# การเรียกใช้เว็บเซอร์วิสเพื่อการถ่ายโอนข้อมูลในระบบกริด

# Web Services Invocation for Data Transfer in Grid System

นายอุดมวิทย์ บุญมาพบ[1], นางสาวอภิญญา เทียมสินสังวร[1], กานดา รุณนะพงศา[1], เจษฎา เพ็งสุวรรณ[2]
[1]ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น
อำเภอเมือง, จังหวัดขอนแก่น 40002, ประเทศไทย
[2]หน่วยปฏิบัติการวิจัยการจำลองขนาดใหญ่ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
อ.คลองหลวง , จ.ปทุมธานี 12120, ประเทศไทย
Email: tong_udomwit@hotmail.com, apinya.jan@gmail.com,
krunapon@kku.ac.th, jedsada.phengsuwan@nectec.or.th

## Abstract

*When organizations do science research, they often need to compute large data or generate large files and transfer them through many sites of which data has different formats.*

*Many people use an open- source middleware that is called "Globus Tookit 4" for building collaborative distributed applications. Globus uses GridFTP (Grid File Transfer Protocol) and RFT (Reliable File Transfer) for transferring large files. Currently, file transfer operates continuously on each node by human beings. In addition, to use file transfer facility in Globus, the users need to type a command line.*

*This paper purposes and develops software for higher efficacious data transfer between multiple nodes. Our software can check a number of active transfers and transfer status by using RFT. It can also notify users when the transfer is finished. The program also has a Graphical User Interface (GUI) mode for ease of use.*

**Keywords:** Grid, RFT, Notification

## บทคัดย่อ

*ในหน่วยงานหรือองค์กรที่มีงานวิจัยทางด้านวิทยาศาสตร์หลายงานที่ใช้ข้อมูลที่มีขนาดใหญ่ หรือตัวโปรแกรมมีเอาท์พุตขนาดใหญ่อยู่ในรูปของไฟล์ ดังนั้นแต่ละองค์กรที่ทำงานร่วมกันจึงจำเป็นต้องมีจึงมีการประมวลผลข้อมูลและการถ่ายโอนไฟล์ ซึ่งในแต่ละองค์กรอาจจะมีรูปแบบข้อมูลหรือไฟล์ที่แตกต่างกัน*

*สำหรับการเชื่อมโยงแต่ละโหนดของโปรแกรมที่ติดต่อกันหรือทำงานร่วมกัน หลายคนได้นิยมใช้มิดเดิ้ลแวร์โกลบัส (Globus) ซึ่งได้นำเทคโนโลยีของเว็บเซอร์วิสเข้ามาใช้ประโยชน์ ทำให้กริดโหนดสามารถสื่อสารกันได้ และใช้ GridFTP (Grid File Transfer Protocol) และ RFT (Reliable File Transfer) ช่วยในการถ่ายโอนไฟล์ ซึ่งปัจจุบันการถ่ายโอนไฟล์ยังต้องมีมนุษย์มาควบคุมการถ่ายโอนให้เป็นไปอย่างต่อเนื่องในแต่ละโหนด และการป้อนข้อมูลเพื่อการถ่ายโอนนั้นยังอยู่ในรูปแบบพิมพ์คำสั่งโหมด*

*ดังนั้นงานวิจัยนี้จึงพัฒนาซอฟต์แวร์เพื่อช่วยให้เกิดกระบวนการถ่ายโอนไฟล์ผ่านหลายกริดโหน คแบบอัตโนมัติ โดยสามารถตรวจสอบการถ่ายโอนที่กำลังดำเนินการอยู่และสถานะการถ่ายโอนของไฟล์ ตลอดจนมีการแจ้งเตือนเมื่อการถ่ายโอนเสร็จในแต่ละโหนด และพัฒนาโปรแกรมให้อยู่ในรูปแบบของกราฟิกโหมดเพื่อความง่ายต่อการใช้*

