บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การคำนวณแบบขนานเป็นการใช้หน่วยประมวลผลจำนวนมากช่วยกันทำงาน โดยมีเป้าหมายเพื่อให้โปรแกรมทำงานได้เร็วขึ้นหรือเพื่อให้สามารถทำงานที่มีขนาดใหญ่ขึ้น โดยความเร็วที่เพิ่มขึ้น (Speedup) วัดได้จาก เวลาที่โปรแกรมทำงานเมื่อใช้หน่วยประมวลผลตัวเดียวหารด้วยเวลาที่โปรแกรมทำงานเมื่อใช้หน่วยประมวลผลหลายตัว (Speedup = Sequential_time / Parallel_time) ดังนั้นเมื่อใช้หน่วยประมวลผล n ตัว เวลาที่ใช้ก็ควรจะลดลง n เท่า (Speedup = n) อย่างไรก็ตามในความเป็นจริงความเร็วที่เพิ่มขึ้นจะมีค่าน้อยกว่า n อันเนื่องมาจากปัจจัยหลาย ๆ อย่าง นั่นหมายความว่า จำนวนหน่วยประมวลผลที่เพิ่มขึ้นไม่สามารถลดเวลาในการทำงานของโปรแกรมได้อย่างเต็มที่ เราจึงวัดประสิทธิภาพ (Efficiency) ของโปรแกรมคำนวณแบบขนานด้วยความเร็วที่เพิ่มขึ้น หารด้วยจำนวนหน่วยประมวลผล (Efficiency = Speedup / n) โดยทั่วไปประสิทธิภาพของโปรแกรมคำนวณแบบขนานจะลดลงไปเรื่อย ๆ เมื่อใช้หน่วยประมวลผลมากขึ้น

ทั้งนี้ ปัจจัยหลักที่ส่งผลต่อประสิทธิภาพของโปรแกรมคำนวณแบบขนาน ได้แก่ วิธีการแบ่งงาน (Partitioning) ความละเอียดในการแบ่งงาน (Granularity) การกระจายงานให้หน่วยประมวลผลอย่างสม่ำเสมอ (Load balancing) การประสานการทำงาน (Synchronization) และการสื่อสารระหว่างหน่วยประมวลผล (Communication) เป็นต้น ซึ่งปัจจัยเหล่านี้มีผลกระทบต่อการทำงานของโปรแกรม เช่น การแบ่งงานเป็นชิ้นใหญ่เกินไปและมีการกระจายงานที่ไม่สม่ำเสมอ (Load imbalance) จะทำให้หน่วยประมวลผลทำงานได้ไม่เต็มที่ (Low utilization) ในทางตรงกันข้ามการแบ่งงานละเอียดเกินไปอาจทำให้สิ้นเปลืองเวลาในการประสานการทำงานและการสื่อสารระหว่างหน่วยประมวลผล ดังนั้น เมื่อใช้หน่วยประมวลผลจำนวนมากขึ้นสัดส่วนของเวลาที่ใช้ในการแก้ปัญหาจริง ๆ จะลดลง ในขณะที่สัดส่วนของเวลาที่เสียไปเนื่องจากค่าใช้จ่ายอื่น (Overhead) จะเพิ่มขึ้น ประสิทธิภาพในการทำงานจึงลดลง การที่จะทำให้ประสิทธิภาพและความเร็วเพิ่มขึ้นจำเป็นต้องปรับตัวแปรต่าง ๆ เช่น จำนวนหน่วยประมวลผล วิธีการแบ่งงาน และความละเอียดในการแบ่งงาน โดยคำนึงถึงการกระจายงานที่สม่ำเสมอ การประสานการทำงานและการสื่อสารระหว่างหน่วยประมวลผล ให้เหมาะกับขนาดของปัญหา ลักษณะการทำงานของโปรแกรม และขีดความสามารถของระบบคอมพิวเตอร์แบบขนานที่ใช้ (หน่วยประมวลผล หน่วยความจำ หน่วยความจำสำรอง และระบบเครือข่าย) เมื่อนำโปรแกรมเดียวกันไปทำงานบน

