

PENINGKATAN AKSES KOLEKSI PERPUSTAKAAN MENGGUNAKAN PERPUSTAKAAN DIGITAL BERBASIS DATA GRID

Arip Mulyanto dan Rochmad Mohammad Thohir Yassin

Jurusan Teknik Informatika, Universitas Negeri Gorontalo
e-mail: arip.mulyanto@ung.ac.id, thohirjassin@gmail.com

Abstrak

Meningkatnya kebutuhan akses koleksi perpustakaan, menjadikan sebuah perpustakaan digital belum sepenuhnya mampu memenuhi kebutuhan pengguna. Pengguna membutuhkan sistem yang mampu menyediakan akses ke banyak koleksi dari berbagai perpustakaan digital. Untuk itu, diperlukan kerja sama antar perpustakaan digital. Namun, kerja sama antar perpustakaan digital tidaklah mudah. Perbedaan format data, hak kepemilikan sampai letak geografis menjadi salah satu penyebabnya. Paper ini bertujuan membuat arsitektur perpustakaan digital berbasis data grid menggunakan *middleware* iRODS, yang diberi nama DLinGrid. Arsitektur DLinGrid, terdiri dari tiga lapisan, yaitu: data, *grid middleware*, dan aplikasi. Arsitektur DLinGrid memungkinkan beberapa perpustakaan digital yang heterogen untuk saling berbagi sumber daya. Dalam arsitektur ini, koleksi perpustakaan disimpan pada server iRODS. Ketika pengguna meminta koleksi ke sebuah perpustakaan digital, kueri akan diberikan pada server-server iRODS. Pengguna dapat memperoleh berbagai koleksi dari beberapa perpustakaan digital. Metodologi yang digunakan dalam penelitian ini adalah metodologi eksperimen. Hasil eksperimen menunjukkan bahwa perpustakaan digital berbasis data grid dapat meningkatkan aksesibilitas terhadap koleksi perpustakaan digital.

Kata kunci : perpustakaan digital, data grid, iRODS, DLinGrid.

PENDAHULUAN

Saat ini, sebagian besar data dan informasi di berbagai bidang dibuat dan disajikan dalam format digital. Dengan format digital, data dan informasi lebih mudah dikelola. Hal yang sama terjadi dalam bidang perpustakaan. Koleksi perpustakaan seperti buku, jurnal, karya ilmiah dan sumber lainnya disajikan dalam format digital. Sebagian besar perpustakaan konvensional telah berkembang menjadi perpustakaan digital. Dengan perpustakaan digital, akses ke koleksi perpustakaan dapat dilakukan di mana dan kapan saja tanpa harus datang langsung ke perpustakaan.

Namun, akses ke koleksi perpustakaan digital di Indonesia belum merata. Hal ini dikarenakan masih kurang memadainya infrastruktur dan minimnya koleksi berbagai perpustakaan di daerah. Selain masalah akses yang belum merata, ukuran koleksi dan kebutuhan akses ke koleksi yang semakin meningkat, menyebabkan sebuah perpustakaan digital belum sepenuhnya mampu memenuhi kebutuhan pengguna. Saat ini, pengguna membutuhkan sebuah sistem yang menyediakan akses ke banyak koleksi

dari berbagai perpustakaan. Untuk itu, sebuah perpustakaan digital perlu bekerja sama dengan perpustakaan digital lainnya untuk saling berbagi sumber daya. Namun, berbagi sumber daya antar perpustakaan digital bukan hal yang mudah. Hal ini disebabkan adanya perbedaan teknologi, protokol, arsitektur, media penyimpanan, dan fitur-fitur yang dimiliki oleh setiap perpustakaan digital.

Untuk mengatasi masalah ini, beberapa penelitian sebelumnya [1][2] mengusulkan penggunaan teknologi grid yang memiliki kemampuan mengintegrasikan platform yang heterogen. Teknologi grid menjanjikan sebuah arsitektur perpustakaan digital yang fleksibel dan lebih terbuka [2]. Dengan bantuan teknologi grid, permasalahan seperti bertambahnya konten yang heterogen, layanan yang kompleks dan ukuran metadata yang besar dapat diatasi.

Paper ini akan membahas tentang usulan arsitektur perpustakaan digital berbasis data grid untuk meningkatkan akses ke koleksi perpustakaan digital dan mengoptimalkan koleksi setiap

perpustakaan digital yang bekerja sama. Perpustakaan digital berbasis data grid ini, mengintegrasikan berbagai perpustakaan digital dengan *datagrid-middleware* iRODS. Dengan iRODS, proses berbagi sumber daya menjadi lebih mudah, sehingga dapat meningkatkan akses ke koleksi perpustakaan digital.