**คำสำคัญ** กริด, การถ่ายโอนไฟล์ที่เชื่อถือได้, การแจ้งเตือน

## 1. บทนำ

จากการใช้งานในระบบกริด การส่งข้อมูลจากโหนดหนึ่งไปยังอีกโหนดนั้นมักจะเกิดขึ้นเสมอในการทำงาน การทำงานในลักษณะนี้เป็นการส่งระหว่าง โหนด ต่อ โหนด โดยใช้ RFT (Reliable File Transfer) ซึ่งเป็นโปรแกรมที่อยู่ใน Globus Toolkit 4 สามารถทำงานได้ ในกรณีของการถ่ายโอนข้อมูลที่ต้องผ่านแบบหลาย ๆ จุด จำเป็นที่จะต้องมีผู้ใช้มาควบคุมในการถ่ายโอนระหว่างโหนด ซึ่งเป็นระบบที่ยังไม่ดีเพราะต้องใช้คนมาคอยจัดการ[1]

ในการถ่ายโอนไฟล์นั้น หลายคนนิยมใช้มิดเดิ้ลแวร์ (Middleware) ที่ชื่อ โกลบัส (Globus) [2] ซึ่งมีพื้นฐานในการทำงานบนเว็บเซอร์วิส และมีบริการพื้นฐาน อย่างเช่น WSRF (Web Services Resource Framework) [3] ซึ่งทำให้ทั้งกริดและเว็บเซอร์วิส ทำงานประสานกันในเฟรมเวิร์คเดียวกันได้ นอกจากนี้ยังมีบริการ WS-Notification ซึ่งเป็นระบบแจ้งเตือนที่โกลบัสสนับสนุน ซึ่งบริการทั้งสองตัวนี้จะช่วยให้กระบวนการโอนถ่ายข้อมูล ตามที่เสนอในงานวิจัยนี้เกิดขึ้นได้

ทางกลุ่มผู้พัฒนาได้นำ RFT และบริการพื้นฐานที่มีอยู่ใน Globus Toolkit 4 (GT4) ซึ่งเป็นโปรแกรมโอเพนซอ ร์ส (Open Source) มาพัฒนาต่อ เพื่อให้มีความสามารถในการจัดการถ่ายโอนข้อมูลแบบผ่านหลายจุดได้อย่างอัตโนมัติ

โดยที่งานวิจัยนี้มีจุดประสงค์เพื่อ พัฒนาระบบการถ่ายโอนข้อมูลที่มีขนาดใหญ่แบบผ่านหลายจุด เพื่อใช้ในระบบกริด โดยที่ระบบ สามารถติดตามดู สถานการณ์ ถ่ายโอนข้อมูล ได้ และผู้ใช้สามารถใช้โปรแกรมใน รูปแบบ กราฟิก โหมด ได้ ประโยชน์ที่คาดว่าผู้ใช้โปรแกรมจะได้รับคือทำให้การถ่ายโอนไฟล์เป็นไปอย่างอัตโนมัติโดยมนุษย์ไม่ต้องมาคอยจัดการ นอกจากนี้ยังทำให้ผู้ใช้โปรแกรมสามารถตรวจดูสถานะของการถ่ายโอนไฟล์ได้สะดวกต่อการใช้งาน และง่ายต่อความเข้าใจ

## 2. เทคโนโลยีและงานวิจัยที่เกี่ยวข้อง

**Globus Toolkit 4** เป็นมิดเดิ้ลแวร์ที่พัฒนาโดยกลุ่ม Globus Alliance [2] ซึ่งสนับสนุนการทำงานร่วมกันของระบบคอมพิวเตอร์ในสภาพแวดล้อมของกริดตามมาตรฐาน OGSA[5] โดย GT4 ประกอบด้วยบริการพื้นฐาน ไลบรารี่ รวมถึงเครื่องมือต่างๆ ที่สนับสนุนการสร้างและใช้งานกริดแอพพลิเคชั่น เช่นบริการการถ่ายโอนข้อมูล (RFT[6]) บริการข้อมูลทรัพยากรระบบ( Index Service[7]) และบริการจัดสรรทรัพยากรในการรันงาน (WS-GRAM[8]) เป็นต้น

**Web Services Resource Framework (WSRF)** [9] เป็นข้อกำหนดมาตรฐานที่พัฒนาโดย OASIS โดยข้อกำหนดนี้ได้เพิ่มความสามารถของเว็บเซอร์วิสทั่วไปซึ่งมีกลไกการทำงานแบบไม่จดจำสถานะ (stateless) ให้มีความสามารถการทำงานแบบจดจำสถานะ (stateful) เพื่อให้เว็บเซอร์วิสสามารถทำงานสอดคล้องกับกลไกการทำงานของกริดแอพพลิเคชั่น

