BAB II
KAJIAN LITERATUR

2.1 Tinjauan Pustaka

Pada bab ini penulis akan menguraikan beberapa dasar teori yang mendukung dalam Visualisasi Log Server Web menggunakan ELK Stack – Studi Kasus Log Akses Elena STT-NF, Berikut akan dijelaskan secara lebih lengkap.

2.1.1 E-Learning STT Terpadu Nurul Fikri (Elena)

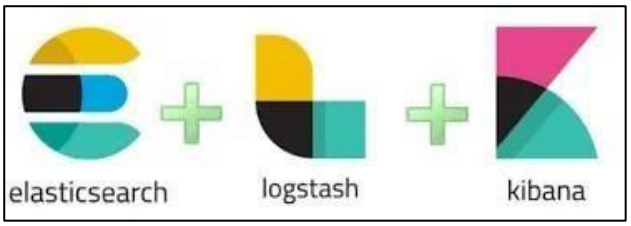
Elena STT-NF adalah media pembelajaran online yang digunakan oleh civitas akademik lembaga pendidikan tinggi STT-NF. Terdapat informasi spesifikasi lainnya mengenai sistem Elena STT-NF, yang dapat dilihat pada tabel berikut :

Table 1.1 Informasi Server Web Sistem Elena STT-NF

Specification	Tool Name & Version
Web Server	Apache versi 2.4.53
Database	Mysql versi 8.0.28
Cloud Storage	Mongodb Atlas

Sistem Elena STT-NF di kembangkan dengan teknologi Moodle. Selain itu, sistem Elena menggunakan database dan server *cloud* sebagai media penyimpanan serta web server. Berdasarkan tabel diatas dapat diketahui secara jelas spesifikasi dari server web sistem elena mulai dari informasi jenis web server, *database* dan *cloud* yang digunakan pada sistem tersebut.

2.1.2 ELK Stack



Gambar 2.1 Diagram ELK Stack (Sumber: www.medium.com)

ELK stack adalah serangkaian *open source project* yang dapat digunakan secara bersamaan. ELK Stack dirancang untuk digunakan sebagai solusi terintegrasi [5]. ELK stack dapat digunakan untuk menganalisis log di lingkungan IT. Komponen ELK terdiri dari Elasticsearch, Logstash, dan Kibana.

2.1.2.1 Elasticsearch

Elasticsearch merupakan aplikasi *open source* yang berfungsi sebagai alat pencarian, penyimpanan dan dapat [8]. Elasticsearch dikembangkan oleh Shay Banon dan dipublikasikan tahun 2010. Elasticsearch juga *real time distributed* dimana dapat menyimpan data dalam beberapa node terpisah serta data yang diinputkan akan langsung masuk ke dalam proses analitiknya. Data yang tersimpan pada Elasticsearch akan disimpan dalam bentuk field data yang mana akan digunakan oleh kibana untuk menampilkan visualisasi. Selain itu, elasticsearch digunakan untuk menyimpan data baik yang terstruktur maupun yang tidak terstruktur dan nantinya saat terindeks, data akan tersaring berdasarkan kategori field datanya. Hal ini karena elasticsearch menyimpan data dengan NoSQL. Elasticsearch telah digunakan oleh beberapa organisasi besar seperti *Stackoveflow* dan Github.

2.1.2.2 Logstash

Logstash adalah sebuah mesin untuk koleksi data berbasis *open-source* dengan kemampuan *pipelining* secara *realtime* yang berfungsi untuk mengumpulkan dan memparsing log dan juga membuat indeks untuk log yang kemudian disimpan pada elasticsearch. Logstash dikembangkan dari satu pengembang yang sama dengan elasticsearch dan kibana, yaitu Elastic.Co.