PERPUSTAKAAN DIGITAL

Perpustakaan digital memiliki konsep yang sama dengan perpustakaan konvensional, yakni menyediakan informasi dan koleksi perpustakaan bagi yang membutuhkan. Perbedaannya terletak dalam hal penyimpanan dan pengaksesan koleksi. Perpustakaan digital adalah sebuah perpustakaan yang informasinya disimpan dalam format digital dan dapat diakses melalui komputer [3]. Format digital disimpan secara lokal dalam sebuah komputer, tetapi dapat diakses oleh pengguna melalui jaringan komputer. Perpustakaan digital didefinisikan sebagai sebuah perpustakaan virtual global yang terdiri dari beberapa jaringan perpustakaan elektronik [4].

Dari beberapa definisi di atas, dan terkait dengan penelitian yang dilakukan, dapat disimpulkan bahwa perpustakaan digital merupakan jejaring perpustakaan elektronik yang informasinya disimpan dalam format digital dan dapat diakses secara universal melalui jaringan komputer. Perpustakaan digital (bisa berupa perpustakaan virtual) terdiri dari beberapa perpustakaan digital yang masing-masing memiliki sumber daya. Masing-masing perpustakaan tersebut dapat saling berbagi informasi dan sumber daya yang dimilikinya.

DATA GRID

Data grid merupakan salah satu tipe dari sistem grid yang menyediakan layanan dan infrastruktur untuk aplikasi distribusi data yang dibutuhkan untuk akses, transfer dan modifikasi dataset berukuran besar yang disimpan dalam sumber daya penyimpanan terdistribusi. Data grid menawarkan infrastruktur terdistribusi dan layanan untuk mendukung aplikasi yang berhubungan dengan ukuran data yang besar dan disimpan dalam sumber daya terdistribusi yang heterogen [5]. Data grid menyediakan layanan yang dibutuhkan pengguna, yaitu: Layanan pencarian: mencari sejumlah data set yang dibutuhkan untuk melindungi sumber daya data; Layanan pengiriman data: mengirim data dalam ukuran besar di antara sumber daya; Layanan replikasi: mengelola berbagai salinan data; dan Layanan akses: mengelola akses terhadap data [6].

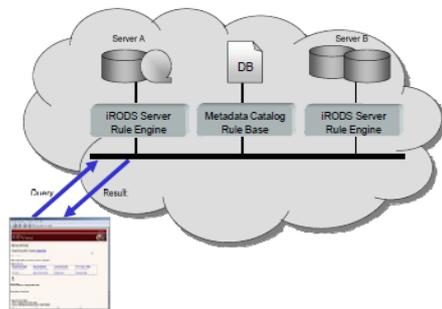
Data grid merupakan penghubung antara sistem penyimpanan, arsip gambar, lingkungan aliran data dan layanan berbasis Web [7]. Data grid bertransaksi dengan data, mengontrol pembagian dan pengelolaan distribusi data dalam jumlah yang besar [8]. Data grid berisi komputasi dan sumber daya penyimpanan yang tersebar di beberapa lokasi yang berbeda yang dapat diakses oleh para pengguna.

iRODS

Untuk memadukan sejumlah variasi sumber daya dalam *grid computing* dibutuhkan *grid-middleware*. *Middleware* adalah sekumpulan perangkat lunak yang mengatur sumber daya tersebut hingga dapat diakses klien tanpa harus mengetahui konfigurasi. Terdapat beberapa jenis *middleware* yang biasa digunakan di lingkungan grid seperti Globus, Gridbus, Alchemi, UNICORE, iRODS, OGSA-DAI dan sebagainya. Di sini hanya akan dibahas *middleware* yang digunakan dalam penelitian ini, yaitu iRODS.

iRODS (*Integrated Rule-Oriented Data Systems*) merupakan sebuah sistem data grid generasi ke-2 yang menyediakan sebuah kesatuan tampilan dan akses untuk mendistribusikan objek digital melalui sebuah jaringan yang luas. iRODS merupakan pengembangan dari *Storage Resource Broker* (SRB) yang dianggap sebagai sistem data grid generasi pertama [9]. iRODS mereplikasi file yang diterima berdasarkan *rule* yang kita definisikan, misalkan ke server mana saja file tersebut akan direplikasi. Informasi mengenai nama file (fisik dan logik) dan di server mana saja file tersebut direplikasi, disimpan dalam server iCAT (*iRODS-catalog*). Server tersebut tidak lain adalah *DBMS-server* yang menyimpan informasi metadata tersebut.

