Jurnal Teknik Informatika (JUTIF)
Vol. 4, No. 4, August 2023, pp. 987-994

DOI: https://doi.org/10.52436/1.jutif.2023.4.4.649
p-ISSN: 2723-3863
e-ISSN: 2723-3871

DATA WAREHOUSE MODELLING INFORMATION SECURITY LOG MANAGEMENT IN BUILDING A SECURITY OPERATION CENTER IN CENTRAL GOVERNMENT AGENCIES WITH KIMBALL METHOD

Maya Asmita[*1], Henny[2], Samidi[3]

[1,2,3]Magister Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur, Indonesia
Email: [1]2111601411@student.budiluhur.ac.id, [2]2111602286@student.budiluhur.ac.id,
[3]samidi@student.budiluhur.ac.id

(Article received: October 19, 2022; Revision: November 21, 2022; published: August 31, 2023)

Abstract

Central Government, is a government agency that manages important and confidential state data and information. The data that is managed needs to be maintained for reliability and security in order to avoid the risk of loss, leakage and misuse of information. To maintain this data, an optimal information security device is needed. Information security tools used today have a variety of functions resulting in many important logs that must be managed, analyzed and evaluated. The log data from each of these information security devices is still separate and must be processed manually to obtain simpler and more efficient data so that it can be monitored and presented to management. The purpose of this research is to make the right data warehouse modeling in order to assist in the process of presenting information quickly and accurately related to the processing of data logs of information security devices as a report that will be given to management in support of the Zero Tollerance data security policy. The method used in designing this data warehouse is using the Kimball 9 step method. The results obtained are in the form of a starflake schema and a data warehouse log of information security devices consisting of a malware fact table, intrusion facts and attack facts that can be used as centralized data monitoring that will be implemented at the Security Operation Center. Testing is done using Pentaho software tools. This data warehouse is expected to provide a quick, accurate, and continuous summary of information so that it can assist management in the decision-making process and policy making for the future.

Keywords: *data warehouse, information security, Kimball, Pentaho, security log.*

PERMODELAN DATA WAREHOUSE PENGELOLAAN LOG KEAMANAN INFORMASI DALAM MEMBANGUN SECURITY OPERATION CENTER INSTANSI PEMERINTAH PUSAT DENGAN METODE KIMBALL

Abstrak

Instansi Pemerintah Pusat merupakan instansi Pemerintahan yang mengelola data dan informasi negara yang penting dan bersifat rahasia. Data yang dikelola tersebut perlu dijaga keandalan dan keamanannya agar terhindar dari risiko kehilangan, kebocoran dan penyalahgunaan informasi. Untuk menjaga data tersebut, dibutuhkan perangkat keamanan informasi yang optimal. Perangkat keamanan informasi yang digunakan saat ini memiliki beragam fungsi sehinggal menghasilkan banyak *log* penting yang harus dikelola, analisis dan dievaluasi. Data *log* dari masing-masing perangkat keamanan informasi tersebut masih terpisah dan harus diolah secara manual untuk mendapatkan data yang lebih sederhana dan efisien agar dapat dipantau dan disajikan kepada manajemen. Tujuan penelitian ini adalah untuk membuat permodelan *data warehouse* yang tepat agar dapat membantu dalam proses penyajian informasi yang cepat dan tepat terkait pengolahan data *log* perangkat keamanan informasi sebagai laporan yang akan diberikan kepada manajemen dalam mendukung kebijakan *Zero Tollerance* keamanan data. Metode yang digunakan dalam perancangan *data warehouse* ini adalah menggunakan metode Kimball 9 *step*. Hasil yang diperoleh adalah dalam bentuk skema *starflake* dan *data warehouse log* perangkat keamanan informasi yang terdiri dari tabel fakta *malware*, fakta *intrusion* dan fakta *attack* yang dapat digunakan sebagai monitoring data terpusat yang akan diimplementasikan pada *Security Operation Center*. Pengujian dilakukan dengan menggunakan *tools software* Pentaho. *Data warehouse* ini diharapkan dapat memberikan ringkasan informasi yang cepat, akurat, dan berkesinambungan sehingga dapat membantu manajemen dalam proses pengambilan keputusan dan penyusunan kebijakan untuk masa yang akan datang.

Kata kunci: *data warehouse, keamanan informasi, Kimball, log keamanan, Pentaho.*

1. PENDAHULUAN

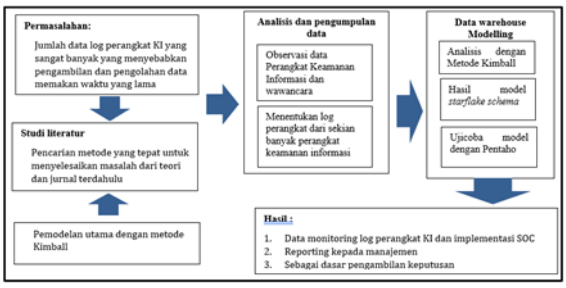
Sebagai tindakan perlindungan data dan informasi negara yang penting dan bersifat rahasia sesuai dengan konsep keamanan informasi yaitu CIA atau *confidentiality*, *integrity*, dan *availability* [1]. Intansi Pemerintah Pusat menggunakan perangkat keamanan informasi yang handal dan optimal. Dalam mewujudkan keamanan tersebut, dilakukan koordinasi pengelolaan data aktivitas (*log*) sistem informasi, serta melakukan analisis dan evaluasi data aktivitas (*log*) perangkat keamanan informasi.

