MarkLogic WORLD 2017

How Aetna Uses an Operational Data Hub to Power Business Success

Michael Fillion, Director of Architecture, Aetna
Tyler Replogle, Principal Consultant, MarkLogic

Aetna Inc.

Quality health plans & benefits
Healthier living
Financial well-being
Intelligent solutions

aetna℠

Enterprise Operational Data Hub – Key Enterprise Use Cases Solved!

Michael G. Fillion

Director of Architecture

About The Presenter

Michael G. Fillion

Architectural Director

Mr. Fillion is the managerial lead for Architecture Delivery covering the International, Insights, Workers Comp, Shared Services, and Finance technology areas. Mr. Fillion is responsible for multiple Teams of Application and Information Architects which support all architectural project deliverables. His Team is part of the Enterprise Architecture area at Aetna under the Chief Technology Officer at Aetna.

Mr. Fillion has been with Aetna for 9 years and has over 34 years total experience in Information Technology most of which supporting Insurance related applications and technologies. Mr. Fillion's past roles include Architectural Director, Management Consultant, and Business Technology Senior Advisor.

Mr. Fillion specializes in insurance technology most specifically Enterprise Data Management, NoSQL, Business Intelligence and Analytics, Data Warehousing, and Information Delivery. Other companies for which Mr. Fillion has worked are The Travelers Insurance Company, Massachusetts Mutual Life Insurance Company, and Cigna.

Mr. Fillion is a 1982 graduate and holds a degree in Business Administration from Springfield Technical College in Springfield Massachusetts and a Certificate in Computer Technology and Programming from the former Computer Processing Institute in East Hartford Connecticut.

In This Session We Will Cover

- *Aetna's Business Data Integration Use Cases:*
  - *Integrated Healthcare Data Search and Display*
  - *Data Governance for Integrated Data*
- *Problematic Traditional Solutions for Enterprise Data Integration*
- *Strategic Enterprise Data Integration Solutions:*
  - *Application Integration*
  - *Enterprise Search*
  - *Operational Data Triage*

Aetna's Business Use Cases

→ *What Is Aetna's Business Trying to Accomplish?*

- *Integrated Healthcare Data Search and Display*
  - *Web Application Allowing for the Search, Query and Display of Operational Healthcare Data:*
    - *Member Data*
    - *Claim Data*
    - *Enrollment Data*
    - *Provider Data*
- *Data Governance for Integrated Operational Data*
  - *Full Data Auditability, Balancing, and Control*
  - *Transparent Data Lineage and Traceability*

Aetna Has Recognized This About Data:

→ *Operational Data Is the Lifeblood of a Healthy Business Operation*

→ *Data Governance Standards for Data Integration:*

- ❑ *Data Must Be **Trusted** (Accurate, Complete, and in Balance)*
- ❑ *Data Must Be **Consistent** (Conformed Across Sources)*
- ❑ *Data Must Be **Timely** (Meeting All SLA's)*

Trusted
Consistent
Timely

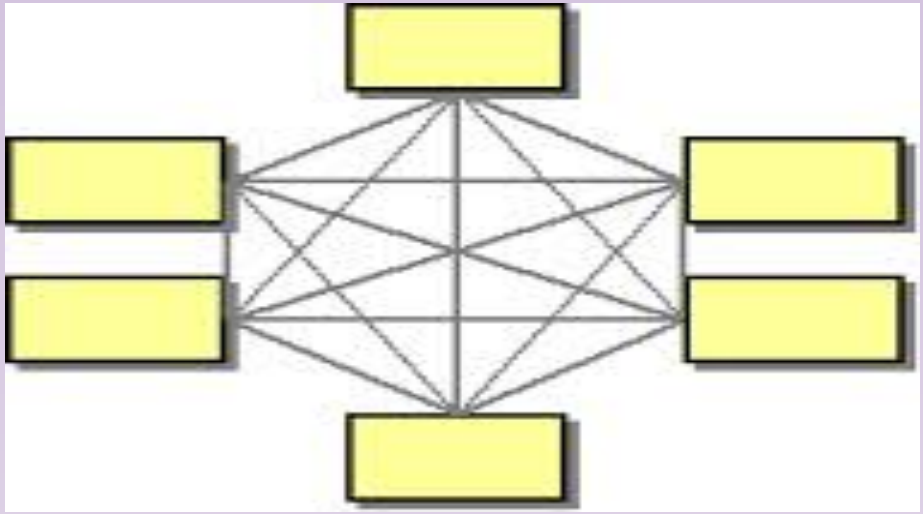
≈

smooth operation

The Pathology of Poor Data Integration – Point to Point

The Profile of Bad Design:

- ***Tight Coupling***
- ***Redundant Code***
- ***Inconsistent Code***
- ***High Maintenance Cost***

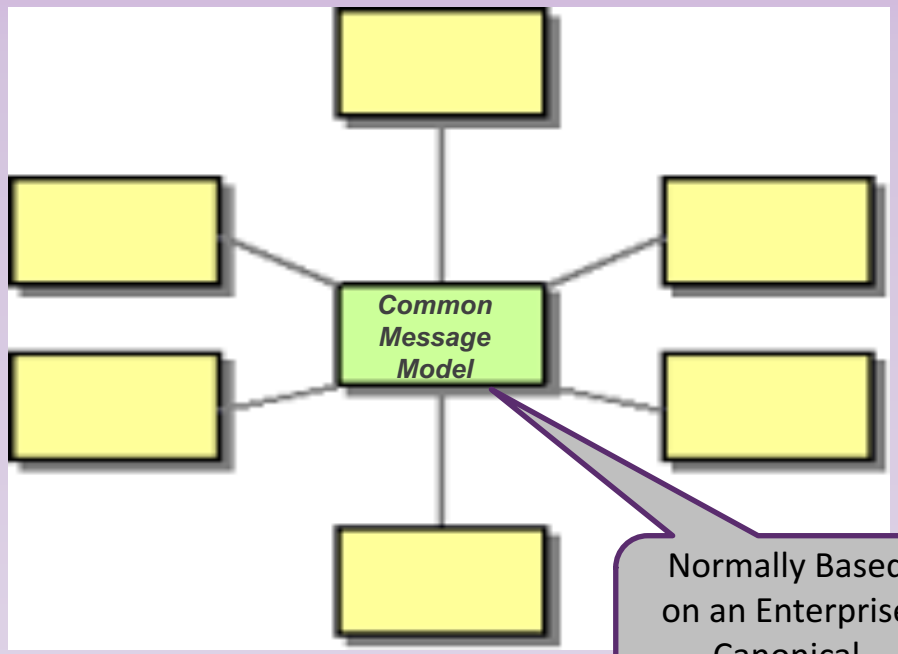


-- High Cost
--Inconsistent
--Timely?

Data Integration – Removing Point to Point

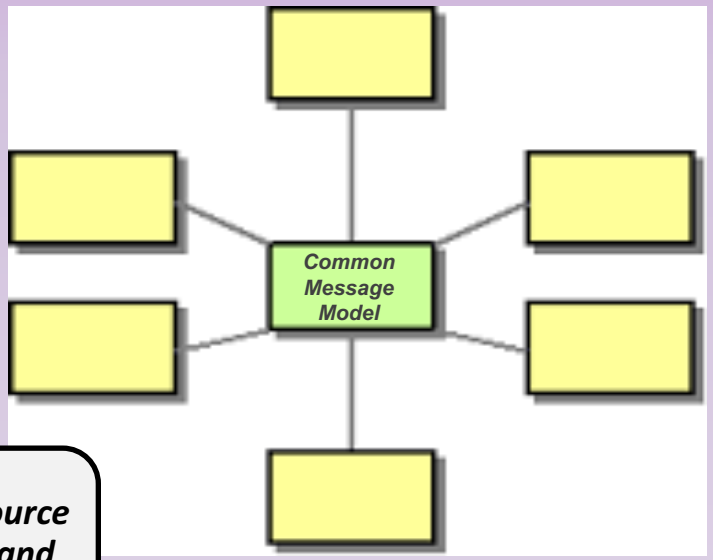
Value Points:

