตรฐานมีอยู่สองระบบคือ ระบบ PVM เป็นระบบที่มีมาก่อน โดยเป็นงานของ Oak Rige National Laboratory และ University of Tennessee at Knoxville และในราวปี ค.ศ. 1994 ได้มีมาตรฐานใหม่เกิดขึ้น คือ MPI ซึ่งเป็นที่ยอมรับกันอย่างกว้างขวาง และจะมาแทนที่ PVM ด้วยโปรแกรม Utility และ Library ต่าง ๆ โปรแกรมเหล่านี้บ้างก็ช่วยให้บริหารระบบได้ดีขึ้น เช่น Library Math บางตัวที่ทำงานแบบขนานได้ เช่น Scalapack , PetSc เป็นต้น หรือ โปรแกรมสำหรับ Graphic Rendering โปรแกรมนี้มีทั้งในระบบลีนุกซ์ และวินโดว์ ซึ่งสามารถทำงานแบบขนานได้โดยใช้เครื่องคอมพิวเตอร์จำนวนมากช่วยกันเรนเดอร์ (Render) [4]

### 2.3.5 การบริหารเครื่องแม่ข่ายโดยใช้ระบบปฏิบัติการ Linux

#### 2.3.5.1 ประวัติความเป็นมาของ Linux

Linux เป็นระบบปฏิบัติการที่คล้าย UNIX แต่ไม่ใช่ UNIX ต่างกันทั้งการพัฒนาการแจกจ่าย Linux จะพัฒนาตามมาตรฐานของ POSIX (Portable Operating System Interface) ซึ่งสามารถทำงานได้บนเครื่องหลายตระกูล ทั้งเครื่องที่เป็นตระกูลอินเทล เช่น 386’S/486’s/Pentium และเครื่องที่ไม่ใช่อินเทล

เช่น Digital Alpha, SUN SPARC Macintosh, Silicon Graphics Workstation เป็นต้น ซึ่งระบบปฏิบัติการ UNIX ส่วนใหญ่จะยึดติดอยู่กับตระกูลใดตระกูลหนึ่ง

Linux เป็นระบบปฏิบัติการที่ไม่เสียค่าใช้จ่ายใดๆในการใช้งานซึ่งใช้ได้ฟรี ซึ่งหมายถึงการมีอิสระในการนำโค้ดต้นฉบับมาทำการแก้ไขปรับปรุงตามความต้องการของผู้พัฒนา และยังสามารถทำสำเนาแจกจ่ายได้ ไม่ว่าจะเพื่อการศึกษา หรือในเชิงพาณิชย์ Linux ได้ถูกพัฒนาขึ้นภายใต้สิทธิ์แบบ GPL (General Public License) ซึ่งเป็นลิขสิทธิ์ที่ให้มีการเปลี่ยนแปลงต้นฉบับ หรือแจกจ่ายโค้ดไม่จำกัดสิทธิ์ แต่ซอฟต์แวร์จะต้องเป็นลิขสิทธิ์แบบ GPL

Linux ถูกพัฒนาโดย Linus Torvalds ซึ่งขณะนั้นยังศึกษาอยู่มหาวิทยาลัย Helsinki ประเทศฟินแลนด์ ในขณะนั้น Linus สนใจระบบปฏิบัติการ Minix จึงเกิดความคิดที่จะพัฒนาระบบปฏิบัติการที่มีความสามารถมากขึ้นตามแบบของ Minix และสามารถทำงานได้บนเครื่องตระกูล PC-Intel ได้ด้วย Linus เริ่มพัฒนาในปี 1991 และได้ออก Kernal เวอร์ชั่น 0.02 มาจนกระทั่งปี 1994 จึงสามารถออก Kernal เวอร์ชั่น 1.0 ได้ จวบจนปัจจุบัน Kernal ได้พัฒนาจนถึงเวอร์ชั่น 2.5.x และยังพัฒนาต่อมา [5]

**2.3.5.2 คุณสมบัติพื้นฐานของ Linux**

1. ความสามารถทางด้านการทำงานหลายๆโปรแกรมในเวลาเดียวกัน (Multitasking) โดยที่โปรแกรมต่างๆมีอิสระในการทำงานต่อกัน
2. Linux ยอมให้ผู้ใช้สามารถเข้าทำงานพร้อมๆกันได้โดยผู้ใช้แต่ละคนทำงานได้เป็นอิสระจากกัน
3. ทำงานได้ทั้ง Text Mode และ Window Mode
4. ทำงานได้ทั้งแบบหน้าเครื่อง (Console) หรือจากเครื่องอื่น (Remote)

**2.3.5.3 คุณสมบัติของเครื่องคอมพิวเตอร์ที่สามารถติดตั้งระบบปฏิบัติการ Linux ได้**

Linux สามารถที่จะติดตั้งใช้งานได้ตั้งแต่เครื่องรุ่นเล็กๆไปจนถึงเครื่องขนาดใหญ่ซึ่งสามารถใช้ Linux ได้ด้วย Memory ขนาด 4 Mb และใช้เนื้อที่บนฮาร์ดดิสก์ขนาดต่ำสุด 15 Mb แต่ยิ่งทรัพยากรยิ่งน้อยเท่าไหร่ ก็ยิ่งเกิดข้อจำกัดในการใช้งาน สำหรับเครื่องที่ Linux สนับสนุนการทำงานมีดังนี้

1. แบบ PC-Base
   - Intel / AMD / Cyrix 386SX / Dx / SL / DXL / SLC
   - Intel / AMD / Cyrix 386SX / Dx / SL / DXL2 / DXL4

- AMD K5, K6, K6-2, K6-3, K7 / Athlon
- Cyrix 6x86 , 6X\x86MX
- Intel Pentium, Pentium Pro, Pentium II, Celeron, Pentium III, Pentium IV
- IDT Winchip C6
- Symmetrical Multiprocessing

2. แบบที่ไม่ใช่ PC-Base
   - Digital Alpha
   - Sun Spare
   - Macintosh Power PC
   - 113M mainframes ( 370 / 390 , As400 )
   - Hewlett-Packard PA-RISC
   - Amiga [5]

## 2.3.6 โครงสร้างของไดเรกทอรี่พื้นฐานบนระบบปฏิบัติการ Linux

### 2.3.6.1 ระบบไฟล์รูท (/)

รูท เป็นระบบไฟล์แรกสุดในระบบไฟล์รูทนี้จะมีไดเรกทอรี่ย่อยต่างๆ ทั้งที่เอาไว้เก็บไฟล์ที่ใช้ในการบู๊ตระบบ หรือเก็บคำสั่งต่างๆของระบบไฟล์รูทนี้ยังเป็นระบบไฟล์เดียวที่สามารถใช้งานได้ เมื่อระบบบู๊ตขึ้นมาแบบใช้งานคนเดียว (Single User Mode) สำหรับรายละเอียดไดเรกทอรี่ย่อยต่างๆ ที่อยู่

### 2.3.6.2 ระบบไฟล์ยูเอสอาร์ (/usr)

ระบบไฟล์ /usr จะเก็บคำสั่งของระบบไว้ด้วยเหมือนกัน รวมทั้งยังเก็บไลบรารีและคู่มือช่วยเหลือ (Man Page) ไว้ด้วยซึ่งไดเรกทอรี่ย่อยต่างๆ

