

CLUSTERING FILE TERDISTRIBUSI

Yenni Astuti, Muhammad Harikem Ali Akbar
Jurusan Teknik Informatika
Sekolah Tinggi Teknologi Adisutjipto Yogyakarta
informatika@stta.ac.id

ABSTRACT

The development of information technology is increasing at this time, certainly makes it easy for users of the technology itself. One of them in transferring files in a distributed manner, there are many types of files that make it difficult admin in selecting and classifying the file and then will be distributed to each - its own servers. By creating a new system that will facilitate Clustering distributed file an admin in file transfer without having to select the files to be distributed, because the system will automatically read from the file types that will be distributed on its own servers - each.

Keywords: *Clustering, files, distributed systems.*

1. Pendahuluan

Perkembangan teknologi informasi yang semakin maju di saat ini, tentunya tidak luput dari kerja keras para pecinta teknologi, yang saling bersaing untuk menciptakan teknologi yang baru, agar dapat memberikan kemudahan serta manfaat bagi pengguna teknologi informasi saat ini. Seperti pada sebuah sistem pengunggah (*upload*) yang bias menyebarkan *file* yang di-*upload* sesuai dengan *server-nya* masing-masing tanpa harus menentukan *server* tujuan *file* yang akan di-*upload* terlebih dahulu. Ketika *file* sudah terkelola dengan teratur di dalam sebuah mesin penyimpanan atau disebut dengan *server*, tentunya memberikan kemudahan bagi pengguna dan pemakai informasi yang di-*upload* tersebut. Pengelompokan *file* dari *server* induk (*server* moderator) ke *slave server* yang menggunakan system terdistribusi tentunya memberikan kemudahan terhadap *admin* untuk mengelola data di dalamnya dan juga memberikan kemudahan bagi *client* dalam mengambil data yang di-*upload* oleh *admin* ke *server*.

Dengan berjalannya waktu, manusia membutuhkan solusi untuk mempermudah dalam hal pengunggah (*upload*) dan unduh (*download*). Oleh karena itu solusinya adalah membuat system yang bias mengelompokkan (*Clustering*) *file* ke *server-nya* masing-masing berdasarkan tipe dari *file* yang akan di-*upload* tersebut. *File* yang berada pada *server-nya* masing-masing tentunya akan memudahkan pengguna dalam mengelola data yang tersimpan di dalamnya serta menghemat waktu yang dibutuhkan dalam *upload*. Menggunakan system terdistribusi tentunya memberikan solusi yang efektif dalam pengiriman data ke dalam *server* tanpa memikirkan tempat yang saling berjauhan dan bias diakses di manapun dan kapanpun selagi terhubung dengan jaringan *Internet*. Pengembangan sistem terdistribusi yang bias mengelompokkan *file* di dalam *server* masing-masing, yang berdasarkan tipe *file-nya*, tentunya akan memudahkan bagi *admin* dalam pengolahan data yang dikirim di dalamnya, sehingga tidak memerlukan banyak system untuk pengiriman sebuah *file* terhadap *server-nya* masing-masing.

