IMPLEMENTASI *SPLUNK* DALAM MEMBANGUN *SECURITY INFORMATION AND EVENT MANAGEMENT* BERDASARKAN *LOG FIREWALL TRAFFIC TYPE* (STUDI KASUS: JARINGAN UII)

Disusun Oleh:

N a m a : Wahluf Abidian

NIM : 17523141

PROGRAM STUDI INFORMATIKA – PROGRAM SARJANA

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS ISLAM INDONESIA

2021

HALAMAN PENGESAHAN DOSEN PEMBIMBING

IMPLEMENTASI *SPLUNK* DALAM MEMBANGUN *SECURITY INFORMATION AND EVENT MANAGEMENT* BERDASARKAN *LOG FIREWALL TRAFFIC TYPE* (STUDI KASUS: JARINGAN UII)

TUGAS AKHIR

Disusun Oleh:

N a m a	: Wahluf Abidian
NIM	: 17523141

Yogyakarta, 11 Januari 2021

Pembimbing,

(Mukhammad Andri Setiawan, S.T., M.Sc., Ph.d.)

HALAMAN PENGESAHAN DOSEN PENGUJI

IMPLEMENTASI *SPLUNK* DALAM MEMBANGUN *SECURITY INFORMATION AND EVENT MANAGEMENT* BERDASARKAN *LOG FIREWALL TRAFFIC TYPE* (STUDI KASUS: JARINGAN UII)

TUGAS AKHIR

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer dari Program Studi Teknik Informatika di Fakultas Teknologi Industri Universitas Islam Indonesia

Yogyakarta, 11 Januari 2021

Tim Penguji

Mukhammad A Setiawan, S.T., M.Sc., Ph.d. _____

Anggota 1

Yudi Prayudi, Dr., S.Si., M.Kom. _____

Anggota 2

Sheila Nurul Huda, S.Kom., M.Kom. _____

Mengetahui,
Ketua Program Studi Informatika – Program Sarjana
Fakultas Teknologi Industri
Universitas Islam Indonesia

(Raden Teduh Dirgahayu, Dr., S.T., M.Sc.)

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Wahluf Abidian

NIM : 17523141

Tugas akhir dengan judul:

IMPLEMENTASI *SPLUNK* DALAM MEMBANGUN *SECURITY INFORMATION AND EVENT MANAGEMENT* BERDASARKAN *LOG FIREWALL TRAFFIC TYPE* (STUDI KASUS: JARINGAN UII)

Menyatakan bahwa seluruh komponen dan isi dalam tugas akhir ini adalah hasil karya saya sendiri. Apabila di kemudian hari terbukti ada beberapa bagian dari karya ini adalah bukan hasil karya sendiri, tugas akhir yang diajukan sebagai hasil karya sendiri ini siap ditarik kembali dan siap menanggung risiko dan konsekuensi apapun.

Demikian surat pernyataan ini dibuat, semoga dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 11 Januari 2021

(Wahluf Abidian)

HALAMAN PERSEMBAHAN

Penelitian ini dipersembahkan untuk Bapak, Ibu, teman-teman, dan semua orang yang masih terus ingin belajar.

HALAMAN MOTO

Jadilah Pembelajar sepanjang hayat (*lifelong learners*).

Coba dahulu, baru bercerita.
Pahami dahulu, baru menjawab.
Pikir dahulu, baru berkata.
Dengar dahulu, baru beri penilaian.
Berusaha dahulu, baru berharap.

- *Socrates*

KATA PENGANTAR

Alhamdulillahirabbil'alamin. Rasa syukur penulis sampaikan kepada Tuhan Yang Maha Esa yaitu Allah Subhanahu Wata'ala atas segala karunia-Nya sehingga penelitian berjudul "isi judul" dapat terlaksana dan tersusun dengan baik. Sholawat beserta salam semoga senantiasa tercurahkan kepada Nabi Muhammad Shallallahu Alahi wa Sallam.

Penelitian ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Strata 1 (S1) pada Program Studi Informatika, Fakultas Teknologi Industri, Universitas Islam Indonesia. Dalam proses penelitian penulis mendapatkan banyak tantangan dan rintangan. Namun pada saat mendapatkan kesulitan penulis juga mendapatkan bimbingan, arahan, bantuan serta dukungan dari berbagai pihak. Oleh karena itu, penulis memberikan penghargaan dan ucapan terimakasih kepada:

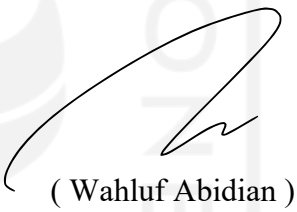
1. Hendrik, S.T., M.Eng. selaku ketua Jurusan Informatika Fakultas Teknologi Industri Universitas Islam Indonesia, Dr. Raden Teduh Dirgahayu, S.T., M.Sc. selaku Ketua Program Studi Informatika - Program Sarjana Fakultas Teknologi Industri Universitas Islam Indonesia, dan Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D. selaku Sekretaris Program Studi Informatika - Program Sarjana Fakultas Teknologi Industri Universitas Islam Indonesia.
2. Mukhammad Andri Setiawan, S.T., M.Sc., Ph.D. selaku Pembimbing penelitian penulis yang disela padatnya kegiatan beliau telah dengan sabar, teliti, menginspirasi, dan ikhlas dalam memberikan bimbingan serta arahan dalam penelitian ini. Semoga Allah Subhanahu Wata'ala melimpahkan rahmat dan hidayah kepada beliau dan keluarga.
3. Kedua orangtua saya tercinta, yang selalu mendoakan, memberikan dukungan moral maupun material sehingga penulis dapat tegar dan bertahan dalam melakukan penelitian sehingga berjalan dengan lancar. Tanpa doa dari beliau penelitian ini tidak akan selesai sebagaimana saat ini.
4. Seluruh staff di Universitas Islam Indonesia yang telah berkenan dalam membantu mempermudah pengerjaan penelitian ini.
5. Teman-teman sekalian dari Program Studi Informatika – Program Sarjana khusunya angkatan 2017 yang telah berkenan sharing dan memberi dukungan dalam membantu saya mengerjakan penelitian ini.

6. Seluruh teman saya yang tidak dapat saya sebutkan namanya satu persatu yang selalu memberikan dukungan, keceriaanm dan berbagi pengalamannya agar selalu berpikir *positif thinking*.

Semoga bagi berbagai pihak yang telah memberikan doa, dukungan, bantuan, dan bimbingan segala urusannya dapat dipermudah oleh Allah Subhanahu Wata'ala. Penulis memohon maaf karena dalam penelitian ini masih terdapat banyak kekurangan. Semoga penelitian ini dapat bermanfaat bagi banyak pihak terutama yang membacanya dan dapat menjadi sumber ilmu untuk menambah wawasan kita semua. *Allahumma aaamiiin*

Wassalamu'alaikum Warohmatullahi Wabarokatuh

Yogyakarta, 11 Januari 2021



(Wahluf Abidian)

SARI

Badan Sistem Informasi Universitas Islam Indonesia saat ini sedang dalam tahap menerapkan teknologi *enterprise*. Pada bidang keamanan, UII telah menerapkan *Next-Generation Firewall* (NGFW) dengan data yang dihasilkan memiliki ukuran yang masuk dalam kategori *big data.* Data yang besar tentunya menyimpan informasi yang besar pula. Sebagai bentuk upaya meningkatkan pengamanan dan informasi pada jaringan UII. Penerapan teknologi SIEM digunakan untuk melakukan penggalian informasi data *Log* yang dihasilkan oleh NGFW. *Tools* yang digunakan merupakan *next-generation* dari SIEM yaitu Splunk. *Log Firewall* dengan *Traffic Type* diolah menjadi visualisasi non-*Cluster* dan *with Cluster* yang dikumpulkan pada sebuah *dashboard*. Diterapkan pula *alert* sistem yang diintegrasikan dengan *bot* Telegram. Tahapan dari penelitian meliputi studi literatur, perancangan, pengumpulan data, implementasi, pengujian, dan analisis. Dari hasil penelitian ditemukan bahwa visualisasi yang dibuat dapat memberikan kemudahaan pembacaan informasi maupun kemudahan mendapat informasi bagi administrator. Selain itu, potensi informasi yang didapat dari *Log Firewall* dapat menjadi acuan bagi pihak-pihak yang memiliki kepentingan dalam menentukan kebijakan.

Kata kunci: *Log*, *firewall*, SIEM, NGFW, *cluster, dashboard*, *alert*.

GLOSARIUM

Log	*File* data yang dihasilkan komputer dan berisikan informasi mengenai pola penggunaan, kegiatan, dan operasi dalam suatu sistem operasi, aplikasi, server, atau perangkat lain.
Firewall	Perangkat keamanan jaringan yang memantau lalu lintas jaringan masuk dan keluar lalu memutuskan apakah akan mengizinkan atau memblokir lalu lintas tertentu berdasarkan seperangkat aturan keamanan yang ditentukan.
Log Firewall	*File* data yang dihasilkan oleh *Firewall* dan berisikan informasi mengenai pola penggunaan, kegiatan, dan operasi.
NGFW	*Next-Generatin Firewall* (NGFW) merupakan perangkat keamanan jaringan yang memberikan kemampuan melebihi kemampuan dari *Firewall* biasa / tradisional.
SIEM	*Security Information and Event Management* (SIEM) merupakan sistem informasi terpusat dan digunakan untuk mengumpulkan *Log* yang nantinya memberikan hasil berupa visualisasi *Log monitoring* untuk mempermudah pembacaan informasi *Log.*
Cluster	Kumpulan dari data yang memiliki kemiripan satu sama lain.
Clustering	Proses memisahkan satu rangkaian objek data ke dalam *Cluster.*
Dashboard	Antarmuka pengguna yang terkait dengan aplikasi dan memiliki satu atau lebih panel visualisasi.
Search	Cara utama pengguna melakukan navigasi data di Splunk.
Alert	Peringatan yang berasal dari hasil *Search* yang disimpan.
Data Mining	Proses dalam menggunakan teknik metematik, statistik, kecerdasan buatan, dan *machine learning* dalam melakukan ekstraksi dan melakukan identifikasi yang bermanfaat serta pengetahuan yang terkait bersumber dari *Big Data* atau maha data.
Event	Sepotong data pada aplikasi Splunk, mirip dengan catatan dalam file *Log* atau masukkan data lainnya.
Field	Pasangan nama / nilai yang dapat dicari pada data *Event* aplikasi Splunk

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PENGESAHAN DOSEN PEMBIMBING ii
HALAMAN PENGESAHAN DOSEN PENGUJI iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR iv
HALAMAN PERSEMBAHAN v
HALAMAN MOTO vi
KATA PENGANTAR vii
SARI ix
GLOSARIUM x
DAFTAR ISI xi
DAFTAR TABEL xiii
DAFTAR GAMBAR xiv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 2
1.3 Tujuan Penelitian 3
1.4 Manfaat Penelitian 3
1.5 Batasan Masalah 3
1.6 Usulan Penyelesaian 3
1.7 Metode Penelitian 4
1.7.1 Studi Literatur 4
1.7.2 Perencanaan 4
1.7.3 Pengumpulan Data 4
1.7.4 Implementasi 5
1.7.5 Pengujian 5
1.7.6 Analisis 5
1.8 Sistematika Penulisan 5
BAB II LANDASAN TEORI 7
2.1 Landasan Teori 7
2.1.1 Keamanan *Cyber* 7
2.1.2 Keamanan Informasi (*Information Security*) 8
2.1.3 Ancaman Jaringan (*Network Threat)* 10

2.1.4 *IT Security Risk Management* 11
2.1.5 *Risk Assesment* 11
2.1.6 *Security Information and Event Management* (SIEM) 13
2.1.7 Splunk 14
2.1.8 *Next-Generation Firewall* (NGFW) 15
2.1.9 *Data Mining* 16
2.2 Penelitian Terkait 17
BAB III METODOLOGI PENELITIAN 20
3.1 Jenis Data 20
3.2 Ukuran Keberhasilan 20
3.3 Metode Penelitian 20
3.3.1 Studi Literatur 20
3.3.2 Perencanaan 21
3.3.3 Pengumpulan Data 21
3.3.4 Implementasi 21
3.3.5 Pengujian 29
3.3.6 Analisis 30
BAB IV HASIL DAN PEMBAHASAN 31
4.1 Visualisasi *Log Firewall non-Clustering* 31
4.2 Visualisasi *Log Firewall with Clustering* 38
4.3 Integrasi *Alert System* dengan *Bot* Telegram 44
4.4 Ananlisis Hasil 48
BAB V PENUTUP 50
5.1 Kesimpulan 50
5.2 Saran 50
DAFTAR PUSTAKA 52
LAMPIRAN 55

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait 17
Tabel 3.1 Kebutuhan Sistem 21
Tabel 4.1 Pengujian *Rule* visualisasi *non-cluster* 37
Tabel 4.2 Pembacaan *Cluster* 42
Tabel 4.3 Pengujian *Rule* visualisasi *with Cluster* 44

DAFTAR GAMBAR

Gambar 2.1 Keamanan Informasi (CIA) 9
Gambar 2.2 *Defense in Depth* 13
Gambar 2.3 *Security Information and Event Management* 14
Gambar 2.4 What is *Splunk?* 14
Gambar 2.5 Cara kerja Splunk 15
Gambar 2.6 *Next-Generation Firewall from Palo Alto Networks* 16
Gambar 2.7 Tahapan *Data Mining* 16
Gambar 3.1 Metode Penelitian 20
Gambar 3.2 Tahapan instalasi *Splunk Web* 22
Gambar 3.3 Aplikasi *Palo Alto Network add-on* untuk Splunk 23
Gambar 3.4 Aplikasi Splunk *Machine Learning Toolkit* 23
Gambar 3.5 Aplikasi *Telegram Alert Action* 24
Gambar 3.6 Sintaks untuk melakukan pemecahan data 24
Gambar 3.7 Pemilihan *Set Source Type* 25
Gambar 3.8 Tampilan *add data* berhasil 26
Gambar 3.9 Menu *Search & Reporting* 26
Gambar 3.10 *Search Language Syntax* 27
Gambar 3.11 Menyimpan hasil *Search & Reporting* ke *dashboard* panel 27
Gambar 3.12 Kegiatan *Clustering* dengan fitur *Smart Clustering* 28
Gambar 3.13 Mengintegrasikan *alert* dengan *bot* Telegram 29
Gambar 4.1 *Rule detection traffic by action* based app:technology. 31
Gambar 4.2 *Rule detection all user by user* 31
Gambar 4.3 *Rule detection Indonesia user from client_location=indonesia.* 31
Gambar 4.4 *Rule detection most popular user by user* 31
Gambar 4.5 *Rule detection most popular country by client_location.* 31
Gambar 4.6 *Rule detection user location by client_location* 32
Gambar 4.7 *Rule detection total users in country by client_location* 32
Gambar 4.8 *Rule detection networking & port scanning by src_ip* 32
Gambar 4.9 *Rule detection brute force attack by user, app* 32
Gambar 4.10 *Rule splunk web logins attemps* 32
Gambar 4.11 *Visualization traffic by Action* 33
Gambar 4.12 *Visualization all user by user* 33

Gambar 4.13 *Visualization user from indonesia by client_location=indonesia* 34
Gambar 4.14 *Visualization most popular user by user* 34
Gambar 4.15 *Visualization most popular country by client_location* 35
Gambar 4.16 *Visualization location User by client_location*..... 35
Gambar 4.17 *Visualization total users in country by client location* 36
Gambar 4.18 *Visualization total users in country by client location* 37
Gambar 4.19 *Elbow Method.* 39
Gambar 4.20 *Silhoutte Method.* 39
Gambar 4.21 *Rule Clustering with K-means algorithm for 3D Scatterplot*. 40
Gambar 4.22 *Visualization Clustering K=4 with 3D Scatterplot.*..... 41
Gambar 4.23 *Rule Clustering with K-means algorithm*. 41
Gambar 4.24 *Visualization Clustering K=4 with Downsampled Line Chart*..... 42
Gambar 4.23 *Rule* pemicu *Alert* bagian pertama..... 45
Gambar 4.24 Integrasi Splunk dengan *bot* Telegram 46
Gambar 4.25 *Alert* dari percobaan bagian pertama 46
Gambar 4.26 *Alert* dari percobaan bagian kedua 47

BAB I
PENDAHULUAN

1.1 Latar Belakang

Universitas Islam Indonesia merupakan salah satu perguruan tinggi yang berkomitmen meningkatkan teknologi informasi dengan mendirikan Badan Sistem Informasi atau disingkat BSI. Secara umum BSI bertugas mengawal perencanaan, pengembangan, operasi, serta layanan sistem dan teknologi informasi untuk kurang lebih 30.000 pengguna (BSI, 2017). Agar dapat memberikan layanan secara profesional, hendaknya BSI berfokus untuk menjamin *confidentiality* (kerahasiaan)*, integrity* (konsistensi, akurasi, dan validitas data)*, availability* (ketersediaan) atau CIA. Ketiga komponen tersebut merupakan dasar dari keamanan informasi (Azaim & Nuryanto, 2017).

Dalam menjaga kerahasiaan data, BSI telah mengimplementasikan *next-generation firewall* (NGFW) dari perusahaan Palo Alto Network. Teknologi *next-generation firewall* (NGFW) merupakan penerus dari *firewall* tradisional yang memiliki banyak kekurangan. Fitur pada *next-generation firewall* (NGFW) mencakup segala fitur yang terdapat pada *firewall* tradisional dan terdapat fitur tambahan seperti *Intrusion Prevention System (IPS), Deep Packet Inspection (DPI), Application Control, Directory Integration*, dan *Encrypted Traffic Inspection* (Cloudflare, 2020). Seiring dengan penambahan mahasiswa setiap tahunnya, jumlah pengguna jaringan UII akan terus mengalami peningkatan. Hal ini menyebabkan catatan *Log Firewall* akan semakin bertambah besar.

Untuk membantu administrator BSI UII dalam membaca *Log Firewall* dengan ukuran yang besar*,* secara umum telah dibantu oleh *dashboard* visualisasi yang diberikan oleh Palo Alto Network. Namun, masing-masing instansi memiliki kebutuhan khusus. Solusi yang dapat ditawarkan adalah dengan membuat *Log monitoring* system untuk mencapai kebutuhan secara khusus tersebut*. Security Information and Event Management* atau disingkat SIEM merupakan sistem informasi terpusat dan digunakan untuk mengumpulkan *Log* yang nantinya memberikan hasil berupa visualisasi *Log monitoring* untuk mempermudah pembacaan informasi *Log* (Perciballi, 2015).

Log yang dihasilkan oleh NGFW memiliki ukuran yang besar. Setiap harinya rata-rata ukuran *Log* yang dihasilkan berkisar antara 30 *gigabyte* hingga 50 *gigabyte*. Solusi untuk menangani jumlah data yang besar atau saat ini biasa disebut *Big Data* salah satunya dengan pendekatan *Data Mining* (Efrain et al., 2005). Data yang besar akan dikelompokkan atau

diklaster dengan hitungan secara matematika dan statistik sebelum data ditampilkan. Oleh karena itu, visualisasi *Log monitoring* lebih mudah dipahami (Erwinsyah, 2019).

Dalam menerapkan *Data Mining* salah satu teknik yang dapat digunakan adalah *machine learning.* Dibutuhkan algoritma sebagai dasar dari penerapan teknik *machine learning.* Pemilihan algoritma salah satunya harus didasarkan dengan *tools* yang digunakan dalam mengimplementasikan teknologi *Security Information and Event Management* (SIEM) (Chancey & Penn, 2018). Pada penelitian kali ini *tools* yang akan digunakan sebagai solusi yang ditawarkan adalah menggunakan aplikasi Splunk. Splunk merupakan *next-generation* SIEM *product.* Splunk mendukung penerapan *Data Mining* dengan teknik *machine learning* yang dibutuhkan untuk mengelola *Log Firewall* dalam melakukan klasterisasi agar visualisasi *Log Firewall* dalam jumlah data yang besar mudah dipahami. Splunk juga memiliki fitur penyaringan, pencarian, masukan, modifikasi, pelaporan, dan penghapusan data (Mock, et al., 2017).