### 2.3.6.3 ระบบไฟล์วาร์ (/Var)

ระบบไฟล์ /var เป็นระบบไฟล์ที่เก็บไฟล์ต่างๆ ที่มักมีการเปลี่ยนแปลงของขนาดไปตามเวลา เช่น ที่เก็บข้อมูลการส่งพิมพ์ของอุปกรณ์ ปริ้นเตอร์ ไฟล์ที่บันทึกการทำงาน หรือ ข้อผิดพลาดของระบบ (Any Log File) เป็นต้น ไดเรกทอรี่และไฟล์ต่างๆ

**2.3.6.4 ระบบไฟล์ โฮม (/home)**

เป็นระบบไฟล์ที่ใช้เป็นโฮมไดเรกทอรี่ของผู้ใช้แต่ละคน แต่ก็ไม่จำเป็นที่จะต้องอยู่ที่ระบบไฟล์นี้ก็ได้ ผู้ดูแลสามารถกำหนดให้ไปอยู่ที่ไดเรกทอรี่ไหนๆก็ได้ โดยจะระบุไว้ไฟล์ /etc/pass แต่ /home นี้ ไดเรกทอรี่ที่ระบบกำหนดไว้อยู่แล้ว ในกรณีที่เราใช้คำสั่งในการสร้างบัญชีผู้ใช้คนใหม่แต่ไม่ได้ระบุโฮมให้ ระบบจะทำการกำหนดโฮมของผู้ใช้คนนั้นไว้ที่ /home เอง

**2.3.6.5 คำสั่งพื้นฐานบนระบบปฏิบัติการ Linux**

นอกจากเราจะใช้งานแบบ GUI ได้แล้ว เราจำเป็นต้องใช้งานแบบคำสั่งได้ด้วย ซึ่งจะมีประโยชน์มากในการทำงานที่ซับซ้อนขึ้น สำหรับคำสั่งพื้นฐานมีดังนี้

- ls เรียกดูรายการข้อมูลของไฟล์และไดเรกทอรี่
- pwd แสดงไดเรกทอรี่ปัจจุบันที่ใช้งานอยู่
- cd เปลี่ยนไดเรกทอรี่ปัจจุบันให้ไปยังไดเรกทอรี่ต่างๆ
- mkdir สร้างไดเรกทอรี่
- rm ลบไฟล์หรือไดเรกทอรี่
- chmod เปลี่ยนสิทธิ์ในการใช้งานไฟล์หรือไดเรกทอรี่
- chown เปลี่ยนการเป็นเจ้าของไฟล์หรือไดเรกทอรี่
- cat , more , head , tail ดูรายละเอียดภายในไฟล์
- file ดูชนิดของไฟล์ต่างๆ
- man เป็นคู่มือช่วยในการดูรายละเอียดคำสั่งต่างๆ
- ps เรียกดูรายการโปรเซสที่กำลังทำงานอยู่

**2.3.6.6 การแก้ไขด้วย vi**

ในการใช้งานระบบปฏิบัติการใดก็ตาม เราจำเป็นต้องมีการแก้ไขไฟล์แน่นอน ซึ่งในระบบปฏิบัติการลีนุกซ์เองก็มีโปรแกรมแก้ไขไฟล์หลายระบบทั้งบน GNOME KDE วินโดว์ที่มีรูปแบบคล้ายกับกราฟฟิกซึ่งง่ายต่อการใช้งาน โปรแกรมแก้ไขไฟล์ที่ทำงานบนโหมดคำสั่ง ซึ่งจะเป็นประโยชน์มากถ้าเครื่องเราเกิดไม่สามารถใช้งานวินโดว์ได้ โปรแกรม vi นี้สามารถใช้งานบนระบบปฏิบัติการยูนิกซ์อื่นๆได้ เช่นกัน

การทำงานบนโปรแกรม vi สามารถแยกเป็น 3 โหมด ด้วยกันคือ

1. โหมดเพิ่มเติมข้อความ ( Insert Mode ) เป็นโหมดที่เราใช้ในการพิมพ์ข้อความนั้นเอง

2. โหมดคำสั่ง (Command Model) โหมดนี้เราใช้ในการแก้ไขข้อความ เช่น ลบข้อความทั้งบรรทัดแก้ประโยค เป็นต้น

3. โหมดคำสั่งเพิ่มเติม (Extended Command Mode) โหมดนี้ก็คือโหมดคำสั่งนั้นเอง แต่จะเป็นคำสั่งที่เพิ่มเติมขึ้นมา ซึ่งก่อนที่จะใช้คำสั่งเพิ่มเติมนั้นต้องกดแป้น ":" ก่อนแล้วจะมีเครื่องหมาย ":" ขึ้นที่มุมล่างซ้าย หลังจากนั้นเราจึงใส่คำสั่งในโหมดนี้ได้ โหมดนี้ก็ใช้สำหรับแก้ไขข้อความเหมือนกัน แต่จะมีความสามารถสูงกว่าในโหมดคำสั่ง เช่น สามารถเปลี่ยนคำเก่าทั้งไฟล์ให้เป็นคำใหม่ได้ซึ่งถ้าเป็นโหมดคำสั่งจะต้องทำทีละตัว เป็นต้น

การเข้าสู่โปรแกรม vi เราทำได้โดยการเรียกคำสั่ง vi ขึ้นมาได้เลย ซึ่งเราแบ่งออกเป็น 2 แบบ คือ

1. การสร้างไฟล์ใหม่ ทำได้โดยเรียกคำสั่ง vi ขึ้นมาเฉยๆ หรือตามด้วยชื่อไฟล์ใหม่ที่เราต้องการ

2. การแก้ไขไฟล์เดิม ทำได้โดยเรียกคำสั่ง vi ตามด้วยชื่อไฟล์ที่ต้องการแก้ไขการเข้าสู่โหมดต่างๆ

โหมดเพิ่มเติมข้อความ (Insert Mode) หลังจากทำการสร้างไฟล์แล้ว เราจะอยู่ในโหมดของคำสั่งก่อน ซึ่งเราสามารถเข้าสู่โหมดเพิ่มเติมข้อความได้ด้วยคีย์ต่างๆดังนี้

| | |
|---|---|
| i | เพิ่มข้อความใต้ ณ ตำแหน่งเคอร์เซอร์ |
| a | เพิ่มข้อความหลังตำแหน่งเคอร์เซอร์ |
| A | เพิ่มข้อความท้ายบรรทัด ณ บรรทัดที่เคอร์เซอร์อยู่ |
| o | เพิ่มบรรทัดใหม่ใต้บรรทัดที่เคอร์เซอร์อยู่ |
| O | เพิ่มบรรทัดใหม่เหนือบรรทัดที่เคอร์เซอร์อยู่ |

ถ้าเราเข้าสู่โหมดเพิ่มเติมข้อความแล้ว จะสังเกตเห็นว่าตรงมุมล่างซ้ายจะมีข้อความ * -INSERT -* ปรากฏอยู่

โหมดคำสั่ง (Command Mode) การเข้าสู่โหมดคำสั่งทำได้โดยการกดแป้น <Ese> ซึ่งเมื่อเข้าสู่โหมดนี้แล้วคีย์ทุกตัวที่เรากดจะนำมาเทียบดูว่าเป็นคำสั่งหรือเปล่าซึ่งไม่มีข้อความใดๆ ปรากฏขึ้นเมื่อเราเปลี่ยนมาอยู่ในโหมดคำสั่งแล้วข้อความ *- INSERT -* ก็จะหายไปนั้นคือเราเปลี่ยนโหมดแล้ว

