

USULAN TUGAS AKHIR

1. IDENTITAS PENGUSUL

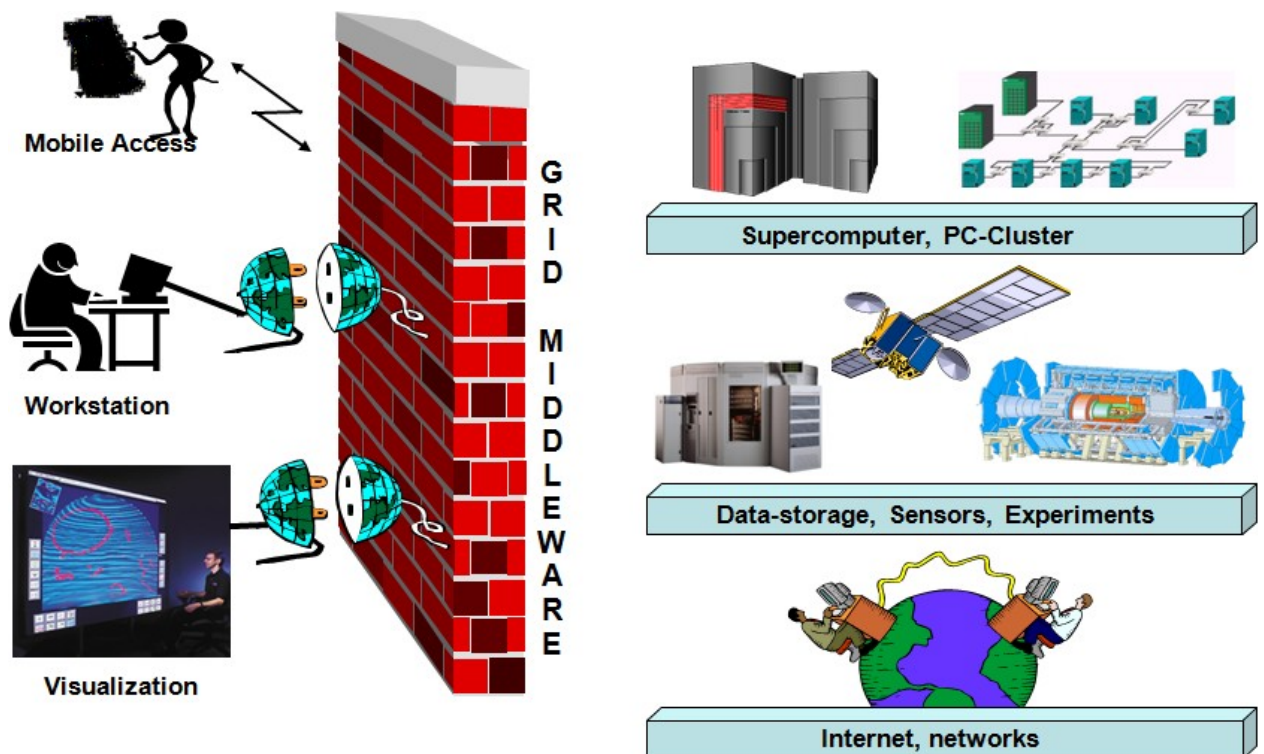
Nama : Benediktus Anindito
NRP : 5108100105
Dosen Wali : Imam Kuswardayan, S.Kom, M.Kom

2. JUDUL TUGAS AKHIR

Analisis Kinerja *Middleware* Komputasi Grid Globus Toolkit dan gLite

3. LATAR BELAKANG

Komputasi grid adalah salah satu jenis sistem terdistribusi dan terpisah secara geografis dimana sumber daya yang ada bisa dimanfaatkan untuk memecahkan persoalan komputasi dalam skala besar. Salah satu komponen penting untuk membangun infrastruktur komputasi grid adalah *middleware*, seperti terlihat pada Gambar 1. [1]



Gambar 1. Posisi *middleware* dalam komputasi grid

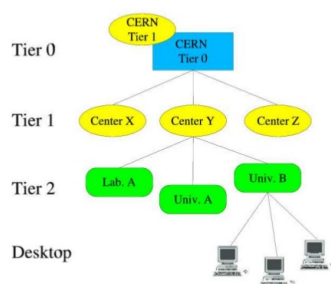
Saat ini ada berbagai macam *middleware* grid yang dibuat oleh institusi pendidikan, maupun oleh *enterprise*, seperti Globus Toolkit, gLite, dan lain-lain. *Middleware* mempunyai peranan yang sangat penting untuk membangun infrastruktur grid diantaranya manajemen

data, manajemen sumber daya melalui perangkat lunak *scheduler*.

