

Arsitektur Data Modern

Konsep Pengembangan Arsitektur Data Masa Kini

Versi 0.1, 14 Juni 2019

Yudha H Tejaningrat

yht@kalamangga.net

Daftar Isi

Abstraksi.....	1
Pendahuluan.....	1
Perkembangan Arsitektur Data.....	2
Basis Data Tradisional.....	2
Data Warehouse.....	2
Masalah Utama Arsitektur Data.....	3
Munculnya Arsitektur Data ‘Modern’.....	3
Perkembangan Teknologi Basis Data.....	4
Teknologi No-SQL.....	4
Teknologi <i>Big Data</i>	4
Solusi yang Hadir.....	4
Kesimpulan.....	5

Abstraksi

Perkembangan teknologi data yang meningkat pesat menimbulkan kegaduhan, utamanya dengan timbulnya 2 (dua) komponen yang menjadi titik pusat yaitu *big data* dan *data science*. Banyak hal menjadi perdebatan mengenai 2 (dua) terminologi baru ini, utamanya mengenai peran dan transformasi yang perlu dilakukan menyikapi tren yang mengikutinya.

Dalam tulisan ini coba diceritakan sejarah perkembangan arsitektur data dan saran mengenai pengembangan arsitektur data menuju arsitektur data modern sehingga peran dan transformasi dalam bisnis bisa berjalan optimal mengikuti tren yang berkembang.

Pendahuluan

Era *Big Data* mengubah banyak hal, utamanya sudut pandang dalam melihat dan merepresentasikan data. Hal ini mengubah pula desain dalam pengelolaan data secara keseluruhan sehingga memunculkan Arsitektur Data Modern yang menawarkan solusi komprehensif dalam pengelolaan data.

Arsitekstur data modern diharapkan dapat memenuhi kebutuhan industri dimana kebutuhan akan data yang besar, mengalir deras dan bermacam bentuk, atau lebih dikenal dengan 3Vs (Volume, Velocity, Variety). Kebutuhan ini seiring dengan perkembangan teknologi analisis yang tidak hanya mampu mengolah data terstruktur namun juga data semi terstruktur dan tidak terstruktur. Baik data dari aplikasi yang dikembangkan sendiri ataupun data yang dikumpulkan dari perangkat/pihak lain dan media sosial yang tersedia umum.

Dengan arsitektur data modern ini diharapkan dapat memenuhi kebutuhan dalam kecepatan dan kelincahan dengan mengakselerasi organisasi dalam menemukan dan meng-integrasi-kan seluruh data dalam berbagai platform yang dimiliki.

Perkembangan Arsitektur Data

Secara sederhana, arsitektur data adalah kebijakan tentang pengelolaan data. Kebijakan mengenai bagaimana data diatur dan dikelola dalam organisasi, baik dalam tataran teknik pengelolaan maupun operasional.

Dalam penerapannya, arsitektur data merupakan seperangkat aturan yang mencakup kebijakan dan model yang menentukan jenis data yang dikumpulkan, dan bagaimana ia digunakan, diproses dan disimpan dalam sistem basis data. Misalnya dalam integrasi data, proses pengolahan akan sangat terpengaruh dari arsitektur yang dimiliki tanpa perubahan paradigma dari pemrograman ke arsitektur data.

Basis Data Tradisional

Pengembangan awal arsitektur data tidak terlepas dari pengembangan aplikasi dalam bisnis itu sendiri. Basis data tradisional adalah sebuah perangkat lunak yang dapat menyimpan seluruh data yang digunakan oleh aplikasi yang dikembangkan untuk memenuhi tujuan bisnis mulai dari antarmuka dengan pengguna dalam hal ini karyawan maupun pelanggan hingga memenuhi kebutuhan eksekutif dalam hal memantau operasional bisnis. Untuk lebih mudahnya dengan melihat bagan di bawah.