Secara umum, iRODS terdiri 4 (empat) komponen. *User-interface* merupakan komponen untuk mengakses data, sedangkan *iRODS-server* adalah komponen yang digunakan untuk menyimpan data. Komponen yang lain adalah *iRODS-rule engine* yang digunakan untuk mengimplementasikan aturan serta komponen *iRODS-metadata catalog* sebagai database dari data.



Gambar 1. Klien meminta file dari iRODS [9]

Gambar 1 menunjukkan sebuah skenario ketika klien meminta sebuah file dari zona iRODS. Langkah pertama, klien menghubungi satu server iRODS (misal, server A) menggunakan sebuah aplikasi klien dan mengirimkan kriteria file yang dibutuhkan. Permintaan tersebut langsung dikirim ke server A yang akan mencari file menggunakan informasi yang disediakan oleh katalog metadata. Server katalog akan memberitahu server iRODS mana yang menyimpan file tersebut (misal, server B). Kemudian server A meminta file dari server B. Server B membuat aturan yang berhubungan dengan permintaan. Aturan akan menentukan apakah klien berhak untuk membaca file dan mengirim file ke klien menggunakan protokol iRODS. Klien tidak mengetahui di mana lokasi file tersebut, karena adanya virtualisasi lokasi yang dilakukan oleh grid.

HASIL DAN PEMBAHASAN

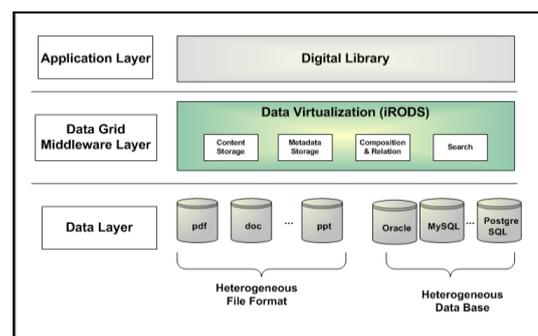
Fungsi utama dari perpustakaan digital adalah menyediakan koleksi beserta informasi yang dimilikinya untuk dapat diakses oleh pengguna yang membutuhkan. Perpustakaan digital akan melakukan proses pencarian informasi yang berhubungan dengan koleksi yang diminta oleh pengguna. Penerapan perpustakaan digital dapat membantu pengguna yang membutuhkan koleksi atau informasi mengenai koleksi perpustakaan kapan saja tanpa harus datang langsung ke perpustakaan dimaksud. Hal ini dapat mengurangi waktu dan tenaga pengguna dibandingkan jika harus mendatangi langsung perpustakaan.

Namun, saat ini pengguna membutuhkan sebuah sistem yang dapat memberikan akses berbagai koleksi dari beberapa perpustakaan digital sekaligus dengan satu kali pencarian. Artinya, dengan satu kali kirim kueri maka sistem tersebut akan melakukan pencarian ke beberapa perpustakaan digital. Hasil pencarian

diberikan kepada pengguna tanpa diketahui dari perpustakaan digital mana koleksi tersebut berasal.

Untuk memenuhi kebutuhan pengguna di atas, diusulkan sebuah arsitektur perpustakaan digital berbasis data grid. Perpustakaan digital akan diintegrasikan dengan *middleware data grid*, seperti iRODS dan OGSA-DAI. Dengan bantuan *middleware data grid*, pengguna hanya melakukan satu kali permintaan kueri terhadap sebuah perpustakaan digital yang bisa saja bertindak sebagai *interface*. Untuk selanjutnya, perpustakaan digital tersebut akan melakukan pencarian terhadap beberapa perpustakaan digital yang telah diregistrasi.

Untuk itu, penulis membuat sebuah desain arsitektur perpustakaan digital berbasis data grid yang diberi nama DLinGrid (Digital Library in Data Grid), seperti terlihat dalam gambar 2.



Gambar 2. Arsitektur DLinGrid

Arsitektur ini terdiri dari tiga layer, *data layer*, *data grid middleware layer* dan *application layer*. Pada *data layer* berbagai jenis sumber daya dengan format dan sistem penyimpanan yang berbeda-beda dapat digabungkan ke dalam satu koleksi data yang besar. Pada *data grid middleware*, data divirtualisasi oleh *data grid middleware*, dalam hal ini adalah iRODS. Pada *application layer*, aplikasi perpustakaan digital dapat menggunakan data virtualisasi dalam menampilkan dan membagi konten objek digital mereka.