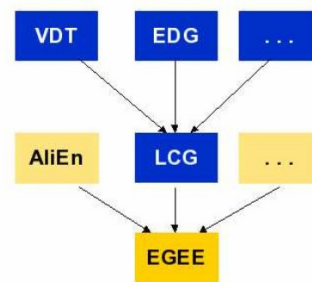
Globus Toolkit adalah *middleware* grid yang paling sering digunakan, dikembangkan oleh para peneliti dari Argonne National Laboratory di University of Illinois. Globus Toolkit terdiri atas GridFTP sebagai manajemen data, Globus Resource Allocation Manager (GRAM) sebagai manajemen sumber daya. Fasilitas lainnya adalah Grid Security Infrastructure yang menyediakan fitur keamanan data dan transaksi pada saat penggunaan infrastruktur grid. Globus Toolkit digunakan oleh Fakultas Ilmu Komputer UI untuk mengembangkan infrastruktur grid, inGRID (Indonesia GRID). Infrastruktur ini dibangun atas dasar kebutuhan akan komputasi berkapasitas besar bagi pengembangan *e-science*, diiringi dengan tersedianya infrastruktur jaringan INHERENT. Dengan inGRID ini, harapannya setiap perguruan tinggi di Indonesia dapat saling berbagi pakai sumber daya komputasi [2].

Selain Globus Toolkit, gLite (dibaca sebagai *gee-lite*), dikembangkan oleh CERN dalam sebuah proyek EGEE (Enabling Grid for E-sciencE) untuk mendukung kegiatan riset pada Large Hadron Collider, sebuah laboratorium fisika partikel yang terletak di perbatasan Swiss dan Perancis.

Setiap tahunnya, seluruh penelitian yang ada di LHC menghasilkan data sebesar 4 petabyte per tahun [6]. Agar dapat memproses data sebesar itu, perlu ada suatu desain yang memenuhi beberapa syarat, antara lain: 1) pengguna tidak harus mempelajari secara mendalam pada saat implementasi dan penggunaan infrastruktur grid; 2) kemampuan untuk menangani akses simultan dan random oleh peneliti dari berbagai belahan dunia; 3) kemudahan dan efisiensi akses data; 4) keamanan dari pihak yang tidak berkepentingan. Solusi dari permasalahan komputasi yang ada pada LHC diawali dengan proposal MONARC (Models of Networked Analysis at Regional Centers). Proposal ini mengembangkan suatu model *multi-tier* yang masih digunakan hingga saat ini. Proposal ini kemudian direalisasikan menjadi LCG (LHC Computing Grid). Pada Gambar 2, Tier-0 terletak di CERN berfungsi sebagai pengambilan data dari LHC, Tier-1 menyediakan sumber daya komputasi berupa penyimpanan dan prosesor, dan Tier-2 melakukan tugas-tugas spesifik seperti simulasi partikel fisika dan tugas-tugas lain. Proyek ini kemudian mengalami perbaikan-perbaikan dan kemudian mejadi EGEE seperti pada Gambar 3.



Gambar 2. Arsitektur MONARC



Gambar 3. Perkembangan Menuju EGEE

Jika pada inGRID menggunakan Globus Toolkit, maka setiap *cluster* yang dimiliki perguruan tinggi harus memiliki *head node*, *storage node*, dan *compute node*. Namun pada gLite, setiap perguruan tinggi bisa memilih untuk menyediakan layanan tertentu saja, sehingga lebih *scalable*.

4. RUMUSAN MASALAH

Rumusan masalah yang akan diangkat dalam tugas akhir ini adalah sebagai berikut:

1. Bagaimana arsitektur infrastruktur grid pada *middleware* gLite?
2. Bagaimana unjuk kerja *middleware* gLite dibandingkan dengan Globus Toolkit?