- ✓ *Decoupled Applications*
- ✓ *Reusable Code*
- ✓ *Consistent Code*
- ✓ *Low Maintenance Cost*



But…

Having A Common Message Model Is Not Enough



Result:

- *Dynamic message conformance can make consistency and reuse of code much harder.* — **Data Conformance**
- *SLA's may require source applications to expand their hardware footprint.* — **Source Data Cost**
- *Additional hardware for message caching and to perform complex integrations centrally may be needed.* — **Central Data Cost**
- *Message data lineage back to source can be difficult or impossible if updated data is gone at source.* — **Data Lineage**
- *Message data is much harder for auditors to use and report on.* — **Data Audit**

Enterprise Data Integration Elements

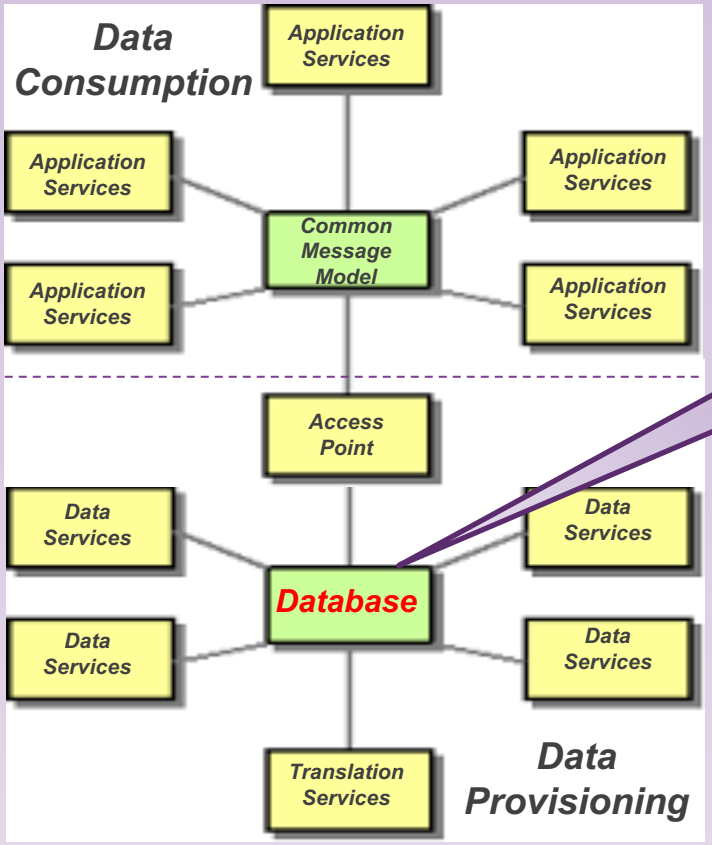
--High Cost?
--Inconsistent?
--Timely?

Does Not Meet Data Governance Standards

Common Message Model Is Not Enough So What's Missing?

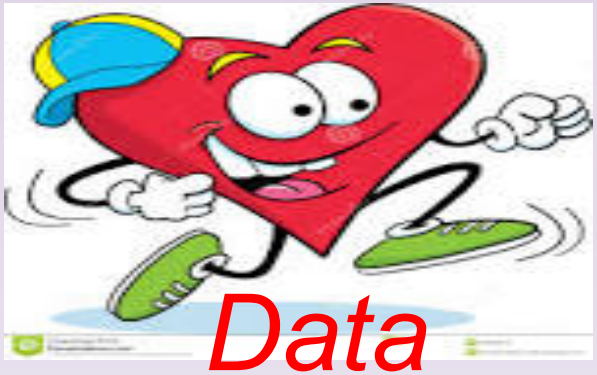
There Is *One* Main Architectural Factor

The Key to Strategic Data Integration?



It's the *database*!