Log atau data aktivitas merupakan catatan peristiwa yang terjadi dalam sistem dan jaringan [2]. *Log* terdiri atas entri *log* dimana setiap entri berisi informasi yang terkait dengan peristiwa spesifik yang terjadi dalam suatu sistem atau jaringan. Pada awalnya *log* dipergunakan terutama untuk mengatasi masalah, tetapi saat ini *log* memiliki banyak fungsi dalam suatu organisasi, diantaranya untuk mengoptimalkan kinerja sistem dan jaringan, mencatat kegiatan pengguna, dan menyediakan data yang berguna untuk menyelidiki aktivitas atau anomali berbahaya. Mengingat pentingnya hal tersebut, maka diperlukan kegiatan analisis dan evaluasi *log* pada perangkat tersebut. Hal ini bertujuan untuk mendeteksi anomali atau ancaman keamanan yang mungkin terjadi.

Sebuah penelitian mengenai desain *data warehouse* penjualan dengan *nine step methodology* untuk *business intelegency* menghasilkan desain *data warehouse* yang lebih baik dengan menggunakan permodelan skema bintang karena *query* yang dihasilkan memiliki *time response* yang lebih cepat [3]. Sementara penelitian lain terkait analisis *data warehouse* dengan menggunakan sembilan langkah Kimball dapat menghasilkan sebuah *data warehouse* yang dapat digunakan sebagai dasar penyajian data yang terintegrasi sebagai pendukung dalam pengambilan keputusan dan diimplementasikan dalam bentuk laporan sehingga memudahkan pengguna dalam melakukan analisis [4]. Agata Filiana dan kawan-kawan dalam penelitiannya menyebutkan bahwa keberadaan *data warehouse* membantu pihak terkait dalam menghasilkan data yang konsisten dan terstruktur [5]. Lalu Novi Sofia Fitriasari, Ishak Ariawan dan Amien Rais menyebutkan dalam penelitan mereka bagaimana implementasi rancangan *data warehouse* dapat digunakan untuk melakukan pengukuran dan evaluasi kinerja proses bisnis [6]. Nurul Hidayat dan kawan-kawan menyebutkan bahwa analisis desain *data warehouse* merupakan sesuatu yang penting untuk dapat meminimalisasi *human error* dan inkonsistensi data [7]. Dua penelitian lain mengatakan bahwa adanya data yang beragam dan variatif dapat diolah dengan menggunakan Pentaho karena dapat menghasilkan bentuk tabel yang lebih teratur sehingga akan mudah untuk diolah [8][9]. Adapun yang membedakan penelitian ini dengan penelitian sebelumnya tersebut adalah *data warehouse* dirancang dengan mendetilkan langkah demi langkah dalam metode Kimball kemudian menghasilkan sebuah *data warehouse* dimana data yang disajikan tidak hanya digunakan sebagai laporan atau data *monitoring* tetapi juga dapat digunakan sebagai sumber data yang akan disajikan dalam pengelolaan sistem yang terpusat yang akan diimplementasikan pada SOC. Pentingnya penelitian ini dilakukan karena diharapkan rancangan *data warehouse* yang dihasilkan dapat membantu pengelolaan data *log* manajemen perangkat keamanan informasi sehingga mampu menyajikan data yang layak dan efisien pada SOC.

2. METODE PENELITIAN

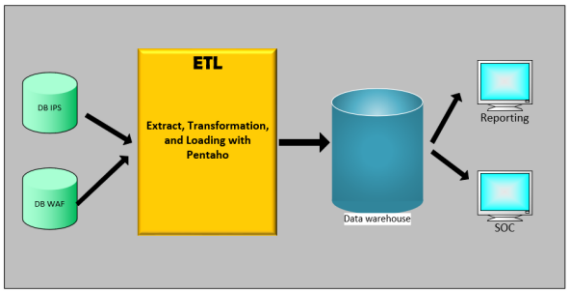
Penelitian dilakukan dengan metode kuantitatif. Langkah penelitian dimulai dengan melakukan pengumpulan data. Data yang dihasilkan diperoleh dari perangkat IPS dan WAF dalam format *log* yang di-*query* setiap hari terdiri dari data *log malware*, *log intrusion*, dan *log attack*. Data tersebut diambil melalui observasi dan wawancara dengan pengelola perangkat keamanan informasi serta melakukan studi dari berbagai sumber untuk mendapatkan informasi yang berkaitan dengan penelitian. Setelah mendapatkan sumber data dengan menggunakan teknik pengumpulan data, dilakukan pemodelan *data warehouse* menggunakan metode Kimball 9 *Step* atau disebut juga metode 9 langkah Kimball. Gambar 1 menunjukkan kerangka konsep penelitian yang dilakukan untuk permodelan *data warehouse log* perangkat keamanan informasi untuk SOC.



Gambar 1 Kerangka Konsep Permodelan *Data Warehouse Log* Perangkat Keamanan Informasi untuk SOC

Pemodelan *data warehouse* digambarkan dengan skema *starflake*. Skema *starflake* merupakan struktur *hybrid* yang berisi kombinasi antara skema bintang yang telah didenormalisasi dan *snowflake* yang telah dinormalisasi [10]. Tabel fakta yang dibuat memiliki beberapa *key* yang merupakan kunci indeks dalam tabel dimensi. Dari hasil pemodelan dengan skema *starflake*, selanjutnya dilakukan proses ETL (*Extract*, *Transform*, *Load*) yaitu dengan mengumpulkan, menyaring, mengolah dan menggabungkan data relevan yang didapatkan

dari berbagai sumber untuk disimpan ke dalam *data warehouse*. Data operasional yang telah dikumpulkan dimasukan ke dalam *data warehouse*. *Tools* yang digunakan dalam proses ETL adalah Pentaho *Data Integration*. Arsitektur rancangan *data warehouse* ini ditunjukkan pada Gambar 2.