Setelah data berhasil diolah dan ditampilkan pada *dashboard Log monitoring system*. Dibutuhkan sistem yang dapat menyiapkan *alert* berdasarkan *critical control* keamanan, agar informasi lebih mudah diketahui oleh administrator (Brown, 2016). Splunk memiliki fitur *alert real-time* yang dapat digunakan untuk membantu memantau *event* atau serangan yang terjadi. Dalam menerapkan *alert* sistem, sebaiknya pemberitahuan *alert* dapat diketahui kapan saja dan di mana saja sebab administrator tidak selalu berada di ruang kerja. Oleh karena itu, dibutuhkan integrasi antara *alert* sistem pada aplikasi Splunk dengan aplikasi yang dapat memberikan kemudahan penerimaan informasi. Telegram merupakan sebuah aplikasi yang dapat diakses pada *smartphone* ataupun perangkat komputer. Telegram memiliki fitur *bot* yang dapat diintegrasikan dengan Splunk untuk membantu kemudahan menerima informasi saat *alert* muncul. Penggunaan *bot* Telegram juga telah diadopsi oleh team BSI UII untuk melakukan pemantauan suhu *server*. Oleh karena itu, penerapan *bot* Telegram sudah sangat familiar di lingkungan BSI UII.

1.2 Rumusan Masalah

Dari latar belakang yang telah dijabarkan sebelumnya, maka rumusan masalah adalah sebagai berikut:

a. Bagaimana visualisasi yang dihasilkan berdasarkan *Log Firewall* tersebut?
b. Bagaimana visualisasi yang dihasilkan berdasarkan *Log Firewall* dengan memanfaatkan teknik *machine learning* menggunakan algoritma *K-means*?

c. Bagaimana *alert* sistem pada *Log Monitoring* bekerja?

1.3 Tujuan Penelitian

Tujuan pencapaian pada penelitian ini adalah:

a. Membangun *dashboard* untuk visualisasi *non-Cluster Log* berdasarkan data *Log Firewall (Traffic Type)* menggunakan *platform* Splunk.
b. Membangun visualisasi *with cluster* berdasarkan data *Log Firewall (Traffic Type)* menggunakan *platform* Splunk dengan *Cluster* memanfaatkan teknik *machine learning* menggunakan algoritma K-means.
c. Mengintegrasikan temuan yang didapat aplikasi Splunk dengan *bot* Telegram secara *real-time* dalam bentuk *alert system.*

1.4 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah:

a. Memudahkan pembacaan *Log Firewall* melalui visualisasi yang telah dibuat.
b. Mempermudah dalam memperoleh informasi dari hasil visualisasi berdasarkan *Log Firewall* berupa *alert* secara *real-time.*
c. Memberikan inspirasi bagi BSI UII untuk menerapkan teknologi mutakhir pada bidang keamanan jaringan, khususnya pada pemanfaatan *Log* yang dihasilkan oleh perangkat-perangkat yang digunakan UII.

1.5 Batasan Masalah

Batasan masalah yang digunakan pada penelitian ini adalah:

a. Tools SIEM yang digunakan adalah Splunk.
b. Penelitian ini menggunakan studi kasus pada jaringan Universitas Islam Indonesia yang dikelola oleh Badan Sistem Informasi.
c. Penelitian tidak masuk dalam ranah insiden respons.
d. Data uji berasal dari *Log Firewall Traffic Type* yang diberikan oleh Badan Sistem Informasi, Universitas Islam Indonesia.

1.6 Usulan Penyelesaian

Usulan penyelesaian yang diajukan adalah membuat *dashboard* visualisasi berdasarkan *Log Firewall Traffic Type* dari Badan Sistem Informasi, Universitas Islam Indonesia

menggunakan teknologi *Security Information and Event Management* dengan *tools* Splunk. Karena data *Log Firewall* yang diolah berasal dari *Next-Generation Firewall* dengan jumlah pengguna jaringan UII lebih dari 30.000. Maka ukuran dari *Log Firewall* masuk dalam kategori *Big Data.* Visualisasi yang akan dibuat terdiri dari visualisasi *non-Cluster* dan visualisasi *with Cluster*. Visualisasi tersebut diharapkan dapat mempermudah dalam hal pembacaan informasi serta menambah informasi baru untuk kebutuhan pengambilan keputusan bagi pemangku kepentingan di Universitas Islam Indonesia.

Dari hasil visualisasi tersebut, untuk mempermudah administrator dalam memperoleh informasi bahaya yang mengancam jaringan UII atau informasi sejenis lainnya serta dapat diakses dengan perangkat apa saja, kapan saja dan di mana saja, maka akan dibuat *alert* sistem yang diintegrasikan dengan *bot Telegram* secara *real-time.*

1.7 Metode Penelitian

Penelitian ini menggunakan metode penelitian sebagai berikut:

1.7.1 Studi Literatur

Tahapan pertama dalam penelitian ini adalah studi literatur untuk pencarian masalah, pencarian referensi, dan landasan teori sebagai dasar untuk melakukan penelitian. Studi literatur dilakukan guna melakukan *review* terhadap makalah, jurnal ataupun hasil penelitian yang sejenis, mengeksplorasi sumber pustaka yang dikutip oleh tulisan yang dibaca dengan bahasan yang sejenis.

1.7.2 Perencanaan

Setelah melakukan studi literatur, dilakukanlah perencanaan secara mendalam untuk solusi ditawarkan. Perencanaan meliputi sumber atau literatur yang akan diadopsi, pemilihan teknologi yang akan digunakan, dan pemilihan *tools* yang akan digunakan.

1.7.3 Pengumpulan Data

Data yang akan digunakan merupakan data yang tercatat pada *Log Firewall* Badan Sistem Informasi, Universitas Islam Indonesia data dalam waktu satu hari secara penuh yang merupakan hasil pilihan dari hari-hari lainnya di Universitas Islam Indonesia. Data *Log Firewall* yang digunakan adalah tipe *Traffic* dan merupakan data sekunder.

1.7.4 Implementasi

Tahapan implementasi dilakukan mulai dari instalasi kebutuhan aplikasi yang akan digunakan untuk penelitian, pelaksanaan *Search* untuk membuat visualisasi, pelaksanaan klasterisasi untuk kebutuhan visualisasi *with Cluster*, dan pembuatan *bot* Telegram hingga mengintegrasikannya dengan aplikasi Splunk.

1.7.5 Pengujian

Pengujian akan dibedakan menjadi tiga bagian. Pada bagian pertama untuk menguji *Rule* yang menjadi hal paling utama dari visualisasi, pengujian akan dilakukan dengan mengujikan *Rule* dengan pilihan data *Log Firewall* di hari-hari lainnya. Pada bagian kedua, pengujian dilakukan dengan menerapkan *Rule* pada data *Log Firewall* di hari-hari lainnya. Pengujian bagian ketiga dilakukan dengan mencoba salah satu *Rule* yang dapat diuji secara *real-time* yaitu bada bagian pengamanan percobaan masuk secara *brute force* ke dalam akun Splunk.

1.7.6 Analisis

Tahapan analisis akan dilaksanakan setelah pengujian dilakukan. Analisis dilakukan untuk mengamati hasil dari visualisasi *non-Cluster* dan visualisasi dengan *Cluster.* Hasil dari analisis ini yang nantinya akan menjadi acuan dalam penarikan kesimpulan.

1.8 Sistematika Penulisan

Pada sistematika penulisan, tulisan dibagi ke dalam lima bab secara berurutan untuk mempermudah pembahasan serta mempermudah pembaca dalam memahami tulisan. Berikut merupakan sistematika penulisan dari hasil penelitian yang telah dilakukan:

BAB I PENDAHULUAN

Berisikan penjelasan dasar dari mengapa penelitian ini dilakukan. Terdiri dari beberapa bagian seperti latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, usulan penyelesaian, metode penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Berisikan teori-teori dasar yang digunakan sebagai acuan pengetahuan dasar dalam membuat rancangan dan juga konsep dari penelitian yang dilakukan sehingga penelitian ini merupakan penelitian yang ilmiah.

BAB III METODOLOGI PENELITIAN

Berisikan bagaimana langkah-langkah yang dilakukan dalam melakukan penelitian. Di dalamnya terdapat penjelasan mengenai jenis data yang digunakan, ukuran keberhasilan penelitian yang dilakukan, dan metode penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bagian ini berisikan bagaimana hasil dari penelitian yang telah dilakukan, kendala yang didapatkan selama penelitian, dan analisis dari hasil yang telah didapat.

BAB V PENUTUP

Berisikan kesimpulan dari hasil penelitian yang telah dilakukan dari bab sebelumnya. Pada bab ini juga terdapat saran yang ditujukan untuk penelitian sejenis agar ke depannya dapat lebih baik lagi.

BAB II
LANDASAN TEORI

2.1 Landasan Teori

Dalam menjalankan operasional kegiatannya, organisasi atau perusahaan tidak akan terlepas dari pemanfaatan infrastruktur information *technology* atau biasa disebut IT. Pemanfaatan IT salah satunya sebagai tempat mengumpulkan sumber informasi serta berbagai data penting yang merupakan aset dari organisasi. Informasi ataupun aset yang terkumpul merupakan sebuah nilai (*value*) yang berharga dari suatu organisasi yang harus dilindungi. Semakin besar sebuah organisasi tentunya akan semakin besar dalam memanfaatkan infrastruktur IT, hal ini akan membuat sistem IT semakin kompleks dan terdistribusi sehingga akan membuat departemen IT mengalami kesulitan dalam melakukan pengelolaan serta kegiatan monitoring. Penerapan teknologi mutakhir tentunya diperlukan untuk mempermudah departemen IT dalam melakukan pekerjaannya.

2.1.1 Keamanan *Cyber*

Keamanan *Cyber* adalah suatu upaya berkelanjutan dalam melindungi jaringan sistem yang digunakan dan seluruh data dari perbuatan penggunaan data yang berbahaya atau tidak sah (Cisco, 2020). Dalam tingkatan secara pribadi, seseorang perlu menjaga keamanan identitas, data, dan perangkat yang terdapat pada komputernya. Pada tingkatan perusahaan, seluruh karyawan bertanggung jawab untuk melindungi reputasi, data, dan pelanggan dari organisasi tersebut. Pada tingkat kenegaraan yang dipertaruhkan adalah tentang keamanan nasional, serta keselamatan dan kesejahteraan dari warga negara.

Dalam tingkatan perusahaan atau organisasi data dibagi menjadi beberapa jenis yaitu

2.1.1.1 Data Tradisional

Yaitu data perusahaan dengan cakupan informasi karyawan, kekayaan intelektual dan data keuangan. Informasi dari karyawan mencakup materi lamaran kerja, gaji, surat-surat, perjanjian karyawan, dan segala informasi yang digunakan dalam membuat keputusan kepegawaian. Jika dalam lingkup organisasi pendidikan seperti halnya universitas. Maka cakupan informasi yang harus dijaga selain dari karyawan adalah dosen dan juga mahasiswa.

2.1.1.2 *Internet of Things* dan *Big Data*

Saat ini, dengan munculnya *Internet of Things* atau biasa disingkat IoT, semakin banyak data yang harus diolah, dikelola, dan diamankan. IoT merupakan jaringan besar benda fisik, contohnya sensor dan peralatan di luar jaringan komputer biasa (Cisco, 2020). Segala hal yang tersambung ini, apabila digabung dengan kenyataan bahwa saat ini selain bidang penyimpanan fisik yang dapat diperluas terdapat penyimpanan baru yaitu penyimpanan melalui *cloud* dan virtualisasi yang akan mengarahkan ke pertumbuhan data yang eksponensial. Data ini telah menciptakan sebuah area dengan kepentingan yang baru dalam ranah teknologi dan bisnis yang saat ini disebut dengan "*Big Data*". Dengan kecepatan, volume, dan ragam dari data yang dihasilkan IoT serta operasional bisnis dalam rentang harian, kerahasiaan, integritas, dan ketersediaan dari data ini sangat penting bagi kelangsungan hidup dalam organisasi.

2.1.2 Keamanan Informasi (*Information Security*)

Keamanan Informasi merupakan bidang serta aktivitas profesional multidisiplin yang berkaitan dengan pengembangan dan juga implementasi mengenai mekanisme keamanan dari berbagai aspek secara teknis maupun organisasional (Dewanto, 2018). Fokus dari keamanan informasi yaitu untuk mengamankan informasi pada bagian infrastruktur *information technology* atau biasa disebut IT maupun sistem informasi dari beragam serangan yang nantinya mengakibatkan informasi tersebut dapat di akses secara tidak sah. Selain itu, serangan juga dapat mengakibatkan penggunaan informasi yang tidak sah, data yang bocor, dan perubahan atau modifikasi data serta perusakan informasi. Tindakan minimal yang harus diterapkan oleh organisasi untuk mengamankan informasi saat ini biasa dikenal dengan kerahasiaan (*confidentiality*), integritas (*integrity*), dan ketersediaan (*availability*) atau biasa disebut dengan CIA. Kerahasiaan bertugas memastikan privasi dari data dengan membatasi akses melalui enkripsi yang terotentikasi. Integritas bertugas menjamin bahwa informasi tepat dan akurat serta terpercaya. Ketersediaan bertugas memastikan informasi selalu dapat diakses oleh pihak atau pengakses yang berwenang (Cisco, 2020). Gambar 2.1 menunjukkan keterkaitan dari ketiga komponen tersebut.

Gambar 2.1 Keamanan Informasi (CIA)

Sumber:(Opentext, 2020)

a. *Confidentiality* (Kerahasiaan)

Kerahasiaan memiliki istilah lain yaitu privasi. Sebuah perusahaan hendaknya memiliki kebijakan untuk membatasi akses ke informasi atau data yang hanya dapat diakses oleh pihak atau staf yang berwenang. Salah satu solusi dari hal tersebut adalah membagi data menurut tingkat keamanan atau sensitivitas dari informasi tersebut. Terdapat metode untuk memastikan kerahasiaan yang mencakup enkripsi data, ID nama pengguna dan kata sandi, otentikasi dua faktor, dan meminimalisir dalam menyebarkan informasi yang sensitif.

b. *Integrity* (Integritas)

Integritas merupakan keakuratan, konsistensi, dan keandalan dari data selama masa pakainya. Data yang dikirimkan tidak boleh berubah dan juga tidak dapat diubah oleh seseorang atau entitas yang tidak berwenang ataupun tidak sah. Perizinan pada file serta kontrol untuk akses pengguna dapat mencegah dari pengakses yang tidak sah. Selain itu, kontrol versi juga dapat digunakan untuk mencegah perubahan dari ketidaksengajaan pengguna yang berwenang. Cadangan data diperlukan untuk membantu apabila data ingin dikembalikan dari risiko kerusakan seluruh data, dan *hashing checksum* dapat digunakan untuk memverifikasi integritas data selama data dalam pengiriman.

c. *Availability* (Ketersediaan)

Memelihara peralatan, memperbaiki perangkat keras, selalu *update* sistem operasi dan perangkat lunak, serta mencadangkan data untuk memastikan ketersediaan jaringan dan

data bagi pengguna yang berwenang. Rencana yang telah disusun hendaknya dijalankan degan cepat misalnya dalam hal memulihkan dari bencana alami atau akibat perbuatan manusia. Peralatan (*tools)* atau perangkat lunak (*software)* keamanan, seperti *firewall* akan berguna untuk melindungi dari gangguan akibat serangan seperti DoS (penolakan layanan). DoS terjadi apabila penyerang mencoba membuat sumber daya kewalahan dalam memberikan respon yang mengakibatkan layanan tidak tersedia bagi pengguna.

2.1.3 Ancaman Jaringan (*Network Threat)*

Komponen infrastruktur dibidang IT salah satunya adalah Jaringan Komputer. Jaringan Komputer menghubungkan dan membuat layanan IT dapat bekerja dalam sebuah instansi. Jaringan merupakan tempat yang mana segala macam layanan IT dapat berjalan untuk mendukung proses pada bisnis yang ada dan menunjang proses bisnis pada instansi untuk tetap berjalan. Hal-hal yang dimiliki oleh instansi merupakan sebuah *value* yang harus dijaga, sehingga segala macam upaya penyerangan agar dapat mengakses *value* salah satunya melalui bagian infrastruktur jaringan (Arfanudin, 2019). Berikut merupakan contoh serangan-serangan yang biasanya menyerang infrastruktur jaringan:

d. *Bruteforce*

   Jenis serangan ini biasanya digunakan untuk meretas *password* dengan mencoba semua kemungkinan dari kombinasi yang biasanya telah disusun berupa *wordlist.* Waktu yang dibutuhkan untuk dapat menemukan *password* yang sesuai didasarkan pada panjang dan juga kombinasi karakter dari *password* yang akan diretas.

e. *Phishing*

   Serangan *phishing* merupakan sejenis serangan yang memanipulasi psikologis di mana penyerang biasanya membuat email, teks, atau situs web palsu untuk mengelabuhi korban agar menyerahkan informasi sensitif misalnya *username* dan *password.* Serangan ini banyak dilakukan karena tergolong serangan yang mudah, efektif, dan mudah untuk dilakukan.

f. DDoS *Attack*

   Distributed Denial of Service atau biasa dikenal dengan istilah DDoS merupakan ancaman dunia maya yang signifikan bagi banyak organisasi atau perusahaan. Serangan ini dirancang untuk membanjiri sumber daya jaringan korban sehingga infrastruktur jaringan korban tidak dapat memproses lalu lintas yang sah di jaringan mereka. Metodologi serangan ini bervariasi dari satu serangan ke serangan berikutnya dan

mungkin juga melibatkan barbagai tingkat kerumitan. Terkadang serangan DDoS digunakan sebagai alat untuk mengalihkan perhatian tim keamanan siber.

2.1.4 *IT Security Risk Management*

Risk Management atau manajemen risiko merupakan proses untuk melakukan identifikasi celah keamanan (*vulnerabilities)* serta ancaman yang mengancam sumber informasi yang biasanya digunakan oleh organisasi dalam upaya mencapai tujuan bisnisnya dan untuk menentukan respon/tindakan pencegahan apa saja yang diperlukan apabila risiko tersebut terjadi atau bahkan dapat mengurangi risiko dengan catatan tingkatnya dapat diterima didasarkan pada nilai dari sumber informasi organisasi tersebut(Arfanudin, 2019). *IT risk management* berupaya untuk mencoba melindugi *Confidentiality, Integrity, dan Availability* (CIA) dengan meminimalisir dampak yang nantinya mungkin akan muncul dan memberikan efek pada *Confidentiality* dari informasi, *Integrity* dari data yang terdapat pada sistem, dan *Availability* yang berasal dari infrastruktur sistem (Haikal, 2017).

Hal utama yang mendasari mengapa suatu organisasi memerlukan penerapan proses manajemen risiko yaitu untuk mendukung misi dari organisasi dan melindungi aset-aset organisasi. *IT Security Risk Management* erat dalam berkaitan dengan bagaimana *security* diimplementasikan pada organisasi sehingga pemahaman terkait proses bisnis organisasi juga kemungkinan risiko yang akan berdampak terhadap proses bisnis tersebut dangan diperlukan. Manajemen Risiko akan sangat membantu dalam memanajemen organisasi yang bertujuan untuk menyeimbangkan antara dampak yang ditimbulkan oleh *risk* (risiko) dan *cost* (biaya) yang dibutuhkan agar *risk* tersebut dapat diminimalisir (Haikal, 2017). Dalam menerapkan *IT Security Risk Management* terdapat beberapa proses yang harus dilakukan seperti *Threat Assessment and Analysis, Asset Identification and Valuation, Vulnarability Analysis, Risk Evaluation, and Interim Report.*

2.1.5 *Risk Assesment*

Risk Assesment atau penilaian risiko adalah tahap melakukan identifikasi risiko pada tahap proses *Risk Management Process.* Organisasi menggunakan *risk assesment* untuk digunakan sebagai penentu bahaya apa sajakah yang berpotensi dan risiko apa saja yang akan ditimbulkan dan berdampak pada sistem IT. Hasil dari proses ini diharapkan dapat membantu untuk mengidentifikasi kontrol yang sesuai, guna mengurangi atau menghilangkan risiko-risiko yang ada selama proses pengurangan risiko (*risk mitigation).*

2.1.5.1 Identifikasi Ancaman (*Threat Identification)*

Ancaman (*threat*) merupakan potensi-potensi yang dapat mengakibatkan bahaya dan biasanya dihasilkan oleh sebuah sumber (*threat source)* yang menyerang celah keamanan dari suatu sistem. Sumber ancaman tidak akan menghasilkan ancaman apabila tidak ditemukan celah keamanan pada sistem tersebut.

2.1.5.2 Mitigasi Risiko (*Risk Mitigation*)

Pada tahapan ini, *risk mitigation* meliputi tahap seperti pemberian prioritas, evaluasi, dan implementasi kontrol untuk pengurangan risiko yang direkomendasikan dan biasanya diperoleh dari tahap *risk assessment.* Dalam proses menghilangkan segala risiko yang nantinya mungkin saja dapat terjadi adalah merupakan hal yang tidak praktis juga mustahil, maka dibutuhkan *senior management / business managers* untuk menggunakan pendekatan *least-cost* dan melakukan implementasi kontrol dengan prioritas yang paling penting guna mengurangi level dari risiko menuju level yang lebih dapat diterima. Hal ini biasanya dilakukan dengan mengedepankan dampak yang minimal dari sumber daya dan tujuan dari organisasi.