2. Metodologi Penelitian

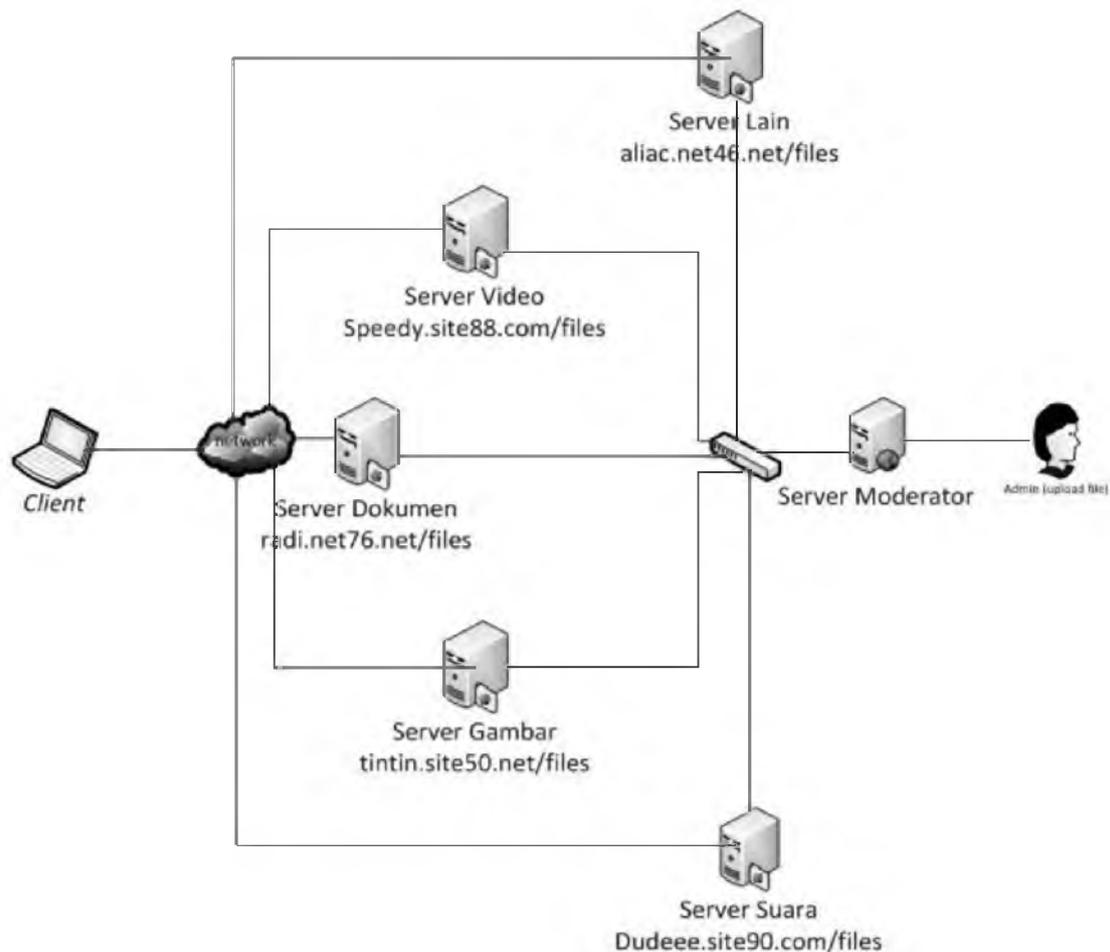
2.1 Pengertian *Clustering*

Clustering merupakan pengelompokan sekumpulan objek sehingga bisa berada dalam satu kelompok yang sama yang disebut *cluster*. Objek-objek dalam sebuah *cluster* memiliki tingkat kemiripan yang tinggi. Dan antar *cluster* memiliki tingkat kemiripan yang rendah. *Clustering* merupakan teknik yang umum digunakan dalam menganalisa data statistik untuk berbagai bidang, misalnya *machine learning*, *pattern analysis*, *image analysis*, *information retrieval* dan *bio informatika*.

Analisa *cluster* bisa dilakukan dengan beberapa algoritma dengan kelebihan dan kekurangan masing-masing. Sebuah algoritma bisa membentuk *cluster* yang detil dan akurat, namun memiliki kekurangan karena memerlukan *resource* komputer yang sangat tinggi dan hal ini berlaku kebalikan untuk jenis algoritma yang lain. Tipe dokumen yang di-*cluster* juga merupakan salah satu faktor untuk menentukan algoritma *clustering* yang digunakan.