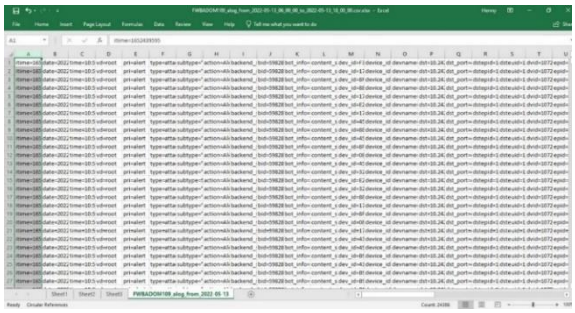


Gambar 2 Arsitektur *Data Warehouse*.

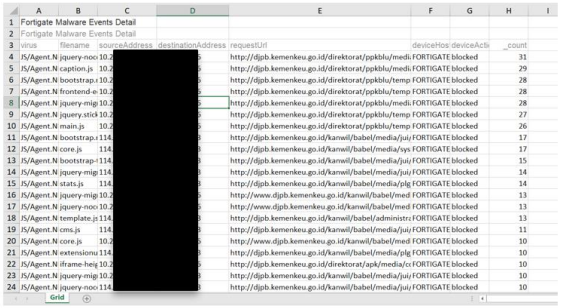
3. HASIL DAN PEMBAHASAN

3.1. Permasalahan

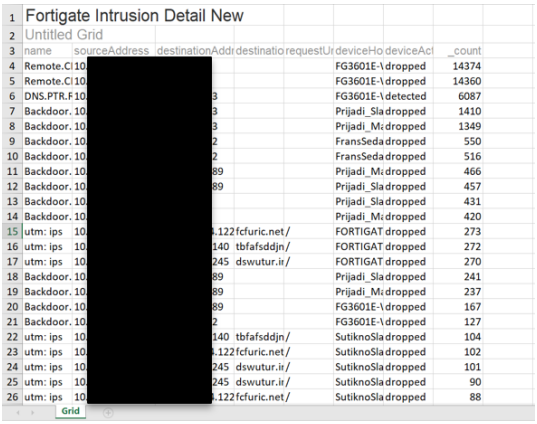
Analisa yang dilakukan pada penelitian ini menggunakan data yang bersumber dari *database* perangkat IPS dan WAF yang digunakan oleh instansi pemerintah yang kemudian dieksport sehingga kemudian diperoleh data dengan extensi .csv. Data yang diperoleh dari *database* perangkat tersebut merupakan kumpulan data *log malware*, data *log intrusion*, dan data *log attack* yang diambil setiap harinya, sebagaimana yang ditunjukkan pada Gambar 3, 4 dan 5. *Malware* atau *malicious software* merupakan perangkat lunak yang dengan sengaja mengeksekusi muatan berbahaya pada sebuah mesin seperti komputer, *smart phone*, jaringan komputer, dan lain-lain) [11]. *Intrusion* dapat diartikan sebagai segala jenis aktivitas yang tidak sah yang menyebabkan kerusakan pada sistem informasi [12]. Sedangkan *attack* merupakan serangan terhadap keamanan sebuah sistem informasi [13].



Gambar 3 Data *Log Attack* Perangkat WAF



Gambar 4 Data *Log Malware* Perangkat IPS



Gambar 5 Data *Log Intrusion* Perangkat IPS

Data di atas diperoleh langsung ketika melakukan kegiatan observasi pada pihak yang terlibat langsung dalam pengolahan data *log* pada perangkat keamanan informasi. Data tersebut sudah melalui sebuah proses *query* yang dilakukan untuk menghasilkan data yang lebih sederhana. Permasalahan yang ditemukan adalah proses *query* masih menggunakan cara manual sehingga menghabiskan waktu yang lebih lama ketika proses pengolahan data untuk menjadi sebuah laporan.

3.2. Studi Literatur

Sebagaimana yang telah dipaparkan pada bagian Pendahuluan, berdasarkan studi literatur dari beberapa contoh penelitian yang telah dilakukan sebelumnya, penelitian rancangan *data warehouse* ini memilih metode Kimball 9 *Step* sebagai model rancangan *data warehouse* yang akan disajikan dalam SOC nantinya.

3.3. Analisis dan Pengumpulan Data

Log yang dianalisa dan dievaluasi pada perangkat keamanan informasi adalah *log* perangkat keamanan informasi pada *firewall* yaitu *Instrusion Prevention System* (IPS) dan *Web Application Firewall* (WAF). IPS adalah perangkat lunak yang bekerja dengan cara mendeteksi akftivitas yang mencurigakan dan melakukan pencegahan terhadap intrusi pada jaringan [14]. WAF adalah suatu metode yang bertujuan untuk pengamanan pada aplikasi web yang terdiri dari beberapa fungsi seperti *monitoring* trafik, *secure directory*, *filter string* dan perlindungan

terhadap serangan seperti SQL *injection*, *cross site scripting*, dan *unrestricted file upload* [15]. Penggunaan perangkat keamanan yang optimal dan handal ini dituntut juga untuk dapat menghindari terjadinya *Single Point of Failure* (SPOF), kebutuhan akan kondisi tersebut membuat Instansi Pemerintahan Pusat mengimplementasikan beberapa perangkat keamanan. SPOF merupakan kegagalan di satu titik sistem sehingga menyebabkan layanan tidak dapat berjalan dengan semestinya [16]. Dari kondisi tersebut ditemukan permasalahan yang dihadapi saat ini, yaitu banyaknya *log* perangkat keamanan yang harus dikelola dan penyimpanan *log* yang saat ini masih terpisah. Tabel 1 menunjukkan daftar *log* yang dihasilkan oleh *log* perangkat keamanan informasi pada *firewall* yaitu IPS dan WAF dalam kurun waktu 3 bulan pada tahun 2022:

Tabel 1 Jumlah Log Perangkat

Perangkat	**Jenis *Log***	**Jumlah *Log* Periode 2022**		
		April	Mei	Juni
IPS	Log Malware	28.578	34.827	68.368
	Log Intrusion	654.559	414.186	349.231
WAF	Log Attack	655.449	752.455	755.854

Banyaknya *log* yang dihasilkan, mengakibatkan banyaknya data yang harus diolah agar dapat menghasilkan laporan yang berisi data sederhana dan efisien ketika dibutuhkan dalam melakukan investigasi adanya anomali atau ancaman dan untuk disajikan kepada level manajemen dalam proses pengambilan keputusan yang cepat, efektif dan tepat. Untuk itu perlu adanya sebuah solusi untuk mengoptimalkan pengelolaan data yang terpusat agar dapat diimplementasikan pada *Security Operation Center* (SOC), dan salah satu solusi yang diusulkan oleh penulis adalah pemodelan *data warehouse* dengan Metode Kimball 9 *Step*.