Berdasarkan penelitian Arifin tentang implementasi logstash untuk menarik data log yang ada pada *web server* untuk membantu sistem administrator dalam memantau kinerja server [7]. Logstash digunakan untuk menarik data log pada web server yang kemudian diuraikan dan diteruskan ke elasticsearch. Setelah itu data ini digunakan untuk menampilkan data, seperti data user yang gagal melakukan login, data user yang dapat melakukan login, dan persentase keseluruhannya. Penggunaan logstash untuk penarikan data log *web server* ini mampu membuat kegiatan penyimpanan dan pemantauan kinerja log *web server* menjadi lebih efektif dan efisien.

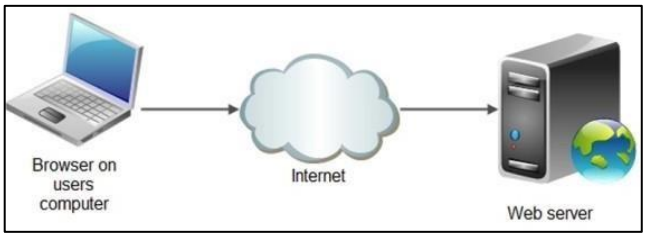
2.1.2.3 Kibana

Kibana merupakan aplikasi *open source* untuk visualisasi dan analisis datayang dirancang agar dapat digunakan bersama dengan Elasticsearch. Kibana dapatmengolah data dalam jumlah yang besar agar mudah dianalisis dan dipahami menggunakan fitur penyajian data bentuk diagram, tabel dan fitur lainnya. Diagram yang terdapat pada kibana dapat di customisasi sesuai dengan kebutuhan visualisasi data pengguna. Visualisasi yang dibuat pada kibana dapat dimasukkan ke dalam dashboard visualisasi untuk penyajian data yg lebih efektif, Penggunaan fitur *browser-based* pada dashboard kibana, memudahkan dalam visualisasi secara cepat dengan menyesuaikan secara *realtime* dari *query* Elasticsearch [7].

2.1.3 Rsync (*Remote Sync*)

Rsync adalah program sinkronisasi *file* dan transfer data yang digunakan untuk menyalin dan sinkronisasi *file* antara mesin lokal dan server secara *remote* melalui jaringan. Sinkronisasi dilakukan dengan cara membandingkan *file* dan direktori yang akan diunduh dan memperbarui *file* yang sudah ada dan menyalin *file* yang tidak ada di lokasi tujuan. *Rsync* juga dapat bekerja dengan protokol *file* remote seperti SSH dan mendukung pengiriman data melalui koneksi enkripsi dengan menggunakan SSH [6]. Penggunaan r*sync* akan membutuhkan otentikasi dan keamanan seperti yang digunakan dalam protokol *Secure Shell* (SSH). Otentikasi pada *rsync* menggunakan SSHpass yaitu program memberikan untuk masuk ke server tanpa memerlukan interaksi dari pengguna(tanpa password).

2.1.4 Web Server



Gambar 2.2 Diagram Web Server. (Sumber: www.classnotes.ng)

Web Server adalah tempat untuk mendapatkan halaman *website* dan data yang berhubungan dengan *website* yang dibuat, sehingga data dapat diakses dan dilihat oleh pengguna. Terdapat beberapa *web server* yang banyak digunakan yaitu *web server* Apache.

2.1.4.1 Apache

Apache adalah salah satu jenis web server yang paling banyakdigunakandi dunia berdasarkan survei yang dilakukan dari *Web Technology* yakni dengan penggunaan persentase sebesar 44,2% [12]. Apache merupakan perangkat lunak berbasis *open source* yang dikelola dan dikembangkan oleh komunitas terbuka dibawah naungan *Apache Software Foundation*. Apache *web server* memproses dan mendistribusikan halaman *website* sesuai dengan permintaan, serta melayani aset dan konten *website* melalui HTTP. Terdapat beberapa tipe darilog file sepertipada tabel berikut :