โหมดคำสั่งเพิ่มเติม (Extended Command Mode) โหมดนี้จะต่างจากโหมดคำสั่ง (Command Mode) ตรงตัวคำสั่งเอง เพราะคำสั่งที่ใช้ในโหมดนี้จะขึ้นต้นด้วย “ : “ ก่อน ดังตัวอย่าง

- set nu แสดงหมายเลขบรรทัด
- set nanu ยกเลิกการแสดงหมายเลขบรรทัด
- set list แสดงให้เห็นสัญลักษณ์ต่างๆ เช่น TAB , Enter เป็นต้น
- set nolist ยกเลิกการแสดงสัญลักษณ์ต่างๆ

**2.3.6.7 คีย์ที่ใช้ในการแก้ไขข้อความบน VI**

การแก้ไขข้อความมีทั้งลบคำ ลบทั้งประโยค แก้คำผิด แก้ประโยคผิด เป็นต้น ซึ่งคีย์ที่ยกตัวอย่างจะเป็นคีย์ที่มีการใช้บ่อยในการใช้ vi แก้ไขคำ

| | |
|---|---|
| r | แก้ 1 ตัวอักษร ณ ตำแหน่งเคอร์เซอร์ |
| cS | แก้คำตั้งแต่เคอร์เซอร์ไปจนสุดบรรทัด |
| cW | แก้ 1 คำที่ถูกคั่นด้วย space |
| ce | แก้คำทั้งบรรทัด |
| dd | ลบข้อความทั้งบรรทัด |
| ndd | ลบข้อความ n บรรทัดตั้งแต่ตำแหน่งเคอร์เซอร์ |
| x | ลบ 1 ตัวอักษร ณ ตำแหน่งเคอร์เซอร์ |
| rx | ลบข้อความ n ตัวอักษร ณ ตำแหน่งเคอร์เซอร์ |
| D | ลบข้อความตั้งแต่เคอร์เซอร์ไปจนถึงท้ายบรรทัด |

**2.3.6.8 คำสั่งที่ใช้ในการค้นหาคำบน vi**

การค้นหาก็เป็นสิ่งที่จำเป็นต้องรู้เพราะช่วยให้ง่ายต่อการอ่านข้อความที่ต้องการและสามารถลบข้อความที่ต้องการออกไปได้เร็วขึ้นด้วย เช่น

| | |
|---|---|
| /word | ค้นหาคำ word จากตำแหน่งเคอร์เซอร์ไปท้ายไฟล์ 1 คำ |
| ?word | ค้นหาคำ word จากตำแหน่งเคอร์เซอร์ไปต้นไฟล์ 1 คำ |
| n | หลังจากค้นเจอคำแรกแล้วต้องการให้ค้นหาต่อไปอีก 1 คำ |

**2.3.6.9 คำสั่งที่ใช้สำหรับบันทึกและออกจาก vi**

หลังจากทำการแก้ไขหรือสร้างไฟล์ใหม่เสร็จแล้วเราก็ต้องรู้จักการออกจาก vi ทั้งแบบที่บันทึกการแก้ไข หรือแบบที่ไม่บันทึกด้วยเช่น

| | |
|---|---|
| ZZ | บันทึกสิ่งที่แก้ไขแล้วออกจาก vi (Command Mode) |
| :wq | บันทึกสิ่งที่แก้ไขแล้วออกจาก vi (Extended Command Mode) |
| :w | บันทึกสิ่งที่แก้ไขโดยยังไม่ออกจาก vi |
| :q! | ออกจาก vi โดยไม่ต้องบันทึกการแก้ไข |

## 2.3.7 การให้บริการของ Linux ในฐานะเครื่องแม่ข่ายประกอบด้วย

### 2.3.7.1 ระบบสืบค้นชื่อโดเมน (DNS)

DNS (Domain Name System) ระบบนี้เป็นระบบที่จำเป็นในการติดต่อสื่อสารไปยังอินเทอร์เน็ต ซึ่งโดยทั่วไปเครื่องคอมพิวเตอร์ที่ต่ออยู่บนอินเทอร์เน็ตที่ต่ออยู่บนเน็ตเวิร์คไม่ว่าจะเป็นแบบ LAN MAN WAN หรือผ่านดาวเทียมก็ตาม จะมีการกำหนดหมายเลขประจำเครื่อง (IP) ไว้ติดต่อกันผ่าน TCP/IP ( Transport Control Protocal / Internet Protocal) ซึ่งการตั้งค่า IP ให้กับเครื่องต่างๆนั้นจะทำให้สามารถติดต่อไปยังเครื่องนั้นๆ ได้ก็สามารถที่จะเข้าใช้บริการของเครื่องต่างๆ ที่เปิดให้บริการ เช่น สามารถเข้าไปดูเว็บไซต์โดยใช้ค่า IP แต่การจดจำ IP นั้นเป็นเรื่องที่ยากต่อการจดจำจึงได้มีการใช้ระบบ DNS เพื่อค้นหา IP ต่างๆทั่วโลกซึ่งเราไม่จำเป็นต้องกำหนดในไฟล์ hosts เลยดังนั้นในกรณีของเว็บเซิร์ฟเวอร์ต่างๆ แม้เราจะไม่รู้ IP ของเครื่องนั้นๆ แต่เมื่อใส่เว็บไซต์แล้วระบบ DNS จะทำการค้นหาเองว่าเว็บไซต์นั้นใช้ IP อะไรจึงทำให้เราติดต่อถึงกันได้

ลักษณะการทำงานของระบบสืบค้นชื่อโดเมนเมื่อทราบว่าในแต่ละโดเมนเนมมีเครื่องที่เป็นเซิร์ฟเวอร์ไว้เป็นแหล่งเก็บข้อมูลของโดเมนและสมาชิก ดังนั้นเมื่อต้องการค้นหา IP เครื่องได้ก็ตามที่มีการลงทะเบียนกับโดเมนไม่ว่าจะเป็นโดเมนได้ถ้าเราไปถามจากเครื่องเนมเซิร์ฟเวอร์ของโดเมนนั้นจะได้ IP เครื่องนั้นมาซึ่งการค้นหาแบ่งเป็น 2 ลักษณะ

1. การค้นหาภายในโดเมนเดียวกันในกรณีที่องค์กรขนาดใหญ่มีเครื่องเซิร์ฟเวอร์และเครื่องลูกข่ายจำนวนมาก เมื่อเครื่องใดเครื่องหนึ่งในเน็ตเวิร์คต้องการติดต่อไปยังเครื่องอื่นแต่ไม่รู้จัก IP รู้เพียงชื่อเครื่อง ซึ่งขั้นตอนการค้นหาจะเริ่มจากไฟล์ hosts ว่ามีเครื่องที่ต้องการอยู่หรือไม่ ถ้าไม่มีจะติดต่อไปยังเนมเซิร์ฟเวอร์แล้วถามดูว่ามีเครื่องที่ต้องการอยู่หรือไม่ ถ้ามีการลงทะเบียนไว้เนมเซิร์ฟเวอร์จะส่ง IP ของเครื่องนั้นกลับไป ซึ่งจะทำให้สามารถติดต่อไปยังเครื่องที่ต้องการได้ ดังนั้นในกรณีของเว็บเซิร์ฟเวอร์ต่างๆ แม้เราจะไม่รู้ IP ของเครื่องนั้นๆ แต่เมื่อใส่เว็บไซต์แล้วระบบ DNS จะทำการค้นหาเองว่าเว็บไซต์นั้นใช้ IP อะไรจึงทำให้เราติดต่อถึงกันได้