SOC adalah sistem pengelolaan keamanan terpusat yang menangani serangan keamanan informasi yang terdistribusi dan bertanggung jawab untuk menghapus dan memblokir serangan [17]. Salah satu bentuk pengelolaan keamanan tersebut adalah *log management* perangkat keamanan informasi. Semua peristiwa dan aktivitas *log* disimpan dalam sebuah *database*. Banyaknya jumlah *log* perangkat dalam hitungan hari bahkan jam membutuhkan kapasitas penyimpanan yang sangat besar.

3.4. *Data warehouse Modelling*

Berikut perancangan *data warehouse* SOC pada perangkat keamanan informasi menggunakan metode 9 langkah Kimball:

3.4.1. Menentukan Proses Bisnis

Berdasarkan hasil observasi dan pengamatan yang dilakukan, maka proses bisnis yang dipilih dalam penelitian ini adalah proses pengolahan data *log* perangkat keamanan informasi IPS dan WAF sebagaimana yang ditunjukkan pada Tabel 2.

Tabel 2 Pemilihan Proses Bisnis

Proses Bisnis	**Deksripsi**
Aktivitas *log* perangkat keamanan informasi IPS dan WAF	*Query* data *log* akses perangkat keamanan informasi IPS dan WAF yang telah dieksport

3.4.2. Menentukan *Granularity*

Pada tahap ini akan diputuskan apa saja yang dapat mewakili atau dipresentasikan oleh sebuah tabel fakta. Sehingga berdasarkan proses bisnis yang telah dipilih sebelumnya, dapat ditentukan *granularity*-nya adalah informasi *log* perangkat keamanan informasi IPS dan WAF seperti yang ditunjukkan pada Tabel 3.

Tabel 3 Menentukan *Granularity*

Grain	**Deksripsi**
Informasi *log* perangkat keamanan informasi IPS dan WAF	Menampilkan informasi *log* perangkat keamanan informasi IPS dan WAF

3.4.3. Identifikasi dan Penyesuaian Dimensi

Langkah berikutnya adalah mengidentifikasi dimensi yang berhubungan dengan tabel fakta. Tabel dimensi biasanya memiliki *record* data yang relatif lebih kecil dibandingkan dengan Tabel Fakta, tetapi setiap *record* mungkin memiliki sejumlah besar atribut untuk mendeskripsikan data [18]. Dari hasil identifikasi maka kemudian dapat ditentukan dimensi yang terlibat adalah dimensi waktu, dimensi server, dimensi *malware*, dimensi *intrusion* dan dimensi *attack* seperti yang terlihat pada Tabel 4.

Tabel 4 Identifikasi Dimensi

Tabel Dimensi	***Field***
Aktivitas *log* perangkat keamanan informasi IPS dan WAF	*Query* data *log* akses perangkat keamanan informasi IPS dan WAF yang telah dieksport.
dimensi_waktu	id_waktu, tanggal, bulan, tahun
dimensi_server	hostname, ip_server, OS
dimensi_malware	Id_malware, nama_malware, action_malware
dimensi_instrusion	Id_intrusion, signature_intrusion, action_intrusion
dimensi_attack	Id_attack, signature_attack, severity_level, action_attack

3.4.4. Menentukan Fakta

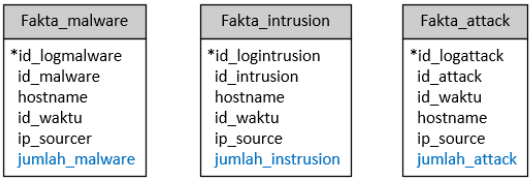
Berdasarkan dimensi yang telah diidentifikasi, fakta yang ditentukan adalah fakta_malware, fakta_instrusion, dan fakta_attack sebagaimana yang tertera pada Tabel 5 berikut.

Tabel 5 Penentuan Fakta

Tabel Fakta	***Field***
Fakta_malware	Id_logmalware, id_malware, hostname, id_waktu, ip_source
Fakta_intrusion	Id_logintrusion, log_intrusion, hostname, id_waktu, ip_source
Fakta_attack	Id_logattack, id_attack, id_waktu, hostname, ip_source

3.4.5. Menyimpan Hasil Perhitungan Sementara pada Tabel Fakta

Pada tahap ini akan dilakukan proses penghitungan dari tabel fakta dan menyimpan hasilnya yaitu pada tiga buah tabel fakta dimana perlu dilakukan penghitungan total *log* yang masuk. Sehingga bentuk tabel fakta_malware, fakta_intrusion, dan fakta_attack dapat digambarkan seperti yang terlihat pada Gambar 6.



Gambar 6 Fakta Setelah Penambahan Penghitungan.

3.4.6. Melengkapi Tabel Dimensi

Pada tahap ini akan ditambahkan atribut selengkap mungkin pada tabel dimensi dan harus dengan mudah dipahami serta dimengerti oleh pengguna. Rincian tabel dimensi ditunjukkan pada Tabel 6 dibawah ini.