Arsitektur DLinGrid terdiri dari beberapa langkah, yaitu:

- Langkah 1 : Setiap perpustakaan digital yang akan menjadi anggota DLinGrid, harus *mensetup* data grid middleware yaitu iRODS.
- Langkah 2 : Masing-masing anggota DLinGrid harus mendigitalkan data koleksi

- perpustakaan yang akan dibagi dengan perpustakaan lain.
- Langkah 3 : Objek digital disimpan dalam berbagai sistem penyimpanan yang dikelola oleh server iRODS.
- Langkah 4 : Setiap server iRODS diregistrasi ke dalam zone DLinGrid, sehingga objek digital dapat direplikasi di antara server-server tersebut.
- Langkah 5 : Setiap perpustakaan digital anggota DLinGrid diregistrasi ke dalam sebuah portal yang mengelola semua sumber daya yang dimiliki oleh koleksi digital.

Berdasarkan arsitektur DLinGrid di atas, langkah penelitian selanjutnya adalah melakukan eksperimen. Eksperimen dilakukan dengan menggunakan dua skenario. Skenario pertama adalah eksperimen pencarian koleksi terhadap satu perpustakaan digital. Skenario kedua adalah pencarian koleksi terhadap perpustakaan digital berbasis data grid menggunakan arsitektur DLinGrid. Dalam eksperimen ini, digunakan dua perpustakaan digital, yakni perpustakaan digital A dan perpustakaan digital B. Kedua perpustakaan digital ini menggunakan sistem yang sama yaitu Lontar.

Pada eksperimen skenario pertama, dilakukan pencarian koleksi terhadap sebuah perpustakaan digital. Ketika sebuah koleksi yang diminta ditemukan, koleksi dalam bentuk digital dikirimkan kepada pengguna. Ketika koleksi tidak ditemukan, pengguna dapat meminta koleksi terhadap perpustakaan digital lain. Hasil eksperimen skenario pertama dapat dilihat dalam gambar 3.

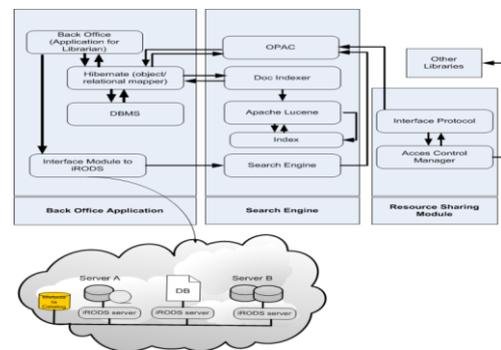


Gambar 3. Hasil Pencarian Koleksi pada Perpustakaan Universitas A

Dengan skenario pertama, pengguna mendapatkan koleksi yang berhubungan dengan kata kunci Grid.

Ketika hasil yang diharapkan belum terpenuhi, pengguna harus melakukan permintaan terhadap perpustakaan digital yang lain. Demikian seterusnya. Bisa dibayangkan apabila koleksi yang diinginkan belum terpenuhi, maka harus melakukan koleksi pencarian pada perpustakaan digital lainnya. Tentu hal ini akan membutuhkan waktu yang lebih lama, karena melakukan pencarian berulang-ulang pada alamat yang berbeda.

Untuk mengatasi permasalahan yang muncul pada skenario sistem perpustakaan digital di atas, diusulkan skenario sistem perpustakaan digital berbasis data grid. Eksperimen skenario kedua dilakukan dengan mengintegrasikan perpustakaan digital ke dalam data grid yaitu iRODS. Pada tahapan ini dilakukan modifikasi terhadap Lontar, yakni data yang biasanya disimpan dalam media penyimpanan Lontar, disimpan di server iRODS seperti terlihat dalam gambar 4. Hal ini dapat dilakukan pada perpustakaan digital lainnya, sehingga terdapat banyak server yang dapat saling berbagi sumber daya. Ketika pengguna mencari koleksi yang dibutuhkan, kueri yang diminta akan dikirim ke seluruh server yang sudah teregistrasi sebagai anggota perpustakaan berbasis data grid yang bernama DLinGrid.



