LENY NOVITA SARI
Computer Engineering, Faculty of Computer Science
Sriwijaya University

Big Data NetFlow Analysis

Big Data adalah istilah yang menggambarkan volume data yang besar, baik yang terstruktur maupun yang tidak terstruktur, tapi itu bukan jumlah data yang penting. Big Data biasanya termasuk set data dengan ukuran diluar kemampuan perangkat lunak yang biasa digunakan untuk menangkap, mengelola, dan memproses data dalam waktu yang telah ditentukan. Ukuran data yang dapat digolongkan sebagai big data dapat berkisar dari beberapa terabyte untuk Yottabytes data dalam satu set data tunggal. Contohnya:

- Sebuah petabyte adalah 1000 terabyte.
- Sebuah exabyte adalah 1000 petabyte.
- Sebuah zettabyte adalah 1000 exabyte.
- Sebuah yoyabyte adalah 1000 zettabytes.

Para pendiri Kentik telah mengamati selama bertahun-tahun untuk melihat efek dari Big Data pada manajemen jaringan. Manajemen jaringan adalah salah satu daerah terakhir dalam operasi digital dan IT yang telah mengambil keuntungan dari teknologi Big Data.

Salah satu hal yang benar-benar berbeda tentang Kentik adalah bahwa Kentik telah memutuskan untuk membangun sebuah arsitektur Big Data pada inti dari solusi pemantauan jaringan. Yang membawa beberapa keuntungan nyata, karena Big Data tidak hanya tentang penanganan volume data yang besar, tetapi juga tentang navigasi dan penjelajahan data yang sangat cepat. Big Data Analysis benar-benar membantu beberapa hal penting, salah satunya adalah visibilitas yang jelas dalam aktivitas saat ini pada jaringan.

Ada tiga item kunci high-level untuk Big Data Analysis :

- Information management : sebuah manajemen dan kontrol proses untuk Big Data Analysis
- High-performance analytics : Kemampuan untuk mendapatkan tindakan lanjutan dari big data dan dan mampu memecahkan masalah yang kompleks menggunakan lebih banyak data.
- Flexible deployment options: Pilihan untuk berbasis lokal atau berbasis cloud, software-as-a-service (SaaS) untuk analisis data besar.

Ada beberapa pendekatan high-level untuk mempercepat Big Data Analysis :

- Grid computing : Sebuah infrastruktur jaringan terpusat untuk analisis dinamis dengan ketersediaan yang tinggi dan pemrosesan paralel.
- Intra-database processing : Performing data management, analytics, and reporting tasks menggunakan arsitektur scalable.
- In-memory analytics: memecahkan masalah yang kompleks secara cepat menggunakan in-memory.
- Support for Hadoop : menyimpan dan memproses volume Big Data pada komoditas perangkat keras.
- Visualizations : memvisualisasikan korelasi dan pola dalam Big Data secara cepat untuk mengidentifikasi peluang untuk analisis lebih lanjut dan untuk meningkatkan pengambilan keputusan.

Big Data Architectural Considerations

Tidak semua arsitektur Big Data secara khusus dirancang untuk kasus penggunaan monitoring jaringan. Banyak sistem Big Data yang bekerja dengan sangat baik pada pengumpulan semua data, tapi kemudian ketika ingin menjalankan report, dapat mengambil puluhan menit, bahkan di kasus terburuk dapat menyita waktu selama berjam-jam, untuk mendapatkan kembali data yang diinginkan. Jadi salah satu tantangan besar dengan solusi yang dibangun di sekitar arsitektur Big Data adalah untuk memberikan kemudahan akses, kecepatan akses, dengan kemampuan terbatas untuk semua data yang tersedia untuk kita.

Itulah salah satu hal yang benar-benar berfokus di Kentik. Kentik menangani volume data yang sangat besar di backend, dengan miliaran catatan yang masuk ke SaaS setiap hari. Tapi hal lainnya adalah secepat mungkin mampu mendapatkan data kembali.

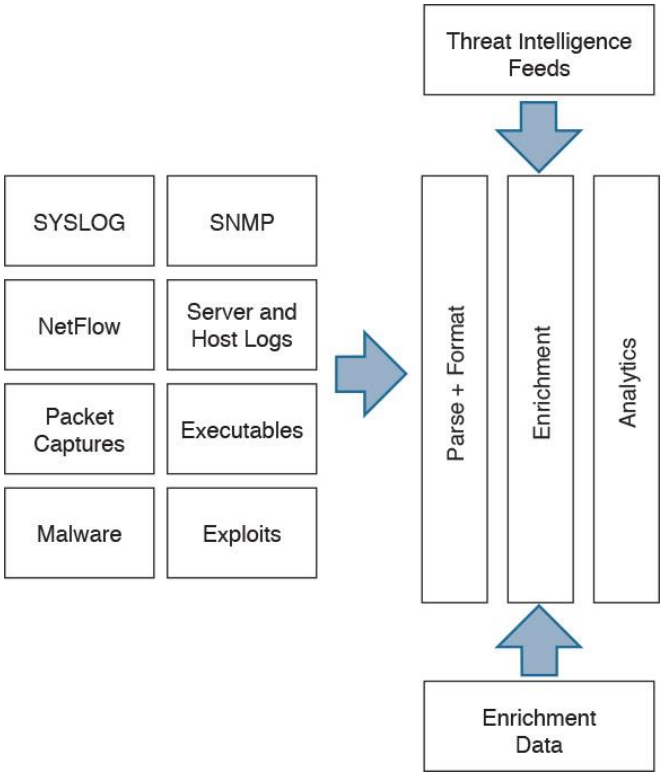
Beberapa tren lain yang benar-benar menarik dan yang Kentik lihat adalah hal-hal seperti SDN karena SDN mampu mengubah perilaku jaringan. Misalnya, bagaimana perilaku sebuah jaringan setelah sebuah perubahan dibuat melalui semacam penegakan kebijakan software-defined.