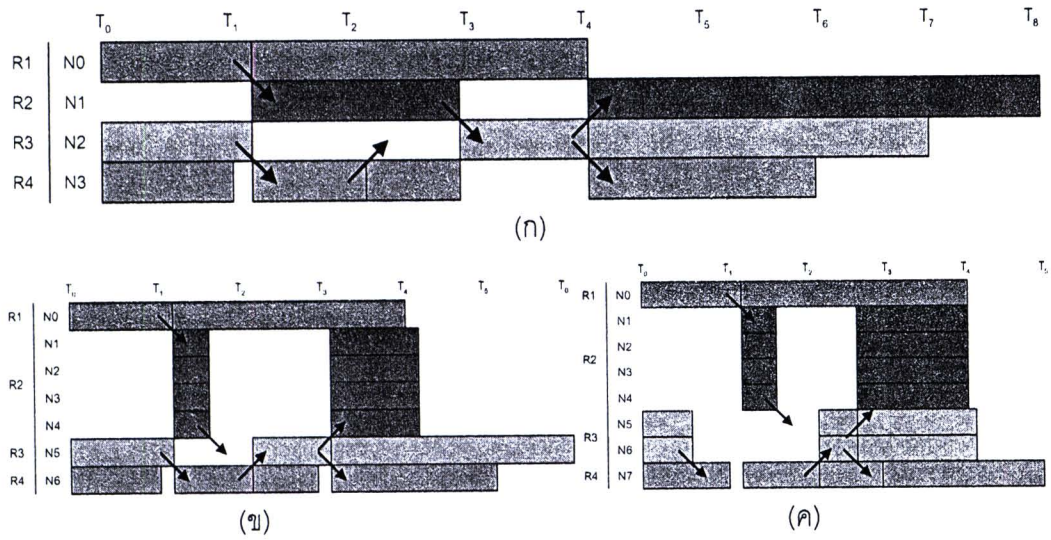
ระบบคอมพิวเตอร์ที่มีสถาปัตยกรรมต่างกันหรือมีประสิทธิภาพต่างกัน ก็จำเป็นต้องปรับตัวแปรเหล่านั้นให้เข้ากับแต่ละระบบด้วย ทั้งนี้ งานวิจัยนี้มุ่งเน้นไปที่ระบบคลัสเตอร์คอมพิวเตอร์เท่านั้น

การปรับปรุงประสิทธิภาพของโปรแกรมโดยขึ้นอยู่กับระบบคอมพิวเตอร์ที่เปลี่ยนไปอาจทำได้โดยการวิเคราะห์การทำงานของโปรแกรมเพื่อหารูปแบบของการคำนวณและการสื่อสาร แล้วนำมาหาวิธีแบ่งงานที่จะได้ประสิทธิภาพสูงสุดโดยนำค่าใช้จ่ายต่าง ๆ ที่จะเกิดขึ้นมาพิจารณาด้วย ซึ่งสามารถทำได้ไม่ยาก ถ้าหากโปรแกรมมีความซับซ้อนไม่มากนัก อย่างไรก็ตาม เมื่อคอมพิวเตอร์มีความเร็วเพิ่มขึ้น ปัญหาที่ถูกนำมาให้คอมพิวเตอร์คำนวณก็ยิ่งทวีความซับซ้อนมากขึ้นไปด้วย ตัวอย่างเช่น ปัญหาที่เกี่ยวกับการจำลองเหตุการณ์ (Simulation) ซึ่งแต่เดิมใช้แบบจำลอง (Model) แยกกันระหว่างการจำลองในระดับมหภาค (Macroscale) และระดับจุลภาค (Microscale) แล้วต่อมามีการนำมารวมกันเรียกว่า ปัญหาที่มีหลายระดับความละเอียด (Multi-scale problem) เป็นปัญหาที่พบมากในการคำนวณทางวิทยาศาสตร์ เช่น แบบจำลองทางฟิสิกส์เพื่อศึกษาสมบัติทางกายภาพ ทางกล ความร้อน ไฟฟ้า และทางแสงของวัสดุ แบบจำลองทางเคมีเพื่อศึกษารายละเอียดของอะตอมในโครงสร้างของสสาร อธิบายกลไกการเคลื่อนที่และการชนกันของอะตอม หรือแบบจำลองทางคณิตศาสตร์ของระบบต่าง ๆ เพื่อเป็นเครื่องมือที่ใช้ในการทำนายเหตุการณ์ล่วงหน้า เป็นต้น ซึ่งแบบจำลองต่าง ๆ ที่กล่าวมานี้ใช้วิธีการคำนวณหลายแบบที่แตกต่างกันมารวมกันตามระดับความละเอียดของปัญหาที่ต้องการศึกษา โดยมีความสัมพันธ์ระหว่างกันในแต่ละระดับความละเอียดด้วย จึงทำให้เกิดเป็นระบบที่มีความซับซ้อนในการคำนวณ ซึ่งเมื่อนำมาพัฒนาเป็นโปรแกรมคำนวณแบบขนานแล้ว จะเป็นการเพิ่มความซับซ้อนในการออกแบบ การวิเคราะห์ และการปรับปรุงประสิทธิภาพของโปรแกรม