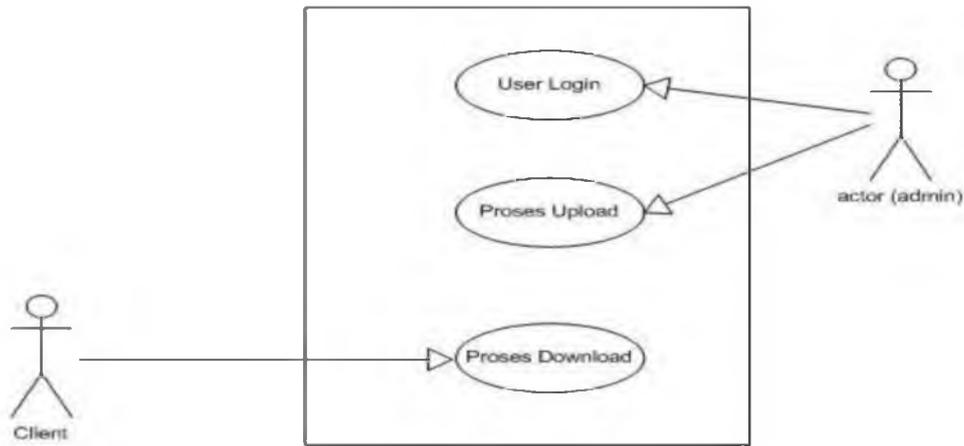
2.2 Perancangan Arsitektur Sistem

Media transmisi sistem yang dibangun menggunakan jaringan *internet* yang terkoneksi secara langsung antar komputer *server* dan komputer *client*, tanpa menggunakan kabel, sehingga *client* dapat mengakses *server* dimanapun dan kapanpun selagi komputer terhubung dengan jaringan *internet*. Terdapat satu buah komputer yang dijadikan sebagai komputer *server* moderator dan lima *server cluster* yang digunakan serta komputer *client*. Komputer moderator berfungsi untuk mengirimkan *file* terhadap lima *server cluster*, *file* yang dikirim melalui komputer moderator akan secara otomatis dipisahkan oleh sistem, dimana sistem membaca setiap tipe *file* yang akan di-*upload* oleh admin komputer moderator yang selanjutnya didistribusikan ke *server cluster*-nya masing-masing. Adapun *server cluster* yang digunakan di antaranya yaitu *server* suara yang berfungsi untuk menampung tipe *file* dari suara, *server* gambar yang berfungsi untuk menampung tipe *file* dari gambar, *server* dokumen yang berfungsi untuk menampung tipe *file* dari dokumen, *server* video yang berfungsi untuk menampung tipe *file* dari video dan *server* lain yaitu berfungsi untuk menampung tipe *file* selain dari tipe *file* dari suara, gambar, dokumen, dan video. Untuk membuka dan mengakses masing-masing server, *client* bisa mengakses dengan nama domain masing-masing server. Untuk nama domain server gambar yaitu dengan nama *tintin.site50.net/files*, server dokumen *radi.net76.net/files*, server suara *dudeee.site90.com/files*, server video *speedy.site88.net/files* dan server lain dengan nama domain *aliac.net46.net/files*.