Tabel 6 Kelengkapan Tabel Dimensi

Tabel Dimensi	**Field**	**Type (Length)**	**Keterangan**
Dimensi_ waktu	id_waktu	Int (8)	PK Waktu
	Tanggal	Int (2)	Tanggal
	Bulan	Int (2)	Bulan
	Tahun	Int (4)	Tahun
Dimensi_ server	hostname	Varchar (10)	PK server
	Ip_server	Varchar (20)	IP server
	OS	Varchar (100)	*Operating system*
Dimensi_ malware	id_malware	Varchar (10)	PK *malware*
	Nama_mal ware	Varchar (200)	Nama *malware*
	action	Varchar (20)	Action/tindakan penanganan *malware*
Dimensi_ intrusion	Id_intrusio n	Varchar (8)	PK *intrusion*
	Signature_ intrusion	Varchar (200)	Nama *intrusion*
	Action_intr usion	Varchar (20)	*Action*/tindakan penanganan *intrusion*
Dimensi_ attack	Id_attack	Varchar (8)	PK *attack*
	Signature_ attack	Varchar (200)	Nama *attack*
	Severity_ level	Varchar (10)	Level *attack*
	Action_ attack	Varchar (10)	*Action*/tindakan penanganan *attack*

3.4.7. Menentukan Durasi Dimensi

Pada tahap ini menentukan durasi waktu data yang diperlukan untuk pembuatan *data warehouse*. Peneliti menetapkan data yang dikumpulkan adalah data satu bulan terakhir. Data tersebut diambil dari database perangkat IPS dan WAF.

3.4.8. Melacak Perubahan Dimensi

Atribut dalam sebuah tabel dimensi tidak selalu memiliki nilai yang tetap atau statis [4]. Pada tahap ini akan dilakukan pengamatan terhadap perubahan dalam tabel dimensi.

Pada penelitian ini tabel yang mengalami perubahan nilai adalah seperti tabel server yang akan mengalami perubahan jika ada penambahan atau pengurangan server, lalu tabel *malware*, *intrusion* dan *attack* juga memungkinkan mengalami perubahan jika nantinya didapatkan perubahan *severity level* hingga *action* yang akan dilakukan pada *malware* tertentu.

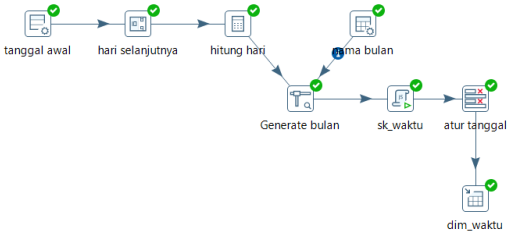
3.4.9. Menentukan *Query* Prioritas dan Tipe *Query*

Pada tahap ini difokuskan pada desain fisik *data warehouse*. Maka pada tahap inilah proses ETL (*Extract*, *Transform*, dan *Loading*) dilakukan. Dalam perancangan ini, jumlah *log* paling banyak tentu akan menjadi isu tersendiri yang akan menarik perhatian manajemen. Hasil *query* seperti inilah yang dapat digunakan oleh manajemen dalam mengambil keputusan dan kebijakan yang strategis.

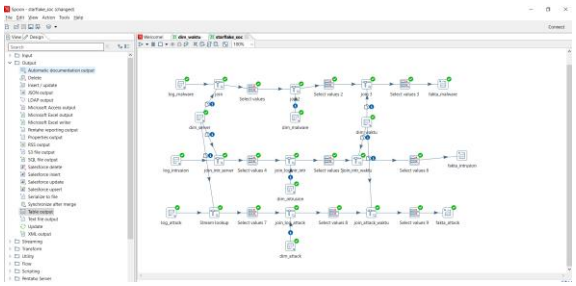
Proses ETL adalah sekumpulan proses untuk mengambil dan memproses dari satu dan atau banyak sumber menjadi sumber baru [4]. Jadi dapat dikatakan bahwa proses ETL merupakan tahap dimana dilakukan pemrosesan data dari sumber data masuk ke dalam *data warehouse*. Proses ETL dilakukan setelah perancangan *data warehouse* selesai dilakukan. Tujuan dari proses ETL adalah untuk mengumpulkan, menyaring, memproses dan menggabungkan atau menyatukan data yang sesuai dari berbagai sumber untuk disimpan di *data warehouse* [17].

Proses *extract* dilakukan dengan tujuan untuk memilih data dan mengambil data dari data yang diperoleh dari *query* perangkat IPS dan WAF. Data yang diperoleh dalam format *log* dan dengan jumlah yang sangat besar sehingga perlu di-*extract* agar dapat ditransformasikan pada aplikasi Pentaho. Data yang diambil dari query perangkat IPS dan WAF adalah data *log malware*, *intrusion* dan *attack* yang jumlahnya mencapai angka ribuan perharinya.

Proses *transform* dilakukan untuk mengubah struktur data dan membersihkannya dari bentuk asli ke dalam bentuk yang lebih sesuai untuk kebutuhan *data warehouse*. Berikut proses *transform* yang dilakukan melalui aplikasi Pentaho *Data Integration*. Sedangkan aplikasi basis data yang digunakan adalah mySql. Pada Gambar 7 ditunjukkan bagaimana proses transformasi dimensi waktu dilakukan.