2.1.5.3 Evaluasi dan Pemantauan (*Evaluation and Monitoring*)

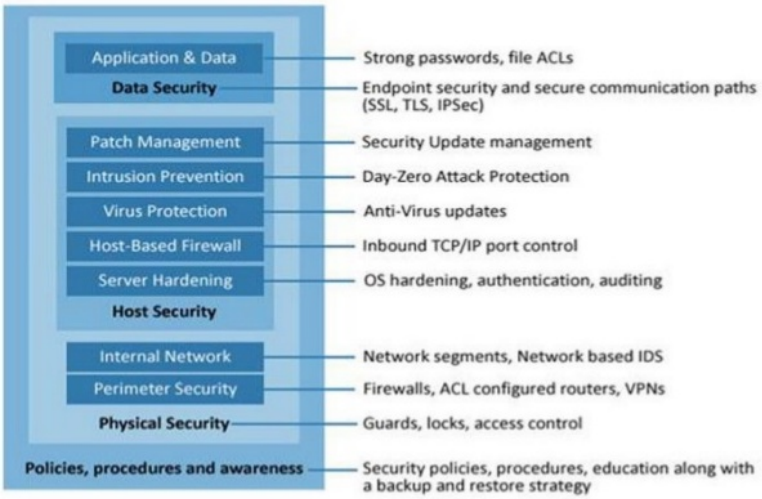
Pada sebagian organisasi atau perusahaan, biasanya jaringan yang dimiliki akan terus-menerus mengalami perkembangan, hal ini meliputi sistem yang dimiliki, berbagai komponen yang dimiliki, dan *software* atau aplikasi yang akan di *update* atau mengalami pergantian dengan versi paling baru. Tak hanya itu, penambahan juga nantinya akan meliputi pembaharuan serta pergantian aturan keamana yang telah dimiliki. Pada tiap-tiap perubahan tersebut akan menimbulkan hal atau permasalahan sebab risiko-risiko yang akan muncul ataupun risiko yang sebelumnya telah berhasil diatasi akan kembali muncul. Maka dari itu, tahapan proses *risk management* harus tetap terus berjalan dan mengalami perkembangan.

2.1.5.4 Pertahanan Strategi Keamanan Secara Mendalam (*Security Strategy Defence in Depth*)

Dalam melakukan implementasi pada sistem keamanan informasi apabila diihat pada sudut pandang kepentingannya, maka tidak jarang organisasi melakukannya dengan pendekatan *deface in depth* untuk upaya pengamanan informasi yang dimiliki. *Deface in depth* adalah mekanisme keamanan yang berlapis guna meningkatkan keamanan dari sistem secara keseluruhan. Apabila satu serangan menyebabkan gagalnya satu mekanisme yang berjalan, mekanisme lain mungkin dapat tetap memberikan keamanan yang dibutuhkan dalam

penjaminan/melindungi sistem. Sebagai contoh, bukan merupakan ide yang bagus apabila benar-benar mengandalkan *firewall* sebagai alat untuk memberikan keamanan yang terdapat pada aplikasi dengan penggunaan internal saja. *Firewall* biasanya dapat dihindari oleh *hacker* (serangan dapat dilakukan melalui serangan fisik atau serangan rekayasa sosial (*social engineering*)). Tentunya perlu menambahkan alat yang digunaan untuk menambah kekuatan mekanisme keamanan seperti CCTV dan *security awareness* serta berbagai macam alat yang dapat digunakan sebagai solusi untuk menangani berbagai serangan dari vektor tertentu. Penggunaan *firewall* juga diharapkan menggunakan variasi terbaru dari *firewall* yaitu *Next-Generation Firewall* yang tentunya lebih aman.

Gambar 2.2 menunjukkan gambaran implementasi *defense ini depth* pada suatu organisasi banyak pilihan solusi sesuai dengan layernya.

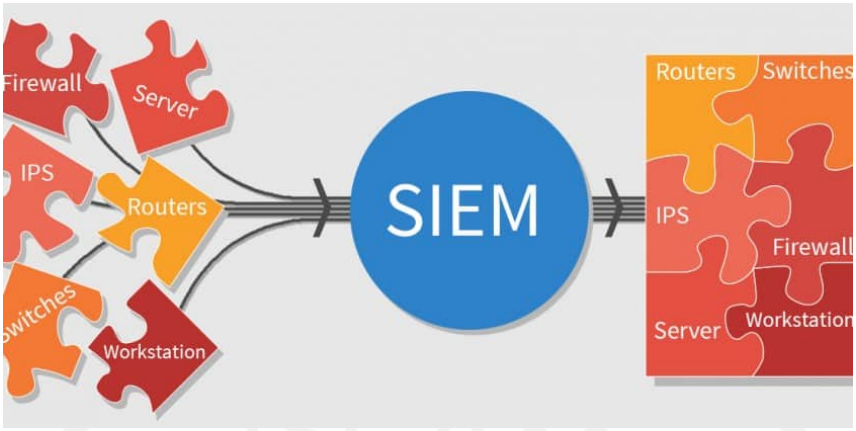


Gambar 2.2 *Defense in Depth*

Sumber: (Kumar, 2015)

2.1.6 *Security Information and Event Management* (SIEM)

Security Information and Event Management atau biasa disebut dengan SIEM merupakan sebuah teknologi yang dapat mendeteksi berbagai ancaman dan insiden dari keamanan dengan mengumpulkan *Log real-time* dan melakukan analisis sejarah *Log* keamanan dari berbagai jenis tipe *Log* dan berasal dari berbagai sumber data dari perangkat yang berbeda-beda (Filkins, 2019). Gambar 2.3 menunjukkan gambaran bagaimana SIEM bekerja.



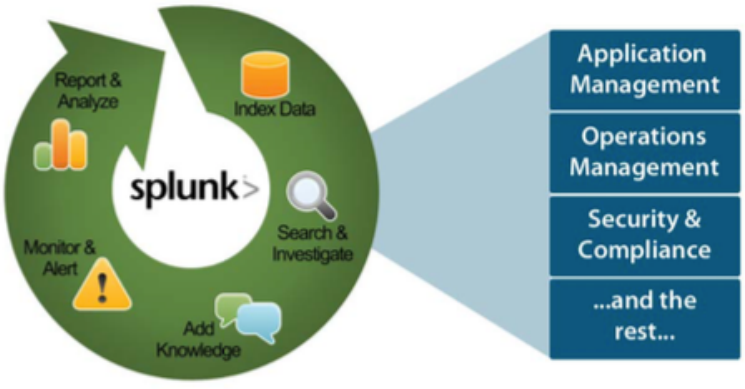
Gambar 2.3 *Security Information and Event Management*

Sumber: (Dqindia, 2017)

SIEM mendukung analisis dan juga investigasi pada insiden melalui analisis dari data historis dan berbagai alat keamananan dari organisasi seperti *Router*, IDS/IPS, *Firewall*, *Server*, dan lain-lain. Inti dari kemampuan teknologi SIEM yaitu cakupan yang luas dalam mengumpulkan *Log* dan juga kemampuan dalam melakukan korelasi dan melakukan analisis kejadian dari berbagai sumber jenis *Log* atau perangkat yang berbeda-beda (Gartner, 2020).

2.1.7 Splunk

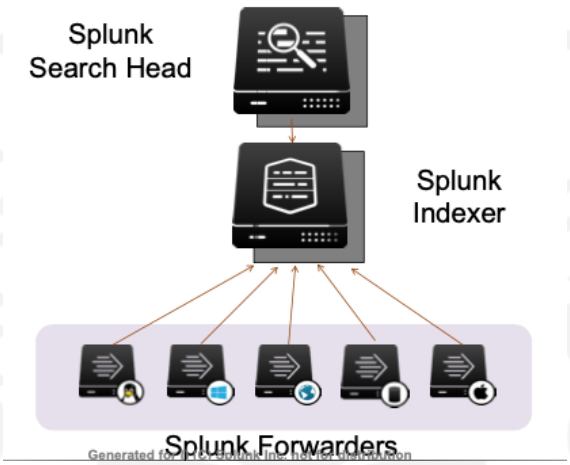
Splunk adalah platform perangkat lunak untuk melakukan pencarian, menganalisis, dan memvisualisasikan data yang dihasilkan mesin dan biasanya berasal dari *computers*, *network device*, *logs*, sensor, *databases*, dan lain-lain yang menyusun infrastruktur IT suatu organisasi (Splunk.inc, 2018). Gambar 2.4 menunjukkan hal-hal yang dapat dilakukan oleh Splunk.



Gambar 2.4 What is *Splunk?*

Sumber: (Splunk.inc, 2018)

Splunk bekerja melalui 3 perangkat penyusun dasarnya, pertama adalah setiap perangkat yang meghasilkan data dipasang alat bernama *Splunk Forwarders* yang berfungsi untuk menghimpun data dari perangkat dan mengirimnya menuju *Splunk Indexer* yang merupakan tempat berkumpulnya data dari berbagai perangkat. Untuk kebutuhan tingkat atas yaitu berkaitan dengan klien dan visualisasi maka alat yang bekerja adalah *Splunk Search Head.* Gambar 2.5 menunjukkan bagaimana Splunk bekerja.



Gambar 2.5 Cara kerja Splunk

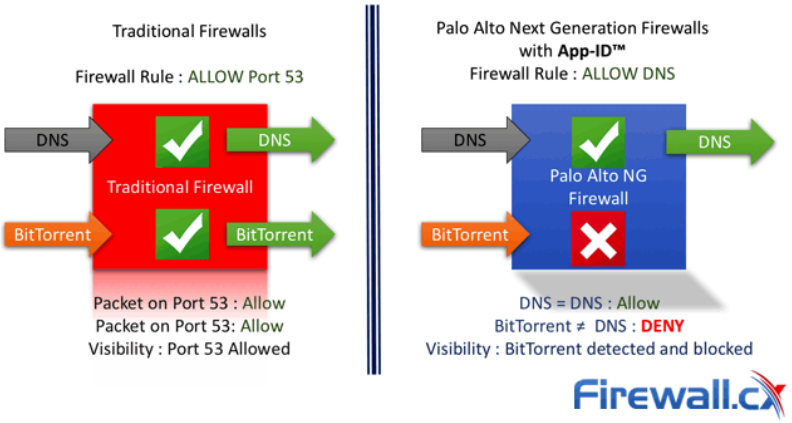
Sumber: (Splunk.inc, 2018)

Splunk memiliki tiga variasi produk yang berbeda, pertama adalah *Splunk Enterprise* yang digunakan oleh bisnis IT besar. Variasi ini membantu organisasi atau perusahaan dalam mengumpulkan dan menganalisis data dari aplikasi, situs web, dll. Kedua *Splunk Cloud* yang merupakan platform yang dihosting. Versi ini memiliki fitur yang sama dengan versi *enterprise,* versi ini dapat diperoleh melalui Splunk atau menggunakan platform *cloud AWS.* Ketiga *Splunk Light* yang merupakan versi gratis. Pada versi ini memungkinkan pencarian, laporan, dan mengubah data *Log* yang digunakan. Namun versi ini memiliki fungsi dan fitur yang terbatas dibanding versi lainnya.

2.1.8 *Next-Generation Firewall* (NGFW)

Next-Generation Firewall atau biasa disebut NGFW lebih kuat dibandingkan dengan *Traditional Firewall.* NGFW memiliki kemampuan *Traditional Firewall* dan juga memiliki sejumlah fitur tambahan untuk menangani lebih banyak variasi kebutuhan organisasi dan memblokir lebih banyak potensi ancaman. *Firewall* ini disebut dengan "*Next-Generation"* untuk membedakan dari *firewall lama (traditional firewall)* yang tidak memiliki kemampuan

tambahan tersebut (Cloudflare, 2020). Segala macam *firewall* mencatat segala macam jenis kegiatan ke dalam *Log Firewall.* Karena terdapat fitur tambahan pada NGFW, maka ukuran dari *Log Firewall* cenderung lebih besar dibanding *Traditional Firewall.* Gambar 2.6 menunjukkan perbedaan antara *Traditional Firewall* dengan *Next-Generation Firewall.*

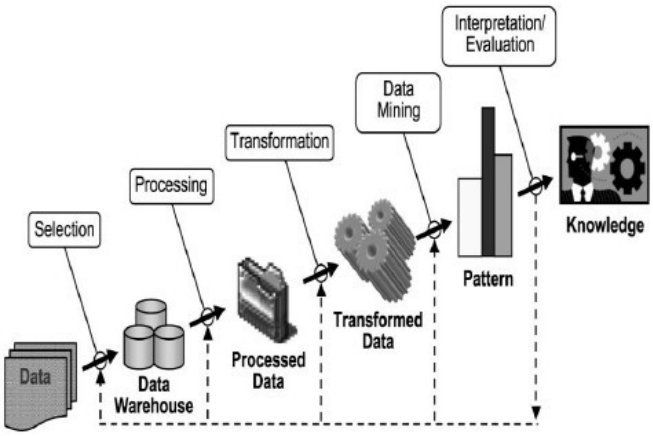


Gambar 2.6 *Next-Generation Firewall from Palo Alto Networks*

Sumber: (Firewall.CX, 2020)

2.1.9 *Data Mining*

Data Mining merupakan proses dalam menggunakan teknik metematik, statistik, kecerdasan buatan, dan *machine learning* dalam melakukan ekstraksi dan melakukan identifikasi yang bermanfaat serta pengetahuan yang terkait bersumber dari *big data* atau maha data (Efrain et al., 2005). Gambar 2.7 menunjukkan bagaimana tahapan dari *Data Mining*.



Gambar 2.7 Tahapan *Data Mining*

Sumber: (Fayyad et al., 1996)

Tujuan dari dilakukannya *Data Mining* adalah untuk memahami lebih jauh terkait perilaku data yang sedang dilakukan pengamatan atau biasa dikenal dengan "deskripsi" dan sebagai acuan untuk memperkirakan kondisi yang nantinya akan dialami atau biasa disebut dengan "prediksi".

2.2 Penelitian Terkait

Terdapat tujuh penelitian yang menjadi referensi dalam membuat perencanaan untuk penelitian ini. Penelitian terkait ini disajikan dalam bentuk tabel agar memudahkan dalam memahami persamaan solusi penelitian dan perbedaan solusi yang ada terhadap problem yang ditemukan.

Tabel 2.1 Penelitian Terkait

No.	Penulis	Uraian Singkat	Persamaan Solusi	Perbedaan Solusi
1.	(Erwinsyah, 2019)	Membuat konsolidasi dan visualisasi *Log server* yang berasal dari BSI UII. Dalam melakukan penelitinnya penulis memanfaatkan *Log Management System* berbasiskan sumber terbuka *(open source)* yaitu *ELK Stack*. Hasil dari penelitian ini adalah *Log* dapat dikonsolidasi dan terkumpul menjadi satu dengan bantuan Filebeat, *Log* dapat dipecah untuk disesuaikan dengan kebutuhan administrator menggunakan Logstash, dan *Log* yang telah dipecah berhasil divisualisasikan dengan memanfaatkan aplikasi Kibana.	Pembuatan *Log Monitoring System* pada BSI UII.	Penerapan teknologi SIEM dengan sumber terbuka. Belum menerapkan teknik *machine learning* dalam pembuatan visualisasi.
2.	(Wulur & Sembiring, 2015)	Penelitian ini melakukan klasifikasi *alert* pada *Intrusion Detection System* menggunakan algoritma *K-means*. Penulis menggunakan *tools* IDS yaitu Snort dalam proses identifikasi serangan dan menggunakan *tools* Matlab untuk proses *Clustering* dengan algoritma *K-means.* Hasilnya peneliti mendapatkan hasil nilai klasifikasi serangan menggunakan algoritma tersebut, namun penulis mencatat beberapa kelemahan dari algoritma *K-means* dalam penelitiannya yaitu kurangnya kejelasan mengenai bagaimana menentukan jumlah klaster (k) yang terbaik.	Penerapan teknik *machine learning* dengan algoritma *K-means* untuk proses Klasterisasi.	Penerapan teknologi yang dipilih. Aplikasi yang digunakan dalam melakukan penerapan. Tujuan dari penerapan teknik *machine learning*.

3.	(Rakhmadhani et al., 2019)	Penelitiannya melakukan integrasi untuk visualisasi *Log* dari honeypot dengan sistem manajemen honeypot paling mutakhir yaitu *Modern Honey Network* (MHN) dengan memanfaatkan *tools* Splunk. Peneliti mendapatkan hasil bahwa MHN mudah diterapkan karena telah memiliki API untuk membaca data penyerang. Hasil yang diperoleh pun mempermudah keterbacaan aktivitas yang telah tercatat pada *Log.*	Pemanfaatan aplikasi *enterprise* Splunk.	Tujuan dari visualisasi yang dilakukan. Teknologi yang digunakan.
4.	(Anastasov & Davcev, 2014)	Kajian literatur dalam penggunaan model dan arsitektur baru dalam melakukan implementasi *Security Information and Event Management* (SIEM) yaitu dengan menggunakan *Hierarchical* SIEM *manager,* sebuah mekanisme yang tujuannya adalah untuk tipe perusahaan dengan tipe organisasi yang telah terdistribusi. Peneliti menyebutkan bahwa telah berhasil dalam mengunakan *ArchSight ESM* dalam melakukan implementasi keamanan jaringan dengan skala besar dan terdistribusi.	Implementasi teknologi *Next-Generation* SIEM pada instansi.	Belum melakukan penerapan secara langsung pada Instansi terkait. Aplikasi untuk menerapkan teknologi SIEM tidak ada penjelasan.
5.	(Vianello et al., 2013)	Melakukan penerapan teknologi *Security Information and Event Management* (SIEM) untuk sebuah infrastruktur IT yang komplek dan banyak jumlah aset yang harus dilindungi. Pada sistem yang lama teknologi *Security Information and Event Management* (SIEM) yang digunakan masih tidak terdistribusi sehingga sulit untuk mengkorelasikan temuan serangan dari sumber yang ada. Penulis melakukan percobaan *SIEM* pada sistem *Olympic Games* dengan menerapkan pendekatan *distributed corelation* serta *query parallezation* dalam melakukan penyerangan. Menggunakan DoS serta Worm sebagai jenis serangan dan hasil monitoring dikirim berupa notifikasi.	Implementasi teknologi *Next-Generation* SIEM pada sebuah instansi dengan menerapkan aplikasi *enterprise*.	Tujuan dari penggunaan teknologi *Next-Generation* SIEM. Aplikasi yang digunakan untuk menerapkan teknologi SIEM.
6.	(Dairinram et al., 2013)	Penelitian ini membahas terkait jumlah *event* yang kian meningkat pada sebuah infrastruktur yang tersusun dengan komplek, akan membuat identifikasi ancaman yang terkumpul menjadi masalah yang harus dihadapi oleh administrator. Peneliti mencoba menggunakan *The Latent Sematic Analysis (LSA)* dalam upaya untuk melakukan analisa terkait	Implementasi teknologi SIEM untuk pengolahan data dengan pertumbuhan *event* yang selalu	Menerapkan jenis teknologi SIEM konvensional. Pendekatan menghadapi pertumbuhan event

		korelasi serangan dari semua *event* yang terkumpul pada SIEM.	mengalami peningkatan.	menggunakan *LSA*.
7.	(Zope et al., 2013)	Penelitiannya melakukan kajian literatur membahas *Security Information and Event Management* (SIEM) dengan pendekatan *Data Mining* telah melihat bahwa berbagai teknik *Data Mining* dapat diterapkan pada SIEM untuk meningkatkan kemampuan dari sistem.	Menerapkan pendekatan *Data Mining* pada teknologi SIEM.	Penerapan belum dilakukan secara langsung pada Instansi terkait. Aplikasi yang digunakan untuk menerapkan teknologi SIEM.

Dari penelitian-penelitian terkait di atas, teknologi *Security Information and Event Management* (SIEM) diadopsi untuk penerapan *Log Monitoring System.* Aplikasi Splunk dipilih untuk digunakan sebagai alat dalam menerapkan teknologi SIEM. Pendekatan *Data Mining* dengan teknik *machine learning* diadopsi untuk meningkatkan keterbacaan informasi dari *Log Firewall* yang berukuran besar. Perbedaan yang sangat menonjol dari penelitian terkait sebelumnya adalah pada bagian data yang diolah, tujuan pengolahan data, serta pendekatan yang digunakan merupakan kombinasi dari penelitian terkait serta saran dari penelitian terkait.

