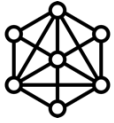


6C697 670
353646203
46203539
204741323
20415344
20363741
46363741
484A48414
446484A4
38394635
353941533
484746474
32336C697
53443536
534446203
484A20474
354620415
444620363
53444636
33344B A
4153444
373638394
444635394
484A48474
44513233
37415344
20415344
4446484A
337446
37415344
46415344
324733344
383941534
41533736
4153446

Why you should design your data hub top-down vs. bottom-up

K2VIEW >> *Data without delay.*

Why you should design your data hub top-down vs. bottom-up



Why are central repositories of data more necessary now than ever?

Every business organization, especially large enterprises, are seeing their pool of applications expand almost exponentially as they respond to market demand. The rate at which new systems are being introduced is very unlikely to slow down. SaaS offerings, full stack containers, and CI/CD initiatives are fueling the rapid, continuous adoption of a vast set of new systems. In this environment, IT organizations quickly realize the value and necessity for central repositories aggregating relevant data from across all these applications for orchestration, business continuity, analytics, auditing, or even sources of new revenue. These central repositories are at the heart of digital transformation initiatives. This paper will not evaluate what type of central repositories organizations should elect to adopt (data warehouses, data lakes, enterprise service buses, data hubs), but



rather, it describes the best practices around designing a data hub, which is the ideal architecture choice for the majority of digital transformation initiatives.

Why are central repositories of data more necessary now than ever?

What is a data hub?

Transforming via traditional architecture stack is a bit like trying to turn your horse into a camel.

Virtualization-only and schema-less architectures are the camel it seems the business needs, but they have inherent flaws.

The ideal data hub design: Stop trying to turn your horse into a camel when what you really need is a really fast llama-corn.



What is a data hub?

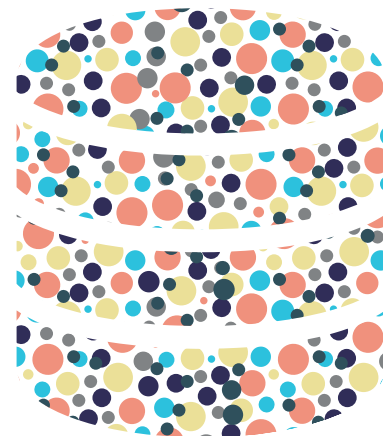
As opposed to data warehouses and data lakes, for which the original function is to store and distribute data, data hubs are defined as a collection of data from multiple sources, optimized not only for storage and distribution, but also for ingestion and exposure.

As such, data hubs can be equated to legacy architecture such as Enterprise Service Bus (ESB). ESBs are obviously not capable of handling the technology and vast amounts of data associated with today's digital transformations, but the comparison is nonetheless valuable, if only to understand the purpose of a data hub.

A data hub should be designed to serve any type of data to any client in any format for any purpose (operational system orchestration, micro service exposure, feed of data for analytics or auditing purposes, etc.). This makes agility, flexibility and rapidness response to ever-changing demands the cornerstone of data hub architecture design. Unfortunately, while designing these data hubs, organization fall into the trap of applying static design to a platform that should honor dynamic design above all.

"The need for increased agility and accessibility for data analysis is the primary driver for data lakes," said Andrew White, vice president and distinguished analyst at Gartner. "Nevertheless, while it is certainly true that data lakes can provide value to various parts of the organization, the proposition of enterprise-wide data management has yet to be realized."

Source: Gartner Says Beware of the Data Lake Fallacy



15 4
15344
16484A
173546
134446
15 4
173334
194153
133736
134446
1 484A
114A51
103741
152043
131446
143637
113753
174641
113247
15144
15344
16484A
173546
134446
15 4
173334
194153
133736
134446
1 484A
114A51
103741
152043
131446
143637
113753
174641
113247
103839
174153
174153
84147
184448
44620
63835
11738
1 344
16361
142047
184A41
154620
101637
163637
184A 8



Transforming via traditional architecture stack is a bit like trying to turn your horse into a camel.