Gambar 7 Transformasi Dimensi Waktu



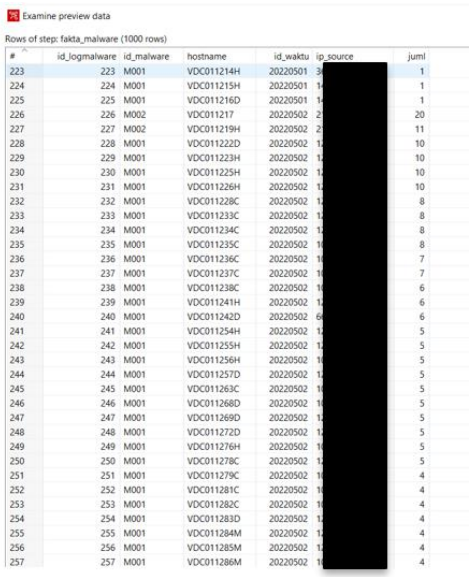
Gambar 8 Transformasi Dimensi *Malware*, *Instrusion*, dan *Attack*

Data *input* pada proses transformasi adalah menggunakan tabel *database* yang telah dibuat terlebih dahulu dengan mengambil data dari *query* perangkat IPS dan WAF. Komponen join tabel pada Gambar 8 menggunakan *stream lookup* untuk menggabungkan dua buah dimensi atau tabel dan mengambil salah satu PK dari tabel yang akan dilakukan penggabungan. Komponen *select values* berfungsi untuk memilih *field* mana saja yang akan ditampilkan dan mengatur urutan *field* berdasarkan kebutuhan. Sedangkan komponen *output* menggunakan *table output* untuk menghasilkan tabel akhir yaitu fakta *malware*, fakta *intrusion* dan fakta *attack* dan dilakukan koneksi ke database yang sudah dibuat sebelumnya.

Proses *loading* merupakan proses dimana dilakukan pengambilan data dari hasil transformasi dimensi untuk dimasukkan ke dalam *data warehouse*. Dalam hal ini dapat dilihat hasil masing-masing *loading* data dari tabel fakta *malware*, fakta *intrusion* dan fakta *attack*..

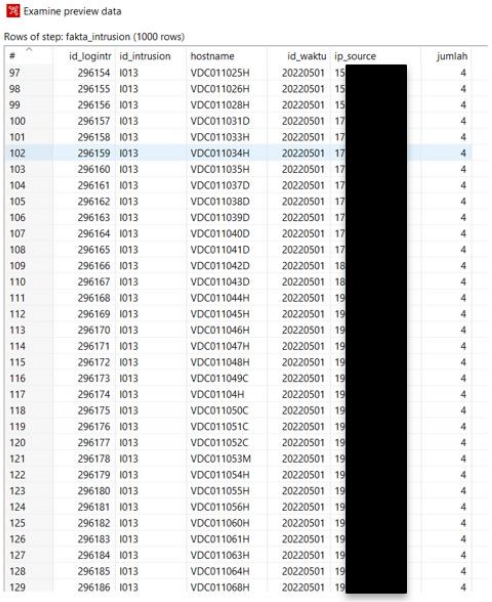
3.5. Hasil

Setelah proses ETL dilakukan maka diperoleh tampilan hasil dari masing-masing tabel fakta *malware*, fakta *intrusion* dan fakta *attack* dimana nantinya tabel-tabel tersebut menjadi data *monitoring* pada SOC. Tampilan hasil dari masing-masing tabel tersebut dapat dilihat pada Gambar 9, 10 dan 11.



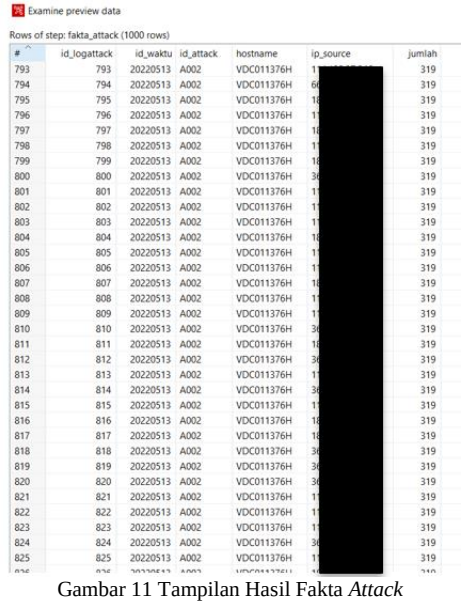
#	id_logmalware	id_malware	hostname	id_waktu	ip_source	juml
223	223	M001	VDC011214H	20220501	[redacted]	1
224	224	M001	VDC011215H	20220501		1
225	225	M001	VDC011216D	20220501		1
226	226	M002	VDC011217	20220502		20
227	227	M002	VDC011219H	20220502		11
228	228	M001	VDC011222D	20220502		10
229	229	M001	VDC011223H	20220502		10
230	230	M001	VDC011225H	20220502		10
231	231	M001	VDC011226H	20220502		10
232	232	M001	VDC011228C	20220502		8
233	233	M001	VDC011233C	20220502		8
234	234	M001	VDC011234C	20220502		8
235	235	M001	VDC011235C	20220502		8
236	236	M001	VDC011236C	20220502		7
237	237	M001	VDC011237C	20220502		7
238	238	M001	VDC011238C	20220502		6
239	239	M001	VDC011241H	20220502		6
240	240	M001	VDC011242D	20220502		6
241	241	M001	VDC011254H	20220502		5
242	242	M001	VDC011255H	20220502		5
243	243	M001	VDC011256H	20220502		5
244	244	M001	VDC011257D	20220502		5
245	245	M001	VDC011263C	20220502		5
246	246	M001	VDC011268D	20220502		5
247	247	M001	VDC011269D	20220502		5
248	248	M001	VDC011272D	20220502		5
249	249	M001	VDC011276H	20220502		5
250	250	M001	VDC011278C	20220502		5
251	251	M001	VDC011279C	20220502		4
252	252	M001	VDC011281C	20220502		4
253	253	M001	VDC011282C	20220502		4
254	254	M001	VDC011283D	20220502		4
255	255	M001	VDC011284M	20220502		4
256	256	M001	VDC011285M	20220502		4
257	257	M001	VDC011286M	20220502		4