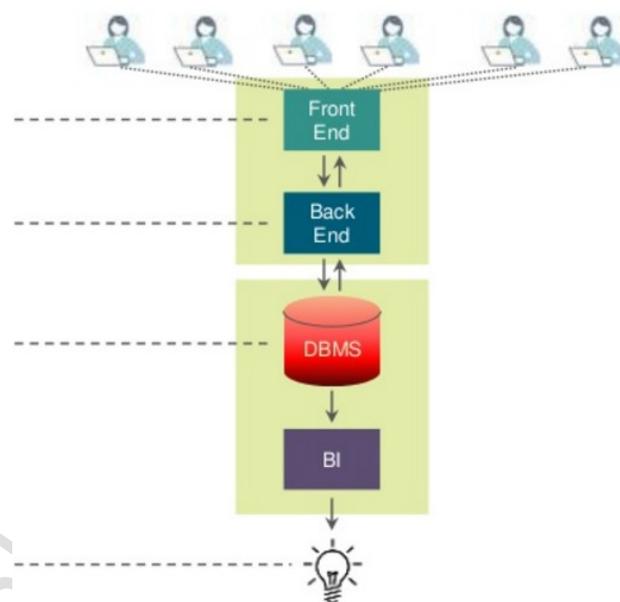


Figure 1: Single DBMS¹

Dengan unit bisnis skala kecil, penggunaan arsitektur ini cukup memadai mengingat kebutuhan atas kemampuan basis data memproses transaksi dan memenuhi kegiatan operasional tidak terlalu besar. Dengan jumlah akses karyawan kurang dari 10 (sepuluh) orang sekali waktu dan pembagian tugas yang tidak banyak tingkat kewenangannya tentunya kebutuhan perangkat tidak terlalu rumit.

Data Warehouse

Setiap orang tentunya ingin bisnis yang dijalankan berkembang. Tak terkecuali kebutuhan perangkat yang digunakan tentunya meningkat seiring dengan perkembangan usaha. Misalkan pemecahan divisi dan pembukaan kantor cabang baru. Selain itu, perubahan aplikasi yang digunakan juga merupakan bagian dari perkembangan bisnis.

1 <https://www.experfy.com/blog/modern-data-architecture>

Dengan perkembangan ini tentunya arsitektur data yang ada harus mampu mengakomodasi kebutuhan semua. Dalam arsitektur data tingkat *enterprise* atau yang lebih dikenal dengan *Data Warehouse* kebutuhan itu diakomodasi dengan membuat perangkat tugas baru yang mengumpulkan semua data dari semua aplikasi yang ada dalam bisnis. Semua data itu dikumpulkan dalam sebuah basis data baru sebagai masukan perangkat *Business Intelligence* dimana eksekutif memantau operasional bisnis.

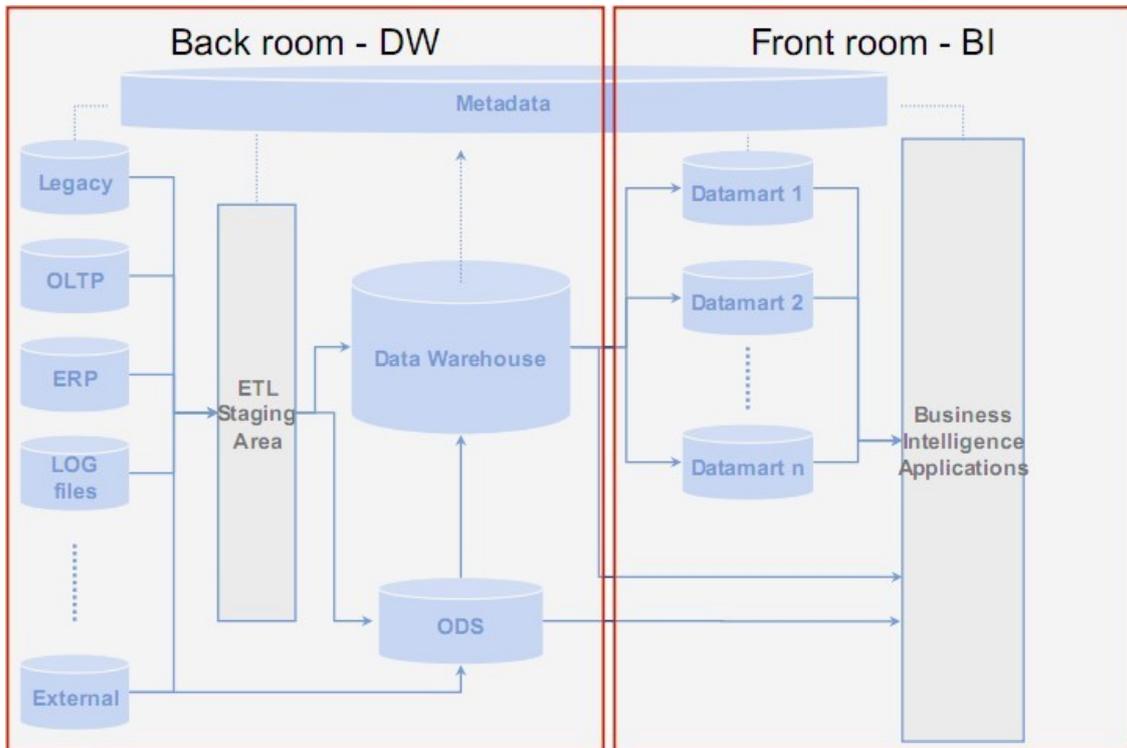


Figure 2: Data Warehouse Architecture²

Pengembangan *Data Warehouse* menjadi suatu hal yang umum. Seringkali dalam sebuah usaha yang telah berkembang kebutuhan atas perangkat / aplikasi yang dibutuhkan antar divisi berbeda. Struktur organisasi juga menjadi hal yang berpengaruh utamanya banyaknya kantor cabang yang tidak mampu terhubung secara langsung / *on-line* / *real-time* dengan sistem utama. Sehingga dibutuhkan sebuah basis data utama yang mencakup semua data tersebut.