**WS-Notification [10]** เป็นมาตรฐานที่ถูกออกแบบบนพื้นฐานของ notification หรือ Observer design pattern [11] ทำให้ไคลเอนท์สามารถได้รับการแจ้งกลับจากเว็บเซอร์วิสเมื่อมีเหตุการณ์ที่ไคลเอนท์นั้นสนใจเกิดขึ้นที่ฝั่งของเว็บเซอร์วิสนั้น

**Grid FTP/RFT** เป็นโปรแกรมอรรถประโยชน์และบริการพื้นฐานของ GT4 ซึ่งทำหน้าที่ในการถ่ายโอนไฟล์ขนาดใหญ่ โดย RFT นั้นถูกพัฒนาตาม

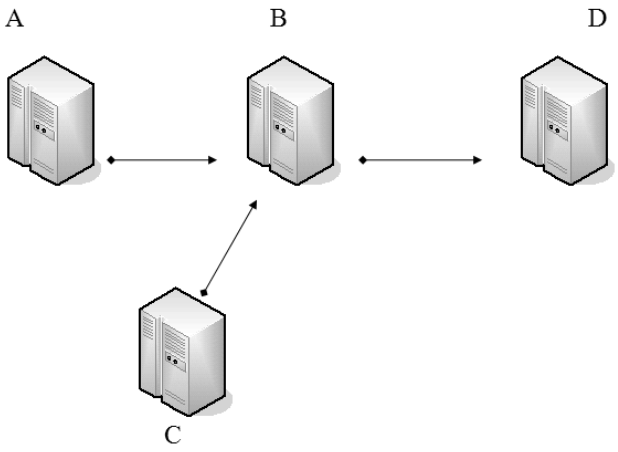
มาตรฐาน WSRF ทำให้สามารถตรวจสอบสถานะการถ่ายโอนข้อมูลได้

**Grid Security Infrastructure (GSI)** [12] เป็นระบบความปลอดภัยภายใต้สภาพแวดล้อมของกริด ซึ่งถูกออกแบบมาเพื่อรองรับการใช้งานทรัพยากรต่างๆ ร่วมกันระหว่างองค์กรหลายองค์กร โดย GSI ถูกพัฒนาบนมาตรฐาน X.509[13] และประกอบด้วยเครื่องมืออรรถประโยชน์ต่างๆ ในการจัดการระบบใบรับรองอิเล็กทรอนิกส์ ( CA) รวมถึง ไลบรารี่ที่สนับสนุนการใช้งาน GSI ร่วมกับเว็บเซอร์วิส

## 3. วิธีการดำเนินงาน

### 3.1 ออกแบบระบบในการควบคุมการถ่ายโอนข้อมูลภายใต้ GridFTP และ RFT

โดยที่ในงานวิจัยนี้ได้จำลองระบบกริดด้วยโหนด 4 โหนดซึ่งเป็น Front-end ดังแสดงในรูปที่ 1




รูปที่ 1 แสดงโหนดต่าง ๆในระบบกริด

จากรูปที่ 1 สมมุติว่ามีองค์กร 4 องค์กรคือ A B C D ซึ่งเชื่อมโยงกันเป็น กริด (ผ่านทาง Front-end ที่ลง GT4) และสมมุติว่า ไฟล์ข้อมูลอยู่ที่ A และ C ซึ่งข้อมูลทั้งสองนี้จะถูกส่งไปประมวลผลที่ B ก่อนเสร็จแล้วจึงจะส่งผลลัพธ์ไปยัง D โปรแกรมที่นำเสนอในงานวิจัยนี้ สามารถที่จะส่ง ข้อมูลจาก A ไป B แล้วมีการแจ้งให้ทราบว่าส่งเสร็จแล้ว และส่งจาก C ไป B แล้วมีการแจ้งให้ทราบว่าส่งเสร็จแล้ว จากนั้นข้อมูลที่ B จะถูกประมวลผล และหลังจากนั้นข้อมูลจะ ถูกส่งไปยัง D เพื่อนำไปใช้ต่อไป อย่างอัตโนมัติ