5. BATASAN MASALAH

Dalam pengerjaan tugas akhir ini, ada batasan-batasan yang perlu ditetapkan, diantaranya adalah sebagai berikut:

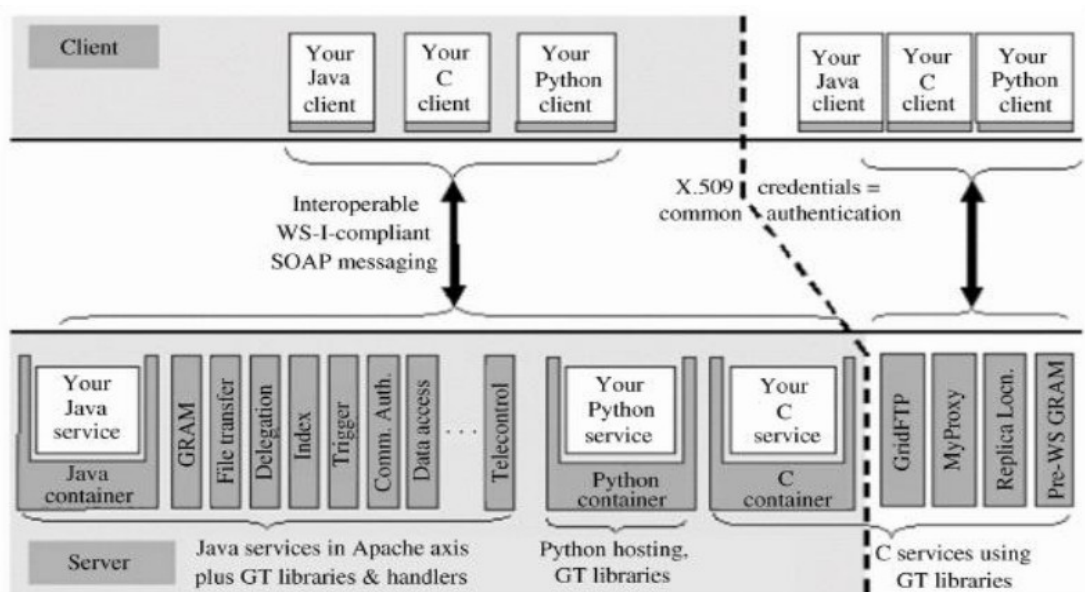
- a) Lingkungan uji coba adalah lingkungan virtualisasi *hardware-assisted*, dengan platform Linux baik pada host OS maupun guest OS.
- b) Pengukuran unjuk kerja meliputi waktu berjalannya program dan kecepatan transfer data dengan program dan ukuran data tertentu.
- c) Pembangunan infrastruktur grid dengan Globus Toolkit mengikuti spesifikasi dan tahapan yang tertulis dalam [3].

6. TUJUAN TUGAS AKHIR

Tujuan pembuatan tugas akhir ini adalah untuk mendapatkan hasil unjuk kerja *middleware* grid selain Globus Toolkit, dalam hal ini adalah gLite. Dengan mengetahui unjuk kerja gLite dan Globus Toolkit, kita bisa menentukan *middleware* mana yang cocok untuk kebutuhan pengguna.

7. RINGKASAN TUGAS AKHIR

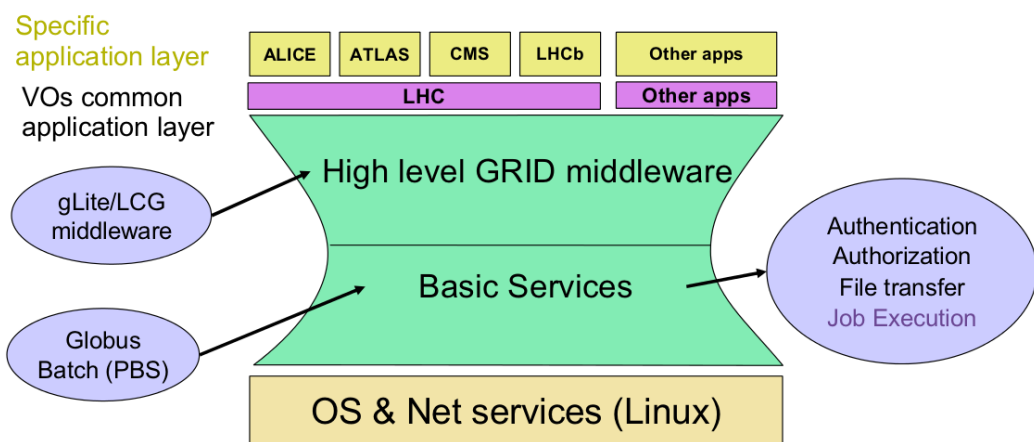
Ada 3 jenis node dalam infrastruktur grid, yaitu *head node*, *storage node*, dan *compute node*. *Head node* adalah sebuah komputer yang merupakan otak dari sebuah *cluster*. Head node berfungsi untuk mengatur compute node mana saja yang akan mengerjakan sebuah job, bagaimana hasil komputasi dari compute node akan disimpan. Storage node adalah sebuah komputer yang berfungsi menyimpan program dan data yang akan dibagi pakai oleh head node dan compute node. Sedangkan compute node adalah unit komputasi terkecil pada sebuah *cluster*.