Masalah Utama Arsitektur Data

Dengan model pengembangan arsitektur yang ada hingga kini terdapat banyak isu utamanya dalam layanan yaitu kualitas, ketersediaan dan tanggung jawab. Beberapa pertanyaan yang selalu muncul adalah:

1. Bagaimana tingkat validitas data yang terkumpul?
2. Seberapa mutakhirkah data yang terkumpul?
3. Berapa orang yang dibutuhkan dalam mengelola dan bagaimana pembagian wewenangnya?

Tidak ada jawaban pasti mengenai pertanyaan yang muncul tersebut, karena menyangkut skala usaha dan seberapa kompleks kebutuhan eksekutif dalam melihat bisnis.

Selain ketiga hal di atas, perkembangan analisis yang tidak hanya mengolah data tabular menjadi tantangan sendiri dalam pengembangan arsitektur data. Misalkan kebutuhan dalam mengolah data teks, gambar, dan suara.

Munculnya Arsitektur Data 'Modern'

Menjawab masalah utama arsitektur data pada bab sebelumnya banyak bermunculan, teknologi mulai berkembang dan beberapa solusi hadir dalam rangka memenuhi kebutuhan yang ada.

² <https://gerardnico.com/data/warehouse/architecture>

Perkembangan Teknologi Basis Data

Perkembangan teknologi basis data dilatar belakangi utamanya kinerja menangani data yang besar dan fleksibilitas dalam penanganan data. Kebutuhan memory membengkak seiring besarnya data yang ditangani dan diolah. Pun dengan model relasi data yang memerlukan penanganan detail agar basis data dapat digunakan secara optimal.

Teknologi No-SQL

Antara akhir abad 19 hingga tahun 2000 muncul pengembangan-pengembangan menyikapi kekurangan basis data relasional. Berbagai *paper* bermunculan sebagai desain pengembangan dalam mencari solusi kekurangan basis data relasional. Puncaknya muncul gerakan “*noSQL Movement*” yang diprakarsai oleh Johan Orkarsson.

No-SQL bukan merupakan bahasa layaknya SQL, namun merupakan tataran konsep mengenai mekanisme untuk menyimpan data dan mengambil data. Mekanisme yang dibangun tidak memerlukan model relasi data dan standar baku bahasa (SQL). Dalam mengelola data, No-SQL menggunakan meta data dan index pada data.

Beberapa mekanisme yang menjadi turunan dalam konsep No-SQL sebagai berikut.

1. Table-oriented
2. Graph-oriented
3. Document-oriented
4. Key-value store

Kelebihan utama yang menonjol dari No-SQL dibandingkan basis data relasional (SQL) adalah skalabilitas, pengelolaan dan skema pengembangan.

Teknologi Big Data

Banyak definisi yang berbeda mengenai terminologi *big data*. Dalam praktik pengembangan pengelolaan data, data tentang banyak hal yang dikumpulkan dengan volume yang besar dan sangat cepat semakin menyesaki ruang hardisk. Dari praktik tersebut muncul “hukum” 3V, yaitu *Variety*, *Volumes* dan *Velocity* yang sering dihubungkan dengan penerapan *big data*.

Dari banyaknya definisi dan “hukum” 3 V tersebut, salah satu penjelasan yang mudah dipahami mengenai *big data* adalah pengumpulan dan penggunaan data / informasi dari berbagai sumber untuk membuat keputusan yang lebih baik. Dengan kata lain, *big data* adalah sebuah konsep mengenai metode pengumpulan data, metode analisis data dan metode memahami data dengan cepat yang datang setiap saat.

Pengembangan teknologi ini berhasil mengurangi biaya penyimpanan dan komputasi data karena dapat menggunakan mesin dan perangkat standar komputasi rumahan. Sehingga dengan murahnya biaya data yang dikumpulkan dapat lebih banyak dan lebih mudah diakses. Dengan menggunakan data yang lebih besar keputusan bisnis yang diambil dapat lebih akurat dan tepat.

Yang perlu diingat, dalam mengambil *insight* dan membuat keputusan bisnis tidak hanya tentang proses analisisnya. Perlu wawasan menyeluruh tentang hal yang berkaitan dengan bisnis, pengenalan pola, pengambilan asumsi dan pendeteksian perilaku. Dan seringkali data / informasi mengenai semua hal ini tersebar dan bervariasi modelnya, sehingga perlu tempat yang dapat mengakomodasi.