Gambar 9 Tampilan Fakta *Malware*



#	id_logintr	id_intrusion	hostname	id_waktu	ip_source	jumlah
97	296154	I013	VDC011025H	20220501	[redacted]	4
98	296155	I013	VDC011026H	20220501		4
99	296156	I013	VDC011028H	20220501		4
100	296157	I013	VDC011031D	20220501		4
101	296158	I013	VDC011033H	20220501		4
102	296159	I013	VDC011034H	20220501		4
103	296160	I013	VDC011035H	20220501		4
104	296161	I013	VDC011037D	20220501		4
105	296162	I013	VDC011038D	20220501		4
106	296163	I013	VDC011039D	20220501		4
107	296164	I013	VDC011040D	20220501		4
108	296165	I013	VDC011041D	20220501		4
109	296166	I013	VDC011042D	20220501		4
110	296167	I013	VDC011043D	20220501		4
111	296168	I013	VDC011044H	20220501		4
112	296169	I013	VDC011045H	20220501		4
113	296170	I013	VDC011046H	20220501		4
114	296171	I013	VDC011047H	20220501		4
115	296172	I013	VDC011048H	20220501		4
116	296173	I013	VDC011049C	20220501		4
117	296174	I013	VDC01104H	20220501		4
118	296175	I013	VDC011050C	20220501		4
119	296176	I013	VDC011051C	20220501		4
120	296177	I013	VDC011052C	20220501		4
121	296178	I013	VDC011053M	20220501		4
122	296179	I013	VDC011054H	20220501		4
123	296180	I013	VDC011055H	20220501		4
124	296181	I013	VDC011056H	20220501		4
125	296182	I013	VDC011060H	20220501		4
126	296183	I013	VDC011061H	20220501		4
127	296184	I013	VDC011063H	20220501		4
128	296185	I013	VDC011064H	20220501		4
129	296186	I013	VDC011068H	20220501		4

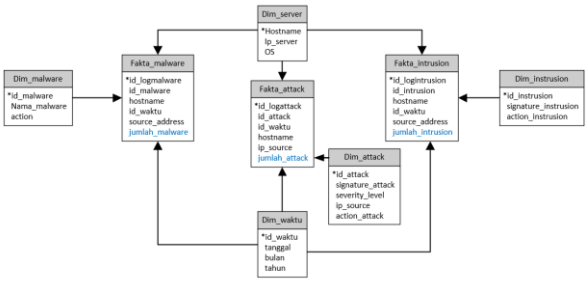
Gambar 10 Tampilan Hasil Fakta *Intrusion*



#	id_logattack	id_waktu	id_attack	hostname	ip_source	jumlah
793	793	20220513	A002	VDC011376H	[redacted]	319
794	794	20220513	A002	VDC011376H		319
795	795	20220513	A002	VDC011376H		319
796	796	20220513	A002	VDC011376H		319
797	797	20220513	A002	VDC011376H		319
798	798	20220513	A002	VDC011376H		319
799	799	20220513	A002	VDC011376H		319
800	800	20220513	A002	VDC011376H		319
801	801	20220513	A002	VDC011376H		319
802	802	20220513	A002	VDC011376H		319
803	803	20220513	A002	VDC011376H		319
804	804	20220513	A002	VDC011376H		319
805	805	20220513	A002	VDC011376H		319
806	806	20220513	A002	VDC011376H		319
807	807	20220513	A002	VDC011376H		319
808	808	20220513	A002	VDC011376H		319
809	809	20220513	A002	VDC011376H		319
810	810	20220513	A002	VDC011376H		319
811	811	20220513	A002	VDC011376H		319
812	812	20220513	A002	VDC011376H		319
813	813	20220513	A002	VDC011376H		319
814	814	20220513	A002	VDC011376H		319
815	815	20220513	A002	VDC011376H		319
816	816	20220513	A002	VDC011376H		319
817	817	20220513	A002	VDC011376H		319
818	818	20220513	A002	VDC011376H		319
819	819	20220513	A002	VDC011376H		319
820	820	20220513	A002	VDC011376H		319
821	821	20220513	A002	VDC011376H		319
822	822	20220513	A002	VDC011376H		319
823	823	20220513	A002	VDC011376H		319
824	824	20220513	A002	VDC011376H		319
825	825	20220513	A002	VDC011376H		319

Gambar 11 Tampilan Hasil Fakta *Attack*

Proses pengujian ETL dari data sumber hingga data akhir untuk *data warehouse* yang dilakukan dengan *tools* Pentaho telah berhasil dilakukan. Berdasarkan hasil pengujian tersebut maka diperoleh model skema *starflake data warehouse* untuk pengelolaan data *log* perangkat keamanan informasi yang akan diimplementasikan pada SOC. Gambar 12 menunjukkan tampilan hasil dari skema *starflake* pengelolaan *log* keamanan informasi untuk SOC.



Gambar 12 Model *Data Warehouse* Skema *Starflake* Pengelolaan *Log* Keamanan Informasi SOC.

4. KESIMPULAN

Berdasarkan pengujian dari proses *exract*, *transformation* dan *loading* yang telah dilakukan dengan metode Kimball dan dengan pengolahan *tools* Pentaho, sebuah *data warehouse* terpusat yang berisi *log* keamanan informasi berhasil dirancang sehingga dapat dimanfaatkan untuk implementasi SOC dan *reporting* terkait pengelolaan *log* perangkat keamanan informasi. *Data warehouse* yang dihasilkan dalam penelitian ini adalah model skema *starflake data warehouse* yang memperoleh hasil akhir berupa tabel fakta *malware*, fakta *intrusion*, dan fakta *attack* dimana merupakan solusi dalam merancang *data warehouse* yang menghasilkan tabel yang lebih teratur sehingga mudah untuk diolah sesuai kebutuhan. Karena pentingnya pengelolaan dan monitoring *log* keamanan informasi ini, saran kepada peneliti berikutnya yang ingin melanjutkan penelitian ini adalah dengan menggunakan pemodelan skema lainnya kemudian membandingkan *response time* yang dihasilkan agar dapat diketahui pemodelan mana yang dapat memberikan informasi yang lebih cepat dalam implementasi SOC.