BAB III
METODOLOGI PENELITIAN

3.1 Jenis Data

Data yang akan digunakan adalah data sekunder. Data yang akan digunakan merupakan data dalam waktu satu hari secara penuh yang merupakan hasil pilihan dari hari-hari lainnya di Universitas Islam Indonesia. Sebagai dasar yang menjadi perhatian dalam memilih data didasarkan pada ukuran data yang lebih besar dibanding hari-hari lainnya (berukuran diatas 40 *gigabyte*). Data *Log Firewall* dari Palo Alto Network yang digunakan adalah bertipe *Traffic.*

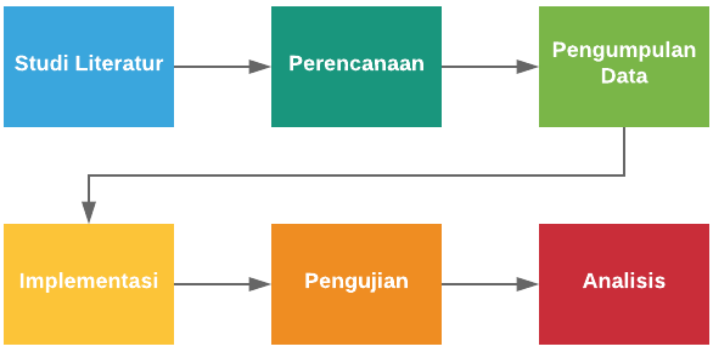
3.2 Ukuran Keberhasilan

Akurasi keberhasilan penelitian ini adalah:

a. *Dashboard* Visualisasi *non-Cluster* dapat memberikan informasi yang mudah untuk dibaca dan dipahami.
b. Algoritma *machine learning* mampu melakukan klastering untuk kebutuhan visualisasi *with Cluster*.
c. *Alert* dapat terintegrasi dengan *bot* Telegram dan mampu memberikan peringatan secara berkala sesuai temuan ancaman pada Splunk.

3.3 Metode Penelitian

Penelitian ini menggunakan metode penelitian yang ditunjukkan pada Gambar 3.1.



Gambar 3.1 Metode Penelitian

3.3.1 Studi Literatur

Tahapan pertama dalam penelitian ini adalah studi literatur untuk pencarian masalah, pencarian referensi, dan landasan teori sebagai dasar untuk melakukan penelitian. Studi

literatur dilakukan guna melakukan *review* terhadap makalah, jurnal ataupun hasil penelitian yang sejenis, mengeksplorasi sumber pustaka yang dikutip oleh tulisan yang dibaca dengan bahasan yang sejenis. Mengadopsi beberapa teknologi yang akan digunakan dalam penelitian dan mengambil beberapa saran penelitian dari studi literatur sebagai masukan dalam meneliti.

3.3.2 Perencanaan

Setelah melakukan studi literatur, hal yang selanjutnya dilakukan adalah melakukan perencanaan secara mendalam untuk solusi ditawarkan. Perencanaan meliputi sumber atau literatur yang diadopsi, pemilihan teknologi yang akan digunakan, dan pemilihan *tools* yang akan digunakan.

3.3.3 Pengumpulan Data

Data yang akan digunakan merupakan data yang tercatat pada *Log Firewall* Badan Sistem Informasi Universitas Islam Indonesia dalam kurun waktu satu hari secara penuh dan merupakan hasil pilihan dari hari-hari lainnya. Data *Log Firewall* yang digunakan adalah tipe *Traffic* dan merupakan data sekunder.

3.3.4 Implementasi

Pada bagian implementasi akan terdiri dari beberapa metode sebagai berikut,

3.3.4.1 Instalasi Splunk Web

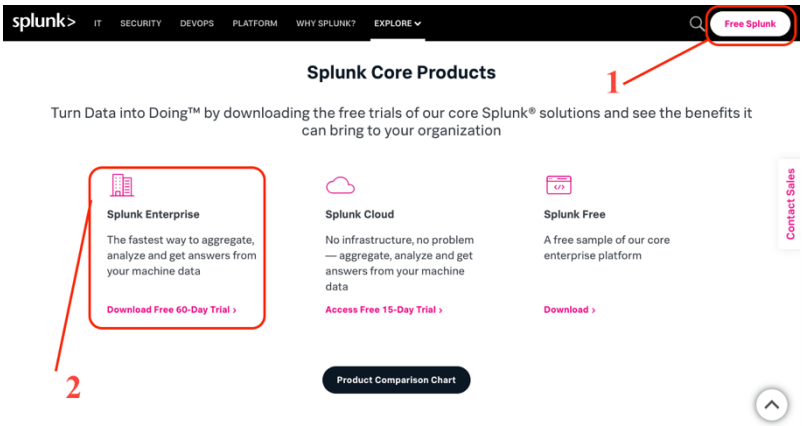
Dalam penelitian kali ini, *tools* yang digunakan merupakan versi *trial* dari Splunk dan merupakan Splunk Web. Penggunaan aplikasi Splunk untuk kebutuhan jangka panjang adalah berbayar. Adapun kebutuhan sistem untuk dapat menggunakan versi *trial* dari *Splunk Enterprise* ditunjukkan pada Tabel 3.1.

Tabel 3.1 Kebutuhan Sistem

Sistem Operasi	Windows / Linux / Mac OS
Processor	Minimal **2-core 64-bit CPU at 2GHz**
RAM	Minimal **4 *gigabyte***
Web Browser	Versi paling *update* dari **Chrome / Firefox / Safari**
Port Splunk Web	8000

Setelah sistem terpenuhi, berikut merupakan tata cara instalasi Splunk Web:

1. Daftar akun Splunk. Pendaftaran akun pada alamat Splunk.com
2. Setelah akun terdaftar, pilih menu "*Free Splunk*" yang terdapat dibagian bawah profile. Kemudian pilih "*Splunk Enterprise"*. Gambar 3.2 menunjukkan gambaran tahapan instalasi Splunk Web.



Gambar 3.2 Tahapan instalasi *Splunk Web*

3. Lalu pilih sistem operasi dari perangkat yang akan digunakan (Pada penelitian ini, sebelumnya menggunakan perangkat dengan sistem operasi Mac OS, namun karena keterbatasan kemampuan dari perangkat tersebut. Penelitian dilanjutkan dengan memanfaatkan perangkat yang berbeda dengan sistem operasi Windows). Kemudian *download* salah satu dari pilihan jenis file instalasi yang tersedia.
4. Setelah *install* selesai, maka setelah muncul tampilan Splunk Web, silahkan untuk membuat akun. Splunk Web yang berjalan bersifat *localhost* sehingga tidak membutuhkan koneksi internet.

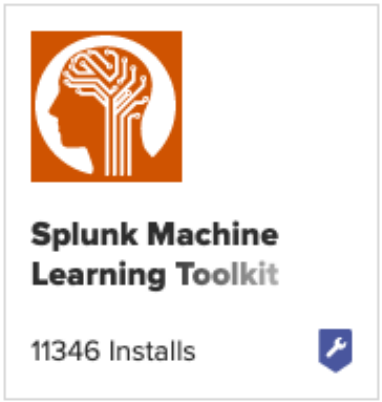
3.3.4.2 Instalasi *Add-on* Aplikasi *Palo Alto Firewall*

Data yang akan diteliti merupakan data yang berasal dari *Firewall* PaloAlto Network. Oleh karena itu, untuk menambah keterbacaan *fields* dari data *Log Firewall Traffic Type* sehingga *fields* yang terdeteksi akan lebih banyak. Maka dibutuhkan instalasi *add-on* Aplikasi Palo Alto Network. Aplikasi dapat diperoleh melalui halaman web Splunkbase.com. Gambar 3.3 menunjukkan logo dari *add-on* aplikasi *Palo Alto Firewall* yang digunakan.

Gambar 3.3 Aplikasi *Palo Alto Network add-on* untuk Splunk

3.3.4.3 Instalasi Aplikasi *Splunk Machine Learning Toolkit*

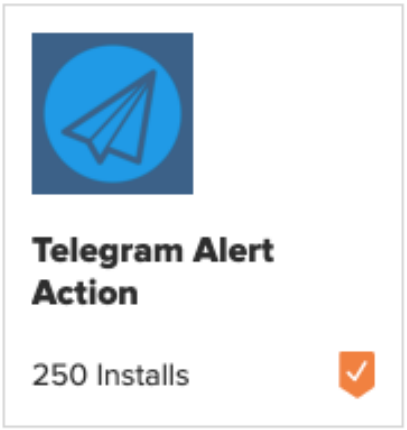
Untuk proses klasterisasi dengan pendekatan *machine learning,* dibutuhkan aplikasi yang dapat melakukan proses komputasi dengan *machine learning*. pada Splunk telah terdapat aplikasi berupa *add-on* untuk pengolahan tersebut yang bernama "*Splunk Machine Learning Toolkit"*. Gambar 3.4 menunjukkan logo dari aplikasi *Splunk Machine Learning Toolkit* yang digunakan.



Gambar 3.4 Aplikasi Splunk *Machine Learning Toolkit*

3.3.4.4 Instalasi Aplikasi Telegram *Alert Action*

Dalam *alert system* yang akan diterapkan nantinya akan diintegrasikan dengan *bot* Telegram. Maka penggunaan aplikasi "*Telegram Alert Action*" bertujuan untuk memudahkan dalam melakukan integrasi tersebut. Gambar 3.5 menunjukkan aplikasi *Telegram Alert Action* yang digunakan.



Gambar 3.5 Aplikasi *Telegram Alert Action*

3.3.4.5 Upload Data

Karena penelitian ini tidak mungkin dilakukan pada *environment* jaringan secara langsung, sebab dapat memberi dampak berbahaya bagi berjalannya *environment* jaringan UII. Maka data yang akan diteliti merupakan data sekunder. Splunk Web versi *trial* memberikan pembatasan untuk upload data yaitu sebesar 500 *megabyte* untuk satu waktu *upload*. Data *traffic type Log Firewall* dalam satu hari memiliki ukuran 30 – 50 *gigabyte*. Maka salah satu solusi agar data tetap dapat diolah pada aplikasi Splunk adalah dengan membagi data menjadi bagian-bagian yang berukuran dibawah 500 *megabyte*.

Pemecahan data dilakukan dengan memanfaatkan *tools* dari linux yaitu *tools* "*Split*". Gambar 3.6 menunjukkan sintaks yang digunakan untuk memecah file menjadi bagian-bagian degan ukuran yang dapat ditentukan.

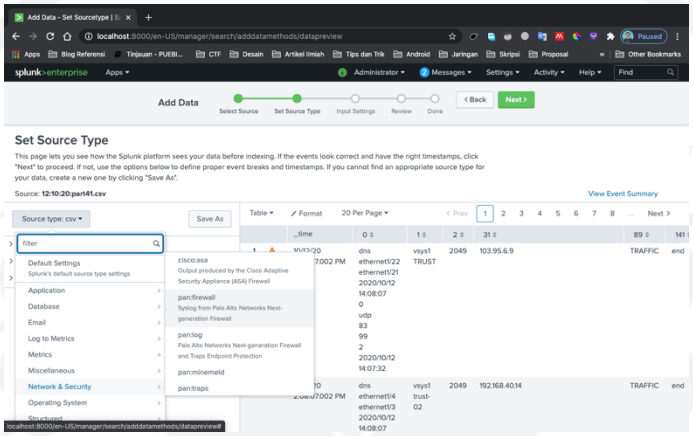
```
Split -b <ukuran file>m –additional-suffix=<ekstensi file> <Nama File>
            <nama file setelah dipecah> --verbose
```

Gambar 3.6 Sintaks untuk melakukan pemecahan data

Setelah data dipecah, maka langkah-langkah untuk melakukan *upload* data pada Splunk Web adalah sebagai berikut:

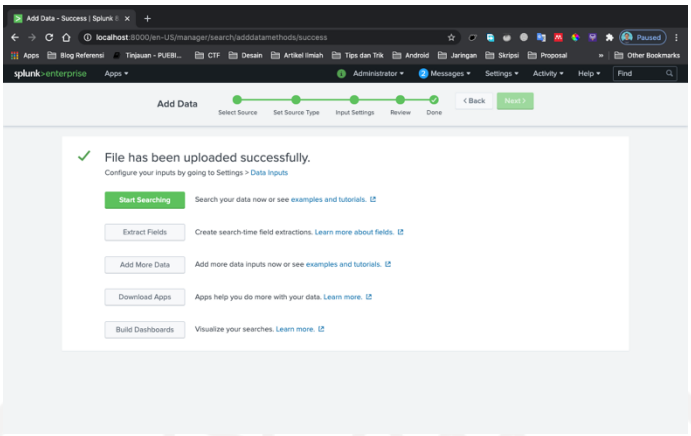
1. Pada menu awal Splunk App, silahkan pilih "*Add Data*", kemudian gulirkan (*scroll)* ke bawah dan pilih "*Upload*",

2. Kemudian pilih ”*Select File*” pada bagian ”*Select Source*”, lalu pilih file yang telah dipecah dari urutan pertama. Setelah data selesai diproses, silakan pilih ”*Next*”,
3. Pada bagian ”*Set Source Type*” pilih bagian ”*Source Type*” kemudian pilih bagian ”*Network & Security*” lalu pilih ”*pan:firewall*”, Gambar 3.7 menunjukkan visualisasi bagaimana pemilihan *Source Type.*



Gambar 3.7 Pemilihan *Set Source Type*

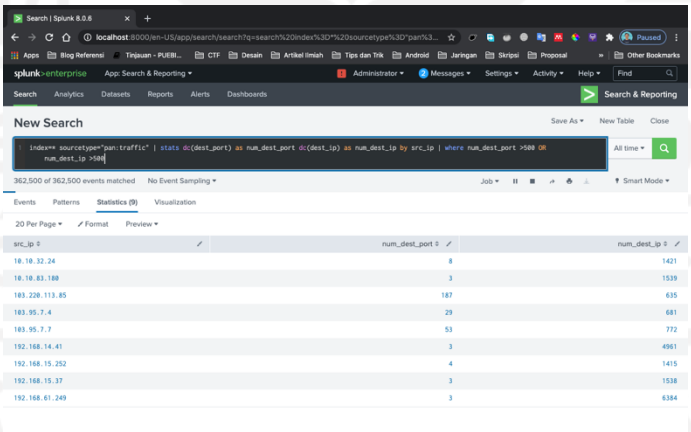
4. Pada bagian ”*Index*” biarkan secara *default* dan cukup lakukan perubahan pada bagian ”*Host Field Value*” sesuai nama yang dibutuhkan,
5. Pada bagian ”*Review*” cek terlebih dahulu apakah data telah sesuai dengan kebutuhan, setelah itu pilih ”*Submit*” maka data akan diunggah.
6. Setelah data selesai diunggah, maka unggah kembali data berikutnya dengan memilih ”*Add More Data*” lalu lakukan kembali langkah-langkah ini dari bagian pertama. Gambar 3.8 menunjukkan tampilan apabila *upload file* telah berhasil.



Gambar 3.8 Tampilan *add data* berhasil

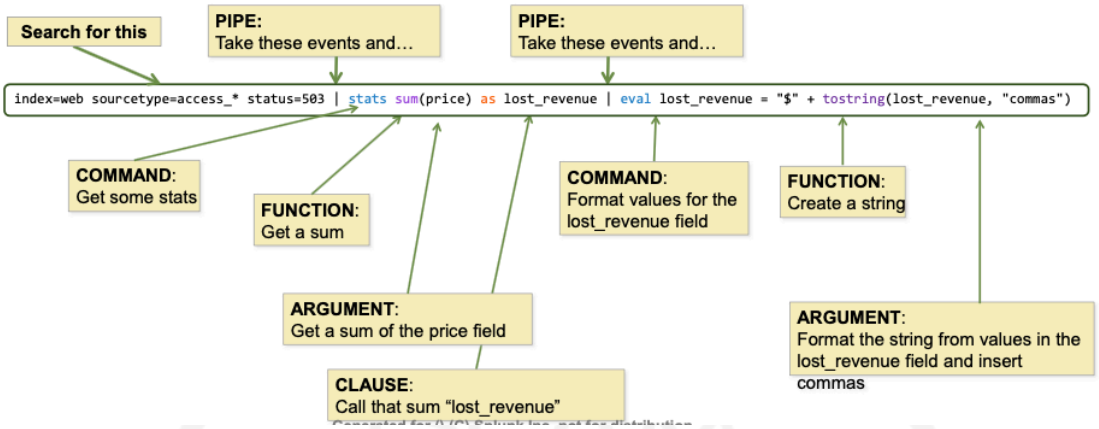
3.3.4.6 Search & Reporting

Pada bagian ini, hal yang dilakukan adalah membuat "*Rule"* untuk menampilkan visualisasi dari data *Log Firewall* agar mudah dipahami atau dibaca. Untuk melakukan hal ini, yang harus dilakukan adalah masuk pada bagian "*Search & Reporting*" dari tampilan awal Splunk Web. Gambar 3.9 menunjukkan tampilan dari menu *Search & Report.*



Gambar 3.9 Menu *Search & Reporting*

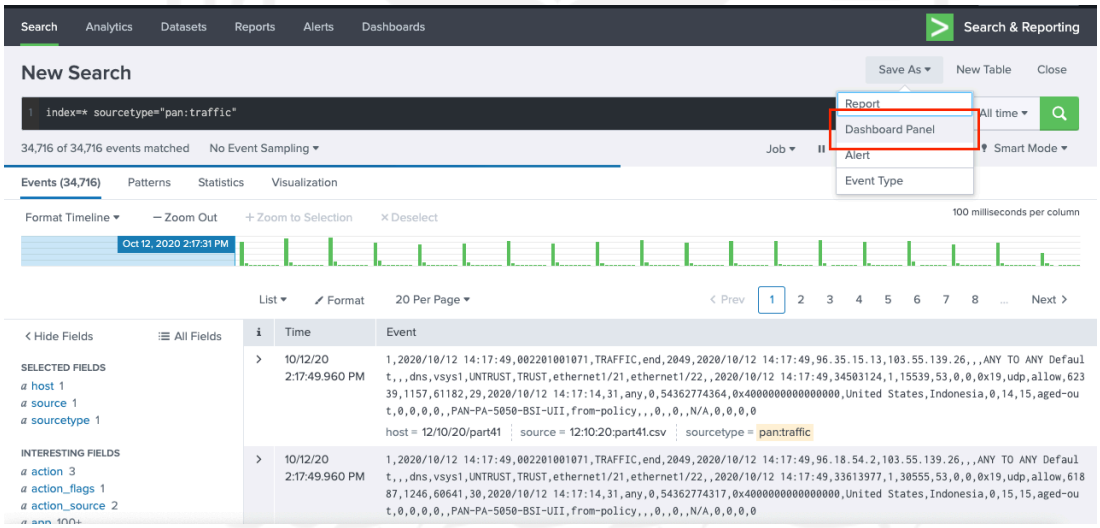
Kemudian untuk melakukan pencarian format penulisan *Rule* adalah mengikuti "*Search Language Syntax*". Masukkan *Search Language Syntax* pada bagian pencarian, pilih rentang waktu dari data yang akan dicari dan divisualisasikan, kemudian klik "*Search*". Adapun komponen dari *Search Language Syntax* ditunjukkan pada Gambar 3.10.



Gambar 3.10 *Search Language Syntax*

Sumber: (Splunk.inc, 2018)

Setelah pencarian dilakukan dan telah didapatkan *Rule* yang sesuai dengan kebutuhan, maka simpan hal tersebut dengan memilih untuk menampilkannya ke dalam *dashboard* panel. Gambar 3.11 menunjukkan tampilan untuk menyimpan hasil temuan ke dalam sebuah *dashboard*.