Solusi yang Hadir

Banyak sekali produk yang muncul dengan mengambil konsep teknologi arsitektur data modern tersebut. Sebagai contoh, berikut adalah *Unified Data Architecture* yang ditawarkan oleh Hortonworks³ sebagai hasil kerjasama dengan Teradata⁴.

3 <https://hortonworks.com>

4 <https://teradata.com>

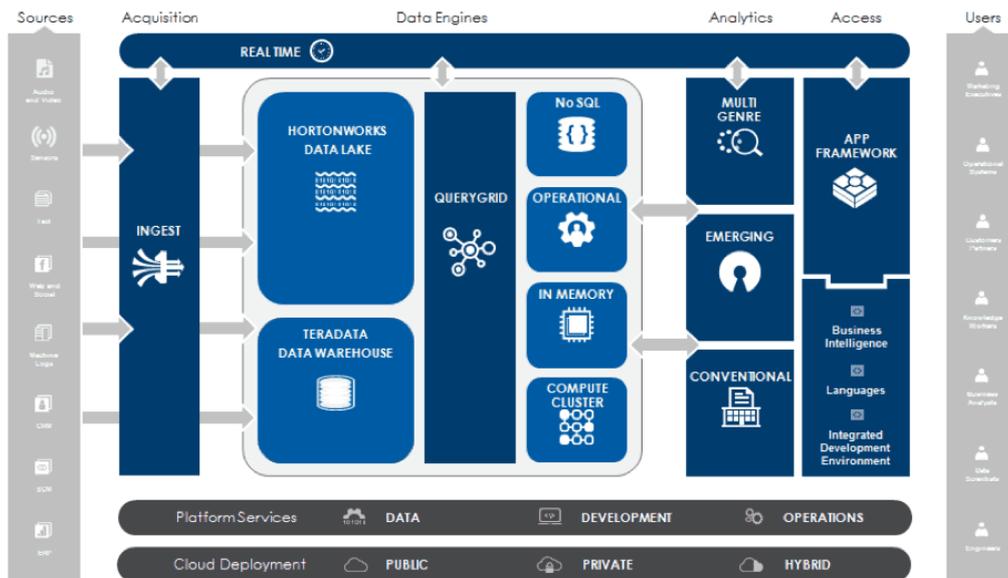


Figure 3: Unified Data Architecture, Hortonworks, Teradata's Partner⁵

Beberapa *vendor* memiliki penggambaran yang berbeda namun dengan pendekatan yang sama, yaitu dalam pengelolaan data terdapat beberapa *stake-holder* yang memiliki wewenang dalam setiap langkah yang ada. Baik itu dari pengumpulan data, pengolahan dan pendistribusian data hingga pengambilan *insight* dari data yang ada.

Basis data diharapkan mampu mengelola seluruh data yang menyangkut kepentingan bisnis di setiap lini sebagai satu sumber terpercaya (*single source of truth*). Sehingga alur data menuju kepada satu tempat dan analisis pendukung kebutuhan bisnis dapat mengambil di satu tempat itu. Tidak ada sumber lain dalam sumber data dari *insight* yang mendasari keputusan bisnis.

Dalam pendekatan digunakan terminologi *ingest* yang berarti menelan. Pendekatan ini dilakukan menyangkut kegiatan akuisisi data dengan mengambil dari sumber yang beraneka ragam dan meletakkannya pada satu tempat.

Data yang telah berada dalam satu tempat tersebut digunakan sebagai sumber informasi dalam melakukan analisis dalam berbagai aspek bisnis. Platform data ini akan melayani permintaan konvensional maupun dalam *business intelligence* dan ruang lingkup aplikasi pendukung keputusan.

Kesimpulan

Pengembangan arsitektur data mengikuti kebutuhan bisnis sebagai pangsa pasar. Berbagai pendekatan dalam penerapan memberikan solusi kepada bisnis untuk dapat membuat keputusan yang tepat dari berbagai data yang dimiliki. Tidak ada standar baku dalam penerapan pendekatan yang bisa diambil karena kebutuhan data setiap bisnis akan berbeda namun hendaknya pendekatan dapat mengakomodasi adanya perubahan dan pengembangan lebih lanjut.

Dalam melakukan pemilihan pendekatan hendaknya mengikuti kebutuhan data utama (primer) dalam bisnis dan akuisisi data sekunder dapat dilakukan kemudian. Hendaknya semua data di berbagai lini bisnis dapat menyatukan datanya dalam suatu wadah. Sehingga keputusan bisnis dapat akurat dan tepat.

Dalam dokumen versi ini jawaban atas masalah masih belum dibahas. Pembahasan ini akan coba dijabarkan analisis dan solusi penerapannya pada rilis dokumen berikutnya.

⁵ <https://hortonworks.com/partner/teradata/>