การจำลองสึนามิเป็นการสร้างแบบจำลองทางคณิตศาสตร์ของการเกิดคลื่นสึนามิเพื่อใช้ประโยชน์ในการเตือนภัยสึนามิในจุดเสี่ยงที่ต้องการความรวดเร็วในการคำนวณ แต่เนื่องจากเป็นปัญหาที่มีการคำนวณที่ซับซ้อนจึงทำให้ต้องใช้เวลานานในการคำนวณ และด้วยลักษณะการทำงานของโปรแกรมที่ใช้ระเบียบวิธีการวิเคราะห์เชิงตัวเลข FDM (Finite Difference Method) ที่มีการแบ่งพื้นที่ออกเป็นหลายระดับความละเอียดที่ซ้อน ๆ กัน โดยมีความละเอียดในแต่ละระดับและวิธีการคำนวณในแต่ระดับที่แตกต่างกัน แต่มีความสัมพันธ์กันในแต่ละระดับความละเอียดด้วย จึงเป็นการเพิ่มความซับซ้อนในการกระจายงานให้สม่ำเสมอ และทำให้มีการสื่อสารกันระหว่างหน่วยประมวลผลมากในการคำนวณแบบนาน

ด้วยลักษณะของปัญหาที่มีหลายระดับความละเอียดที่มีความขึ้นต่อกันของข้อมูลระหว่างระดับความละเอียด เมื่อนำมาพัฒนาการทำงานเป็นการคำนวณแบบขนานแล้วพบว่าจำเป็นต้องมีการรอรับข้อมูลระหว่างระดับความละเอียดที่มีความขึ้นต่อกัน ซึ่งรูปแบบการแบ่งงานที่แตกต่าง

กันมีผลต่อเวลาที่ใช้ในการรอรับข้อมูล โดยเวลาในการรอคิดจาก เวลาที่ผู้ส่งทำงานก่อนการส่งข้อมูล ลบด้วยเวลาที่ผู้รับทำงานก่อนการับข้อมูลในช่วงการรอ ที่สัมพันธ์กับขนาดของปัญหาที่ได้รับการแบ่งให้แต่ละหน่วยประมวลผล ดังนั้นการหารูปแบบการแบ่งงานจึงต้องทำการวิเคราะห์เวลาที่ใช้ในการทำงานของแต่ละส่วนการทำงานที่สัมพันธ์กับขนาดของปัญหา และนำมาใช้ประมาณเวลาที่ใช้ในการทำงานด้วยรูปแบบการแบ่งงานต่าง ๆ เพื่อหารูปแบบการแบ่งงานที่ทำให้โปรแกรมสามารถทำงานได้เร็วที่สุดตามเวลาที่กำหนดไว้โดยใช้จำนวนหน่วยประมวลผลรวมน้อยที่สุด หรือสามารถหารูปแบบการแบ่งงานที่ทำให้โปรแกรมสามารถทำงานได้เร็วที่สุดโดยใช้จำนวนหน่วยประมวลผลรวมตามที่กำหนดไว้ ดังรูปที่ 1-1 แสดงให้เห็นถึงการเปลี่ยนแปลงของเวลาที่ใช้เมื่อทำการเปลี่ยนรูปแบบการแบ่งงานในแต่ละระดับความละเอียดสำหรับการคำนวณแบบขนาน