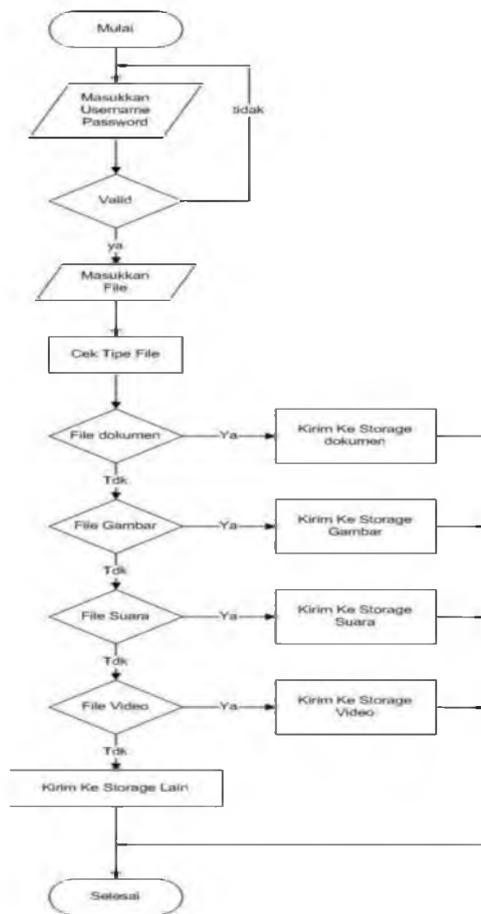
Gambar 1 Perancangan Arsitektur Sistem

Pengiriman, pengelompokan dan pengambilan *file* yang berasal dari computer pengguna dilakukan melalui FTP *server*. Sehingga *file* yang akan di-*upload* dan di-*download* berada pada FTP *server*. Berdasarkan pada pengidentifikasiian sistem, bahwa dalam system ini hanya memiliki satu aktor atau yang memiliki *username* sendiri. *File* yang akan di-*download* melalui komputer pengguna terlebih dahulu harus dikirim ke FTP *server* agar dapat langsung di-*download* oleh *client* di manapun dan kapanpun selagi terkoneksi dengan jaringan *internet*. Dari ilustrasi tersebut dapat diketahui obyek-obyek yang terlibat beserta atribut dan operasi setiap obyeknya, obyek-obyek ini kemudian dibangun dalam diagram *high-level-class*, seperti pada gambar 2 berikut:



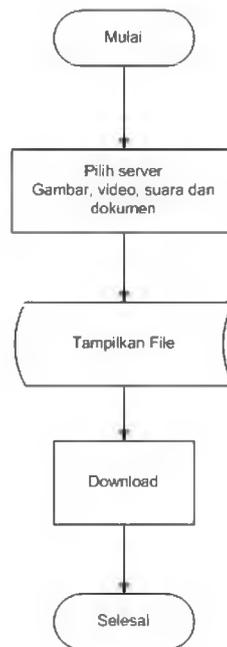
Gambar 2 Diagram *High-level-use-case*

2.2.1 Flowchart System



Gambar 3 *Flowchart System Upload File*

Dari gambar 3 *Flowchart System Upload File* menjelaskan bahwa, pertama *admin* melakukan proses *login* setelah berhasil *login* seorang *admin* melakukan *upload file*, dan kemudian system akan melakukan pengecekan pada *file* yang di-*upload* jika *file* yang di-*upload* merupakan tipe dari *file* gambar, video, dokumen, dan suara, maka *file* tersebut akan menuju *server* masing-masing, dan jika bukan dari keempat dari tipe *file* tersebut maka *file* akan berada pada *server* lain.