Table 2.2 Tipe Log Server

Tipe log	Fungsi	Format	Informasi Log
Accesslog	Merekan semua *request* dari pengguna untuk di proses oleh server dan memuat informasi *user*	*[Wed Oct 11 14:32:52 2000] [error] [Client 127.0.0.1] client denied by server configuration: /export/home/live/ap/ htdocs/test*	• Profil pengguna • Pola yang sering • muncul • Bandwidth Penggunaan

Table 2.2 Tipe Log Server

Error log	Daftar *error* yang berisi *request* oleh pengguna yang dibuat oleh server	*[Wed Oct 11 14:32:52 2000] [error] [Client 127.0.0.1] client denied by server configuration: /export/home/live/ap/ htdocs/test*	• Jenis *error* • *Errors* IP • *address.* • Tanggal dan • waktu dari *error* yang muncul
Agent log	Merekam *browser* dan versi *browser* yang dipakai Pengguna	*"Mozilla/4.0 (compatible; MSIE 4.01; Windows NT)"*	• Versi *Agent* • Sistem Operasi • yang digunakan • *Browser* yang Digunakan
ReferrerLog	Merekam informasi tentang *link* dan *redirect*s dari sebuah situs yang dikunjungi *User*	*"http://www.google.com/ search?q= keyword", "/page.html"*	• *Browser* yang digunakan • Kata kunci *Redirect link*

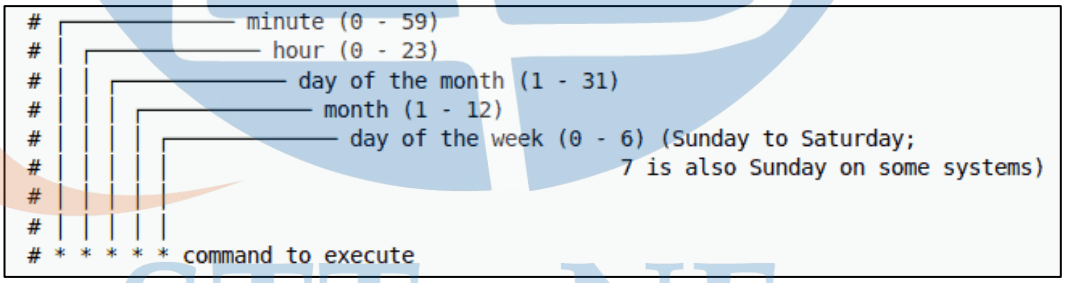
Access log Apache web server bertugas untuk mencatat semua proses permintaan yang terjadi di server dan menyimpan informasi pengunjung, seperti alamat IP, halaman yang diakses pengguna, informasi browser, dan waktu akses. Pada linux ubuntu, access log terdapat pada direktori */var/log/apache2/access.log*. Apache pertama kali dikembangkan untuk bekerja dengan sistem operasi Linux/Unix, namun kemudian diadaptasi untuk bekerja di sistem lain, termasuk Windows dan Mac [13].

2.1.5 Shell Script

Shell script adalah serangkaian perintah atau kode yang dieksekusi oleh*shell* di sistem operasi dan umumnya ditulis dengan menggunakan bahasa *shell*, seperti Bash [6]. *Shell script* berupa file yang berisi serangkaian perintah atau daftar tugas yang dieksekusi secara manual pada sistem operasi, seperti menjalankan *command line* linux secara berurutan, mengambil *input* dari pengguna, membuat dan mengelola *file*, dan lainnya. *shell script* juga dapat digunakan sebagai alat transfer *file* yang digunakan dalam skenario tertentuseperti membuat *backup* data dari server ke komputer lokal dan lainnya.