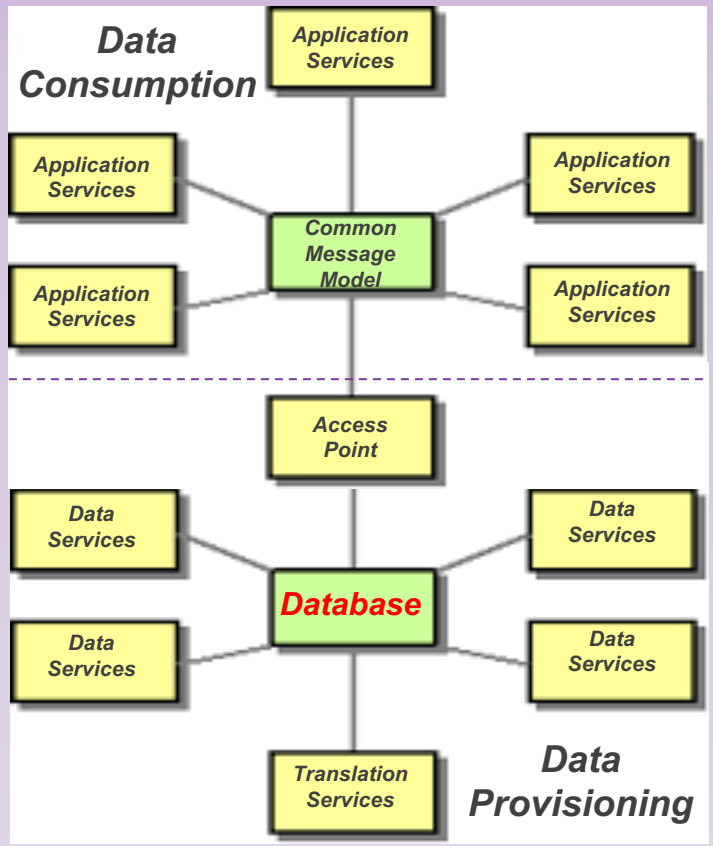
…aka An Enterprise Operational Data Hub



The Pathology of Strategic Data Integration

Value Points:

- ***Data-centric data conformance ensures consistency and enables code reuse.*** — ***Data Conformance***
- ***Source applications are minimally impacted as prefetch of data replaces real-time calls.*** — ***Source Data Cost***
- ***Message caching size can be cut back saving cost.*** — ***Central Data Cost***
- ***Data lineage back to source data state is ensured through the storage of raw data.*** — ***Data Lineage***
- ***Normalized data facilitates auditor data ease-of-use.*** — ***Data Audit***

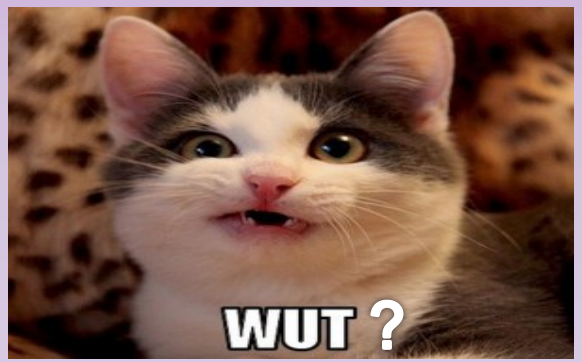


A Database Promotes the Above Elements

--Low Cost
--Consistent
--Timely

One More Critical Ingredient for Strategic Data Integration

What Kind of Database Are You Using?