### 3.2 ติดตั้ง Testbed

ติดตั้ง Globus Toolkit และปรับแต่งค่าต่าง ๆ เพื่อให้แต่ละเครื่องทำงานร่วมกันได้เพื่อใช้ในการพัฒนาระบบ โดยนำเครื่องคอมพิวเตอร์มาติดตั้งระบบปฏิบัติการลินุกซ์ ซึ่งที่เลือกให้ลินุกซ์ เพราะเป็นระบบปฏิบัติการที่ฟรีและรองรับการติดตั้งของโกลบัส

### 3.3 พัฒนาระบบ

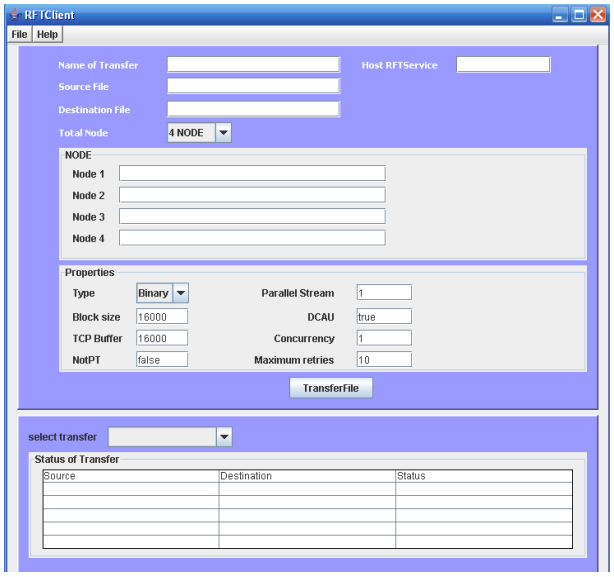
เริ่มต้นด้วยการทดสอบโปรแกรม RFT ในการถ่ายโอนไฟล์ว่าสามารถทำงานได้ถูกต้องหรือไม่ ซึ่งผลจากการทดสอบปรากฏว่าโปรแกรม RFT ทำงานได้ถูกต้อง จากนั้นจึงได้พัฒนาโปรแกรมเพื่อเรียกใช้ RFT ซึ่งเป็นเว็บเซอร์วิสโดยการศึกษาไฟล์ WSDL ของ RFT ซึ่งอธิบายเว็บเซอร์วิส RFT ว่ามีโอเปอเรชันอะไรบ้าง และในแต่ละโอเปอเรชันจะต้องส่งข้อความเป็นแบบใด และจะได้รับข้อความเป็นแบบใด ซึ่งทำให้ได้โปรแกรมที่สามารถส่งต่อเนื่องผ่านไปยังอีกหลาย ๆ โหนดได้โดยอัตโนมัติ (ไม่ต้องใช้คนในการพิมพ์คำสั่ง)

### 3.4 ทดสอบระบบ

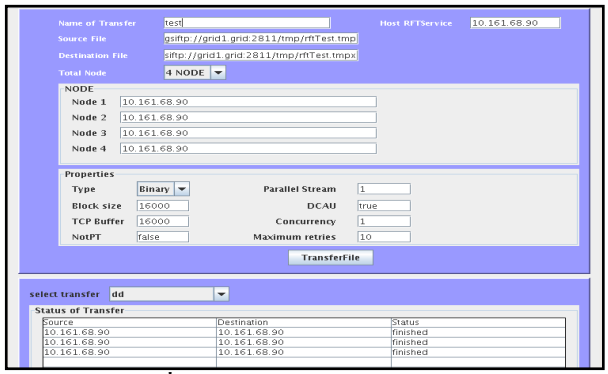
จากนั้นทดลองให้ส่งไฟล์แบบผ่านหลายโหนดบน Testbed ทำการทดสอบและแก้ไขในส่วนที่ผิดพลาดของโปรแกรม ติดตั้งและทดสอบระบบ