2.1.6 Cronjob

Cronjob adalah metode untuk menjalankan sebuah tugas berulang dengan penjadwalan [6]. Tugas dijadwalkan di crontab *file* yaitu *file* yang berisi perintah tugas yang akan dijalankan, misalnya menjadwalkan baris perintah pada *file shell script* untuk melakukan *backup* data dari server dan kemudian dieksekusi oleh cronjob pada waktu tertentu yang telah ditentukan. Waktu yang dijalankan pada penjadwalan bisa diatur sesuai dengan kebutuhan tugas seperti dijadwalkan setiap beberapa menit, jam, hari atau tanggal. Berikut adalah sintaks penulisan tugas untuk mengatur penjadwalan tugas di *file* crontab.



Gambar 2.3 Sintaks Cronjob (Sumber:www.medium.com)

Pada gambar diatas terdapat 5 field waktu yang dapat di *custom* penulisannya, berikut penjelasan secara rinci :

1. *Minute* adalah waktu dalam menit yang ditunjukan saat perintah tugas di jalankan yaitu rentang 0 – 59 menit.
2. *Hour* adalah waktu dalam jam yang ditunjukan saat perintah tugas dijalankan yaitu rentang 0-23 jam.
3. *Day of the moth* adalah hari dalam suatu bulan yang dapat diatur oleh user saat perintah tugas dijalankan dengan rentang 1 - 31 hari.

4. *Month* adalah bulan yang dapat diatur oleh user pada saat perintah tugas dijalankan yaitu dengan rentang 1-12 untuk bulan januari sampai desember.
5. *Day of the week* adalah hari dalam satu minggu yang dapat diatur oleh user daat perintah tugas dijalankan dengan rentang 0-6 untuk satu minggu yaitu minggu sampai sabtu.

2.1.7 Penelitian Terkait

Pada penelitian ini, penulis melakukan perbandingan terhadap beberapa penelitian terkait yang berkaitan dengan masalah yang penulis ambil. Pada dasarnya penelitian terkait memiliki kesamaan dalam pandangan mengenai visualisasi log server web menggunakan ELK Stack, namun menggunakan studi kasus yang berbeda-beda. Berikut adalah tabel perbandingan beberapa penelitian terkait yang peneliti gunakan :

2.1.7.1 Tabel Penelitian Terkait

Berikut adalah tabel yang membahas penelitian terkait dengan jumlah empat penelitian yang peneliti ambil untuk dijadikan referensi:

Table 2.3 Penelitian Terkait

No	Judul Penelitian	Peneliti	Topik	Alat	Kesimpulan
1.	Sistem Pengawasan Kinerja Jaringan Server Web Apache dengan Log Management System ELK (Elasticsearch, Logstash, Kibana)	Claudia Tarigan, Ventje Jeremias Lewi Engel, Dina Angela. 2018 Institut Teknologi Harapan Bangsa. Bandung.	Monitoring Log	ELK Stack, Apache Web Server (akses log dan error log).	Pada jurnal ini data log diolah dan disimpan menggunakan ELK. Fokus pembahasan lebih kepada pengujian kinerja log server apache dibandingkan dengan visualisasi log.
2.	*Visualizing Web Server Logs Insights with Elastic Stack – a Case Study of UMMAIL'S Access Logs*	Harni Yusnidar Muhammad dan Jasni Mohammad Zain. 2018 UiTM Shah Alam, Selangor, Malaysia.	Visualisasi Log	ELK Stack, Nginx Web Server, Virtual Machine, AWS Cloud.	Pada jurnal ini, data log tersebut diolah dan disimpan menggunakan ELK serta menggunakan OS virtual machine dan AWS Cloud. Pada jurnal ini lebih banyak menjelaskan visualisasi dan analitik log.