In many design initiatives, data hubs are traditionally architected around existing and well-understood technologies. They usually boil down to a few key elements:

- Central Data Repository: At the center, a data warehouse/lake to store and distribute the data.
- Data Ingestion/Extraction: A set of ETL jobs feeding and extracting data to this central data repository.
- Data Exposure: A middleware/service layer exposing data from the central data repository.
- Data Governance: A set of external components to ensure security, audit, compliance, and other data governance imperatives.

This type of architecture stack, which can be referred to as bottom up because it's taking existing technologies and stacking them up in order to create a data hub, is unfortunately brittle and, therefore, lacking dynamic responsiveness. Indeed, it usually starts with designing the schema to be used in the central data repository. Depending on the industry, multiple standards of design can be used: eTom, NGOSS and the like.

These schemas are very powerful because they guide IT organizations to think about their data in terms of business entities instead of being shackled by the representation of data as it stands in operational applications. The goal is to be simultaneously broadly defined to encapsulate any new application schema, as well as to cater to a specific industry to provide best practices and ease of understanding for the consumer of the data being stored.

This is a very ambitious goal, and whether or not it is actually achieved, it paralyzes the central data repository by shackling it to an immovable structure. From there, every other layer is made to fit into this structure – a bit like fitting square pegs in round holes or trying to turn your horse into a camel.



15344
15344
16884A
173546
134446
115 4
173384
184153
133736
134446
17484A
184A51
183711
152041
134446
143637
113753
174641
113247
115344
15344
16884A
173546
134446
115 4
173384
184153
133736
134446
17484A
184A51
183711
152041
134446
143637
113753
174641
113247
183839
174153
174153
184147
184848
144620
163835
11713
1344
16361
142047
184A41
154620
183637
163637
184A 8

This creates serious responsiveness issues. First, the onerous work that is put into designing the schema, adapting data to it and customizing it for any data demand is a continuous effort. Every new source must be ingested and transformed to fit into this schema, and every data request is then decoupled from the schema to expose data to the services that the business needs.

The end result of this type of design is very long initial setup timelines, and more importantly, lack of the agility and flexibility needed to meet ever-changing user demands, which goes against the grain of the fundamental principles of a data hub.



Virtualization-only & schema-less architectures seem to be the camel you think the business needs, but have inherent flaws.

Aware of these flaws, one market response has been to move to data virtualization and or schema-less architecture. The idea behind these architectures is that as data is needed, a temporary set of data needed to provide the requested answer is created, followed by the creation of the service.

The advantage of this type of architecture is that it is often holistic, ensuring the encapsulation of every layer of a data hub, from ingestion to exposure. Moreover, it has a tendency to drive top-down design: data consumers request the data they need, the virtualization platform

picks the elements needed to fulfill the request, and then the data is exposed. This, in turn, enables the essential agility features that are at the heart of any data hub.



However, the volatile aspect of this type of solution is a massive pitfall. First, by virtualizing data per request, a lot of effort is duplicated. But perhaps more importantly, since the data is not stored in a distributed fashion, identifying sources of truth, or enabling

proper data governance is close to impossible.

15344
15344
16884A
173546
134446
115 4
173384
184153
133736
134446
17484A
184A51
183741
152041
134446
143637
113753
174841
113247
115344
115344
16884A
173546
134446
115 4
173384
184153
133736
134446
17484A
184A51
183741
152041
134446
143637
113753
174841
113247
183839
174153
174153
184147
184848
144628
163835
11753
11344
163641
142047
184A41
154628
101637
163637
184A 8