Gambar 4. Arsitektur Globus Toolkit

Gambar 4 adalah arsitektur *middleware* Globus Toolkit. *middleware* ini terdiri atas beberapa komponen, yaitu:

- Globus Resource Allocation Manager, berfungsi untuk manajemen eksekusi. GRAM akan mengatur compute node mana yang akan menjalankan suatu *job*.
- GridFTP, berfungsi untuk mengatur akses data ke elemen storage.
- MyProxy, berfungsi untuk mengatur hak akses dan manajemen sertifikat pengguna.
- Komponen yang bisa diprogram oleh pengguna, yaitu *container*. *Container* ini menyediakan programabilitas layanan dalam hal keamanan, manajemen penelusuran, manajemen *state*, dan mekanisme lain yang sering dibutuhkan ketika mengembangkan layanan. *Container* tersedia dalam bahasa Java, C, dan Python.
- Pustaka yang digunakan sebagai client atas layanan yang diprogram dan ditaruh di dalam *container*.



Gambar 5. Arsitektur gLite [7]

Pada gLite, ada sedikit perbedaan terminologi untuk elemen-elemen dalam *cluster*. Head node dalam gLite disebut sebagai Computing Element, storage node disebut sebagai Storage Element, dan compute node disebut sebagai Worker Node. Jika dalam Globus Toolkit, pengembang harus memanfaatkan portlet atau aplikasi berbasis web (disebut sebagai portlet) yang terhubung ke *web service* yang disediakan agar pengguna bisa memakai infrastruktur grid, dalam gLite ada program *user interface* berbasis desktop untuk pengguna. Gambar 5 adalah arsitektur *middleware* gLite.

Middleware gLite terdiri atas beberapa komponen, antara lain:

- VOMS (Virtual Organization Membership Service)
Virtual organization terdiri dari individu-individu yang membentuk satuan unit yang mempunyai aturan-aturan dalam pembagipakaian sumber daya yang dimiliki. VOMS menyediakan layanan administrasi virtual organization, manajemen pengguna. Komponen ini juga dapat mengatur pengguna ke dalam suatu struktur hirarkis.
- MyProxy
Komponen ini fungsinya sama seperti pada Globus Toolkit.

c) Information System

Komponen ini mengumpulkan informasi dari beberapa *cluster* grid: jumlah host yang menyala, jumlah host yang dapat dipakai, penggunaan memori. Komponen ini diimplementasikan berdasarkan konsep basis data relasional yang terdistribusi R-GMA (Relational Grid Monitoring Architecture).

d) Workload Management Service

Komponen ini menentukan sebuah job akan dikerjakan oleh worker node mana saja. WMS menentukannya berdasarkan informasi dari komponen Information System.

e) FiReMan Service

FiReMan adalah kependekan dari File Replication Manager. Komponen ini mengatur *identifier* setiap file yang menjadi data. Setiap file yang masuk akan memiliki UUID (universally unique identifier), dan dapat diakses melalui URL yang unik.

f) File Transfer Service

Perbedaan komponen ini dengan FiReMan Service adalah komponen ini mengatur transfer. File transfer ini mendukung konkurensi, pengaturan bandwidth, otentikasi dengan MyProxy.

Tiga poin penting dalam pengerjaan tugas akhir ini adalah:

- a) Pembangunan *cluster* grid dengan Globus Toolkit
- b) Pembangunan *cluster* grid dengan gLite
- c) Mengirimkan 1 job dengan ukuran data tertentu, kemudian diukur kecepatan transfer datanya dan kecepatan pemrosesannya.