## 4. ผลการทดลอง

ผลทดสอบการใช้งานของโปรแกรม ได้แสดงไว้ในรูปที่ 2, 3, และ 4 โดยที่รูปที่ 2 แสดงอินเตอร์เฟสของโปรแกรมแบบกราฟิกโหมด รูปที่ 3 แสดงตัวอย่างการป้อนข้อมูลในการถ่ายโอนไฟล์ และรูปที่ 4 แสดงผลรันของโปรแกรมที่แสดงสถานะของการถ่ายโอนไฟล์



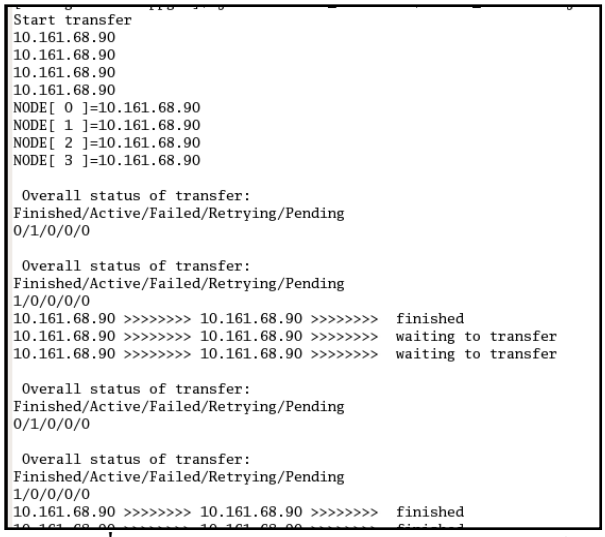

รูปที่ 2 โปรแกรมถ่ายโอนข้อมูลแบบกราฟิก




รูปที่ 3 ตัวอย่างการถ่ายโอนไฟล์

จากรูปที่ 2 จะเห็น โปรแกรมถ่ายโอนข้อมูล ในกราฟิกโหมดซึ่งมีช่องให้ ผู้ใช้กรอกชื่อและที่อยู่ของไฟล์ และตำแหน่งปลายทางที่จะเก็บไฟล์ จำนวนโหนดที่ผ่านเพื่อจะทราบสถานการณ์ส่งสูงสุด 4 โหนด ส่วนการปรับคุณสมบัติของการถ่ายโอนไฟล์ปกติค่าที่ใส่ให้เป็นค่าเริ่มต้นอยู่แล้วไม่จำเป็นต้องเปลี่ยนแปลง และในส่วนสุดท้ายคือการแสดงสถานะการถ่ายโอนไฟล์ระหว่างโหนดเสร็จสิ้น

รูปที่ 3 เป็นการถ่ายโอนไฟล์ชื่อ rftTest.tmp เพื่อส่งไปยังปลายทางชื่อ rftTest.tmpx ผ่านโหนดจำนวน 4 โหนด โดยจำลอง Testbed เครื่องเดียวเสมือนเป็น 4 โหนด ดังนั้น IP ของโหนดทั้ง 4 จึงเหมือนกันคือ 10.161.68.90 ส่วนคุณสมบัติใช้ค่าเริ่มต้นที่ให้มา หลังจากนั้นกดปุ่ม Transfer File เพื่อเริ่มการถ่ายโอนและจะมีหน้าจอแสดงสถานการณ์ถ่ายโอนดังรูปที่ 4




รูปที่ 4 แสดงสถานะขณะทำการถ่ายโอนไฟล์

จากรูปที่ 4 จะเห็นได้ว่ามีการแจ้งการถ่ายโอนไฟล์จากโหนดที่ 1 ไปยังโหนดที่ 4 แบ่งได้ 3 ช่วง ซึ่งจะแสดงสถานะของการถ่ายโอนว่ากำลังถ่ายโอนไฟล์อยู่ หรือไฟล์ถูกถ่ายโอนระหว่างโหนดนั้นๆ เสร็จเรียบร้อยแล้ว