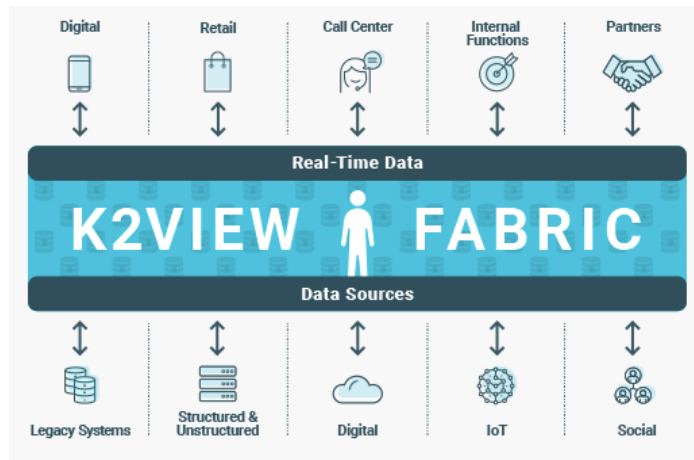
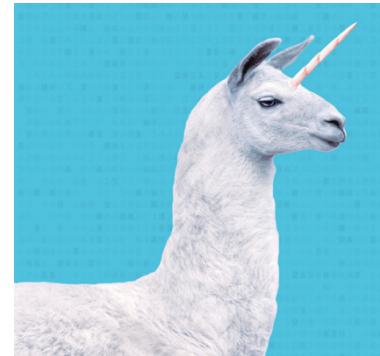
The ideal data hub design: Stop trying to turn your horse into a camel when what you really need is a really fast llama-corn.

A proper data hub should allow both tremendous agility and reusability. It should have a business-driven flexible schema and be a data governance driver, designed around business entity representation - a single, holistic stack, designed from the top-down.

So how do you take all your legacy systems, designed for business-critical operations, and transform them into this ideal solution stack? In short, you don't: You quit trying to turn your horses into camels and just solve the data problem.

This is entirely possible with K2View Fabric, a real-time, end-to-end 360 Data as a Service (DaaS) platform designed specifically with the dynamic, real-time operational needs of digital transformation in mind:

- **AGILITY:** Enable top-down design, eliminate laborious manual processes and increase speed to market using K2View Fabric's business-entity oriented, easy-to-deploy, flexible-on-the-fly schema.
- **PERFORMANCE:** Respond to real-time needs inherent to data hubs through K2View's linear, scalable architecture.
- **HOLISTIC FRAMEWORK:** Simplify your data hub architecture stack using K2View Fabric's end-to-end design.
- **DATA GOVERNANCE:** Ensure full security, audit trail and data governance best practices with K2View Fabric's automated audit, data lineage and governance capabilities.



K2View Fabric has been the fuel of digital transformation for many Fortune 100 brands, accelerating data hub deployment by incrementally building services using top-down design without wasted or duplicated effort, resulting in the robust and comprehensive data hub your digital transformation efforts really need – and in a fraction of the time required for implementing other solutions, allowing you to bring solutions to market faster than ever before.

15 3 4
15348
164848
173546
134448
15 4
173334
184153
133736
134446
1 884A
104A51
103741
152041
131486
143637
113753
174641
113247
103839
174153
174153
84147
184848
44620
63835
11758
1 344
636 1
142047
184A41
154620
101637
163637
184A 8
113753
174641
113247
103839
174153
174153
84147
184848
44620
63835
11758
1 344
636 1
142047
184A41
154620
101637
163637
184A 8

Real-time 360 data In weeks.

Sound too good to be true? It's true anyway.

Let us show you.

Give us a couple hours of your time, and if you qualify, we'll invest 2 weeks to show you something you've never been able to do before.

Simply visit www.k2view.com and request a demo.

Or call us at 1(855)4K2VIEW.

For more information about how K2View Fabric can help, view our videos:

- [Intro to K2view](#)
- [K2View Fabric for Customer 360](#)
- [K2View Technical Overview](#)

K2VIEW
Data without delay.

1.855.4K2VIEW

www.k2view.com

info@k2view.com

115344
115344
16484A
178546
134448
115 4
178334
194153
133736
134446
1 484A
114A51
103741
152041
134448
143637
113753
174641
113247
115344
115344
16484A
178546
134448
115 4
178334
194153
133736
134446
1 484A
114A51
103741
152041
134448
143637
113753
174641
113247
103839
174153
174153
184147
184848
144620
163835
1117 3
1 344
163641
142047
184A41
154620
101637
163637
184A 8