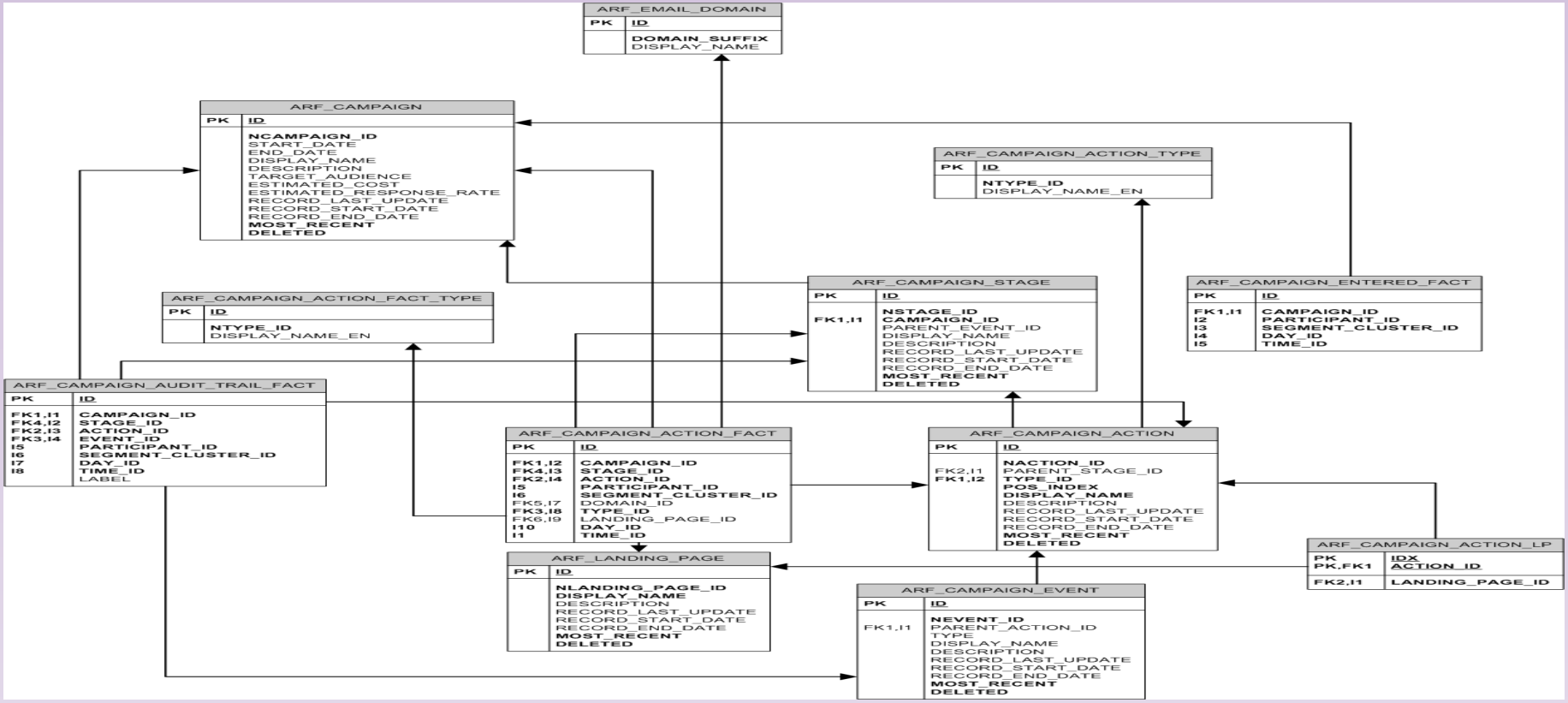
An RDBMS of Course! – Right?

Wrong! - You Have An *Architectural Mismatch*

Why Can't an RDBMS Fulfill Our Needs?

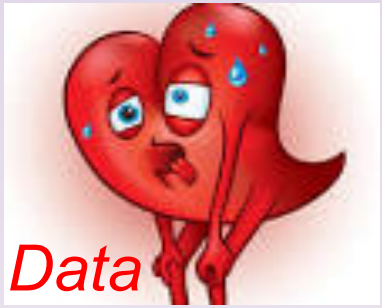
A Central Relational Database Is Impossible!

RELATIONAL DATA MODEL



- ***Developed Mostly Entirely Up Front – Usually Takes Many Months***
- ***Incremental Changes Are Difficult – Cost Is High***
- ***Rigid Across Disparate Data – Operational Conformance Is Hard***
- ***Speed to Market – Very Slow***