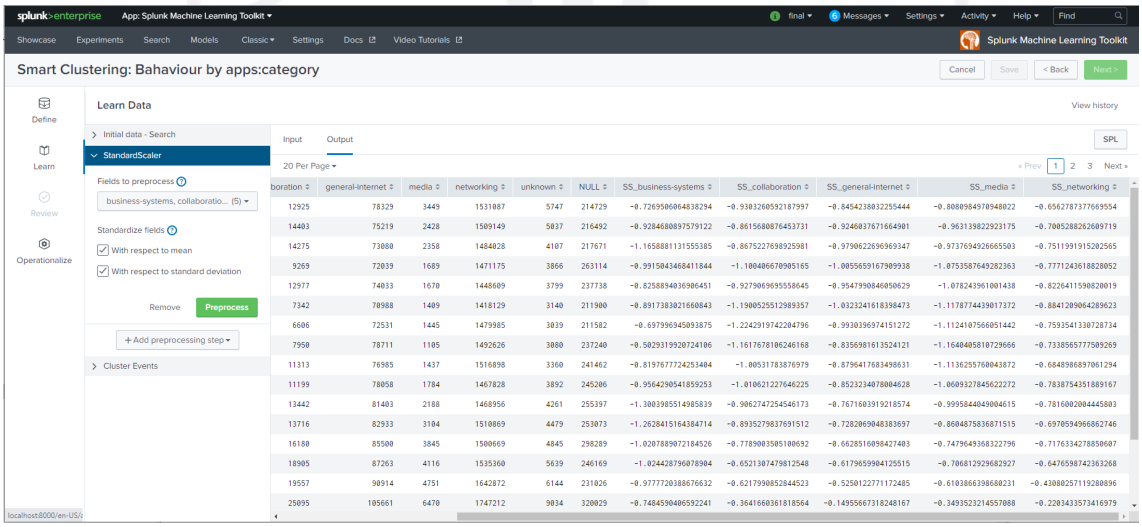
Gambar 3.11 Menyimpan hasil *Search & Reporting* ke *dashboard* panel

3.3.4.7 Klasterisasi

Pada bagian ini, hal yang dilakukan adalah mencoba melakukan klasterisasi dari data "*Log Firewall*". Kegiatan ini dilakukan dengan menggunakan aplikasi *Splunk Machine Learning Toolkit* yang sebelumnya telah dilakukan instalasi. Pada aplikasi *Splunk Machine Learning Toolkit* terdapat dua jenis fitur pilihan untuk melakukan *clustering* yaitu *Smart*

Clustering dan *Cluster Numeric Event.* Fitur *Smart Clustering* dipilih untuk melakukan *Clustering* karena hasil visualisasi lebih mudah disesuaikan dibanding fitur *Cluster Numeric Event.* Algoritma *clustering* yang digunakan adalah *K-means*. Karena nilai data pada masing-masing *field* yang akan diklaster memiliki *range yang* cukup jauh, maka diperlukan tahapan *preprocessing* untuk menormalisasi data. Untuk menghindari perubahan makna pada *fields* yang akan diklaster, maka algoritma *preprocessing* yang digunakan adalah *Standard Scaller*. Gambar 3.12 menunjukkan bagaimana tampilan dari Smart Clustering dari aplikasi *Splunk Machine Learning Toolkit.*

Untuk membantu analisis dalam menentukan nilai K, terdapat dua metode populer sebagai rujukan untuk menentukan nilai klaster yang optimal. Kedua metode tersebut adalah *Elbow Method* (mencari skema klaster dengan SSE (*Sum Square of Errors*) atau WSS (*Within-Cluster-Sum of Squared errors)* yang relatif kecil dan jumlah klaster yang sedikit) dan *Silhoutte Method* (untuk menentukan seberapa tepat setiap objek berada dalam klasternya dan lebar rata-rata siluet yang tinggi menunjukkan pengelompokkan yang baik). Karena pada aplikasi Splunk tidak terdapat fitur untuk menerapkan metode tersebut. *Tools* RStudio digunakan sebagai media untuk menerapkan perhitungan nilai klaster yang optimal dengan ketiga metode tersebut.

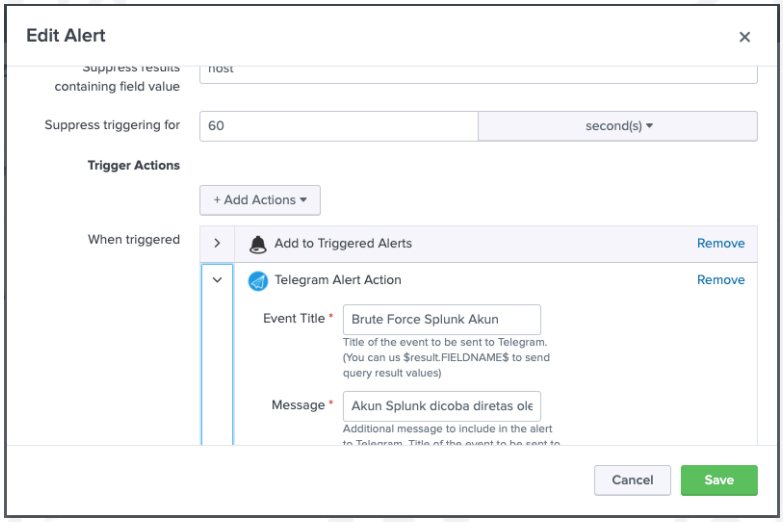


Gambar 3.12 Kegiatan *Clustering* dengan fitur *Smart Clustering*

3.3.4.8 Membuat Alert dan Integrasi ke Telegram

Setelah tahapan-tahapan sebelumnya selesai dilakukan, tahapan paling akhir adalah membuat laporan berdasarkan temuan-temuan yang didapatkan dari hasil visualisasi.

Laporan dibuat diawali dengan melakukan *Search* dan *Reporting.* Setelah itu pada pilihan *Save As*, pilih bagian *alert* kemudian akan muncul pengaturan *alert*, pengaturan *Rule* pemicu *alert* disesuaikan dengan ketentuan dari laporan yang akan dibutuhkan. Pembuatan *bot* pada aplikasi Telegram memanfaatkan *FatherBot* dari Telegram. Kemudian konfigurasi untuk kebutuhan integrasi dengan Splunk adalah dengan memasukkan informasi berupa "*Token ID*" yang didapat saat membuat *bot* dan "*Chat ID*" yang didapat saat *bot* telah di start oleh pengguna. Gambar 3.13 menunjukkan tampilan dalam melakukan konfigurasi untuk integrasi *alert* Splunk dengan *bot* Telegram.



Gambar 3.13 Mengintegrasikan *alert* dengan *bot* Telegram

3.3.5 Pengujian

Kegiatan utama dari penelitian ini adalah menghasilkan tiga macam keluaran, hal ini disesuaikan dengan poin **3.2 Ukuran Keberhasilan**. Mengacu pada poin tersebut, maka pengujian akan dibedakan menjadi tiga bagian. Pada bagian pertama untuk menguji *Rule* yang menjadi hal paling utama dari visualisasi, pengujian akan dilakukan dengan mengujikan *Rule* dengan pilihan data *Log Firewall* dihari-hari lainnya. Apabila visualisasi tetap dapat berjalan dengan baik, maka *Rule* yang digunakan bersifat universal dan dapat digunakan pada data-data dihari lainnya.

Pada bagian kedua, pengujian dilakukan dengan menerapkan *Rule* pada data *Log Firewall* dihari-hari lainnya. Apabila visualisasi tetap dapat menampilkan hasil dengan baik tanpa ada suatu permasalahan. Maka *Rule* yang digunakan bersifat universal dan dapat digunakan pada data-data di hari lainnya.

Pengujian bagian ketiga dilakukan dengan mencoba salah satu *Rule* yang dapat diuji secara *real-time* yaitu pada bagian pengamanan percobaan masuk secara *brute force* ke dalam akun Splunk. *Rule* yang lainnya tidak dapat diujikan secara langsung karena data yang dibutuhkan untuk melakukan pengujian harus bersifat *Real-Time,* namun untuk memastikan apakah *Rule alert* dapat berjalan atau tidak. Maka pengujian dilakukan dengan sedikit memodifikasi dari *Rule* yang telah ditetapkan, hal ini bertujuan hanya untuk memastikan apakah *alert* telah terintegrasi dengan *bot* Telegram.

3.3.6 Analisis

Tahapan analisis akan dilaksanakan setelah pengujian dilakukan. Analisis dilakukan untuk mengamati hasil dari visualisasi *non-Cluster* dan visualisasi *with Cluster.* Hasil dari analisis ini yang nantinya akan menjadi acuan dalam penarikan kesimpulan.

BAB IV
HASIL DAN PEMBAHASAN

4.1 Visualisasi *Log Firewall non-Clustering*

Dalam membuat beberapa jenis visualisasi untuk visualisasi *non-Cluster* dari *Log Firewall,* berikut merupakan kumpulan *Rule* yang telah ditetapkan untuk melakukan deteksi:

```
1. index=* sourcetype="pan:traffic"     // Basic Query data yang diteliti
2. | chart count over app:technology by action  // untuk menampilkan informasi
   action berdasarkan fields app:category
```

Gambar 4.1 *Rule detection traffic by action* based app:technology.

```
1. index=* sourcetype="pan:traffic" // Basic Query data yang diteliti
2. | stats  dc(user) as total_pengguna // Perhitungan seluruh jumlah pengguna
   berdasarkan fields total_pengguna
```

Gambar 4.2 *Rule detection all user by user*.

```
1. index=* sourcetype="pan:traffic" client_location=Indonesia // Basic Query
   data yang diteliti didasarkan hanya pada lokasi negara Indonesia
2. | stats  dc(user) as total_pengguna  // Perhitungan seluruh jumlah pengguna
   berdasarkan fields total_pengguna
```

Gambar 4.3 *Rule detection Indonesia user from client_location=indonesia*.

```
1. index=* sourcetype="pan:traffic"  // Basic Query data yang diteliti
2. | stats count by user             // Perhitungan jumlah user
3. | sort -count          // Visualisasi didasarkan pada nilai paling tinggi
```

Gambar 4.4 *Rule detection most popular user by user*.

```
1. index=* sourcetype="pan:traffic"   // Basic Query data yang diteliti
2. | stats count by client_location // Perhitungan jumlah user pada setiap lokasi
   pengguna (lokasi = negara)
3. | sort -count       // Visualisasi didasarkan pada nilai paling tinggi
```

Gambar 4.5 *Rule detection most popular country by client_location*.

```
1. index=* sourcetype="pan:traffic"    // Basic Query data yang diteliti
2. | iplocation client_ip            // Query alamat ip yang merupakan syarat untuk
   visualisasi jenis Cluster Map
3. | geostats dc(client_ip) by client_location    // Pemanggilan comand untuk
   visualisasi jenis Cluster Map berdasakan lokasi pengguna
```

Gambar 4.6 *Rule detection user location by client_location*.

```
1. index=* sourcetype="pan:traffic"    // Basic Query data yang diteliti
2. | iplocation client_ip              // Query alamat ip yang merupakan syarat
   untuk visualisasi jenis Cloropeth Map
3. | stats dc(user) as Pengguna by client_location  // Menghitung jumlah
   pengguna berdasarkan lokasi pengguna
4. | geom geo_countries featureIdField=client_location  // Pemanggilan command
   untuk visualisasi jenis Cloropeth Map berdasakan bagian pada lokasi pengguna (Bagian =
   Negara)
```

Gambar 4.7 *Rule detection total users in country by client_location*

```
1. index=* sourcetype="pan:traffic"    // Basic Query data yang diteliti
2. | stats dc(dest_port) as num_dest_port dc(dest_ip) as num_dest_ip by
   src_ip      // Menghitung jumlah destinasi port yang dituju berdasarkan ip destinasi
   yang dalam tabel dituliskan num_det_ip berdasar Lokasi pengakses
3. | where num_dest_port >500 OR num_dest_ip >500    // Persyaratan kapan sebuah
   kegiatan akan dicurigai
```

Gambar 4.8 *Rule detection networking & port scanning by src_ip*

```
1. index=* sourcetype="pan:traffic" user=* user!=""    // Basic Query data yang
   diteliti dengan syarat pengguna memiliki username
2. | stats count(eval(action="success")) as successes
   count(eval(action="failure")) as failures by user, app    // Menghitung aksi
   yang sukses dan aksi yang gagal berdasarkan percobaan dari pengguna
3. | where successes>0 AND failures>100  // Peryaratan kapan sebuah kegiatan akan
   dicurigai
```

Gambar 4.9 *Rule detection brute force attack by user, app*

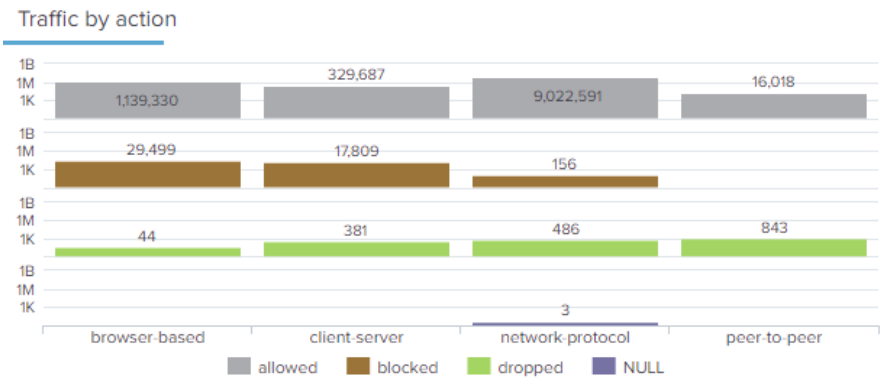
```
1. index=_audit action="login attempt" info=failed user=*     // Query untuk
   menampilkan informasi setiap aksi pengguna yang gagal masuk ke Splunk Web
```

Gambar 4.10 *Rule splunk web logins attemps*

Setelah melakukan banyak percobaan, pertimbangan, serta studi literatur untuk pemilihan visualisasi yang akan ditampilkan pada *dashboard* panel guna memudahkan

pembacaan informasi serta variasi dari visualisasi. Dibawah ini merupakan hasil dari visualisasi yang akhirnya secara final ditetapkan pada *dashboard* panel:

Rule pada Gambar 4.1, dengan memilih jenis visualisasi "*Column Chart",* memberikan hasil visualisasi seperti pada Gambar 10. Jenis visualisasi ini dipilih karena dapat memudahkan dalam pembacaan dan membandingkan data bagian satu dengan bagian lainnya. Pemisahan visualisasi dibedakan berdasarkan *technology* dari aplikasi yang terdeteksi.



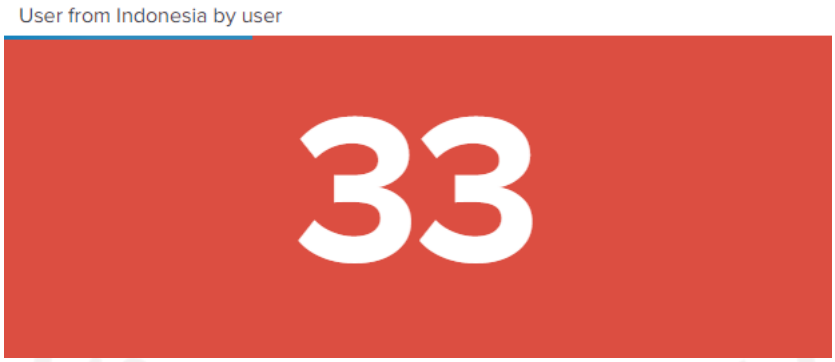
Gambar 4.11 *Visualization traffic by Action*

Rule pada Gambar 4.2, dengan memilih jenis visualisasi "*Single Value",* memberikan hasil visualisasi seperti pada Gambar 4.12. Jenis visualisasi ini dipilih karena informasi yang dibutuhkan hanya berupa angka jumlah dari seluruh pengguna.

Gambar 4.12 *Visualization all user by user*

Rule pada Gambar 4.3, dengan memilih jenis visualisasi "*Single Value",* memberikan hasil visualisasi seperti pada Gambar 4.13. Jenis visualisasi sama dengan jenis visualisasi pada Gambar 4.12, yang membedakan hanya pada background dari visualisasi ini yang berwarna merah. Warna merah dimaksudkan untuk menunjukkan identitas Indonesia. Pemilihan jenis

visualisasi ini karena informasi yang dibutuhkan hanya berupa angka dari jumlah seluruh pengguna yang tercatat berasal dari Indonesia.



Gambar 4.13 *Visualization user from indonesia by client_location=indonesia*

Rule pada Gambar 4.4, dengan memilih jenis visualisasi "*Word Cloud",* memberikan hasil visualisasi seperti pada Gambar 4.14. Untuk mengetahui pengguna dengan akses terbanyak ditunjukkan kata dengan ukuran paling besar. Misalnya pada Gambar 4.14, kata "14421057" adalah kata dengan ukuran paling besar. Maka pengguna dengan nama 14421057 merupakan pengguna dengan jumlah akses terbanyak di jaringan UII. Jenis visualisasi ini dipilih karena visualisasi cenderung lebih modern, simpel, dan mudah untuk dipahami.

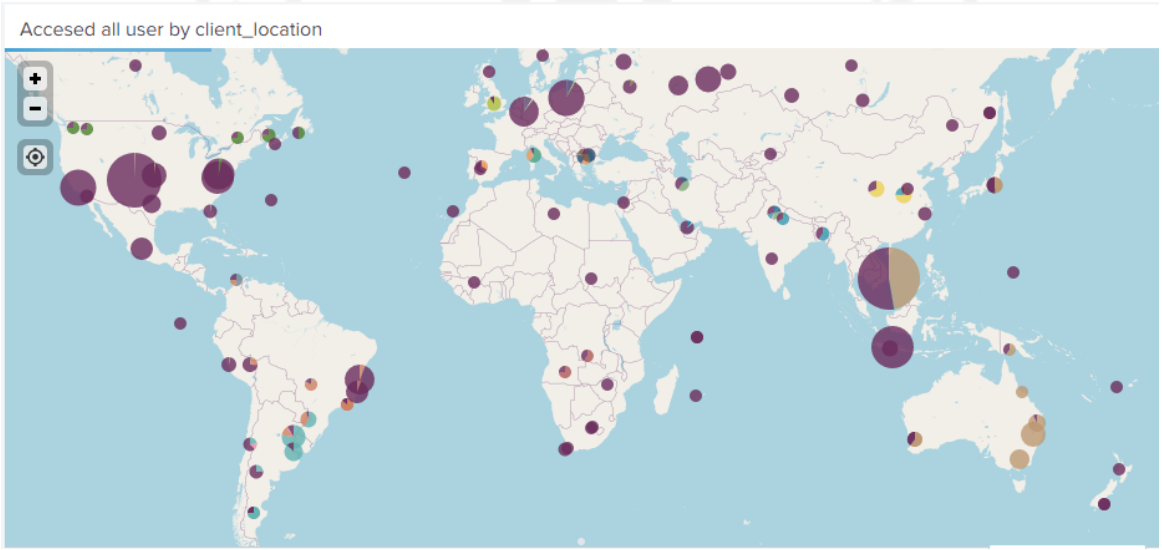
Gambar 4.14 *Visualization most popular user by user*

Rule pada Gambar 4.6, dengan memilih jenis visualisasi "*Word Cloud",* memberikan hasil visualisasi seperti pada Gambar 4.15. Kata "splunk>" pada visualisasi tersebut menunjukkan aplikasi yang digunakan untuk menampilkan visualisasi. Untuk mengetahui negara pengakses terbanyak ditunjukkan dengan kata dengan ukuran paling besar. Misalnya

pada Gambar 4.15, kata “Indonesia” adalah kata dengan ukuran paling besar. Maka negara Indonesia merupakan negara dengan pengakses terbanyak di jaringan UII. Jenis visualisasi ini dipilih karena visualisasi cenderung lebih modern dan mudah untuk dipahami.

Gambar 4.15 *Visualization most popular country by client_location*

Rule pada Gambar 4.7, dengan memilih jenis visualisasi “*Cluster Map”,* memberikan hasil visualisasi seperti pada Gambar 4.16. Bentuk lingkaran yang terdapat pada wilayah negara menunjukkan lokasi dari pengguna. Bentuk lingkaran kecil menunjukkan jumlah pengguna dari sedikit, sedangkan ukuran yang lebih besar menunjukkan jumlah pengguna yang lebih banyak. Visualisasi ini dipilih karena visualisasi cenderung lebih modern, mudah untuk dibaca dan dipahami.



Gambar 4.16 *Visualization location User by client_location*

Rule pada Gambar 4.8, dengan memilih jenis visualisasi "*Cloropeth Map",* memberikan hasil visualisasi seperti pada Gambar 4.17. Visualisasi yang menunjukkan informasi akan muncul apabila salah satu bagian wilayah ditunjuk dengan kursor dari perangkat yang digunakan. Warna dari wilayah menunjukkan jumlah pengguna pada setiap negara. Informasi warna dapat dilihat pada bagian *Legend* yang terdapat pada sebelah kanan bawah visualisasi. Visualisasi ini dipilih karena visualisasi cenderung lebih modern, mudah untuk dibaca dan dipahami.