รูปที่ 1-1 ลักษณะปัญหาในการแบ่งงานสำหรับการคำนวณแบบขนาน

งานวิจัยนี้จึงได้นำโปรแกรมจำลองสึนามิมาเป็นกรณีศึกษาและทดสอบเพื่อหาระเบียบวิธีการจัดการงาน (Methodology) ที่ทำให้โปรแกรมคำนวณแบบขนานสามารถทำงานได้มีประสิทธิภาพเพิ่มขึ้น และใช้เวลาในการคำนวณลดลง ด้วยการปรับโครงแบบข้อมูลของโปรแกรม (Data configuration) ที่มีผลต่อประสิทธิภาพในการคำนวณ และตามขีดความสามารถของระบบคลัสเตอร์คอมพิวเตอร์ที่ใช้ เพื่อนำไปใช้ในระบบเตือนภัยสึนามิที่มีต้องการความรวดเร็วในการคำนวณ และเพื่อนำระเบียบวิธีการจัดการงานที่มีประสิทธิภาพมาประยุกต์ใช้กับโปรแกรมคำนวณแบบขนานสำหรับปัญหาที่มีหลายระดับความละเอียดอื่น ๆ ต่อไป โดยสามารถกำหนดจำนวนหน่วยประมวลผลเพื่อหาเวลาที่ใช้ในการทำงานที่เร็วที่สุดได้ และสามารถกำหนดเวลาที่ใช้ทำงานเพื่อหาจำนวนหน่วยประมวลผลที่ใช้ให้น้อยที่สุดในการทำงานให้ทันเวลาที่กำหนดได้ ทั้งนี้

แนวคิดในการแก้ปัญหาของงานวิจัยนี้คือ ทำการเก็บข้อมูลการทำงานของกรณีศึกษาด้วยปัญหาเพียงส่วนหนึ่ง เพื่อนำมาวิเคราะห์ได้เป็นรูปแบบในการทำงานของกรณีศึกษา และใช้รูปแบบที่ได้มาทำนายการทำงานของกรณีศึกษากับปัญหาทั้งหมดเพื่อเปรียบเทียบผลที่ได้จากการประเมินด้วยรูปแบบการทำงานที่วิเคราะห์ได้

1.2 วัตถุประสงค์ของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบระเบียบวิธีการจัดการงานของการคำนวณแบบขนานสำหรับปัญหาที่มีหลายระดับความละเอียดบนระบบคลัสเตอร์คอมพิวเตอร์ ซึ่งใช้การจำลองสึนามิเป็นกรณีศึกษา ด้วยวิธีการแบ่งงานและการกระจายงานอย่างสม่ำเสมอ และการลดปริมาณการสื่อสารและการประสานการทำงาน ตามข้อกำหนดในเรื่องของจำนวนหน่วยประมวลผลที่ใช้ และข้อกำหนดในเรื่องของเวลาที่ใช้ โดยสามารถปรับเปลี่ยนโครงแบบข้อมูลของโปรแกรมในแต่ละระดับความละเอียดตามขีดความสามารถของระบบคลัสเตอร์คอมพิวเตอร์ที่ใช้ เพื่อให้ได้ประสิทธิภาพและความเร็วในการคำนวณแบบขนานเพิ่มขึ้น และเพื่อนำระเบียบวิธีการจัดการงานมาประยุกต์ใช้กับโปรแกรมคำนวณแบบขนานสำหรับปัญหาที่มีหลายระดับความละเอียด

1.3 ขอบเขตของการวิจัย

1. งานวิจัยนี้ทำการศึกษาและพัฒนากับโปรแกรมจำลองสึนามิด้วยการคำนวณแบบขนานเท่านั้นซึ่งเป็นปัญหาที่มีหลายระดับความละเอียด เพื่อประโยชน์ในการนำไปใช้งานในระบบเตือนภัยสึนามิ ซึ่งมีข้อกำหนดในเรื่องของเวลาที่ใช้ในการคำนวณ
2. การทดสอบความเร็วในการคำนวณแบบขนานด้วยโปรแกรมจำลองสึนามิ เป็นการทดสอบการคำนวณบนระบบคลัสเตอร์ TERA และระบบคลัสเตอร์ TSUBAME เท่านั้น ซึ่งเป็นระบบที่เปิดให้ใช้งานจริง ไม่สามารถควบคุมสภาพแวดล้อมที่ใช้ในการทำงานได้