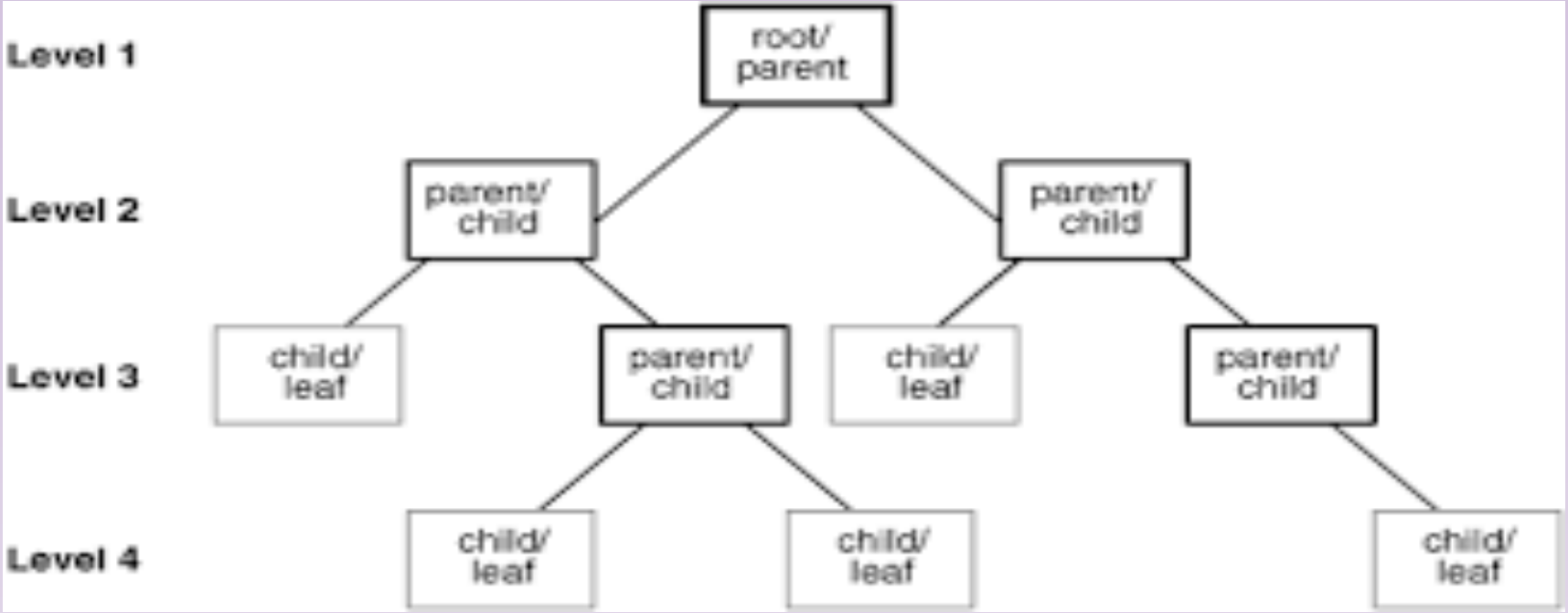
--High Cost
--Slow
--Inflexible



Note: Constantly Changing and Disparate Data Is Becoming Today's Reality

A NoSQL Database Solves The Riddle

CANONICAL DATA MODEL



- ***Database Model Developed with <u>Agility</u>***
- ***<u>Minimal Rework</u> Upon Data Model Changes***
- ***Hierarchically <u>Relates All Classes of Data</u> Logically <u>Without Duplication</u>***
- ***<u>Flexible</u> to Support And Relate Disparate Source Data***
- ***Faster <u>Speed to Market</u>***

--Agile
--Fast
--Flexible

Use Case #1: Application Integration Requirements

Overall Business Needs:

- *Real-time and Batch Data Integration*
- *Increased Speed to Market Over Past Programs*
- *Stringent Response Time for Services*

Business Value:

- *Real-time Data Search*

Architectural Principles Required to Be Fulfilled:

- *Data Decoupling Source Data from Target Data*
- *Code Reuse Is High with Little Duplication*
- *Development Is Agile*
- *Transactions Are ACID Compliant*
- *Consistent Data Access for Services*
- *Minimal Resource Impact to Source Applications*

BUSINESS USE CASE

- *Web Application Allowing for The Query And Display of Operational Healthcare Data:*
  - *Member Data*
  - *Claim Data*
  - *Enrollment Data*
  - *Provider Data*
- *Data Governance Program Standards:*
  - *Full Data Auditability, Balancing, and Control*
  - *Transparent Data Lineage and Traceability*

Use Case #1 – Application Integration
How Does The Data Look In The Database?