## 4. บทสรุปและข้อเสนอแนะ

ผู้เขียนบทความ นี้ได้พัฒนาซอฟต์แวร์ที่สามารถควบคุมการทำงานช่วยในการตรวจสอบสถานะของการถ่ายโอนข้อมูลหรือไฟล์ ซึ่งสามารถนำไปใช้ในการพิจารณาในเรื่องของการถ่ายโอนข้อมูลที่มีขนาดใหญ่แบบผ่านหลายกริดโหนดเป็นไปอย่างอัตโนมัติ โดยผู้ใช้สามารถที่จะสั่งงานโปรแกรมบนเครื่องที่ติดตั้งมิดเดิ้ลแวร์เพียงครั้งเดียวแล้วระบบการถ่ายโอนไฟล์จะเป็นไปอย่างอัตโนมัติในแต่ละโหนด โดยไม่ต้องมีคนมาคอยสั่งการในโหนดนั้นๆ และสามารถนำไปประยุกต์เพื่อเสริมประสิทธิภาพการใช้งานบนเว็บเพจได้อีกเพื่อความสะดวกและรวดเร็วต่อการใช้งาน หรือจะนำไปประยุกต์กับการเปลี่ยนแปลงข้อมูลระหว่างโหนดได้ เนื่องจากโปรแกรมนี้สามารถที่จะเก็บสถานะของข้อมูลเมื่อมีการหยุดการถ่ายโอนข้อมูลชั่วคราวหรือการถ่ายโอนล้มเหลว เมื่อจะทำการถ่ายโอนข้อมูลต่อสามารถทำได้โดยไม่จำเป็นต้องไปเริ่มถ่ายโอนข้อมูลใหม่ทั้งหมดเหมือนที่ใช้งานอยู่ในปัจจุบัน

งานที่เกี่ยวข้องที่ควรจะทำต่อในอนาคตคือการพัฒนาซอฟต์แวร์นี้ให้เป็นเว็บเซอร์วิสเพื่อให้โปรแกรมอื่นสามารถเรียกใช้ได้โดยอัตโนมัติในการถ่ายโอนข้อมูลผ่านโหลดหลายโหลดในระบบกริด

## 6. กิตติกรรมประกาศ

ขอขอบคุณนายสุริยะ อุรุเอกโอฬาร ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติที่ได้ช่วยแนะนำเกี่ยวกับการจัดการความปลอดภัยในระบบกริด

## 7. บรรณานุกรม

1] Borja Sotomayor, The Globus Toolkit 4 Programmer's Tutorial, 2004-2005 , Retrieved Jan. 17, 2007, from http://gdp.globus.org/gt4-tutorial/multiplehtml/index.html
[2] Globus Alliance , Globus, Retrieved Jan. 17, 2007, from http://www.globus.org
[3] OASIS ,Web Services Resource Framework (WSRF) TC , Retrieved Jan. 17, 2007, from http://www.oasis-open.org/committees/wsrf/
[4] Globus Alliance ,GridFTP and RFT Doc , Retrieved Jan. 17, 2007, from http://www.globus.org/toolkit/docs/4.0/data/
[5] I. Foster,H. Kishimoto, A. Savva,D. Berry, A. Djaoui, A. Grimshaw, B. Horn, F. Maciel, F. Siebenlist, R. Subramaniam, J. Treadwell, J. Von Reich, "The Open Grid Services Architecture, Version 1.0", 29 January 2005, from http://www.gridforum.org/documents/GWD-I-E/GFD-I.030.pdf,
[6] GT 4.0 Reliable File Transfer (RFT) Service, from http://www.globus.org/toolkit/docs/4.0/data/rft/index.pdf
[7] GT 4.0: Information Services: Index, from http://www.globus.org/toolkit/docs/4.0/info/index/index.pdf
[8] GT 4.0 WS_GRAM, from http://www.globus.org/toolkit/docs/4.0/execution/wsgram/index.pdf
[9] “WS-Resource specification”, 1.2 Working Draft 03, Web Services Resource Framework (WSRF) TC, OASIS, March 8 2005
[10] “WS-BaseNotification specfification”, 1.2 Working Draft 01, Web Services Notification(WSN) TC. OASIS, June, 2004
[11] Observer pattern, from http://en.wikipedia.org/wiki/Observer_pattern
[12] Overview of the Grid Security Infrastructure, from http://www.globus.org/security/overview.html
[13] RFC 3820 – Internet X.509 Public Key Infrastructure (PKI) Proxy Certificate Profile, from http://www.faqs.org//rfcs/rfc3820.html, June, 2004