2. การค้นหาเครื่องที่อยู่ต่างโดเมนในกรณีที่ติดต่อไปยังเครื่องที่อยู่ต่างโดเมนกันซึ่งมีการกำหนดไว้ในไฟล์ hosts ด้วยแล้วขั้นตอนการค้นหาจะไปถามยังเนมเซิร์ฟเวอร์ของโดเมนเราก่อนถ้าเนมเซิร์ฟเวอร์เราไม่รู้จัก ก็จะไปถามยังเนมเซิร์ฟเวอร์ภายนอกจนกว่าจะได้ IP เครื่องที่ต้องการมา

**2.3.7.2 ระบบเว็บเซิร์ฟเวอร์ (Web Server)**

เมื่ออินเทอร์เน็ตเข้ามามีบทบาทในชีวิตประจำวันมากขึ้นไม่ว่าจะเป็นการให้บริการข้อมูลข่าวสารบันเทิง แม้แต่การใช้ในเชิงธุรกิจทั้งการขายสินค้าผ่านอินเทอร์เน็ต การเปิดประมูลและบริการอื่นๆ แต่สิ่งต่างๆเหล่านี้จะเกิดขึ้นไม่ได้เลยถ้าขาดเว็บเซิร์ฟเวอร์ที่เป็นเหมือนองค์กรบนอินเทอร์เน็ตคอยทำหน้าที่ เชื่อมต่อเว็บเพ็จให้คนทั่วโลกสามารถเข้ามาอ่าน หรือทำกิจกรรมบนอินเทอร์เน็ตได้ Apache เว็บเซิร์ฟเวอร์เริ่มต้นจากการนำซอฟต์แวร์ NCSA HTTP Server มาพัฒนาโดยกลุ่มผู้พัฒนาได้รวมตัวกันในโครงการ “Apache Software Foundation” จุดมุ่งหมายเพื่อจัดทำเว็บเซิร์ฟเวอร์ที่เป็นมาตรฐานขึ้น โดยไม่ปกปิดต้นฉบับ

**2.3.7.3 ระบบการขนส่งไฟล์ผ่านเน็ตเวิร์ค (FTP Server)**

FTP (File Transfer Protocal) เป็นการให้บริการกับไฟล์ผ่านทางเน็ตเวิร์ค ซึ่ง FTP จะแชร์เพื่อให้เครื่องต่างๆ เข้ามาคัดลอกไฟล์ไปกลับ ถ้ามีการแชร์ไฟล์ที่ Execute ไฟล์สามารถที่จะเรียกใช้ไฟล์นั้นให้ทำงานผ่านเน็ตเวิร์คได้เลยที่สำคัญคือผู้ที่จะมาก๊อบปี้ไฟล์จาก FTP Server นั้นไม่จำเป็นจะต้องมีบัญชี (Account Name) อยู่บนเซิร์ฟเวอร์ก็ได้เพราะ FTP สามารถกำหนดให้ ล็อกอินโดย Anonymous หรือ FTP ได้ ซึ่งเปรียบเสมือน User guest นั้นเอง โดยทั่วไปนอกจากจะนำมาใช้ภายในเน็ตเวิร์คแบบ LAN แล้วยังสามารถให้บริการผ่านอินเทอร์เน็ตได้ ซึ่งจะให้บริการผ่านทางโปรโตคอล ftp โดยระบุเว็บไซต์ขึ้นต้นด้วย “ ftp:// ” แทน “ http:/

**2.3.7.4 ระบบการ Login ไปยังเครื่องอื่น (Telnet)**

Telnet เป็นคำสั่งในการ Login เข้าไปใช้งานเครื่องอื่นที่มีการเชื่อมต่อกับเครื่องที่ใช้งานอยู่ซึ่งการ Login เข้าไปใช้งานอาจจะ Login ด้วยชื่อเครื่องที่เรียกว่า Hosts Name หรืออาจจะ Login ด้วย IP ของเครื่องนั้นๆ โดยตรงซึ่งรูปแบบการใช้งานและคำสั่งต่างๆ เกี่ยวกับการใช้ Telnet นั้น

**2.3.7.5 ระบบแจกจ่าย IP Address อัตโนมัติ (DHCP Server)**

ในเครือข่ายที่มีคอมพิวเตอร์มากๆ เช่นห้องอบรมคอมพิวเตอร์ ห้องปฏิบัติการตามสถานศึกษาต่างๆ ผู้ที่ทำหน้าที่ดูแลระบบอาจจะพบความยุ่งยากในการตั้งค่า IP Address และค่าอื่นๆทางระบบเครือข่าย เช่น Gateway, Netmask, DNS ให้กับแต่ละเครื่องและจะต้องคอยระวังไม่ให้ค่า IP เหล่านั้นเกิดการซ้ำ

กันซึ่งปัญหาเหล่านี้สามารถแก้ได้ โดยใช้ระบบ DHCP Server แจกจ่าย IP Address โดยอัตโนมัติ โดย DHCP จะมีระบบควบคุม IP อัตโนมัติไม่ให้เกิดการซ้ำกัน นอกจากนี้แล้วยังเพิ่มความสะดวกสบายไม่ต้องติดตั้งค่าในเครื่องพีซีให้ยุ่งยาก

DHCP (Dynamic Host Configuration Protocol) เป็นโปรโตคอลที่เกี่ยวข้องกับการตั้งค่าต่างๆ ของเครือข่ายบนเครื่องคอมพิวเตอร์ การใช้งานจะแบ่งออกเป็น 2 ส่วนคือ DHCP Server และ DHCP Client โดย DHCP Server จะเป็นตัวกำหนดค่าต่างๆให้กับ DHCP Client ในการกำหนดค่าต่างๆให้กับ Client เช่น IP Address ตัว Server สามารถกำหนดระยะเวลาให้กับเครื่อง Client ว่าสามารถใช้ได้เป็นเวลานานเท่าไหร่ซึ่งอาจจะ ใช้ได้แค่ 10 นาที แล้วขอใหม่หรือใช้ได้ตลอดไปตามปกติแล้วในเครือข่ายหนึ่งๆควรมี DHCP Server ตัวเดียว แต่สามารถมีเครื่อง Client หลายๆ เครื่องได้ซึ่ง DHCP Server จะต้องเปิดทำงานตลอดเวลาเพื่อที่ Client ตัวใดๆ สามารถที่จะใช้บริการได้ตลอดเวลา

### 2.3.8 ข้อแตกต่างของระบบ Cluster Computing เมื่อเปรียบเทียบกับระบบอื่น

การเชื่อมต่ออุปกรณ์คอมพิวเตอร์เข้าด้วยกันหลายๆเครื่องนั้นจะทำให้เกิดเป็นระบบคอมพิวเตอร์ระบบหนึ่ง และในแต่ละระบบมีโครงสร้างการทำงานที่แตกต่างกันออกไปขึ้นอยู่กับผู้ที่ออกแบบและนำไปใช้ว่าจะนำไปใช้ให้เหมาะสมกับองค์กร หน่วยงานอย่างไร ระบบ Cluster Computing เป็นระบบที่เชื่อมต่ออุปกรณ์คอมพิวเตอร์เข้าด้วยกันเพื่อให้การบริการของเครื่องแม่ข่ายนั้นมีความต่อเนื่อง รวดเร็วทั้งยังรองรับผู้ใช้งานจำนวนมากได้