- ***Full Data Conformance to Enterprise Model:***

```
<provider_at_location id="23CBW">
          <src_pvd_id>"123"</src_pvd_id>
          <src_loc_id>"24"</src_loc_id>
          <pvd_name>"St Mary Hospital"</pvd_name>
          <pvd_at_loc_name>"Bucks County"</pvd_at_loc_name>
          <pvd_at_loc_products>
                    <pvd_at_loc_product>"Medical"</pvd_at_loc_product>
                    <pvd_at_loc_product>"BH"</pvd_at_loc_product>
          </pvd_at_loc_products>
          <pvd_at_loc_specialties>
                    <pvd_at_loc_specialty>"Hospital"</pvd_at_loc_specialty>
                    <pvd_at_loc_specialty>"ER"</pvd_at_loc_specialty>
                    <pvd_at_loc_specialty>"Outpatient"</pvd_at_loc_specialty>
          </pvd_at_loc_specialties>
          <pvd_at_loc_type>"Facility"</pvd_at_loc_type>
          <pvd_at_loc_addr>
                    <street>"123 Main St."</street>
                    <city>"Langhorne"</city>
                    <state>"PA"</state>
                    <zip>"19053"</zip>
          </pvd_at_loc_addr>
          <pvd_at_loc_contract>"ABC234"</pvd_at_loc_contract>
          <pvd_at_loc_phone>"215-555-1212"</pvd_at_loc_phone>
          <pvd_at_loc_enrich>
                    <pvd_at_loc_id>"101"</pvd_at_loc_id>
                    <pvd_at_loc_src_id>"S1"</pvd_at_loc_src_id>
                    <pvd_at_loc_tax_id>"A123"</pvd_at_loc_tax_id>
          </pvd_at_loc_enrich>
</provider_at_location>
```

Use Case #2: Enterprise Search Requirements

Overall Business Needs:

- *Real-time Data Search*
- *Increased Speed to Market Over Past Programs*
- *Stringent Response Time for Services*

Business Value:

- *Real-time Data Search*

Architectural Principles Required to Be Fulfilled:

- *No Data Modeling Needed to Load Raw Data*
- *Development Is Agile*
- *Consistent Data Access for Services*

BUSINESS USE CASE

- *Web Application Allowing for The Search And Display of Operational Healthcare Data:*
  - *Member Data*
  - *Provider Data*
- *Data Governance Program Standards:*
  - *Full Data Auditability, Balancing, and Control*
  - *Transparent Data Lineage and Traceability*

Use Case #2 – Enterprise Search
How Does The Data Look In The Database?

- ***Raw Operational Data Decorated With Conformed Enterprise Dimensions:***

```
<provider_at_location id="123CBX">
        <provider_id>"123"</provider_id>
        <location_id>"24"</location_id>
        <prvd_descr>"St Mary Hospital"</prvd_descr>
        <prvd_location>"Bucks County"</prvd_location>
        <prvd_at_loc_specialties>
                <prvd_loc_specialty>"112"</prvd_loc_specialty>
        </prvd_at_loc_specialties>
        <provider_type_cd>"25"</provider_type_cd>
        <location_zip>"19053"</location_zip>
        <location_state>"PA"</location_state>
        <location_address>"123 Main St".</location_address>
        <location_city>"Langhorne"</location_city>
        <contract_cd>"ABC234"</contract_cd>
        <phone>"215-555-1212"</phone>
        <pvd_at_loc_srch_dim>
                <pvd_at_loc_state>"PA"</pvd_at_loc_state>
                <pvd_at_loc_county>"Bucks"</pvd_at_loc_county>
                <pvd_at_loc_zip>"19053"</pvd_at_loc_zip>
                <pvd_at_loc_specialties>
                        <pvd_at_loc_specialty>"Hospital"</pvd_at_loc_specialty>
                        <pvd_at_loc_specialty>"ER"</pvd_at_loc_specialty>
                        <pvd_at_loc_specialty>"Outpatient"</pvd_at_loc_specialty>
                </pvd_at_loc_specialties>
        </pvd_at_loc_srch_dim>
</provider_at_location>
```