Table 2.3 Tabel Penelitian Terkait

3.	Penerapan Log Analyzer untuk Mengetahui Lalu Lintas Jaringan Berbasis Elasticsearch, Logstash dan Kibana	Putra, MuhammadJavier Rama. 2021 Sekolah Tinggi Teknologi Terpadu Nurul Fikri.	Analisis Logs	Web Log *Analyzer*, *Traffic network*log, ELK Stack	Pada jurnal ini, data log yang digunakan ada 3 yaitu apache log, *syslog* dan *network traffic* dan Data log tersebutdiolah menggunakan ELK Stack dan fokus pembahasan analisa pemrosesandengan log berbeda dan tidak secara rinci membahas visualisasi
4.	Implementasi Elasticsearch Logstash Kibana Stack pada Sistem Portal Pengembangan dan Pembinaan Sumber DayaManusia	Elang Putra Sartika, Andhik Budi Cahyono, 2020 Universitas Islam Indonesia Yogyakarta, Indonesia.	Visualisasi Log	ELK Stack, Apache Web Server (akses log dan error log).	Pada jurnal ini, data log yang digunakan hanya data log E-Learning (Portal PPSDM) dan data log diolah dan divisualisasikan menggunakan ELK (, Elasticsearch, Logstash, Kibana) pada Linux ubuntu. Pengujian dilakukan dengan cara API ELK Stack untuk mencari sebuah data dan kibana untuk membuat visualisasi yang diperlukan untuk melihat data log.

Table 2.3 Penelitian Terkait

5.	Visualisasi Log Server Web menggunakan ELK STACK - studi kasus Log Akses Sistem Elen STT Terpadu Nurul Fikri	Hera Karmila, 2022 STT Terpadu Nurul Fikri, 2021.	Visualisasi Log	ELK Stack, Apache Web Server (akses log sistem Elena STT-NF).	Pada Penelitian ini data akses log sistem Elena akan diolah dan disimpan menggunakan elasticsearch dan logstash, kemudian divisualisasikan dengan kibana. Fokus peneliti akan lebih membahas visualisasi log.

Pada penelitian pertama, penelitian hanya berfokus pada pengawasan kinerja jaringan dari server apache dan implementasi manajemen log menggunakan *tools* yang sama yaitu ELK Stack. Kinerja dari server pada penelitian ini lebih ditekankan pengawasannya dari pada visualisasi log yakni menggunakan indikator - indikator performa tertentu seperti waktu, *utilization* dan *error* untuk memonitoring kinerja server apache yang bekerja.

Selanjutnya pada penelitian kedua, lebih membahas visualisasi akses log dengan studi kasus akses log yang berbeda dengan studi kasus dari penelitian yang penulis lakukan. Log yang dipakai untuk divisualisasikan pada penelitian tersebut adalah akses log dengan studi kasus website resmi kampus UITM Shah Alam sedangkan pada penelitian yang penulis lakukan menggunakan studi kasus log akses Elena STT-NF.

Selanjutnya pada penelitian ketiga, terdapat tiga masukkan data log yang berbeda yaitu apache log, *syslog* dan *network traffic*, yang mana ketiga log tersebut nantinya diproses dengan menggunakan ELK Stack. Pada penelitian ini tidak membahas visualisasi dari log yang diproses secara menyeluruh dan terperinci. Visualisasi yang dilakukan pada penelitian terkait tersebut hanya memvisualkan data yang sudah terindeks di elasticsearch dan memeriksa apakah data tersebut dapatdivisualisasikan di kibana. Jadi fokus penelitian pada penelitian ini lebih membahas pemrosesan ketiga log tersebut sampai ke ELK Stack dan bagaimana tahapannya.

Selanjutnya pada penelitian keempat, penelitian ini berfokus pada pemrosesan data log dari sistem *E-Learning* PPSDM dengan mengunakan metode dan alat pemrosesan data log yang berbeda serta tidak berfokus pada visualisasi log. Pada penelitian ini, penulis akan memilih tema yang berkaitan namun dengan studi kasus yang berbeda dengan judul "Visualisasi Log Server Web menggunakan ELK Stack - Studi Kasus Log Akses Sistem Elena STT Terpadu Nurul Fikri".