Pembangunan *cluster* grid akan dilaksanakan pada lingkungan virtualisasi. Platform *host* OS mendukung *hardware-assisted virtualization* dengan aplikasi virtualisasi VirtualBox.

8. METODOLOGI

Pengerjaan tugas akhir ini meliputi beberapa tahap, antara lain:

a) **Studi Literatur**

Penulis mengumpulkan informasi dan tahapan-tahapan terkait pembangunan *cluster* grid baik menggunakan Globus Toolkit maupun gLite.

b) **Perancangan**

Tahapan ini meliputi pemilihan perangkat lunak yang digunakan untuk mengukur kinerja *middleware* grid.

c) **Implementasi**

Pada tahapan ini, instalasi dan *deployment* perangkat lunak ke lokasi yang telah ditentukan. Jika ada masalah pada saat instalasi dan *deployment*, segera dilakukan *troubleshooting*.

d) **Uji Coba dan Evaluasi**

Pengujian dilakukan dengan beberapa aplikasi grid beserta datanya. Satu job yang

terdiri atas program dan data dikirim ke *cluster* Globus Toolkit dan *cluster* gLite. Parameter diukur menggunakan perangkat lunak yang telah dipilih pada tahap perancangan.

e) Penyusunan Buku Tugas Akhir

Pada tahapan ini, penulis akan mendokumentasikan segala hal yang telah dilakukan ke dalam sebuah buku yang terdiri atas bab-bab berikut:

1. Bab I. Pendahuluan: berisi latar belakang, rumusan masalah dan tujuan pembuatan tugas akhir, metodologi pengerjaan, dan sistematika penulisan buku.
2. Bab II. Landasan Teori: mencakup dasar ilmu yang mendukung pembahasan tugas akhir.
3. Bab III. Perancangan Infrastruktur.
4. Bab IV. Implementasi: berisi tentang tahapan-tahapan instalasi dan *deployment*.
5. Bab V. Uji Coba dan Evaluasi: berisi penjelasan mengenai program dan data yang akan diujicobakan, bagaimana program dan data tersebut dikirim ke *cluster*, kemudian pengambilan hasil uji coba. Setelah itu hasil uji coba dianalisis.
6. Bab VI. Penutup: memuat kesimpulan dari hasil uji coba, dan saran untuk riset selanjutnya.

9. JADWAL Pengerjaan Tugas Akhir

No	Kegiatan	Bulan											
		Maret				April				Mei			
1	Studi Literatur												
2	Perancangan												
3	Implementasi												
4	Uji Coba dan Evaluasi												
5	Penyusunan Buku TA												

10. DAFTAR PUSTAKA

- [1] Xay, Xon et al. *Grid Computing Middleware*. <http://www.dit.hcmut.edu.vn/~ptvu/gc/GRIDComputingMiddleware.ppt>. Diakses tanggal 2 Maret 2012.
- [2] Bobby Nazief, 2007. *InGRID: Pengalaman Mengembangkan Infrastruktur Komputasi Grid*. Yogyakarta: UGM.
- [3] Fauzi, Zulkifli. 2009. *Pemanfaatan Infrastruktur Grid ITS untuk Aplikasi Paralel Berbasis MPI*. Surabaya: Jurusan Teknik Informatik ITS.
- [4] Silva, Vladimir. 2006. *Grid Computing for Developers (Programming Series)*. Hingham: Charles River Media, Inc.
- [5] Anonim. 2011. *History of EGI*. http://www.egi.eu/about/history_of_EGI.html. Diakses tanggal 26 Februari 2012.
- [6] Berlich, Rüdiger et. al. 2005. *Grid Computing in Europe: From Research to Deployment*. <http://crpit.com/confpapers/CRPITV44Berlich.pdf>. Diakses pada tanggal 2 Maret 2012
- [7] Garcia, Ariel. 2005. *The gLite middleware: architecture and components*.

http://gks05.fzk.de/upload/lectures/Garcia_GLite-Introduction.pdf. Diakses pada tanggal 27 Februari 2012.

LEMBAR PENGESAHAN

Surabaya, 5 Maret 2012

Menyetujui,
Dosen Pembimbing,

Ir. F. X. Arunanto, M.Sc