#### 2.3.8.1 ระบบ Cluster Computing กับ ระบบ Lan (Local Area Networking)

ระบบ Cluster Computing มีส่วนสำคัญ 3 อย่างคือ เครือข่ายความเร็วสูง ระบบซอฟต์แวร์ที่สนับสนุนระบบคลัสเตอร์ และโปรแกรมประยุกต์ที่ใช้ขีดความสามารถของการประมวลผลแบบขนานหรือแบบกระจาย ส่วนระบบ LAN เครื่องทุกเครื่องที่อยู่บนระบบ LAN เป็นอิสระต่อกันไม่มีระบบซอฟต์แวร์ที่นำความสามารถของการประมวลผลแบบขนานและแบบกระจายมาใช้

#### 2.3.8.2 ระบบ Cluster Computing กับ ระบบ Grid (Grid Computing)

ระบบ Cluster Computing เป็นการเชื่อมต่อเพื่อเพิ่มสมรรถนะของการประมวลผลด้วยเครื่องคอมพิวเตอร์ที่มีแพลตฟอร์ม (Platform) เดียวกันอยู่ในพื้นที่จำกัด ส่วน Grid Computing นั้นจะเชื่อมต่อได้ทุกแพลตฟอร์ม ไม่ว่าแต่ละแพลตฟอร์มจะห่างไกลกันเท่าไร ระบบ Cluster Computing กับ ระบบ Load balancing ระบบ Cluster Computing มีการจัดกลุ่มของคอมพิวเตอร์หลายตัวเพื่อให้สามารถทำงานได้เหมือนกับเป็นคอมพิวเตอร์ตัวเดียวกัน ดังนั้นไม่ว่า ผู้ใช้เข้ามาใช้งานเครื่องใด

ภายในกลุ่มก็จะรู้สึกเหมือนใช้งานเครื่องเดียวกัน คุณสมบัติของการทำ Clustering คือ การทำรีพลิเคท (Replication) โดยในด้านของ Web Application คือการทำ Session Replication ซึ่งตามปกติแล้ว Session ของผู้ใช้เก็บใน Web Server เครื่องที่ผู้ใช้ใช้งานอยู่เท่านั้นแต่การทำ Clustering จะเป็นการคัดลอก Replicate Session นั้นไปยัง Web Server อื่นภายในกลุ่มด้วย ทำให้ไม่ว่าผู้ใช้จะเข้าไปใช้งานใน Server เครื่องใดก็จะมี Session ของผู้ใช้อยู่ด้วยเสมอ ส่วน Load Balancing คือการจัดกลุ่มของคอมพิวเตอร์หลายตัวเพื่อแบ่งงานกัน หรือกระจาย Load การใช้งานของผู้ใช้ไปยังคอมพิวเตอร์ภายในกลุ่ม เพื่อให้สามารถรับจำนวนผู้ใช้ที่เข้ามาใช้งานได้มากขึ้น หรือสามารถรับงานที่เข้ามาได้มากขึ้น นอกจากนั้นยังมีคุณสมบัติของ Fail Over คือหากมีคอมพิวเตอร์ใดภายในกลุ่มมีปัญหาไม่สามารถทำงานได้ ตัว Load Balancer ที่เป็นตัวแจก Load ให้คอมพิวเตอร์ภายในกลุ่มก็จะส่ง Load ไปยังเครื่องอื่นแทน จนกว่าเครื่องนั้นจะกลับมาใช้งานได้ดังเดิม การทำ Cluster ไม่จำเป็นต้องพึ่ง Feature ของ Server เป็นหลัก แต่สามารถ Develop ตัว Application ให้เป็น Cluster ได้โดยไม่ต้องพึ่ง Feature ของ Server เช่น การใช้หลักการของ File Sharing หรือ Database สามารถทำงานได้เหมือนกัน เช่นเดียวกับการทำ Load Balance ไม่ต้องหา Hardware หรือ Software พิเศษที่จะทำหน้าที่เป็น Load Balancer แต่เขียน Application เพื่อทำการกระจาย Traffic ไปยัง Server ได้เหมือนกัน โดยใช้หลักการของ Redirection เป็นต้น การประมวลผลแบบขนาน (Parallel Processing) คือ การแบ่งงานออกเป็นชิ้นเล็กให้แต่ละงานประมวลผลหลาย ๆ ตัวในเวลาพร้อมกัน ประโยชน์ของการใช้วิธีการนี้ คือ แก้ปัญหาขนาดใหญ่ได้ ในเวลาที่เร็วขึ้น ลดค่าใช้จ่าย ซึ่งสามารถใช้เครื่องพีซี โดยเชื่อมต่อกันเป็น ระบบคลัสเตอร์ แทนการใช้เครื่องเมนเฟรม หรือ ซูเปอร์คอมพิวเตอร์ [1]

## 2.3.9 ซอฟต์แวร์และเทคโนโลยีที่เกี่ยวข้อง

### 2.3.9.1 ระบบคอมพิวเตอร์แบบขนาน (Parallel Computer)

นับเป็นเวลาหลาย 10 ปี ที่นักวิจัยและนักวิทยาศาสตร์ต่างประสบปัญหาในการประมวลผลข้อมูลที่มีขนาดใหญ่และมีจำนวนมากส่งผลให้ต้องใช้เวลาที่มากกว่าจะได้ผลลัพธ์เพื่อนำมาใช้ประโยชน์ อีกทั้งเครื่องที่มีสรรถนะสูงในอดีตนั้นมีราคาแพงมาก ดังนั้นกลุ่มนักวิจัย และนักวิทยาศาสตร์จึงได้มีการร่วมกันคิดค้นวิธีการพัฒนาสมรรถนะของเครื่องคอมพิวเตอร์ให้มีประสิทธิภาพสูงสามารถประมวลผลได้รวดเร็วโดยใช้หลักการนำหน่วยประมวลผลหลายๆหน่วยมาทำงานร่วมกันเพื่อให้สามารถประมวลผลข้อมูลที่มีจำนวนมากให้เสร็จลงได้ภายในเวลาที่รวดเร็ว รูปแบบระบบนี้เรียกว่าระบบคอมพิวเตอร์แบบขนาน (Parallel Computer) การออกแบบระบบคอมพิวเตอร์แบบขนานจะมีการนำไปประยุกต์ใช้ 2 ด้านใหญ่ๆคือ

1. ระบบคอมพิวเตอร์แบบขนานที่ถูกออกแบบมาใช้สำหรับงานเฉพาะด้าน

2. ระบบคอมพิวเตอร์แบบขนานที่นำประยุกต์ใช้สำหรับงานทั่วไปสำหรับการออกแบบระบบ

คอมพิวเตอร์แบบขนานโดยทั่วไปนั้นสามารถแบ่งได้เป็น 2 แบบ ได้แก่

1. แบบที่ใช้หน่วยความจำร่วม (Shared Memory Multiprocessors)
2. แบบที่ใช้หน่วยความจำแยก (Distributed Memory Multicomputer)