1.4 คำจำกัดความที่ใช้ในการวิจัย

คำจำกัดความที่ใช้ในการวิจัย มีดังต่อไปนี้

1. ปัญหาที่มีหลายระดับความละเอียด (Multi-scale problem) เป็นปัญหาที่พบมากในการคำนวณเพื่อสร้างแบบจำลองต่าง ๆ ทางด้านวิทยาศาสตร์และคณิตศาสตร์ โดยลักษณะของปัญหาประเภทนี้คือ เป็นการวิเคราะห์และการคำนวณเชิงตัวเลขด้วยข้อมูลหลาย ๆ ระดับที่มีความละเอียดของข้อมูลต่างกัน และใช้วิธีการคำนวณที่แตกต่างกันตามระดับความละเอียดของข้อมูลนั้น ๆ เพื่อให้ได้ผลการคำนวณที่ถูกต้องแม่นยำขึ้น โดยตัวอย่างของปัญหาประเภทนี้ที่นำมาใช้ในการสร้างแบบจำลองต่าง ๆ

ได้แก่ แบบจำลองทางฟิสิกส์เพื่อศึกษาสมบัติทางกายภาพ ทางกล ความร้อน ไฟฟ้า และทางแสงของวัสดุ แบบจำลองทางเคมีเพื่อศึกษารายละเอียดของอะตอมในโครงสร้างของสสาร อธิบายกลไกการเคลื่อนที่และการชนกันของอะตอม หรือแบบจำลองทางคณิตศาสตร์ด้านต่าง ๆ เพื่อเป็นเครื่องมือที่ใช้ในการทำนายเหตุการณ์ล่วงหน้า เป็นต้น ในงานวิจัยนี้จึงนำการจำลองสีนามิมาเป็นกรณีศึกษาในการออกแบบระเบียบวิธีการจัดการงานที่มีประสิทธิภาพ ซึ่งโปรแกรมจำลองสีนามิที่นำมาศึกษาประกอบด้วย โปรแกรมจำลองสีนามิด้วยการคำนวณแบบลำดับ และโปรแกรมจำลองสีนามิด้วยการคำนวณแบบขนาน

2. การปรับปรุงประสิทธิภาพได้ (Adaptive performance tuning) เป็นการศึกษาและวิเคราะห์การทำงานของโปรแกรม เพื่อหาปัจจัยต่าง ๆ ที่มีผลต่อประสิทธิภาพและความเร็วในการคำนวณ ซึ่งในงานวิจัยนี้เป็นการศึกษาเพื่อหาความสัมพันธ์ของการแบ่งงานที่มีผลต่อเวลาที่ใช้ในการคำนวณที่สามารถนำมาใช้ในการปรับปรุงประสิทธิภาพการทำงานของโปรแกรมคำนวณแบบขนานได้

1.5 ขั้นตอนและวิธีดำเนินการวิจัย

1. ศึกษาหลักการทำงานของโปรแกรมจำลองสีนามิด้วยการคำนวณแบบลำดับและแบบขนาน เพื่อให้ทราบถึงค่าตัวแปร และส่วนการคำนวณต่าง ๆ ของโปรแกรมที่สามารถปรับแต่งได้
2. ดำเนินการพัฒนาและปรับปรุงโปรแกรมจำลองสีนามิด้วยการคำนวณแบบขนานเพื่อให้ได้การคำนวณที่รวดเร็วและถูกต้องแม่นยำมากขึ้น
3. ศึกษางานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ประสิทธิภาพการคำนวณ และการจัดการงานที่คำนวณแบบขนาน โดยหาปัจจัยที่มีผลต่อการแบ่งงานและการกระจายงานของโปรแกรมจำลองสีนามิด้วยการคำนวณแบบขนาน และหาความสัมพันธ์ของปัจจัยต่าง ๆ ที่มีผลต่อประสิทธิภาพการทำงาน
4. ทำการวางแผนและออกแบบระเบียบวิธีการจัดการงาน โดยเน้นที่วิธีการแบ่งงานและการกระจายงานอย่างสม่ำเสมอ และการลดปริมาณการสื่อสารและการประสานการทำงาน โดยพิจารณาถึงประสิทธิภาพของระบบที่นำมาใช้ทำการคำนวณด้วย โดยจะทดสอบกับระบบคลัสเตอร์ TERA และระบบคลัสเตอร์ TSUBAME
5. นำระเบียบวิธีการจัดการงานที่ได้ออกแบบไว้ มาทำการทดสอบและวัดประสิทธิภาพจากอัตราเร็วที่เพิ่มขึ้นของแต่ละระเบียบวิธีการจัดการงานที่ได้ออกแบบไว้ และเปรียบเทียบผลออกมาเป็นกราฟที่เข้าใจได้ง่าย