Gambar 4 *Flowchart System Download File*

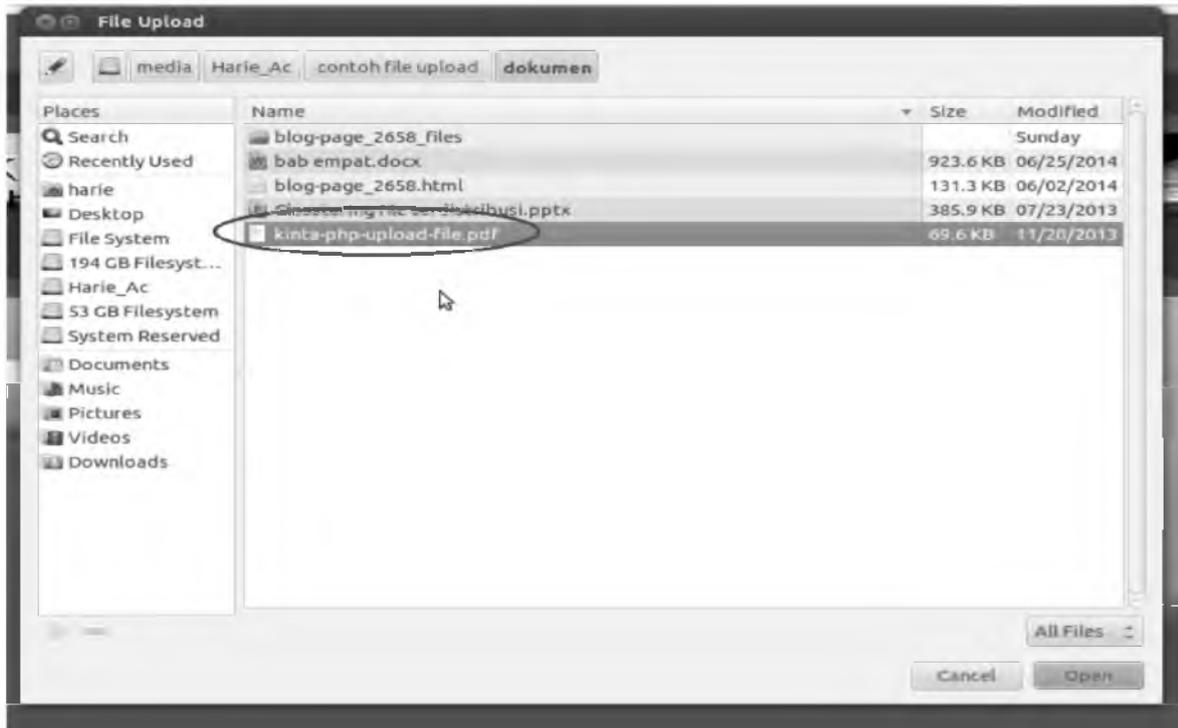
Dari gambar 4 *Flowchart System Upload File* menjelaskan *client* yang akan *men-download file*, akan memilih terlebih dahulu *server* dari *file* yang akan *di-download*. Setelah memilih salah satu dari *server* maka *file* dari *server* tersebut bisa untuk *di-download*.

3. Hasil Dan Pembahasan

Uji coba sistem ini dilakukan cukup mengandalkan jaringan internet dimanapun dan kapanpun selagi terkoneksi dengan internet baik menggunakan *modem* untuk akses *internet* maupun jaringan WIFI. Uji coba yang dilakukan yaitu meng-*upload* sejumlah *file* untuk memastikan *file* berada pada *server* *shuster*-nya masing-masing, serta *download file* yang telah di-*upload* pada server-nya masing-masing.

3.1 Uji Coba Upload File

Upload file yang dilakukan yaitu dengan menggunakan komputer moderator karena komputer moderator yang hanya memiliki andil untuk melakukan *upload file*.



Gambar 5 upload file

Gambar 5 merupakan contoh *upload* dari salah satu tipe *file* dari dokumen yang dilakukan melalui komputer server moderator yang terhubung dengan jaringan *internet*.

INFO_FILE

Table **Data** Indexes Model Constraints Grants Statistics UI Defaults Triggers Dependencies SQL

Query Count Rows Insert Row

EDIT	NAMA_FILE	OWNER_NAME	SERVER_NAME	TGL_UPLOAD	STATUS_FILE	FILE TYPE
	bab empat.docx	admin	radi.net76.net	2014-07-05 18:37:02	private	application/vnd.openxmlformats-officedocument.wordprocessingml.document
	Clasering file terdistribusi.pptx	admin	radi.net76.net	2014-07-05 18:37:35	private	application/vnd.openxmlformats-officedocument.presentationml.presentation
	kinta-php-upload-file.pdf	admin	radi.net76.net	2014-07-05 18:38:11	private	application/pdf
	IMG_198418948193440.jpeg	admin	tntn.site50.net	2014-07-05 18:27:12	private	image/jpeg

[Download](#)

Gambar 6 Hasil Record Pada Database

Gambar 6 merupakan hasil *record* di dalam *database file* berada pada lingkaran yang berwarna merah merupakan hasil *upload* dari komputer moderator yang telah dilakukan pada gambar 5 sebelumnya.