ระบบคอมพิวเตอร์แบบขนานที่ใช้หน่วยความจำร่วม ระบบคอมพิวเตอร์รูปแบบนี้ถูกออกแบบให้ประกอบด้วยหน่วยประมวลผลจำนวนมาก ซึ่งถูกเชื่อมต่อด้วยเครือข่ายภายในความเร็วสูง (High Speed Interconnection Network) โดยหน่วยประมวลผลทั้งหมดใช้หน่วยความจำกลางร่วมกัน ตัวอย่างของระบบคอมพิวเตอร์รูปแบบนี้ได้แก่ เครื่องคอมพิวเตอร์ส่วนบุคคล (Personal Computer)

ข้อดีของระบบคอมพิวเตอร์รูปแบบนี้คือ สามารถพัฒนาโปรแกรมทำงานกับระบบได้ง่ายเนื่องจากใช้หน่วยความจำร่วมกัน อีกทั้งระบบมีความเร็วค่อนข้างสูง ตั้งแต่ระบบขนาดเล็กไปจนระบบขนาดกลางแต่ข้อเสียที่สำคัญคือระบบคอมพิวเตอร์รูปแบบนี้ไม่สามารถขยายให้เป็นระบบคอมพิวเตอร์ขนาดใหญ่ได้ เนื่องจากถ้าหน่วยประมวลผลมีปริมาณการสื่อสารระหว่างกันเป็นจำนวนมากก็จะทำให้เครือข่ายที่เชื่อมต่อกันจะถึงจุดอิ่มตัวอย่างรวดเร็วจึงทำให้ประสิทธิภาพที่ได้ค่อนข้างถูกจำกัดมาก

ระบบคอมพิวเตอร์แบบขนานที่ใช้หน่วยความจำแยก ระบบคอมพิวเตอร์รูปแบบนี้จะประกอบด้วยหน่วยประมวลผลจำนวนมากที่มีการเชื่อมต่อกันผ่านเครือข่ายความเร็วสูง (High Speed Interconnection Network) โดยที่แต่ละหน่วยประมวลผลมีหน่วยความจำเป็นของตนเองและหน่วยประมวลผลดังกล่าวแยกการทำงานจากกันโดยเด็ดขาด การติดต่อสื่อสารกันระหว่างหน่วยประมวลผลในระบบคอมพิวเตอร์แบบขนานนี้จะใช้รูปแบบการสื่อสารกันโดยการส่งข้อความ (Message Passing) ผ่านเครือข่ายในระบบจึงทำให้บางครั้งเรียก ระบบคอมพิวเตอร์แบบขนานรูปแบบนี้ว่า ระบบคอมพิวเตอร์แบบส่งข้อความ (Message Passing Parallel Computer) ตัวอย่างของระบบคอมพิวเตอร์รูปแบบนี้ เช่น เครื่องซูเปอร์คอมพิวเตอร์ของ IBM รุ่น IBMSP และ ระบบคอมพิวเตอร์คลัสเตอร์เป็นต้น

ข้อดีของระบบคอมพิวเตอร์ที่มีโครงสร้างรูปแบบนี้คือสามารถขยายระบบจนเป็นระบบคอมพิวเตอร์แบบขนานขนาดใหญ่มากได้ เนื่องจากปริมาณการสื่อสารข้อมูลผ่านเครือข่ายของโครงสร้างระบบรูปแบบนี้จะน้อยกว่าแบบที่ใช้หน่วยความจำร่วมมาก

**2.3.9.2 Linux Virtual Server Project (LVS)**

The Linux Virtual Server Project พัฒนาเพื่อรองรับการทำงานแบบ High-Performance และ Highly Available ของเครื่องแม่ข่ายสำหรับระบบปฏิบัติการลีนุกซ์ระบบ LVS นั้นจะทำหน้าที่กระจายภาระงานไปยังเครื่องแม่ข่ายหรือโหนด และตรวจสอบการบริการของโหนดในระบบคลัสเตอร์ เพื่อทำการปรับเปลี่ยนการแจกงานในกรณีที่โหนดในระบบคลัสเตอร์ ไม่สามารถให้บริการได้ โดยจะไม่แจกงานไปยังโหนดที่ไม่สามารถให้บริการได้ หลักการทำงานของระบบ LVS จะใช้ advance IP load balancing software (IPVS) ซึ่งก็คือโปรแกรม ipvsadm ทำหน้าที่เป็น Layer 4 Switching ที่รับส่งโปรโตคอลทีซีพี (TCP) และยูดีพี (UDP) และยังสามารถเลือกใช้ algorithm ได้เช่น Round-Robin, Weighted Round-Robin, Least-Connection, Weighted Least-Connection ฯลฯ ในการกระจายภาระงานได้ เราสามารถใช้ระบบ LVS สำหรับงานบริการด้าน web, e-mail, media, และ VoIP

หลักการทำงาน Linux Virtual Server (LVS) Linux Virtual Server อาศัยหลักการเปลี่ยนทิศทางของ TCP หรือ UDP connection จากเครื่อง Load Balancerไปยังเครื่อง Linux เซิร์ฟเวอร์โดยไม่ให้ผู้ใช้รับรู้ ทำให้ผู้ใช้ไม่จำเป็นต้องเปลี่ยนแปลงโปรแกรมหรือเรียนรู้วิธีการใช้งาน Linux Virtual Server นอกจากนั้นการเปลี่ยนทิศทางของTCP/UDP connection กระทำ ในระดับชั้นการทำงานของ IP Protocol ทำให้บริการหรือโปรแกรมที่ทำงานอยู่บนเครื่องLinux เซิร์ฟเวอร์ไม่จำเป็นต้องแก้ไขเพื่อที่จะให้ทำงานได้กับ Linux Virtual Server การเชื่อมต่อ IP Load balance clustering ขึ้นอยู่กับวิธีส่งผ่านข้อมูล (Packet Forwarding Methods) ที่เลือกใช้โดยการติดตั้งบนตัว Load balancer ที่มีหน้าที่กระจายภาระงาน หรือเรียกอีกอย่างว่า Linux Director ซึ่งมี 3 รูปแบบด้วยกันคือ

1. Virtual Server แบบ NAT ใช้กระบวนการเปลี่ยนแปลงแพ็คเกจของ IP โดยเฉพาะในส่วนของต้นทาง (SOURCE) และปลายทาง (DESTINATION) เพื่อจะเปลี่ยนทิศทางของการเชื่อมต่อจากเครื่องหนึ่งไปยังอีกเครื่องหนึ่งการเชื่อมต่อ Virtual Server แบบ NAT สามารถใช้ Real Server ที่เป็นแพลตฟอร์มแบบใดก็ได้ ขอเพียงให้สนับสนุนการเชื่อมต่อแบบ TPC/IP เป็นพอ ส่วน IP ที่จะต้องติดตั้งให้แก่ Real Server นั้น สามารถใช้ได้ทั้งแบบ private และ public IP แต่มีข้อเสีย คือ เมื่อมีการเชื่อมต่อเครื่องคอมพิวเตอร์มากกว่า 20 ตัวขึ้นไป อาจทำ ให้ระบบเกิดคอขวด (Bottleneck) ขึ้นได้ เป็นผลให้ประสิทธิภาพของระบบ Cluster ลดลง เนื่องจากตัว Load Balancer ต้องทำการแก้ไข

Header ของ Packet ที่ส่งผ่านตัวมันตลอดเวลา ทั้งกับ Request และ Response Packet เหมือน NAT Gateway ทั่วไป ที่ต้องมีการเปลี่ยน IP ของต้นทางและปลายทาง