Gambar 4. Integrasi Perpustakaan Digital Lontar ke dalam iRODS

Pada skenario kedua ini, beberapa perpustakaan digital bekerja sama satu sama lainnya. Skenario ini memanfaatkan teknologi data grid yaitu iRODS sebagai media penyimpanan seluruh perpustakaan digital yang bekerja sama. Sebagai contoh, dua perpustakaan digital yaitu perpustakaan digital A dan perpustakaan digital B bekerja sama dan menjadi anggota sebuah sistem yang bernama DLinGrid, sebuah perpustakaan berbasis data grid. Gambar 5 menunjukkan hasil pencarian koleksi menggunakan perpustakaan digital berbasis data grid.



Gambar 5. Hasil Pencarian Koleksi pada DLinGrid melalui Perpustakaan Universitas A

Hasil eksperimen menunjukkan bahwa sistem DLinGrid dapat meningkatkan akses ke koleksi perpustakaan digital. Jika pada perpustakaan digital pencarian koleksi terhadap 2 perpustakaan dilakukan sebanyak 2 kali, maka pada perpustakaan DLinGrid, pencarian dapat dilakukan hanya satu kali saja. Dapat dikatakan bahwa akses ke koleksi meningkat, karena dengan usaha lebih minimal (1 kali akses), mendapatkan hasil maksimal yaitu mendapatkan banyak koleksi dari 2 perpustakaan. Peningkatan akses ini akan lebih terasa apabila DLinGrid terdiri dari banyak perpustakaan digital.

SIMPULAN

Salah satu permasalahan dalam perpustakaan digital adalah terbatasnya akses ke koleksi perpustakaan. Dalam penelitian ini diusulkan sebuah arsitektur perpustakaan digital berbasis data grid yang disebut DLinGrid. Dengan mengimplementasikan perpustakaan digital ke dalam data grid yakni iRODS, memberikan kemudahan mekanisme kerja sama antar perpustakaan digital. Sistem iRODS merupakan solusi penyimpanan data grid yang bersifat *open-source* pada sebuah arsitektur klien-server terdistribusi.

Berdasarkan arsitektur DLinGrid, dilakukan eksperimen dengan dua skenario, yaitu skenario sistem perpustakaan digital dan skenario sistem perpustakaan digital berbasis data grid. Hasil eksperimen menunjukkan bahwa mekanisme perpustakaan digital berbasis data grid lebih memudahkan pengguna untuk mendapatkan koleksi yang dicari. Dengan satu kali pengiriman kueri,

pengguna dapat mengakses berbagai koleksi dari berbagai perpustakaan digital.

UCAPAN TERIMA KASIH

Paper ini merupakan hasil penelitian yang dibiayai oleh DP2M Dikti tahun 2011 dan 2012. Untuk itu, penulis ucapkan terima kasih kepada DP2M Dikti atas biaya penelitian yang diberikan.

REFERENSI

- [1] Chao TY. and Hsin CH. Using Data Grid Technologies to Construct a Digital Library Environment. *Proceeding of the 3rd International Conference on Information Technology Research and Education (ITRE)*.388-392. 2005.
- [2] Sebestyen GP, Banciu D, Balint T, Moscaiu B, and Sebestyen AP. Towards a GRID-based Digital Library Management System. *Distributed and Parallel Systems*. 77-90. Springer-Verlag. 2008.
- [3] Greenstein D, and Thorin SE. *The Digital Library: A Biography*. Washington, DC: Digital Library Federation. 2002
- [4] Larson R and Sanderson R. Grid-based Digital Libraries: Chesire3 and Distributed Retrieval. *JCDL*. 2005
- [5] Chervenak A, Foster I, Kesselman C, Salisbury C, and Tuecke S. The Data Grid: Towards an Architecture for The Distributed Management and Analysis of Large Scientific Datasets. *Journal of Network and Computer Applications*, 187-200. 2000
- [6] Venugopal S, Buyya R, and Kotagiri R. (2006). *A Taxonomy of Data Grids for Distributed Data Sharing Management and Processing*. Melbourne: ACM Computer Survey. 2006
- [7] Moore R, Rajasekar WA and Wan M. Data Grids, Digital Libraries, and Persistent Archives: An Integrated Approach to Sharing, Publishing, and Archiving Data. *Proceedings of The IEEE*. 93: . 2005.
- [8] Joshi H, and Jakharia JC. Digital Library Grid: A Roadmap to Next Generation Digital Libraries Using Grid Technologies, *4th International Convention CALIBER*, Gulbarga.2006.
- [9] iRODS. (n.d.). *iRODS*. Retrieved April 10, 2010, from iRODS: <http://www.irods.org>