Jadi Arsitektur Big Data yang dibangun akan dirancang untuk menjadi sangat fleksibel, dan memiliki kehidupan yang sangat panjang. Karena kita benar-benar menggunakan API untuk memungkinkan terhubungnya ke sistem lain. Hal ini dapat terhubung dengan sangat sederhana dan mudah dengan jenis sumber data lain.

Big data NetFlow Visibility

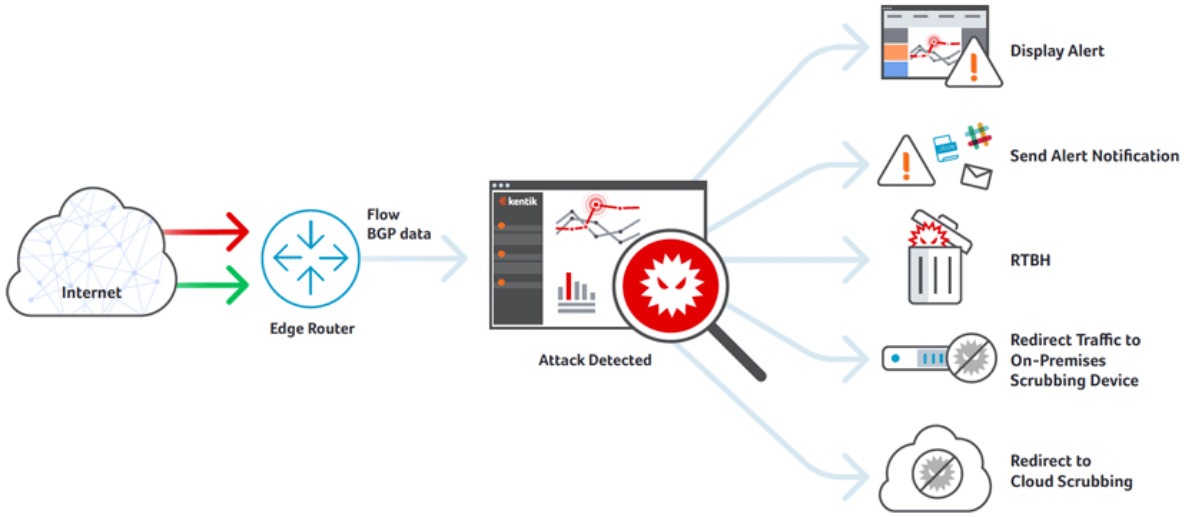
NetFlow memberikan rincian telemetri jaringan yang memungkinkan administrator untuk :

- Melihat apa yang sebenarnya terjadi di seluruh jaringan
- Mendapatkan kembali kontrol dari jaringan, dalam kasus serangan denial-of-service (DoS)
- Secara cepat mengidentifikasi endpoints dan perangkat infrastruktur jaringan
- Memonitor penggunaan jaringan
- Memperoleh telemetri jaringan saat penanggapan insiden keamanan dan forensik
- Mendeteksi kesalahan konfigurasi firewall dan akses yang tidak pantas



NetFlow and Other Telemetry Sources

Beberapa hal yang kita lihat sekarang adalah untuk menambahkan jenis keamanan yang lebih dalam, lebih kaya, dan lebih spesifik, peringatan, dan deteksi anomali. Dengan mengakui bahwa DDoS adalah masalah Big Data dan menghapus kendala arsitektur skala-up. Faktanya adalah bahwa ada miliaran catatan arus lalu lintas untuk menelan jutaan IP yang dilacak secara individual dan diukur untuk anomali. Satu pelanggan Kentik, melaporkan peningkatan yang lebih besar dari 30 persen dalam menangkap dan menghentikan serangan DDoS karena menerapkan built-in deteksi dan peringatan dari deteksi Kentik.



Jim Frey dan team sedang dalam proses meningkatkan kemampuan untuk secara otomatis mengenali penyimpangan, sehingga kita dapat memahami anomali ketika mereka datang. Juga meningkatkan kemampuan untuk langsung mengintegrasikan dengan sistem eksternal untuk memicu respons jika terjadi masalah.