3.2 Uji Coba Download File

Proses *download* bisa dilakukan oleh *client* tanpa harus memiliki hak akses untuk melakukan *download file* tersebut. *Client* cukup mengetahui domain yang sudah ditentukan oleh komputer moderator atau pembuat sistem untuk melakukan *download* cukup mengandalkan jaringan *internet* dan mengakses nama *domain* dari *server* yang diinginkan maka *download* bisa dilakukan. Berikut contoh *download file* yang sudah di-*upload* oleh komputer moderator sebelumnya.



3.3 Pembahasan

Untuk mendapatkan hasil perbandingan kecepatan pada saat *download file* maka dilakukan uji fungsi antara *server file* yang di dalamnya sudah di *cluster* dengan *server file* yang tidak di *cluster*. Berikut dipaparkan dalam bentuk table perbandingan.

Tabel 1 Perbandingan *Download file*

NO	Namaserver	Namafile	Size	Kecepatan rata-rata	Waktuselesai
1	Campur	Bab 1.vv.docx	5.3 MB	10.000 KB/sec	9 min 4sec
2	Dokumen	Bab 1.vv.docx	5.3 MB	10.000 KB/sec	8 min 4sec

Pada table 1 hasil pengujian yang diperoleh untuk perbandingan antara kedua *server* campur dan *server cluster* maka terdapat selisih waktunya itu berkisar 1 *min* 3 *sec* dengan kecepatan rata-rata 10.000 KB/sec.

Dari hasil semua percobaan dapat dianalisa dari implementasi dan simulasi percobaan adalah tentang proses mekanisme kerja dari sistem secara keseluruhan. Adapun hasil dari percobaan pada sistem yang telah dilakukan sebelumnya dapat dilihat pada tabel2 berikut:

Tabel 2 Hasil Uji Sistem

No	Nama File	Tipe File	Hasil Upload	Keterangan
1	Clusstering.pptx	Dokumen	Radi.net76.net	Berhasil
2	Kampus.tif	Gambar	Tintin.site50.net	Berhasil
3	Destop.fly	Video	Speedy.site88.net	Berhasil
4	Aline.amr	Suara	Dudeee.site90.com	Berhasil
5	Dsc01231.pcx	Gambar	Tintin.site50.net	Berhasil
6	Drawing8.vsd	Dokumen	Radi.net76.net	Berhasil
7	Movie.wmv	Video	Speedy.site88.net	Berhasil

8	mapping.dotx	Dokumen	<i>Radi.net76.net</i>	Berhasil
9	panorama.gif	Gambar	<i>Tintin.site50.net</i>	Berhasil
10	BondanPrakoso -	Suara	<i>Dudeee.site90.com</i>	Berhasil
11	film.avi	Video	<i>Speedy.site88.net</i>	Berhasil
12	Wajah.jpg	Gambar	<i>Tintin.site50.net</i>	Berhasil
13	Hiburan.mpg	Video	<i>Speedy.site88.net</i>	Berhasil
14	Stta.bmp	Gambar	<i>Tintin.site.50.net</i>	Berhasil
15	Ben E King - Stand By Me.mmf	Suara	<i>Dudee.site90.com</i>	Berhasil
16	Ultah.vob	Video	<i>Speedy.site88.net</i>	Berhasil
17	Bab empat.dock	Dokumen	<i>Radi.net76.net</i>	Berhasil
18	Dua.tiga	Gambar	<i>Tintin.site50.net</i>	Berhasil
19	Keramaian.MOV	Video	<i>Speedy.site88.net</i>	Berhasil
20	14 where.m4a	Suara	<i>Dudeee.site90.com</i>	Berhasil
21	Camp.mp4	Vidio	<i>Speedy.site88.net</i>	Berhasil
22	Gambar.png	Gambar	<i>Tintin.site50.net</i>	Berhasil
23	Ewel tugas.rtf	Dokumen	<i>Radi.net76.net</i>	Berhasil
24	Laguislami	Suara	<i>Dudeee.site90.com</i>	Berhasil
25	Konser.3gp	Video	<i>Speedy.site88.net</i>	Berhasil
26	Data.xlsx	Dokumen	<i>Radi.net76.net</i>	Berhasil
27	BARA-.MP3	Suara	<i>Dudeee.site90.com</i>	Berhasil
28	Harie.txt	Dokumen	<i>Radi.net76.net</i>	Berhasil
29	Lagumengejar mimpi.m3u	Suara	<i>Dudeee.site90.com</i>	Berhasil
30	Kinta-php.pdf	Dokumen	<i>Radi.net76.net</i>	Berhasil
31	FIRMAN ~.wma	Suara	<i>Dudeee.site90.com</i>	Berhasil
32	7crack.rar	Lain	<i>Aliac.net46.net</i>	Berhasil