6. วิเคราะห์ประสิทธิภาพของแต่ละระเบียบวิธีการจัดการงานที่ผ่านการทดสอบ เพื่อหาระเบียบวิธีการจัดการงานที่ทำให้ได้ประสิทธิภาพการทำงานเพิ่มขึ้น
7. สรุปผลการทดลองและนำเสนอระเบียบวิธีการจัดการงานที่มีประสิทธิภาพ มาเป็นระเบียบวิธีมาตรฐานที่ใช้ในการจัดการงานของการคำนวณแบบขนานสำหรับปัญหาที่มีหลายระดับความละเอียด และจัดทำวิทยานิพนธ์

1.6 ประโยชน์ที่คาดว่าจะได้รับ

ได้ระเบียบวิธีการจัดการงานของการคำนวณแบบขนานสำหรับปัญหาที่มีหลายระดับความละเอียด ซึ่งใช้โปรแกรมจำลองสึนามิด้วยการคำนวณแบบขนานเป็นกรณีศึกษา โดยสามารถปรับเปลี่ยนโครงแบบข้อมูลของโปรแกรมได้ตามขีดความสามารถของระบบคอมพิวเตอร์ที่ใช้ เพื่อให้ได้ความเร็วในการคำนวณและประสิทธิภาพในการทำงานเพิ่มขึ้น โดยสามารถนำระเบียบวิธีที่ได้ไปประยุกต์ใช้งานกับโปรแกรมคำนวณแบบขนานสำหรับปัญหาที่มีหลายระดับความละเอียดอื่น ๆ ได้ อีกทั้งได้ใช้ประโยชน์จากการใช้งานทรัพยากรคอมพิวเตอร์ที่มีอยู่อย่างมีประสิทธิภาพ และส่วนของกรณีศึกษาสามารถนำมาพัฒนาเป็นระบบเตือนภัยที่จำลองสึนามิแบบเรียลไทม์ได้

1.7 ลำดับการจัดเรียงเนื้อหาในวิทยานิพนธ์

วิทยานิพนธ์นี้แบ่งเนื้อหาออกเป็น 6 บทดังต่อไปนี้ บทที่ 1 เป็นบทนำซึ่งกล่าวถึงความเป็นมาและความสำคัญของปัญหา รวมถึงวัตถุประสงค์ของการวิจัย บทที่ 2 กล่าวถึงทฤษฎีพื้นฐานและงานวิจัยที่เกี่ยวข้องในงานวิจัยนี้ บทที่ 3 กล่าวถึงแนวคิดของงานวิจัย บทที่ 4 กล่าวถึงการทดลองและวิเคราะห์ผลการวิจัย บทที่ 5 กล่าวถึงการประเมินผลงานวิจัย และบทที่ 6 กล่าวถึงสรุปผลการวิจัยและข้อเสนอแนะ

1.8 ผลงานที่ได้รับการตีพิมพ์

ส่วนหนึ่งของวิทยานิพนธ์นี้ได้ตีพิมพ์เป็นบทความทางวิชาการในหัวข้อเรื่อง "Parallel Program Development for Tsunami Simulaiton with the Message Passing Interface" โดย สิทธิกร ถาวรรัตนวณิช, กิตติพัฒน์ วิโรจน์สิริ, วีระ เหมืองสิน และอาณัติ เรืองรัศมี นำเสนอในงานประชุมวิชาการ "14th International ANnual Symposium on Computational Science and Engineering (ANSCSE 14)" ณ มหาวิทยาลัยแม่ฟ้าหลวง จังหวัดเชียงราย ระหว่างวันที่ 24 - 25 มีนาคม 2553 และได้นำเสนอส่วนหนึ่งของวิทยานิพนธ์นี้ในการสัมมนาเชิงปฏิบัติการ First ECTI Graduate Workshop on High Performance Distributed Computing ณ มหาวิทยาลัยเกษตรศาสตร์ จังหวัดกรุงเทพมหานคร ในวันที่ 16 กรกฎาคม 2553