Saran dan rekomendasi untuk instansi pemerintahan pusat ini adalah bahwa perlunya sebuah *data warehouse* untuk pengolahan data *log* keamanan informasi yang lebih baik, cepat dan efisien sehingga dapat digunakan untuk *monitoring* data terpusat dalam mendukung kebijakan *zero tolerance* keamanan informasi dan dapat membantu manajemen dalam pengambilan keputusan berdasarkan data yang akurat dan relevan. Hal ini juga dapat dilakukan untuk instansi pemerintahan lainnya yang melakukan pengolahan data dengan jumlah dan kapasitas yang sangat besar.

DAFTAR PUSTAKA

[1] P. D. Pandit, "An Analysis of Computer Security , Attack Models and Defensive Mechanisms," no. September, 2021, doi: 10.13140/RG.2.2.34616.06406.

[2] M. Siwach and S. Mann, "Anomaly Detection for Web Log based Data: A Survey," *2022 IEEE Delhi Sect. Conf. DELCON 2022*, vol. 13, no. 1, pp. 129–148, 2022, doi: 10.1109/DELCON54057.2022.9753130.

[3] M. Akbar and Y. Rahmanto, "Desain Data Warehouse Penjualan Menggunakan Nine Step Methodology Untuk Business Intelegency Pada Pt Bangun Mitra Makmur," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 1, no. 2, pp. 137–146, 2020, doi: 10.33365/jatika.v1i2.331.

[4] I. G. W. Darma, K. S. Utami, and N. W. S. Aryani, "Data Warehouse Analysis to Support UMKM Decisions using the Nine-step Kimball Method," *Int. J. Eng. Emerg. Technol.*, vol. 4, no. 1, pp. 1–5, 2019.

[5] A. Filiana, A. G. Prabawati, M. N. A. Rini, G. Virginia, and B. Susanto, "Perancangan Data Warehouse Perguruan Tinggi untuk Kinerja Penelitian dan Pengabdian kepada Masyarakat," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 174–183, 2020, doi: 10.28932/jutisi.v6i2.2557.

[6] N. S. Fitriasari, I. Ariawan, A. Rais, T. E. Ahmad, and R. D. Azhari, "View of Rancangan Dan Implementasi Modul Data Warehouse Dan Data Mining Sebagai Kritikal Sukses Faktor Pada Enterprise," *Pros. Semin. Nas. Ilmu Komput. Vol. 1, No 1.*, vol. 1, no. 1, pp. 41–52, 2021, [Online]. Available: https://proceeding.unived.ac.id/index.php/snasikom/article/view/51/45

[7] N. Hidayat *et al.*, "Analysis and Design of Data Warehouse Based on Sndikti Using Data Warehouse Life Cycle Method At Unsoed Engineering Analisis Dan Perancangan Data Warehouse Berdasarkan Sndikti Menggunakan Metode Data Warehouse Life Cycle Di Fakultas," vol. 3, no. 3, pp. 797–805, 2022.

[8] R. J. Salaki, J. Waworuntu, and I. R. H. T. Tangkawarow, "Extract transformation loading from OLTP to OLAP data using pentaho data integration," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 128, no. 1, 2016, doi: 10.1088/1757-899X/128/1/012020.

[9] A. D. Barahama and R. Wardani, "Utilization Extract, Transform, Load for Developing Data Warehouse in Education Using Pentaho Data Integration," *J. Phys. Conf. Ser.*, vol.

2111, no. 1, pp. 0–8, 2021, doi: 10.1088/1742-6596/2111/1/012030.

[10] A. Vaisman, *Data-Warehouse-Systeme*. 2007. doi: 10.1007/978-3-8350-9178-8_2.

[11] O. Aslan and R. Samet, "A Comprehensive Review on Malware Detection Approaches," *IEEE Access*, vol. 8, pp. 6249–6271, 2020, doi: 10.1109/ACCESS.2019.2963724.

[12] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, 2019, doi: 10.1186/s42400-019-0038-7.

[13] D. Juardi, "Kajian vulnerability keamanan data dari eksploitasi hash length extension attack vulnerability data satisfaction study from exploitation hash length extension attack," vol. 6, no. 1, 2017.

[14] A. Rahmatillah *et al.*, "Implementasi Intrusion Prevention System (IPS) Pada Keamanan Jaringan Dengan Notifikasi Berbasis Telegram di Jurusan Teknik Komputer," 2021.

[15] J. K. Anggraenni, "Simulasi Keamanan Pada Aplikasi Web Dengan Web Application Firewall," *Ilm. Komput.*, pp. 45–50, 2013.

[16] L. Z. A. Mardedi, "Analisa Kinerja System Gluster FS pada Proxmox VE untuk Menyediakan High Availability," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 19, no. 1, pp. 173–185, 2019, doi: 10.30812/matrik.v19i1.473.

[17] A. Madani, S. Rezayi, and H. Gharaee, "Log management comprehensive architecture in Security Operation Center (SOC)," *Proc. 2011 Int. Conf. Comput. Asp. Soc. Networks, CASoN'11*, pp. 284–289, 2011, doi: 10.1109/CASON.2011.6085959.

[18] K. A. Shobirin, A. P. S. Iskandar, and I. B. A. Swamardika, "Data Warehouse Schemas using Multidimensional Data Model for Retail," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, p. 84, 2017, doi: 10.24843/ijeet.2017.v02.i01.p17.