Gambar 4.17 *Visualization total users in country by client location*

Rule pada Gambar 4.8 dan 4.9 menerapkan jenis visualisasi "*Statistic Table". Rule* pada Gambar 4.8 memberikan hasil visualisasi seperti pada Gambar 4.18. Jenis visualisasi ini dipilih karena kebutuhan informasi *Rule* ini harus didapat secara detail, sehingga jenis tabel sangat cocok untuk menampilkan visualisasi tersebut. Hasil percobaan pada tiga hari yang berbeda menunjukkan "*data not found*" saat menerapkan *Rule* Gambar 4.9. Hal ini menunjukkan bahwa dalam tiga hari pengujian, tidak ada *Bruto Force Attack* pada jaringan UII.

Detecting Networking and Port Scanning by src_ip

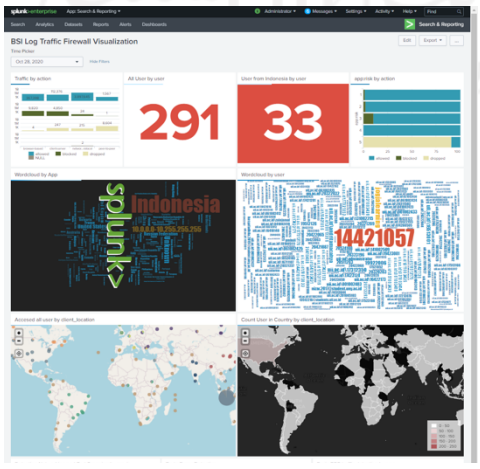
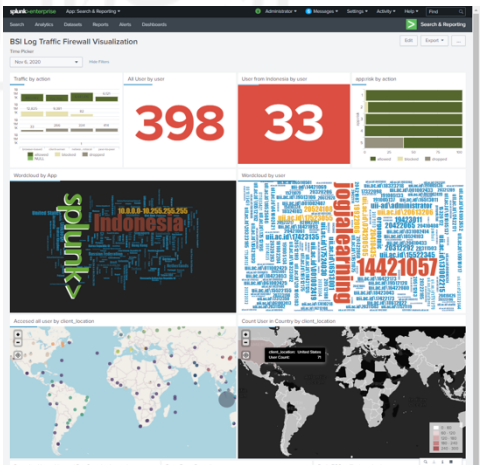
src_ip	num_dest_port	num_dest_ip
10.0.44.18	1620	795
10.10.81.162	15	555
10.10.81.167	9	521
10.10.81.175	8	933
10.10.81.194	8	531
10.10.81.197	10	562
10.10.81.203	11	658
10.10.83.37	9	178745
10.100.240.162	28	1013
10.100.240.198	18	736

« Prev 1 2 3 4 5 6 7 8 9 10 Next »

Gambar 4.18 *Visualization total users in country by client location*

Dalam proses pengujian, dilakukan pengujian pada *Rule* yang telah ditetapkan. Pengujian dilakukan untuk menilai apakah *Rule* telah bersifat universal sehingga dapat digunakan pada data *Log Firewall* di setiap harinya atau tidak. Karena keterbatasan waktu dan *space* penyimpanan data *Log Firewall* setiap harinya berukuran lebih dari 30 *gigabyte*. Maka pengujian hanya dilakukan pada dua hari yang berbeda. Hari ditentukan berdasarkan ukuran *Log Firewall* yang lebih besar dibanding hari-hari lainnya. Hari yang pertama adalah pada Rabu, 28 Oktober 2020. Kemudian hari yang kedua adalah pada Jumat 6 November 2020. Berikut tabel perbandingan dari kedua hari tersebut:

Tabel 4.1 Pengujian *Rule* visualisasi *non-cluster*

Waktu	Rabu, 28 Oktober 2020	Jumat, 6 November 2020
Visualisasi		

Ukuran File	40.8 *gigabyte*	43.3 *gigabyte*
Total *Event*	104.484.528 *events*	111.866.547 *events*
Problem	Tidak ada	Tidak ada

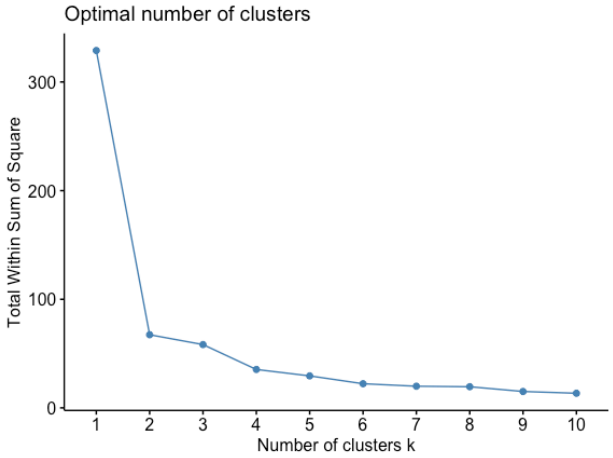
Dari hasil pengujian, berdasarkan Tabel 4.1 dapat disimpulkan bahwa *Rule* yang telah ditetapkan telah bersifat universal. Karena dapat diintegrasikan pada hari yang berbeda dan tidak memiliki masalah. Hal ini terjadi karena proses yang dilakukan hanya sekedar melakukan pencarian data yang telah ada kemudian ditampilkan dalam bentuk visualisasi, tidak ada proses perhitungan matematika ataupun statistika. *Rule* tidak akan ada masalah apabila data yang digunakan merupakan data *Log Firewall* dengan *Traffic Type* dari Palo Alto *Firewall.* Selain data tersebut, maka *Rule* tidak dapat digunakan, sebab penulisan *Rule* dari setiap perangkat memiliki konvensi masing-masing.

4.2 Visualisasi *Log Firewall with Clustering*

Pembuatan klasterisasi untuk visualisasi *with cluster* dari *Log Firewall* dilakukan berdasarkan pada *field* "_time" dan "app:category". Kedua *field* tersebut dipilih dengan harapan untuk mendapatkan informasi terkait kecenderungan pengguna saat membuka aplikasi berdasarkan waktu akses. *Field app:category* dipilih karena *Field* tersebut memberikan informasi secara umum tentang aplikasi yang digunakan oleh pengguna sehingga cakupan informasi lebih luas. Di dalam *field "*app:category" terdapat enam bagian *fields* yang menunjukkan kategori dari aplikasi yang telah ditetapkan oleh Palo Alto *Firewall*. Keenam *fields* tersebut yaitu "networking", "general-internet", "collaboration", "business-systems", "media", dan "unknown". *Field "*unknown*"* berisikan aplikasi yang terdeteksi, namun tidak masuk dalam daftar kategorisasi dari Palo Alto *Firewall.* Karena terdapat batasan pada jumlah *fields* yang dapat dipilih untuk melakukan *clustering* menggunakan Splunk. Maka semua *fields* dipilih kecuali *field "*unknown*",* karena batas maksimal *fields* yang dapat dipilih hanya lima.

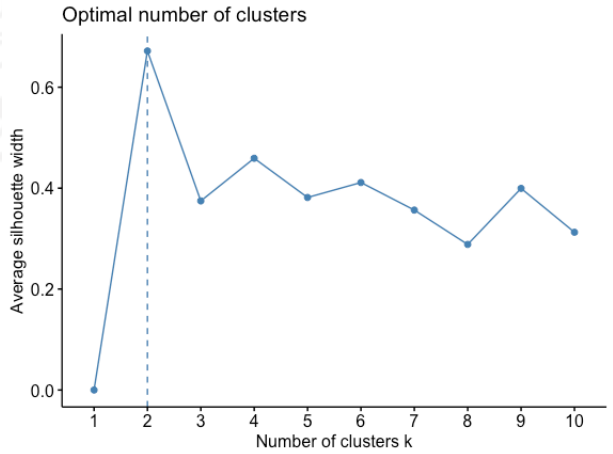
Sebelum tahapan *clustering* dilakukan, tahapan *preprocessing* terlebih dahulu dilakukan. Hasil dari tahapan *preprocessing* menggunakan algoritma *Standard Scaller with respect to Mean and Standart Deviation* membuat *fields* baru dengan aturan penamaan SS_<nama *fields* awal>. Aturan tersebut merupakan aturan baku dari Aplikasi *Splunk Machine Learning Toolkit* sehingga tidak dapat diubah. Kemudian dalam melakukan *clustering,* hasil dari penerapan dengan menggunakan *Elbow Method* menunjukkan bahwa nilai klaster atau K yang paling

optimal ditunjukkan pada K = 2. Nilai K = 2 memiliki nilai sudut terkecil (nilai paling optimal) selanjutnya diikuti oleh nilai K = 4 yang merupakan nilai optimal kedua. Gambar 4.19. menunjukkan hasil dari perhitungan dengan *Elbow Method.*



Gambar 4.19 *Elbow Method.*

Merujuk pada bagian **Bab III** poin **3.3.4.7 Klasterisasi** maka metode kedua yang digunakan sebagai rujukan untuk menentukan nilai klaster yang optimal adalah dengan *Silhoutte Method.* Hasil dari penerapan metode *Silhoutte Method* menunjukkan bahwa nilai klaster atau K = 2 memaksimalkan nilai rata-rata siluet (nilai paling optimal) dan K = 4 merupakan nilai kluster optimal kedua. Gambar 4.20. menunjukkan hasil dari perhitungan dengan *Silhoutte Method*. (*Source Code* yang digunakan untuk melakukan komputasi *Elbow Method* dan *Silhoutte Method* dalam bahasa R pada RStudio dapat dilihat pada lampiran C).



Gambar 4.20 *Silhoutte Method.*

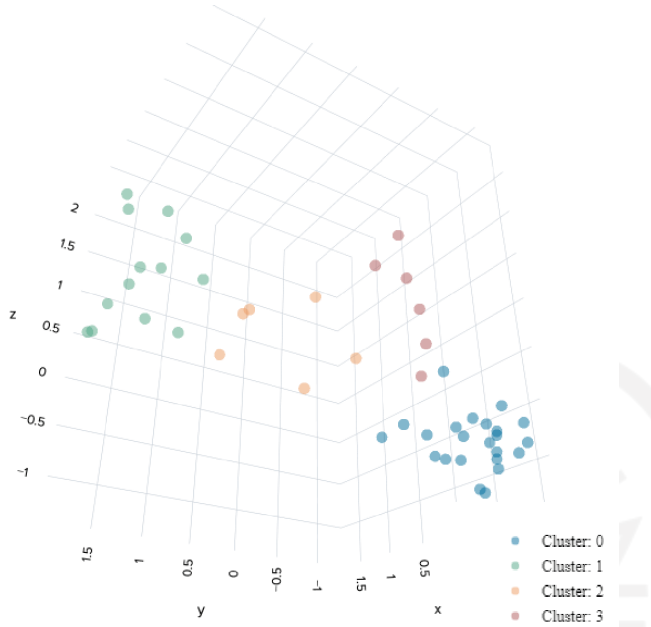
Dalam menentukan nilai klaster atau K yang akan digunakan dalam klastering ini, merujuk pada hasil dari *Elbow Method* dan *Silhoutte Method* yang menujukkan K = 2 merupakan nilai klaster paling optimal, kemudian K = 4 merupakan nilai klaster optimal kedua. Dalam kegiatan penelitian ini, ditetapkanlah nilai klaster yang akan digunakan yaitu K = 4. Selain didasarkan dari hasil evaluasi *Elbow Method* dan *Silhoutte Method,* penggunaan nilai K = 4 juga didasarkan pada variasi klaster yang lebih beragam dibanding K = 2. Variasi informasi yang lebih beragam serta tetap memperhatikan hasil yang optimal akan memberikan banyak informasi yang lebih akurat.

Secara teori memang tidak ada perhitungan secara khusus tentang penentuan nilai K yang paling tepat pada algoritma K-means. Namun, menurut hasil studi literatur yang telah dilakukan, nilai K dapat ditetapkan sebagai acuan apabila setiap *cluster* telah berisi anggota sehingga terdapat persebaran dan tidak ada *cluster* yang kosong. Selain itu, apabila perhitungan dengan aturan yang sama dilakukan secara berulang-ulang, maka hasil cenderung menunjukkan kemiripan yang tinggi (Guttag, 2017). Setelah dilakukan beberapa kali percobaan, nilai K = 4 memenuhi syarat-syarat tersebut. Gambar 4.21. menunjukkan *Rule* yang digunakan.

```
1. index=* sourcetype="pan:traffic"    // Query basic dari data yang diteliti
2. | timechart count by app:category  // Menampilkan timechart untuk visualisasi
   berdasarkan app:categroy dan perhitungan untuk setiap kategori
3. | fit StandardScaler "business-systems" "collaboration" "general-
   internet" "media" "networking" with_mean=true with_std=true    // Tahapan
   Preprocessing dengan Standart Scaler dengan Field yang telah dipilih
4. | fit KMeans "SS_business-systems" "SS_collaboration" "SS_general-
   internet" "SS_media" "SS_networking" k=4    // Tahapan Klusterisasi dengan
   algoritma K Means K = 4 dengan Field hasil Preprocessing
5. | eval clusterId= "Cluster: " + cluster, x='SS_general-internet',
   y='SS_collaboration', z='SS_business-systems'  // Penetapan fields untuk
   informasi dan syarat visualisasi
```

Gambar 4.21 *Rule Clustering with K-means algorithm for 3D Scatterplot.*

Rule pada Gambar 4.21. merupakan *Rule* yang dibuat khusus untuk jenis visualisasi "*3D Scatterplot"* sebab jenis visualisasi tersebut mengharuskan untuk menentukan nilai x, y, dan z secara langusng. Hasil dari visualisasi ditunjukkan pada Gambar 4.22. Jenis visualisasi ini dipilih karena visualisasi dari setiap *cluster* dapat dengan mudah dilihat. Persebaran *cluster* dapat dilihat dan dibaca dengan jelas. Apabila kursor diarahkan pada salah satu *cluster,* maka akan muncul informasi dari nilai x, y, dan z.



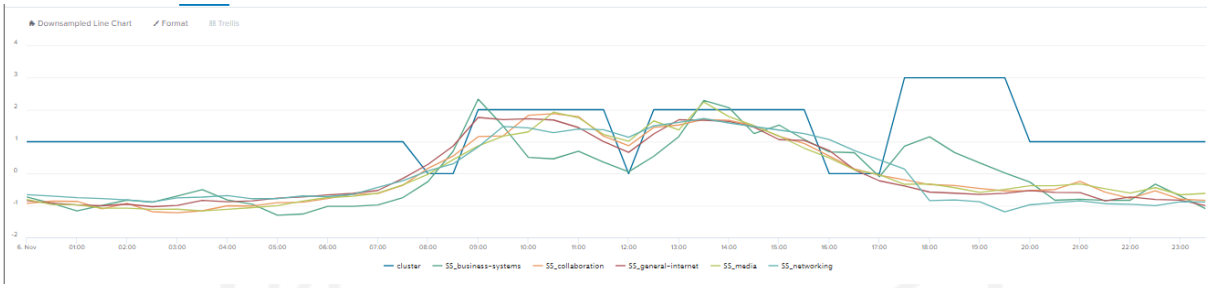
Gambar 4.22 *Visualization Clustering K=4 with 3D Scatterplot.*

Nilai x merupakan nilai dari "SS_general-internet", y merupakan nilai "SS_collaboration", dan z merupakan nilai "SS_business-system". Pemilihan nilai x, y, dan z dipilih hanya untuk tujuan mempermudah dalam pembacaan informasi dengan visualisasi jenis *3D Scatterplot.* Setelah dilakukan banyak percobaan dan juga mencoba memahami visualisasi, kelemahan dari visualisasi jenis *3D Scatterplot* ini adalah informasi yang disajikan tidak dapat dikombinasikan dengan waktu (*field "*_time*"*). Sehingga untuk mendapat informasi yang merupakan tujuan utama dari penerapan klaster yaitu melihat kecenderungan pengguna saat membuka aplikasi berdasarkan waktu akses, sulit didapatkan.

```
1. index=* sourcetype="pan:traffic" "app:category"="*" // Query basic dari data
   yang diteliti dengan menyertakan seluruh category
2. | timechart count by app:category    // Menampilkan timechart untuk visualisasi
   berdasarkan app:categroy dan perhitungan untuk setiap kategori
3. | fit StandardScaler "business-systems" "collaboration" "general-
   internet" "media" "networking" with_mean=true with_std=true // Tahapan
   Preprocessing dengan Standart Scaler dengan Field yang telah dipilih
4. | fit KMeans "SS_business-systems" "SS_collaboration" "SS_general-
   internet" "SS_media" "SS_networking" k=4     // Tahapan Klusterisasi dengan
   algoritma K Means K = 4 dengan Field hasil Preprocessing
5. | table _time, cluster, SS_business-systems, SS_collaboration,
   SS_general-internet, SS_media, SS_networking  // Menetapkan informasi yang akan
   ditampilkan pada visualisasi dan tabel visualisasi
```

Gambar 4.23 *Rule Clustering with K-means algorithm*.

Gambar 4.23 menunjukkan *Rule clustering* yang dapat digunakan secara umum untuk berbagai jenis visualisasi. Dengan memanfaatkan jenis visualisasi "*Downsapled Line Chart",* maka Rule pada Gambar 4.23 menghasilkan visualisasi seperti pada Gambar 4.24.



Gambar 4.24 *Visualization Clustering K=4 with Downsampled Line Chart*

Visualisasi jenis *Downsapled Line Chart* dipilih karena dapat dikombinasikan dengan *field "*_time" sehingga untuk pembacaan informasi kecenderungan aplikasi yang digunakan dalam waktu akses tertentu, dapat dengan mudah dilihat dan dibaca. Kelemahan dari visualisasi jenis ini adalah *cluster* tersebar secara acak, namun informasi *cluster* tetap dapat dilihat pada visualisasi secara langsung.

Setelah *clustering* berhasil dilakukan, dari Gambar 4.22, Gambar 4.24, dan dibantu oleh tabel informasi dari hasil *clustering* yang berada di bawah visualisasi tersebut (nilai dari tabel terdapat pada lampiran A). Dapat ditarik kesimpulan bahwa informasi pada hari Jumat, 6 November 2020 ditunjukkan pada Tabel 4.2. Informasi yang didapat pada hari Rabu, 28 Oktober 2020 juga memiliki hasil yang tak jauh berbeda dengan hasil dari hari Jumat, 6 November 2020. Hanya saja saat pukul 00.00 tercatat *field* "business-system", "networking*",* dan *"*general-internet" menunjukkan aktivitas di atas rata-rata. Kemudian pukul 00.30 *field* "networking*"* dan *"*general-internet" menunjukkan aktivitas di atas rata-rata. Hal ini merupakan kondisi normal dari perkuliahan dengan sistem *full online* karena pandemi.