Raw

Conformed Dimensions

Use Case #3: Operational Data Triage

Overall Business Needs:

- *Real-time and Batch Data Query and Profiling*
- *Zero Data Loss Due to Low Quality Data*
- *Stringent Response Time for Data Quality Reporting*

Business Value:

- *Ensuring that the quality of data is reported on in a timely manner allowing for appropriate treatment*

Architectural Principles Required to Be Fulfilled:

- *No Data Modeling Needed to Load Raw Data*
- *Development Is Agile*
- *Transactions Are ACID Compliant*
- *Consistent Data Access for Services*

BUSINESS USE CASE

- *Data Quality Triage Application Allowing for The Profiling of Data Regardless of Quality*
- *Data Governance Program Standards:*
  - *Full Data Auditability, Balancing, and Control*
  - *Transparent Data Lineage and Traceability*

Note: Can This Be Done Easily Without NoSQL?

Use Case #2 – Operational Data Triage
How Does The Data Look In The Database?

- ***Raw Operational Data Decorated with Enriched Descriptive Data:***

Business Control Is Gained

```
<claim_raw>
        <claim_id>"123"</claim_id>
        <claim_dt>"02/01/2017"</claim_dt>
        <paid_dt>"00/00/0000"</paid_dt>
        <paid_amt>"500.00"</paid_amt>
        <adj_paid_amt>"525zz"</adj_paid_amt>
        <source_data_descriptive_info>
                <originating_source_app>"CLAIMAPP"</ originating_source_app>
                <update_system_nm>"Sally McBad"</update_system_nm>
                <extract_date>"02/01/2017"</extract_date>
                <batch_job_num>"19053"</batch_job_num>
        </source_data_descriptive_info>
</claim_raw>
```

Raw

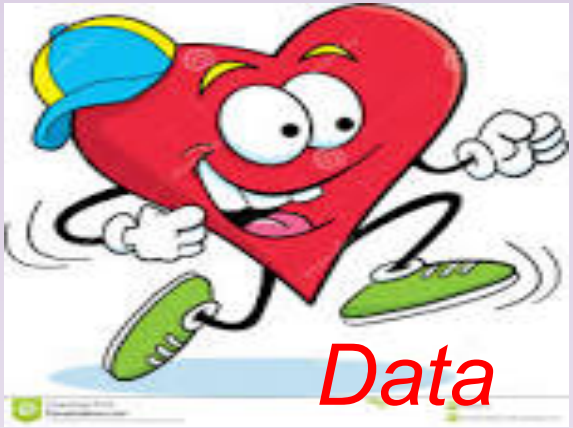
Dirty Data

Descriptive

Can You Load Dirty Data into a RDBMS?

Technology Consideration Summary

- *Make Sure You Consider Your Use Cases Together in Choosing the Proper Solution and Technology*
- *Also, Consider Your Future Needs as Much as Can be Anticipated (aka "Future Proofing")*
- *In Aetna's Case, A Single Solution Can Meet All Current Strategic Enterprise Data Integration Use Cases:*
  - *Application Integration*
  - *Enterprise Search*
  - *Operational Data Triage*
- *Positioned Well for the Future*

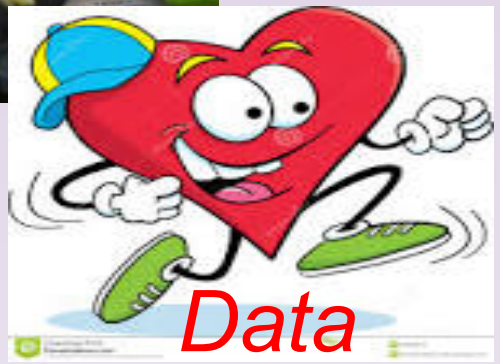


Conclusion

A Healthy Business Operation Needs Strategic Data Integration AND A NoSQL Central Database to Underpin It to Effectively Work

Koji, data is the lifeblood of the Business!

Thanks Papi, A NoSQL database was the key to integrating my data!!!!



Thank you

aetna℠