2. Virtual Server แบบ IP Tunneling คือ การห่อหุ้ม (Encapsulated) แพ็คเกจของ IP ด้วยแพ็คเกจ IP อีกครั้งหนึ่ง โดยข้อมูลที่ถูกห่อหุ้มจะถูกส่งไปตามปลายทาง (Destination) อันนอกและอันในตามลำดับ

ข้อดีของวิธีการส่งผ่านข้อมูลแบบ IP Tunneling คือ ช่วยลดภาระงานของ Load balancer เพราะเครื่องเซิร์ฟเวอร์จะทำการตอบสนองกับผู้ใช้โดยไม่ผ่านกระบวนการอื่นใด ด้วยเหตุนี้จึงสามารถมี Cluster ขนาดใหญ่ได้ นอกจากนั้นเครื่องเซิร์ฟเวอร์แต่ละเครื่องสามารถอยู่ต่างเครือข่ายกันได้

3. Virtual Server แบบ Direct Routing ตัว load balancer จะทำการประมวลผลแค่การเชื่อมต่อระหว่างไคลเอนท์ กับตัวมันเองเท่านั้น ส่วนตัว request packet จะเป็นการตอบกลับจาก real server ไปยัง client โดยตรง ผ่านการเชื่อมต่อแบบ TCP/IP ในแบบปกติ แต่การเชื่อมต่อแบบนี้ ไม่มีการทำ tunneling ทำให้ไม่เสียเวลา

#### 2.3.9.3 Network File System (NFS)

เน็ตเวิร์คไฟล์ซิสเต็มส์ (Network File System) พัฒนาขึ้นโดย บริษัท ซัน ไมโครซิสเต็มส์ ปัจจุบันมีการพัฒนาเป็น NFS เวอร์ชั่น 4 (NFSv4) แต่เนื่องจากมีการปรับแต่งระบบที่ซับซ้อน จึงเลือกใช้เป็น NFS เวอร์ชั่น 3 (NFSv3) ซึ่งถูกระบุเป็นมาตรฐานเปิด RFC 1813 ของ Internet Engineering Task Force (Ietf) โดยระบบปฏิบัติการ ลีนุกซ์และ ยูนิกซ์ (Unix) สามารถใช้งานได้อย่างสะดวก รูปแบบการทำงานจะเป็นการให้เครื่องลูกข่าย (Client) เข้าถึงแฟ้มข้อมูลจากระยะไกลด้วย buffer cache ของเครื่องลูกข่ายผ่านทางระบบเครือข่าย TCP/IP และมี block size ขนาด 4069 byte โดยจะทำงานร่วมกับเครื่องแม่ข่ายที่เปิดให้บริการ NFS ซึ่งเปรียบเหมือนกับว่าแฟ้มข้อมูลนั้นอยู่บนเครื่องลูกข่าย และการทำงานของเครื่องแม่ข่ายจะทำงานผ่านทาง RPC (remote Procedure Call) ทำให้เครื่องลูกข่ายสามารถที่จะ mount ไว้ที่ใดก็ได้ใน directory trees ของระบบปฏิบัติการ [6]

#### 2.3.9.4 Heartbeat

วิธีการแก้ปัญหาการหยุดทำงานของเครื่องแม่ข่ายที่ให้บริการจะใช้วิธีการติดตั้งเครื่องแม่ข่ายอีกเครื่องหนึ่งเป็นเครื่องสำรอง (Backup) และทำงานร่วมกันแบบ Active/Standby โดยวิธีการนี้ใช้โปรแกรมมอนิเตอร์ชื่อ heartbeat ที่อยู่บนเครื่องทั้งสองทำหน้าที่ตรวจสอบการทำงานซึ่งกันและกันในช่วงเวลาหนึ่งๆ เมื่อใดก็ตามที่เครื่องแม่ข่ายที่ทำหน้าที่เป็นเครื่องสำรองไม่ได้รับสัญญาณตอบรับจากเครื่องแม่ข่ายที่เป็นเครื่องหลัก (Master) ในช่วงเวลาที่กำหนดไว้ เครื่องแม่ข่ายสำรองจะดึงหมายเลขไอพีบริการ

ที่เครื่องหลักถืออยู่มาเป็นของเครื่องสำรองแทน (Take Over Resource) และจะทำหน้าที่แทนเครื่องหลักต่อไป เมื่อเครื่องหลักกลับทำงานได้เป็นปกติ สามารถกำหนดการทำงานได้เป็นสองวิธีคือ เครื่องหลักที่กลับมาทำงานได้นั้นกลายเป็นเครื่องสำรอง หรือ เครื่องหลักกลับมาให้บริการเช่นเดิม

#### 2.3.9.5 Piranha

เป็นโปรแกรมหนึ่งที่อยู่บนระบบปฏิบัติการ Linux มีความสามารถในการจัดการในฟังก์ชั่นของ Heartbeat, Linux Virtual Server, Load Balancing เป็นต้น มีหน้าต่างติดต่อกับผู้ใช้งาน (GUI) เพื่อใช้งานได้อย่างสะดวกง่ายต่อการจัดการ [3]

## 2.4 งานวิจัยที่เกี่ยวข้อง

ฆอซาลี หีมสุหรี และคณะ [4] Cluster Computing คือระบบคอมพิวเตอร์ที่ประกอบด้วยคอมพิวเตอร์มากกว่า 1 เครื่องต่อเชื่อมกัน และแต่ละเครื่องอาจมีมากกว่า 1 หน่วยประมวลผล (CPU) โดยสามารถจัดสรรให้ใช้กับ CPU, ROM, RAM ร่วมกันได้ ทำให้ได้ระบบที่มีประสิทธิภาพสูงและง่ายต่อการขยาย เพื่อการใช้ทรัพยากรการคำนวณและเข้าถึงข้อมูล ที่อยู่กระจัดกระจายอย่างมีประสิทธิภาพและรวดเร็วทันต่อเหตุการณ์ ปัจจุบันมีการแข่งขันเพื่อนำเทคโนโลยีใหม่เข้าสู่ตลาดรุนแรงขึ้น เพื่อให้สินค้าสามารถขายได้ จึงต้องเพิ่มคุณสมบัติเข้าไปในระบบของตนเพื่อความได้เปรียบ เช่น การใส่ Feature การทำ Load Balance รวมเข้ากับการทำ Clustering เข้าไปในสินค้าของตัวเอง ทำให้เครื่องคอมพิวเตอร์ส่วนบุคคล (Personal Computer : PC) มีความสามารถสูงขึ้นไม่ต่างจากเครื่องซูเปอร์คอมพิวเตอร์ เมื่อเทียบระหว่างราคากับประสิทธิภาพที่ได้รับส่งผลให้ประสิทธิภาพสูงขึ้นมาก ในราคาที่เท่าเดิม ดังนั้นการเชื่อมระบบพีซีเข้าด้วยกันเพื่อทำงานเป็นซูเปอร์คอมพิวเตอร์ จึง สามารถทำได้เรียกว่า “Beowulf Cluster” ปัญหาอีกอย่างหนึ่งที่พบเมื่อใช้ระบบราคาแพง คือ ค่าบำรุงรักษาที่สูงมาก ส่วนระบบ PC เป็นเทคโนโลยีที่คนส่วนใหญ่คุ้นเคย ทำให้สามารถบำรุงรักษาระบบได้ง่ายกว่า นอกจากนั้น เมื่อเทคโนโลยีนั้นเก่า หรือ ช้าไปแล้ว การหาทุนเพิ่มระบบจะเป็นไปได้ยาก ในขณะที่ในระบบ PC Cluster การเพิ่มความสามารถทำได้ทีละน้อยในราคาที่ถูกกว่า นอกจากนั้นเครื่องที่นำออกจากระบบยังเอาไปใช้ต่อได้ รวมถึงความก้าวหน้าของ Software เช่น ลีนุกซ์ (Linux) ที่เป็นระบบปฏิบัติการฟรี (Open Source) ที่มีประสิทธิภาพสูง, ระบบโปแกรมแบบขนาน MPI (Message Passing Interface) และ PVM (Parallel Virtual Machine) ทำให้สามารถสร้างและใช้ขีดความสามารถของระบบคลัสเตอร์ได้เพิ่มมากขึ้นด้วย