2.1.7.2 Posisi Penelitian

Tabel berikut bertujuan untuk menentukan posisi pada penelitian ini dari penelitian terkait yang sebelumnya telah dijelaskan:

Table 2.2 Posisi Penelitian

No.	Penelitian	Visualisasi Log	Akses Log	Apache Log	ELK	Pengem-bangan	Pengujian Fungsional
1.	Claudia Tarigan, Ventje Jeremias Lewi Engel, Dina Angela, 2018 Sistem Pengawasan Kinerja Jaringan Server Web Apache denganLog Management System ELK (Elasticsearch, Logstash, Kibana)						
2.	Muhammad Harni Yusnidar dan Jasni Mohamad Zain, 2018 Visualizing Web Server Logs Insights with Elastic Stack –a CaseStudy of UMMAIL'S Access Logs						
3.	Putra, Muhammad Javier Rama. 2021 Penerapan Log Analyzer untuk Mengetahui Lalu Lintas JaringanBerbasis Elasticsearch, Logstash dan Kibana						
4.	Elang Putra Sartika, Andhik Budi Cahyono, 2020 Implementasi Elasticsearch Logstash Kibana Stack pada SistemPortal Pengembangan dan Pembinaan Sumber Daya Manusia						

Table 2.4 Tabel Posisi Penelitian

5.	Hera Karmila, 2021 Visualisasi Log Server Web Menggunakan ELK (Elasticsearch, Logstash, Kibana) Stack - Studi Kasus Log Akses Elena STT NF						

Pada tabel posisi penelitian diatas penulis menggunakan enam parameter perbandingan berdasarkan alat dan bahan serta metode yang digunakan pada masing-masing penelitian untuk menentukan posisi penelitian penulis terhadap penelitian terkait. Keenam parameter tersebut adalah Visualisasi Log, Akses Log, Apache Log, ELK, Metode Pengembangan serta Pengujian Fungsional. Parameter-parameter tersebut dibuat untuk mengetahui lebih rinci mengenai perbandingan posisi penelitian yang dilakukan penulis dengan beberapa penelitian terkait yang menjadi rujukan pada topik penelitian yang penulis lakukan, serta untuk mengetahui perbedaan alat, bahan dan metode apa saja yang dipakai pada tiap - tiap penelitian.

Pada penelitian pertama, pembahasan penelitian mencakup empat parameter, yaitu Visualisasi Log, Akses Log, Apache Log, ELK, Metode Pengembangan sedangkan untuk pengujian tidak menggunakan metode pengujian fungsional/ keefektifan visualisasi log, namun menggunakan pengujian penggunaan *space memory*, RAM, CPU dan lainnya. Sehingga fokus pengujian yang dilakukan pada penelitian pertama adalah menguji performa *hardware* pada saat sistem ELK di implementasikan dalam penelitian tersebut.

Selanjutnya pada penelitian kedua, pembahasan mencakup empat parameter yaitu lebih membahas Visualisasi Log, Akses log, ELK dan Metode Pengembangan. Pada penelitian kedua ini tidak menggunakan inputan berupa Apache Log namun menggunakan inputan dari log Nginx dan tidak melakukan pengujian fungsionalitas.

Selanjutnya pada penelitian ketiga, pembahasan mencakup lima parameter yaitu lebih membahas Visualisasi Log, Apache log, ELK, Metode Pengembangan dan Pengujian Fungsional. Pada penelitian ketiga tidak menggunakan inputan Akses log, namun hanya inputan file log yang berasal dari komputer lokal saat implementasi.

Selanjutnya pada penelitian keempat, pembahasan mencakup empat parameter yaitu lebih membahas Visualisasi Log, ELK dan Metode Pengembangan. Data log yang digunakan berupa data log dri Apache log dan pengujian yang dilakukan adalah pengujian *hardware* terhadap implementasi sistem. Pada penelitian ini, penulis mengggunakan ke-enam parameter tersebut sehingga berbeda dengan penelitan terkait yang sudah dijelaskan.