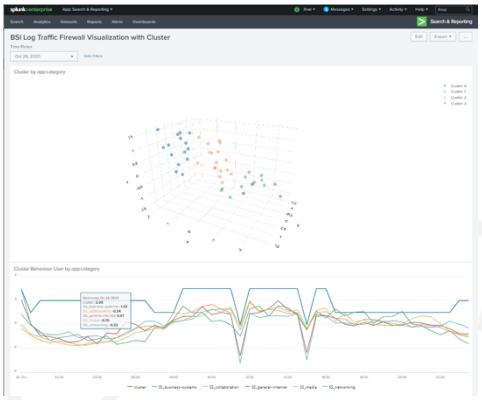
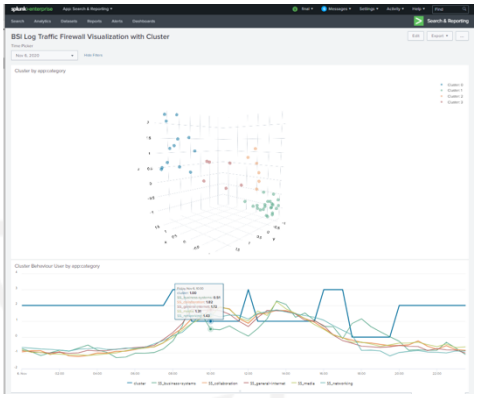
Tabel 4.2 Pembacaan *Cluster*

Cluster	**Informasi yang diperoleh**
0	Menunjukkan bahwa segala macam aktivitas yang masuk dalam kategori *cluster* 0 adalah dominan di atas rata-rata. Hal ini memberi informasi bahwa pada bagian *cluster* 1, akses jaringan UII masuk kategori **tinggi**. Tercatat hanya *field* "*business-system*" dan *field "media"* yang terkadang berada di bawah rata-rata. Apabila dikaitakan dengan waktu

	akses, maka hal ini sangat dimungkinkan karena waktu *cluster* 0 ditunjukkan pada pukul 08:00 – 08.30 WIB, 12.00 WIB, dan 16.00 – 17.00 WIB. Waktu tersebut merupakan waktu perkuliahan sedang dilakukan namun saat waktu perpindahan perkuliahan.
1	Menunjukkan bahwa segala macam aktivitas yang masuk dalam kategori *cluster* 1 adalah di bawah rata-rata. Hal ini memberi informasi bahwa pada bagian *cluster* 1, akses jaringan UII masuk kategori **sangat rendah**. Apabila dikaitakan dengan waktu akses, maka hal ini sangat dimungkinkan karena waktu *cluster* 1 ditunjukkan pada pukul 00:00 – 07.30 WIB dan 20.00 – 23.30 WIB. Waktu tersebut merupakan waktu perkuliahan sedang tidak dilakukan.
2	Menunjukkan bahwa segala macam aktivitas yang masuk dalam kategori *cluster* 2 adalah di atas rata-rata. Hal ini memberi informasi bahwa pada bagian *cluster* 2, akses jaringan UII masuk kategori **sangat tinggi**. Tercatat bahwa seluruh *fields* menunjukkan aktivitas di atas rata-rata. Apabila dikaitakan dengan waktu akses, maka hal ini sangat dimungkinkan karena waktu *cluster* 2 ditunjukkan pada pukul 09:00 – 11.30 WIB dan 12.30 – 15.30 WIB. Waktu tersebut menunjukkan waktu perkuliahan sedang berjalan. Terjadi penurunan pada pukul 12.00, namun penurunan tersebut tidak menunjukkan hingga angka di bawah rata-rata. Namun hanya menunjukkan akses yang sedikit berkurang saja. Hal ini mungkin terjadi karena pada waktu tersebut adalah waktu istirahat.
3	Menunjukkan bahwa segala macam aktivitas yang masuk dalam kategori *cluster* 3 adalah telah dekat dengan di bawah rata-rata. Hal ini memberi informasi bahwa pada bagian *cluster* 1, akses jaringan UII masuk kategori **rendah.** Tercatat hanya akses "*business-system*" (aplikasi yang berjalan adalah aplikasi yang biasanya dominan digunakan oleh *team* BSI UII) yang berada di atas rata-rata, selain itu seluruh *fields* menunjukkan hasil berada di bawah rata-rata. Apabila dikaitkan dengan waktu akses, maka hal ini sangat dimungkinkan karena waktu *cluster* 3 ditunjukkan pada pukul 17.30 - 19.30 WIB. Waktu tersebut menunjukkan waktu perkuliahan telah usai.

Pada bagian pengujian, prinsipnya adalah sama seperti pengujian pada bagian visualisasi *non-Cluster*. Tabel 4.3 menunjukkan perbandingan dari kedua hari tersebut.

Tabel 4.3 Pengujian *Rule* visualisasi *with Cluster*

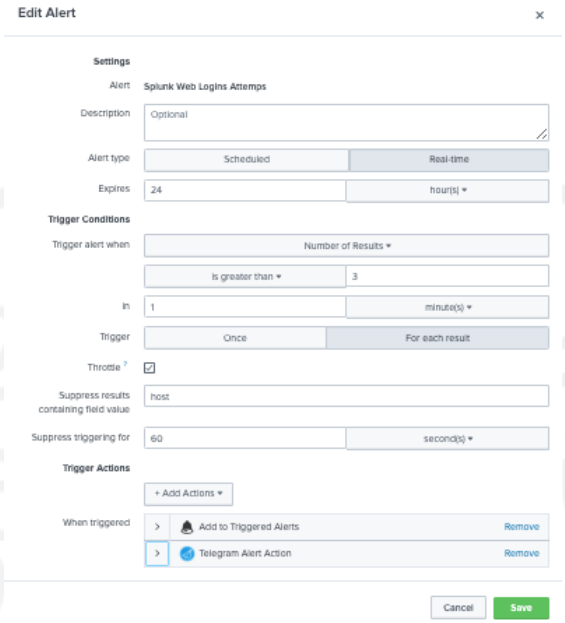
Waktu	Rabu, 28 Oktober 2020	Jumat, 6 November 2020
Visualisasi		
Algoritma *Preprocessing*	*Standard Scaler with respect to Mean and Standart Deviation*	*Standard Scaler with respect to Mean and Standart Deviation*
Algoritma *Clustering*	*K-Means with K = 4*	*K-Means with K = 4*
Problem	Tidak ada	Tidak ada

Berdasarkan Tabel 4.3 dapat disimpulkan bahwa *Rule* yang telah ditetapkan telah bersifat universal. Karena dapat diintegrasikan pada hari yang berbeda dan tidak memiliki masalah. Hal ini terjadi karena pemilihan *fields* yang digunakan sudah tepat. Sebab dalam pengamatan data yang telah dilakukan, *fields "*app:category" memiliki anggota tetap dan tidak pernah berubah-ubah setiap harinya. Hal ini merupakan salah satu keuntungan dari *fields* yang menampilkan informasi bersifat umum. Untuk mengetahui informasi aplikasi secara lebih lanjut, *fields "*app:category" dikolaborasikan dengan *fields "*app:subcategory" dan "application*"* (hasil dapat dilihat pada lampiran B).

4.3 Integrasi *Alert System* dengan *Bot* Telegram

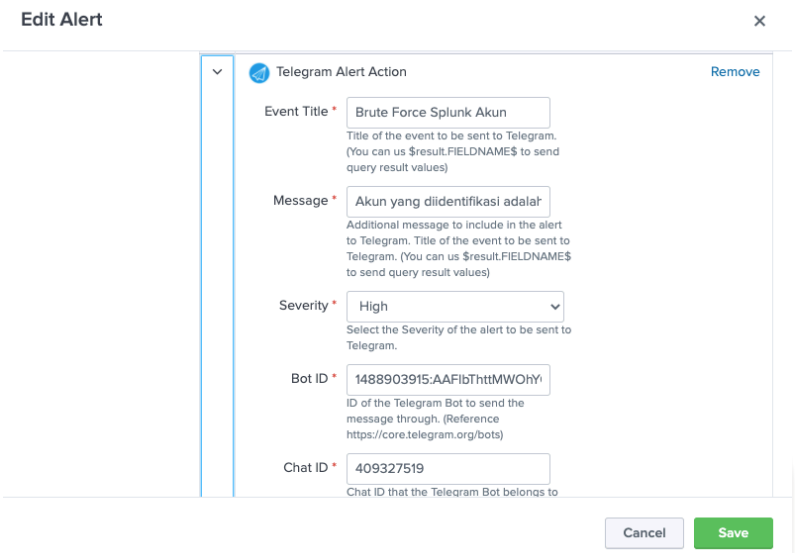
Untuk integrasi *alert* dengan *Bot* Telegram, telah dibuat tiga Jenis *alert.* Bagian *alert* yang pertama adalah mendeteksi upaya masuk yang tidak sah pada akun Splunk Web. Bagian *alert* yang kedua adalah memberitahu informasi jumlah pengguna apabila melebihi batas dari pengguna seharusnya dalam kurun waktu yang telah ditentukan. Bagian *alert* yang ketiga adalah mendeteksi percobaan *brute force attack* pada jaringan UII.

Gambar 4.11 menunjukkan *Rule* pada bagian *Search & Reporting,* kemudian Gambar 4.23 menunjukkan *Rule* yang memicu kapan *alert* akan dikirimkan pada percobaan *testing alert* bagian pertama.



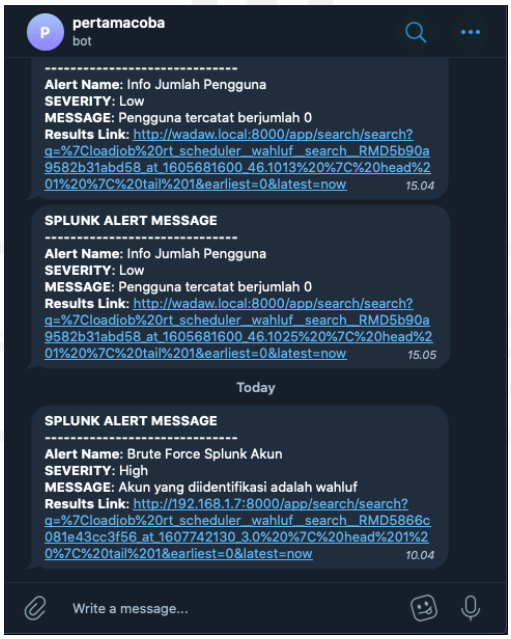
Gambar 4.25 *Rule* pemicu *Alert* bagian pertama

Gambar 4.24 menunjukkan konfigurasi yang dilakukan ketika mencoba menghubungkan Splunk dengan *bot* Telegram. Untuk menambah informasi yang diterima, dalam penulisan Judul *Event* dan pesan yang dikirimkan dapat berisi nilai yang diambil dari *field* saat pembuatan *Search & Reporting.* Contohnya adalah informasi *username* pengguna yang gagal dalam melakukan *log in* pada aplikasi Splunk dengan menuliskan $result.user$ pada bagian pesan ataupun judul *Event.* Pada bagian level *severity* biasanya ditetapkan melalui kesepakatan oleh organisasi yang terkait karena ada beberapa hal yang perlu dipertimbangkan seperti ukuran dari tim, frekuensi insiden, tinggi dan rendahnya *traffic* dari layanan (Atlasian, 2019).



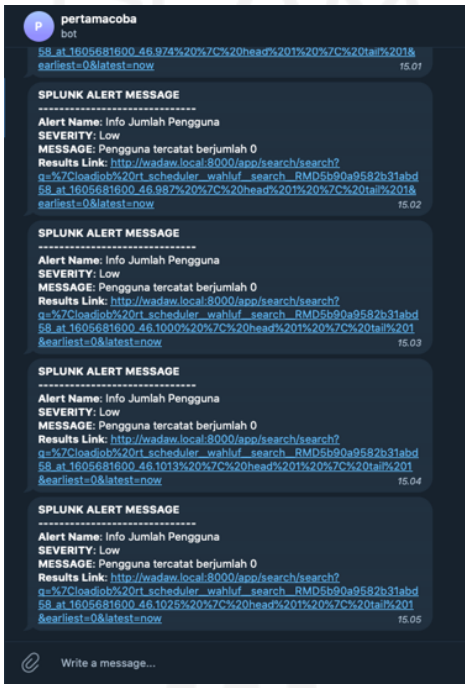
Gambar 4.26 Integrasi Splunk dengan *bot* Telegram

Untuk mencoba mendapatkan pemberitahuan berupa *alert* pada *bot* Telegram. Dilakukan percobaan dengan melakukan kesalahan selama empat kali dalam kurun waktu satu menit saat melakukan *Log*-in untuk melanggar *Rule* Pemicu agar notifikasi dikirimkan, Gambar 4.25 menunjukkan hasil setelah percobaan tersebut.



Gambar 4.27 *Alert* dari percobaan bagian pertama

Pada bagian pengujian *alert* yang kedua, hasilnya tidak muncul. Hal ini terjadi karena data yang akan dibaca pada *Rule pemicu* bersifat sedang berlangsung (*present*), untuk data yang telah lalu (*past*) data hanya akan terbaca sebatas pada bagian *Search & Reporting* saja. Karena data *Log Firewall* yang digunakan merupakan data yang telah lalu, maka dilakukanlah perubahan pada menu *setting* bagian "*Trigger Alert When*" menjadi "*Per-Result*" untuk memastikan apakah integrasi dari *Rule* tersebut dapat berjalan. Gambar 4.26 menunjukkan hasil dari perubahan tersebut.



Gambar 4.28 *Alert* dari percobaan bagian kedua

Alert memberikan pemberitahuan ke *bot* Telegram namun jumlah *user* yang terdeteksi adalah nol. Sebab dalam waktu *real-time,* tidak ada data yang dibaca sehingga data *user* berjumlah nol. Untuk pengujian pada bagian *alert* yang ketiga adalah sama seperti kasus pada bagian kedua. Apabila administrator yang akan melakukan akses *bot* Telegram berupa *team*. Maka solusi terbaik adalah mengundang *bot* Telegram kedalam *group* yang berisikan seluruh administrator. Keuntungan di dalam *group* adalah administrator dapat berkomunikasi secara langsung dengan lainnya di *group* Telegram apabila terdapat *alert* yang membutuhkan penanganan secara *team*. Untuk konfigurasi kedalam Splunk adalah dengan memasukkan "*Chat ID*" dari *group* Telegram pada bagian saat penerapan *Rule* pemicu Gambar 4.24.

4.4 Ananlisis Hasil

Hasil visualisasi *non-Cluster* menunjukkan banyak informasi yang dapat diperoleh dari *Log Firewall* untuk kebutuhan informasi bagi administrator. Sumber informasi berasal dari *field* atau gabungan dari beberapa *field* pada tiap-tiap *event* kemudian dikombinasikan dengan jenis visualisasi yang sesuai dengan karakter informasi. Kemampuan untuk memahami jenis visualisasi yang sesuai dengan karakter informasi merupakan poin penting untuk menghasilkan visualisasi yang mudah dipahami. Kebutuhan pembuatan visualisaisi secara khusus bagi masing-masing instansi, secara dasar dapat diperoleh melalui hasil visualisasi *non-Cluster* ini.

Hasil visualisasi *with Cluster* menunjukkan informasi yang lebih mendalam dapat diperoleh melalui rangkaian proses penerapan teknik *machine learning* dengan pendekatan *Data Mining.* Kemampuan memahami hubungan sebab akibat pada tiap-tiap *field* dan pemilihan teknik pengolahan data menjadi poin penting untuk menghasilkan informasi yang lebih mendalam. Informasi yang diperoleh dari visualisasi *with Cluster* dapat menjadi acuan untuk kebutuhan pengamanan sistem, kebutuhan pengambilan kebijakan, dan kebutuhan instansi lainnya. Dalam kebutuhan pembuatan visualisasi secara khusus bagi masing-masing instansi, secara lebih lanjut dapat diperoleh melalui hasil visualisasi *with Cluster* ini.

Hasil pembuatan *alert* sistem menunjukkan bahwa kebutuhan

Secara umum penerapan teknologi *Security Information and Event Management* dengan menggunakan aplikasi *enterprise* Splunk pada tahapan fungsi untuk visualisasi *Log Firewall* dan *alert* sistem cukup mudah diterapkan. Namun, pada tahapan penerapan secara langsung di lingkungan jaringan UII, perlu untuk mempersiapkan komputer dengan spesifikasi yang harus disesuaikan dengan kebutuhan pengolahan *Big Data* yang akan diproses secara *real-time.* Penelitian ini hanya memanfaatkan sebagian dari fungsi yang ada pada aplikasi Splunk. Oleh karena itu, sangat memungkinkan apabila nantinya dilakukan optimalisasi seluruh fitur-fitur yang terdapat pada aplikasi Splunk.

Saat ini Badan Sistem Informasi sedang gencar dalam menerapkan sistem dengan sangat memperhatikan sudut pandang *enterprise* (menerapkan teknologi-teknologi mutakhir). Hal tersebut tentunya sangat tepat dilakukan karena sebagai sebuah instansi, UII harus berpikir secara visioner dalam menentukan keputusan. Dengan sudut pandang *enterprise*, tentunya merupakan keputusan yang tepat apabila BSI UII menerapkan teknologi *Security Information and Event Management* dengan menggunakan aplikasi Splunk. Selain karena konfigurasi yang dapat dilakukan secara fleksibel menyesuaikan kebutuhan khusus dari BSI UII. Penerapan

Splunk juga akan memberikan keseimbangan dalam proses integrasi dengan sistem-sistem enterprise lainnya.

BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan hasil yang telah dipaparkan, penelitian berhasil dilakukan dan sesuai dengan tujuan awal yaitu membuat visualisasi *non-cluster,* visualisasi *with cluster,* dan membuat *alert* yang terintegrasi dengan *bot* Telegram. Diharapkan hasil dari visualisasi yang telah dibuat dapat memberikan kemudahan bagi administrator dalam pembacaan informasi untuk menambah informasi yang terjadi pada *traffic* jaringan UII. Bahkan untuk kebutuhan lebih lanjut, identifikasi data *Log Firewall* dengan pendekatan *data mining* ini dapat digunakan untuk keperluan investigasi.

Dapat ditarik kesimpulan bahwa penerapan teknologi *Security Information and Event Management* dengan *tools* Splunk pada level organisasi yang telah menerapkan sistem *enterprise* sangatlah dibutuhkan. Selain dapat memudahkan administrator dalam pembacaan dan kemudahan di dalam mendapatkan informasi pada *traffic* jaringan UII yang terjadi secara *real-time.* Penerapan teknologi SIEM akan dapat memberi banyak masukan bagi pihak-pihak yang memiliki kepentingan dalam menentukan kebijakan karena tren penggunaan perangkat pintar semakin meningkat. Informasi-informasi seperti kebiasaan pengguna aplikasi atau informasi penting lainnya akan dapat diperoleh dari *Log Firewall* jaringan UII.

5.2 Saran

Setelah dilakuakn penarikan kesimpulan dari hasil penelitian, adapun beberapa saran untuk penelitian terkait berikutnya adalah sebagai berikut:

1. Besar harapannya untuk penelitian selanjutnya adalah menambah data dari perangkat penghasil *Log* yang berbeda, sehingga data tidak hanya bersumber dari *Log Firewall*. Oleh karena itu, penerapan *Security Information and Event Management* (SIEM) lebih dapat bermanfaat.
2. Melakukan eksplorasi lebih jauh terkait penerapan pendekatan *Data Mining* untuk memberikan visualisasi yang lebih memudahkan pengguna dalam membaca *big data* menggunakan teknik ataupun algoritma *machine learning* yang lainnya.

3. Pada bagian *bot* Telegram dapat dilakukan penambahan fitur untuk melakukan *request* informasi agar administrator tidak hanya menerima informasi apabila terdapat *alert* saja.

DAFTAR PUSTAKA

Anastasov, I., & Davcev, D. (2014). SIEM implementation for global and distributed environments. *2014 World Congress on Computer Applications and Information Systems, WCCAIS 2014*. https://doi.org/10.1109/WCCAIS.2014.6916651

Arfanudin, C. (2019). *Real Time Analisis Keamanan Router di Jaringan dengan Security Information and Event Management (SIEM) dan Implikasinya pada Indeks Keamanan Informasi (KAMI) Studi Kasus : Dinas Komunikasi dan Informatika Kota Tegal* [Universitas Islam Indonesia]. https://dspace.uii.ac.id/handle/123456789/15873

Atlasian. (2019). *Incident Management*. https://www.atlassian.com/incident-management/kpis/severity-levels

Azaim, H., & Nuryanto, W. (2017). *Mengenal Confidentiality, Integrity, dan Availability Pada Keamanan Informasi*. Netsec.Id. https://netsec.id/confidentiality-integrity-availability-keamanan-informasi/

Brown, D. (2016). Finding Bad with Splunk. *SANS.Edu*.

BSI, A. (2017). *Sekilas Tentang BSI*. https://bsi.uii.ac.id/sekilas-bsi/

Chancey, J., & Penn, J. (2018). *Analyticsdriven Security Power Next Generation SIEMS*.

Cisco, N. (2020). *Pendahuluan Tentang Keamanan Cyber*. http://static-course-assets.s3.amazonaws.com/CyberSec2.1/id/index.html#1.2.1.2

Cloudflare. (2020a). *What Is a Next-Generation Firewall (NGFW)? | NGFW vs. FWaaS*. CloudFlare. https://www.cloudflare.com/learning/cloud/what-is-a-next-generation-firewall/

Cloudflare. (2020b). *What Is a Next-Generation Firewall (NGFW)? | NGFW vs. FWaaS*. https://www.cloudflare.com/learning/cloud/what-is-a-next-generation-firewall/

Dairinram, P., Wongsawang, D., & Pengsart, P. (2013). SIEM with LSA technique for Threat identification. *IEEE International Conference on Networks, ICON*. https://doi.org/10.1109/ICON.2013.6781951

Dewanto, A. P. (2018). *Penetration Testing Pada Domain Uii.Ac.Id Menggunakan OWASP 10*.

Dqindia. (2017). *RSA Releases NetWitness Suite to Optimize Security Operations*. Dataquest. https://www.dqindia.com/rsa-releases-netwitness-suite-to-optimize-security-operations/

Efrain, T., Aronson, J. E., & Lian, T. P. (2005). *Decision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1*. Andi.

Erwinsyah, Y. B. (2019). *Konsolidasi dan Visualisasi Log Server BSI UII Menggunakan Elk Stack*.

Fayyad, U., Smyth, P., Piatetsky-Shapiro, G., & Uthurusamy, R. (1996). *Advances in Knowledge Discovery and Data Mining* (First). AAAI Press.