วัลลภา วัชรากร และคณะ [7] ได้ทำการวิจัยเรื่อง "การเปรียบเทียบวิธีการกระจายภาระงานโดย ระบบแม่ข่ายเสมือนบนระบบปฏิบัติการลีนุกซ์” เป็นการวิจัยในเรื่องของการทดลองวัดประสิทธิภาพและ

เสถียรภาพของการกระจายภาระงานทั้ง 3 รูปแบบของ Linux Virtual Server โดยใช้อัตราการตอบกลับของข้อมูล (Reply Rate) เป็นตัววัดประสิทธิภาพของระบบ และใช้ค่าสูงสุดของอัตราการตอบกลับข้อมูลเป็นตัววัดความเสถียรภาพของระบบ สิ่งที่สามารถวิเคราะห์ได้จากการทดลอง คือ วิธีการเชื่อมต่อแบบ Direct routing นั้นมีประสิทธิภาพมากกว่าแบบ Tunneling แต่มีเสถียรภาพที่ใกล้เคียงกัน

ชาญวิทย์ แก้วกสิ [9] ได้ทำการวิจัยเรื่อง “รายงานการวิจัยการออกแบบคลัสเตอร์สำหรับกลุ่มเครื่องแม่ข่ายโปรแกรมประยุกต์” ได้ให้จำกัดความและผลการวิจัยไว้ดังนี้ ความคงทนต่อความผิดพลาด หมายถึง ค่า HA ที่สูงถึงแม้ว่าข้อมูลที่อยู่ในการบริการที่ HA สูงจะไม่มีความจำเป็นว่าจะถูกต้องแม่นยำ แต่ว่าบริการที่คงทนต่อการล่มจะประกันความปกติของพฤติกรรมมของระบบและรู้แน่ชัดต่อจำนวนครั้งและชนิดของความผิดพลาดที่จะเกิดขึ้น นั้นคือในบางระบบจึงต้องการเพียงความเป็น HA การกระจายภาระงาน หรือ Load Balancing เป็นการทำงานในคลัสเตอร์เพื่อให้ได้ระบบประสิทธิภาพสูง โดยการจัดการการร้องขอไปยังแต่ละโหนดต่างๆอย่างรวดเร็วแต่การกระจายภาระงานที่ดีไม่ได้หมายความว่าระดับความทนทานต่อการล่มหรือการให้บริการได้ของระบบจะสูงขึ้น ตัวอย่างเช่น เว็บไซต์หนึ่งอาจใช้ฟาร์มของเครื่องแม่ข่ายเพื่อสร้างหน้าเวปโดยมีข้อมูลปกติเก็บอยู่ในระบบฐานข้อมูล ถ้าไม่เกิดคอขวดขึ้นที่ระบบฐานข้อมูล การกระจายภาระงานจะให้ประสิทธิภาพสูงอย่างชัดเจน อย่างไรก็ตาม ระบบฐานข้อมูลเดี่ยวจะทำให้เกิด "จุดล่ม" การพังของเครื่องสำหรับฐานข้อมูลเพียงเครื่องเดียวจะทำให้แม่ข่ายทั้งหมดไร้ประโยชน์ทันที ในกรณีนี้ การเพิ่มจำนวนเครื่องแม่ข่ายในฟาร์มไม่ได้เพิ่มค่า HA

จากการทดสอบระบบคลัสเตอร์สรุปได้ว่าการเชื่อมต่อและการสื่อสารภายในระบบคอมพิวเตอร์คลัสเตอร์เป็นชนิด n-n นั้นคือ เมื่อโหนดใดโหนดหนึ่งในระบบคอมพิวเตอร์คลัสเตอร์มีการติดต่อจากลูกข่ายเพื่อปรับปรุงวัตถุข้อมูล โหนดนั้นจะทำการติดต่อไปยังทุกโหนดที่เหลือในระบบเพื่อทำการปรับปรุงสถานะสำเนาของวัตถุ ทำให้การแสดงผลนั้นออกมาจากวัตถุที่เดียวกันหรือรูปแบบเดียวกัน

อุดม สุธาคำ และคณะ [8] ได้ทำการวิจัย เรื่อง “ออกแบบและสร้างกลุ่มตัวประมวลผลกราฟิกขนาดเล็ก สำหรับการประมวลผลสมรรถนะสูง” คอมพิวเตอร์ขนานหรือคอมพิวเตอร์สมรรถนะสูง ถูกสร้างขึ้นใช้ในงานวิจัยทางวิทยาศาสตร์และงานวิจัยทางวิศวกรรม หรือในสาขาอื่นที่ต้องการคอมพิวเตอร์ช่วยแก้ปัญหาที่ซับซ้อนและกระทำกับข้อมูลจำนวนมาก ซุปเปอร์คอมพิวเตอร์ถูกจัดเป็นคอมพิวเตอร์ขนานกลุ่มดังกล่าวเช่นกัน หากแต่ในอดีตนั้น มีต้นทุนหรือราคาที่สูงมากๆมีแต่หน่วยงานที่มีงบประมาณสูงๆเท่านั้นจึงจะมีโอกาสได้สัมผัสหรือนำไปใช้ในงานวิจัย หากแต่ความก้าวหน้าของ

เทคโนโลยีปัจจุบัน ทำให้ผู้วิจัยหรือมหาวิทยาลัยมีโอกาสได้สัมผัสกับเทคโนโลยีการประมวลผลแบบขนานได้ด้วยคอมพิวเตอร์คลัสเตอร์หรือ คอมพิวเตอร์ที่ประมวลผลแบบกระจายงาน (Distributed Proccessing) ซึ่งมีลักษณะทำงานแบบขนานเช่นกัน ยิ่งกว่านั้นเทคโนโลยีซอฟแวร์ช่วยเสริมให้วิศวกรคอมพิวเตอร์หรือผู้สนใจสามารถสร้างคอมพิวเตอร์ขนานขึ้นเองได้จากคอมพิวเตอร์ส่วนบุคคลในราคาตลาด งานวิจัยจึงได้ออกแบบสร้างคอมพิวเตอร์คลัสเตอร์ขนาดเล็กขึ้น มุ่งหวังที่จะนำมาใช้ในงานวิจัยด้านการประมวลผลเมตริกมากเลขศูนย์ (Sparse Matrix) ขนาดใหญ่ซึ่งเป็นงานที่เกี่ยวข้องกับสมการเชิงเส้นและเป็นพื้นฐานในงานวิศวกรรมส่วนใหญ่รวมถึงงานวิจัยด้านประมวลผลภาพ (Image Processing)