4. Penutup

4.1 Kesimpulan

Dari hasil uji coba yang telah dilakukan dapat ditarik beberapa kesimpulan:

1. Sistem *clustering file* terdistribusi telah berhasil digunakan untuk *upload*, dimana *file* terlebih dahulu di-*filter* oleh sistem untuk dikelompokkan berdasarkan tipe *file*-nya masing-masing.
2. Jumlah *server* yang digunakan adalah satu yang bertindak sebagai *server* moderator dan lima *slave server* yang masing-masing bertindak sebagai *server* dari gambar, suara, dokumen, video dan lain.
3. *File* yang telah berada di dalam *server* bisa diakses *secara real time* oleh *client* kapanpun dan di manapun selagi terhubung dengan jaringan *internet*.
4. Dari hasil uji fungsi yang dilakukan bahwa antara server campur dan yang di *cluster* maka terdapat selisih waktu *download* yang lebih diunggulkan oleh server *cluster* dengan menggunakan kecepatan yang sama dan kecepatan dari internet yang digunakan mempengaruhi hasil akhir *download*.

4.2 Saran

Adapun beberapa saran diperoleh dari tugas akhir ini adalah :

1. Tampilan dari sistem *Clustering File* terdistribusi dirancang lebih apik dan lebih menarik lagi.
2. Pada sisi *server* menggunakan ruang kapasitas yang lebih besar jika akan diterapkan pada sebuah kantor maupun sebuah perusahaan yang berskala besar.

3. Menambahkan auto *rename* pada sistem yang akan di-*upload* jika pada *server* sudah ada *file* yang sama.
4. Menambahkan *server* untuk *file* yang bertipe lainnya jika dibutuhkan.

Daftar Pustaka

- Kadir, Abdul. 2009. *Membuat Aplikasi Web Dengan PHP + Database MySQL*. Yogyakarta. Andi Offset.
- Lysy C, Moleong. 2013. *Implementasi Cluster Computing Untuk Render Animasi Jurusan Teknik Elektro-FT. UNSRAT*, Manado.
- Munandar. Ai Tb. 2013. *Clustering Data Nilai Mahasiswa Untuk Pengelompokan Kosentrasi Jurusan Menggunakan Fuzzy Cluster Mean*. Banten. Jurnal Teknik Informatika Serang Raya Banten.
- Samsiar Evara. 2008. *Belajar Sendiri Administrasi Database Oracle 10g*. Jakarta. Elex Media Komputido.
- Sutedjo, Budi. 2003. *Konsep & Aplikasi Client Server dan Sistem Terdistribusi*. Yogyakarta. Andi Offset.
- Syafrizal, Melwin. 2005. *Pengantar Jaringan Komputer*. Yogyakarta. Andi Offset.
- Ursula Maier dan Georg Stellner. 2008. *Distributed Resource*