Filkins, B. (2019). *An Evaluator's Guide to NextGen SIEM*.

Firewall.CX. (2020). *Palo Alto Firewall Application-Based Policy Enforcement (App-Id), User Identification (User-Id) And Application Control Centre (Acc) Features For Enterprise Networks*. Palo Alto Firewall Application-Based Policy Enforcement (App-Id), User Identification (User-Id) And Application Control Centre (Acc) Features For Enterprise Networks

Gartner. (2020). *Security Information And Event Management (SIEM)*. Gartner Glossary. https://www.gartner.com/en/information-technology/glossary/security-information-and-event-management-siem

Guttag, J. (2017). *Introduction to Computational Thinking and Data Science | Lecture 12: Clustering*. MITOpenCourseWare. https://www.youtube.com/watch?v=esmzYhuFnds

Haikal, A. (2017). *Pentingnya IT Risk Management Dalam Mendukung Keberlangsungan Bisnis*. https://netsec.id/pentingnya-it-risk-management/

Kumar, P. N. (2015). *SIEM Architecture*. https://pt.slideshare.net/pathinishanth/siem-architecture/4

Opentext. (2020). *Information Security and Privacy*. https://www.opentext.com/products-and-solutions/business-needs/information-governance/ensure-compliance/information-security-and-privacy

Perciballi, S. (2015). *Design Correlation Rules to Get the Most Out of Your SIEM*. Paloalto Network. https://blog.paloaltonetworks.com/2015/08/design-correlation-rules-to-get-the-most-out-of-your-siem/

Rakhmadhani, Syaifuddin, & Sari, Z. (2019). Integrasi Visualisasi Modern Honey Network (MHN) dengan Splunk. *SENTRA 2019*, 167–172.

Splunk.inc. (2018). Splunk Fundamentals 1. In *Manual Book* (Copyrigth, p. 5). Splunk.Inc.

Splunk. (2020). *Splunk® Machine Learning Toolkit User Guide 5.2.0*. Splunk Inc. https://docs.splunk.com/Documentation/MLApp/5.2.0/User/WelcometoMLTK

Vianello, V., Gulisano, V., Jimenez-Peris, R., Patiño-Martínez, M., Torres, R., Díaz, R., & Prieto, E. (2013). A Scalable SIEM Correlation Engine and Its Application to The Olympic Games it Infrastructure. *Proceedings - 2013 International Conference on*

Availability, Reliability and Security, ARES 2013, 625–629. https://doi.org/10.1109/ARES.2013.82

Wulur, F. C., & Sembiring, I. (2015). *Klasifikasi Alert pada Intrusion Detection System Menggunakan Algoritma K-Means*. *December*, 2–4.

Zope, A. R., Vidhate, A., & Harale, N. (2013). Data Mining Approach in Security Information and Event Management. *International Journal of Future Computer and Communication*, *2*(2), 80–84. https://doi.org/10.7763/ijfcc.2013.v2.126

LAMPIRAN

Lampiran A

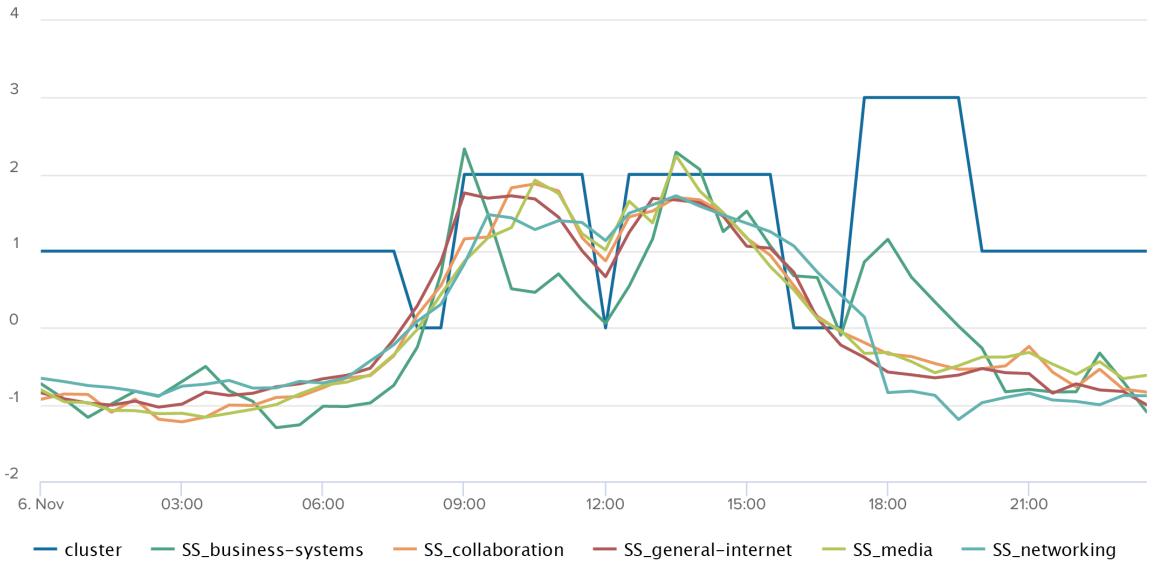
New Search

```
index=* sourcetype="pan:traffic" "app:category"="*" | timechart count by app:category | fit StandardScaler
    "business-systems" "collaboration" "general-internet" "media" "networking" with_mean=true with_std
    =true | fit KMeans "SS_business-systems" "SS_collaboration" "SS_general-internet" "SS_media"
    "SS_networking" k=4 | table _time, cluster, SS_business-systems, SS_collaboration, SS_general-internet
    , SS_media, SS_networking
```

during Fri, Nov 6, 2020

✓ 97,744,177 events (11/6/20 12:00:00.000 AM to 11/7/20 12:00:00.000 AM) No Event Sampling

Visualization



_time	cluster	SS_business-systems	SS_collaboration	SS_general-internet	SS_media	SS_networking
2020-11-06 00:00:00	1	-0.7269506064838294	-0.9303260592187996	-0.8454238032255444	-0.8080984970948022	-0.6562787377669554
2020-11-06 00:30:00	1	-0.9284680897579122	-0.8615680876453731	-0.92460376716649	-0.9631398229231751	-0.7005288262609719
2020-11-06 01:00:00	1	-1.1658881131555383	-0.8675227698925981	-0.9790622696969348	-0.9737694926665504	-0.7511991915202565
2020-11-06 01:30:00	1	-0.9915043468411844	-1.1004066670905165	-1.0055659167909938	-1.0753587649282363	-0.7771243618828052
2020-11-06 02:00:00	1	-0.8258894036906451	-0.9279069695558644	-0.9547990846050628	-1.078243961001438	-0.8226411590820019
2020-11-06 02:30:00	1	-0.8917383021660843	-1.1900525512989355	-1.032324161839847	-1.1178774439017372	-0.8841209064289623
2020-11-06 03:00:00	1	-0.6979969450938749	-1.2242919742204796	-0.9930396974151272	-1.1124107566051442	-0.7593541330728734

_time	cluster	SS_business-systems	SS_collaboration	SS_general-internet	SS_media	SS_networking
2020-11-06 03:30:00	1	-0.5029319920724106	-1.1617678106246168	-0.8356981613524121	-1.1640405810729666	-0.7338565777509269
2020-11-06 04:00:00	1	-0.8197677724253404	-1.00531783876979	-0.8796417683498631	-1.1136255760043872	-0.6848986897061294
2020-11-06 04:30:00	1	-0.9564290541859252	-1.010621227646225	-0.8523234078004628	-1.0609327845622272	-0.7838754351889167
2020-11-06 05:00:00	1	-1.300398551498584	-0.9062747254546172	-0.7671603919218574	-0.9995844049004616	-0.7816002004445803
2020-11-06 05:30:00	1	-1.2628415164384714	-0.8935279837691512	-0.7282069048383697	-0.8604875836871515	-0.6970594966862746
2020-11-06 06:00:00	1	-1.0207889072184526	-0.7789003505100692	-0.6628516098427403	-0.7479649368322796	-0.7176334278850607
2020-11-06 06:30:00	1	-1.024428796078904	-0.6521307479812548	-0.6179659904125515	-0.706812929682927	-0.6476598742363268
2020-11-06 07:00:00	1	-0.9777720388676632	-0.6217990852844523	-0.5250122771172485	-0.6103866398680231	-0.43080257119280896
2020-11-06 07:30:00	1	-0.7484590406592241	-0.3641660361818564	-0.14955667318248167	-0.3493523214557088	-0.2203433573416979
2020-11-06 08:00:00	0	-0.24946336779006675	0.16780108488859938	0.2895484194507945	-0.0190732972086550262	0.08687384470605676
2020-11-06 08:30:00	0	0.697900432889242	0.5495985630368487	0.8623174383509402	0.4352691580302877	0.30633717903475416
2020-11-06 09:00:00	2	2.3310523847763283	1.1593487209616815	1.7574329278158165	0.8677448641607537	0.8367432105998491
2020-11-06 09:30:00	2	1.4841164267467473	1.1836326595011462	1.689608032658685	1.1720571236710955	1.4760115598835402
2020-11-06 10:00:00	2	0.5089571111330823	1.824272731583459	1.7201343274384902	1.3064465197123396	1.4337543188428965
2020-11-06 10:30:00	2	0.4627967024028121	1.8759575126511698	1.6794241144669235	1.9250932987767773	1.2804180198495911
2020-11-06 11:00:00	2	0.7030293671926053	1.7844507940551415	1.44259709691751	1.7509185674103291	1.3966466117581842
2020-11-06 11:30:00	2	0.3603834658292018	1.174002821804462	1.0049432126265598	1.2270277014868358	1.372901874692291
2020-11-06 12:00:00	0	0.058438139905391435	0.8723609492184685	0.666124254386655	1.0148898638940478	1.1369652760400166
2020-11-06 12:30:00	2	0.5427088078390863	1.446476055570064	1.2428904611770673	1.650544114547886	1.4940359372553391
2020-11-06 13:00:00	2	1.1586772727236596	1.5221656494469011	1.6866037767921154	1.3673393420993891	1.6030394486194173
2020-11-06 13:30:00	2	2.2891936628811367	1.6986661529310558	1.6662359404085925	2.2415537522795486	1.7213577064811416

_time	cluster	SS_business-systems	SS_collaboration	SS_general-internet	SS_media	SS_networking
2020-11-06 14:00:00	2	2.0648441494824046	1.667776238773576	1.6302867091916744	1.7837186911898868	1.585555641204606
2020-11-06 14:30:00	2	1.2539761810700238	1.499509944334412	1.4548177987476238	1.5002102138915792	1.4708196678339642
2020-11-06 15:00:00	2	1.5206807648449183	1.1686994329280271	1.0637298803885031	1.1775238109676889	1.3632321270288614
2020-11-06 15:30:00	2	1.0694999956426	0.9462827468031606	1.0386774416367701	0.7985001584039094	1.2533108569691516
2020-11-06 16:00:00	0	0.6788737411187005	0.5518315688795581	0.7249363819440776	0.4958582755675264	1.0678550206808004
2020-11-06 16:30:00	0	0.6553799130194232	0.1531469840 4581906	0.1244416457 6686127	0.12473094909882593	0.7263015411039345
2020-11-06 17:00:00	0	-0.09509899020819544	-0.05173130202276754	-0.22435755230096985	-0.03319557280274877	0.42914539482280173
2020-11-06 17:30:00	3	0.8593791386983592	-0.19152677196738666	-0.3844996658664193	-0.3328004071410246	0.14173362713199614
2020-11-06 18:00:00	3	1.1542101363949238	-0.34025426528284336	-0.5751426144161944	-0.3197410985991636	-0.8429972486563301
2020-11-06 18:30:00	3	0.6605088473227866	-0.3733306643279762	-0.6113209837924272	-0.4394008094245875	-0.8241660506453206
2020-11-06 19:00:00	3	0.3373859862108952	-0.4642791314633274	-0.6508600461719412	-0.5847235800590173	-0.8791569412642357
2020-11-06 19:30:00	3	0.023528296743789253	-0.5400617672502773	-0.6153181716826935	-0.492701010566369	-1.193371296971656
2020-11-06 20:00:00	1	-0.2643538222191861	-0.5314088696097784	-0.5301042362131292	-0.3794191015869702	-0.9751928248168434
2020-11-06 20:30:00	1	-0.8333346309052048	-0.4957272970814847	-0.5838753242656299	-0.3807857734111184	-0.9059272564475972
2020-11-06 21:00:00	1	-0.8014028786294265	-0.2418624453384608	-0.5953576920268411	-0.3195892461742582	-0.8496353670460559
2020-11-06 21:30:00	1	-0.8343273278671146	-0.5763015912392485	-0.8513813953677251	-0.472504638053956	-0.9391380189191701
2020-11-06 22:00:00	1	-0.8335000803988617	-0.7701078900044009	-0.7291998368620664	-0.605679214695957	-0.9577997847455559
2020-11-06 22:30:00	1	-0.3275555287961152	-0.5383870128682453	-0.8087616977735203	-0.4395526618494929	-1.0016787357657163
2020-11-06 23:00:00	1	-0.6996514400304438	-0.7977878582629861	-0.8281111422995499	-0.6606497925116974	-0.8806354404376386
2020-11-06 23:30:00	1	-1.094579381389422	-0.8366328557351183	-1.0032236156068888	-0.6175237038385751	-0.8843629526783598

Lampiran B

18,417,902 of 18,417,902 events matched No Event Sampling

Statistics (6)

app:category	subkategori teknologi	aplikasi
business-systems	auth-service	360-safeguard-update
	database	adobe-creative-cloud-base
	erp-crm	adobe-update
	general-business	aerospike
	management	amazon-aws-console
	marketing	amqp
	office-programs	apache-solr
	software-update	appdynamics
	storage-backup	apple-update
		apt-get
		arcgis
		avast-av-update
		avira-antivir-update
		bacnet-read-prop-multiple
		bacnet-read-property
		bitbucket-base
		bitdefender
		datadog
		dynatrace-app-monitoring
		echo
		eset-update
		eventbrite
		firefox-update
		github-base
		gitlab-base
		google-docs-base
		google-update
		hp-jetdirect
		hp-update
		hubspot
		kerberos
		ldap
		liveperson
		lpd
		mailchimp
		mcafee-update
		memcached
		metatrader
		ms-delve
		ms-ds-smb-base
		ms-ds-smbv1
		ms-ds-smbv3
		ms-office365-base
		ms-product-activation
		ms-spynet
		ms-teams
		ms-update
		mssql-mon
		mysql
		new-relic
		oracle
		oracle-eloqua
		paloalto-updates
		paloalto-wildfire-cloud
		pan-db-cloud
		postgres
		radius
		rmcp
		rmi-iiop
		rsync
		salesforce-base
		samsung-updates
		sap
		soap
		sway
		symantec-av-update
		synology-cloudstation
		syslog
		tacacs-plus
		trello-base
		windows-azure-base
		yahoo-calendar
		yum

app:category	subkategori teknologi	aplikasi
collaboration	email	chatango
	instant-messaging	discord
	internet-conferencing	disqus
	social-business	facebook-base
	social-networking	facebook-chat
	voip-video	facetime
	web-posting	foursquare
		gmail-base
		google-classroom
		google-hangouts-audio-video
		google-hangouts-base
		google-meet
		ichat-av
		icloud-mail
		irc-base
		jabber
		linkedin-base
		mail.ru-base
		naver-line
		netlog
		outlook-web
		outlook-web-online
		pastebin-base
		pinterest-base
		plurk
		qq-audio-video
		qq-base
		reddit-base
		sharepoint-online
		sip
		skype
		slack-base
		snapchat
		telegram
		tinder
		tumblr-base
		twitter-base
		wechat-base
		whatsapp-base
		wire
		yy-voice-base
		zalo-base
		zello

app:category	subkategori teknologi	aplikasi
general-internet	file-sharing	adobe-cloud
	internet-utility	adobe-flash-socketpolicy-server
		amazon-cloud-drive-base
		apache-jserv
		apple-maps
		apple-push-notifications
		apple-siri
		ares
		bing-maps
		bittorrent
		cloudup
		dropbox-base
		flash
		ftp
		google-analytics
		google-app-engine
		google-base
		google-cloud-storage-base
		google-drive-web
		google-keep
		google-maps
		google-play
		google-translate-base
		icap
		icloud-base
		mediafire
		mega-base
		mendeley-base
		ms-onedrive-base
		opera-mini
		ping
		redis
		rss
		scribd-base
		simplenote
		slideshare-base
		solidfiles
		sourceforge-base
		speedtest
		tftp
		traceroute
		waze
		web-browsing
		web-crawler
		websocket
		weiyun-base
		whatsapp-file-transfer
		winbox
		windows-push-notifications
		yahoo-web-analytics
media	audio-streaming	adobe-media-player
	gaming	apple-game-center
	photo-video	bet365
		cloudinary-base
		dailymotion
		facebook-video
		flickr-base
		http-audio
		http-video
		imgur-base
		instagram-base
		itunes-base
		origin
		ppstream
		rtcp
		rtmp
		rtp-base
		soundcloud-base
		spotify
		steam
		twitch
		vimeo-base
		xbox-live
		youtube-base
		youtube-uploading

app:category	subkategori teknologi	aplikasi
networking	encrypted-tunnel	airdroid
	infrastructure	anydesk
	ip-protocol	apple-airport
	proxy	capwap
	remote-access	chrome-remote-desktop
	routing	citrix
		cotp
		dhcp
		dns
		dnscrypt
		dtls
		gre
		gtp-u
		gtpv1-c
		hola-unblocker
		hotspot-shield
		http-proxy
		icmp
		ike
		ipsec-esp
		ipsec-esp-udp
		ipv6
		isatap
		llmnr
		mobility-xe
		ms-rdp
		netbios-dg
		netbios-ns
		netbios-ss
		ngrok
		ntp
		ocsp
		open-vpn
		panos-global-protect
		pcanywhere-base
		portmapper
		pptp
		psiphon
		quic
		rip
		rpc
		sctp
		slp
		snmp-base
		snmp-trap
		snmpv1
		snmpv2
		socks
		splashtop-remote
		ssdp
		ssh
		ssl
		stun
		t.120
		teamviewer-base
		teamviewer-sharing
		teamviewer-web
		teredo
		tor
		ultrasurf
		upnp
		vnc-base
		xdmcp
unknown		unknown-tcp
		unknown-udp

Lampiran C

```r
library(tidyverse)  # Library data manipulation
library(cluster)    # Library clustering algorithms
library(factoextra) # Lirary clustering algorithms & visualization

df <- read.csv('Data 06 November 2020.csv')  # Untuk Membaca data uji
df <- na.omit(df)   # Untuk Menghapus Missing Value
df <- scale(df)     # Untuk Preprocessing Data
head(df)            # Untuk melakukan Preview Data

# ====== 1. Elbow Method ======
set.seed(1000)       # Random Number Generator
# function untuk komputasi total nilai wss (within-cluster sum of square)
wss <- function(k) {
  kmeans(df, k, nstart = 25 )$tot.withinss
}
# komputasi dan pemetaan wss untuk k = 1 to k = 15
k.values <- 1:15

# ekstraksi wss untuk 1-15 clusters
wss_values <- map_dbl(k.values, wss)

# Visualisasi hasil perhitungan
plot(k.values, wss_values,
     type="b", pch = 19, frame = FALSE,
     xlab="Number of clusters K",
     ylab="Total within-clusters sum of squares")

set.seed(1000)    # Random Number Generator
# menerapkan fviz_nbclust yang merupakan fungsi elbow method
fviz_nbclust(df, kmeans, method = "wss")


# ====== 2. Silhouette Method ======
# function untuk menghitung rata-rata nilai silhouette untuk k clusters
avg_sil <- function(k) {
  km.res <- kmeans(df, centers = k, nstart = 25)
  ss <- silhouette(km.res$cluster, dist(df))
  mean(ss[, 3])
}

# melakukan komputsi dan pemetaan wss untuk k = 2 hingga k = 15
k.values <- 2:15

# Ekstraksi nilai rata-rata silhouette untuk 2-15 clusters
avg_sil_values <- map_dbl(k.values, avg_sil)

plot(k.values, avg_sil_values,
     type = "b", pch = 19, frame = FALSE,
     xlab = "Number of clusters K",
     ylab = "Average Silhouettes")
# sama dengan sebelumnya menerapkan fviz_nbclust
# yang fungsinya juga dapat digunakan untuk Silhouette Method
fviz_nbclust(df, kmeans, method = "silhouette")
```

Alamat akses *Code* : https://github.com/Wahluf/R/blob/master/K%20means%20Cluster%20Analysis.R