Data Clouds: One Small Step for IT, One Giant Leap for Modern Enterprises

ABSTRACT

The failure to connect information buried within silos has proven to be a costly and dangerous problem. Sad experience has shown that organizations' are not marshaling their siloed information to support enterprise goals. Data clouds are an innovative IT concept that provides businesspeople with self-service access to integrated and mashable information from siloed information systems. Data clouds offer a way for enterprises to connect the data within their silos without having to eliminate them. The technologies that enable data clouds are relatively new but readily available. A few pioneering enterprises are already implementing their own data clouds.

A data cloud is a new IT concept that puts the full power of an enterprise's computer-based information into the hands of the people who carry out its mission. Data clouds provide self-service access to integrated and mashable information from siloed information systems. Data clouds liberate information from application silos and make it appropriately accessible to the businesspeople who need it. Forward-thinking enterprises are already implementing data clouds, and the technologies to build them are readily available. Enterprises can implement data clouds in a step-by-step, cost-effective manner, using a small collection of innovative tools and techniques.

BIOGRAPHY

Lyn Robison
Vice President and Research Director
Data Management Strategies
Gartner, Inc.

Lyn Robison is the Research Director for Gartner's Data Management Strategies service. Lyn directs the research of analysts who provide guidance for Global 2000 corporations on data management issues. He is the author of two books and numerous articles on information technology, and is a frequent speaker at IT industry conferences.

Data Clouds: One Small Step for IT, One Giant Leap for Modern Enterprises

Lyn Robison

VP and Research Director

Data Management Strategies

Gartner

A new IT concept called "data clouds" offers an innovative, systematic way to integrate information buried in silos and marshal it to support enterprises' missions. Data clouds connect information that is scattered in silos, and put all appropriate information at the fingertips of the people who need it most: the people responsible for the work of the enterprise.

1

Gartner

Agenda

- The Pervasive Silo Problem
- Silos, Cloud Computing, and Data Clouds
- Implementing a Data Cloud
- Data Cloud Infrastructure
- Recommendations
- Conclusion

2

Gartner

Agenda

- The Pervasive Silo Problem
- Silos, Cloud Computing, and Data Clouds
- Implementing a Data Cloud
- Data Cloud Infrastructure
- Recommendations
- Conclusion

3

Gartner

The Pervasive Silo Problem

- Silos of information exist in large enterprises today and will continue to exist, and the failure to connect the information buried within them is a dangerous problem.

4

Gartner

The Pervasive Silo Problem

- The U.S. intelligence community is experiencing continued difficulty in accurately and appropriately identifying terrorists in homeland security

  - "Michael Leiter (Director of the National Counterterrorism Center) spends much of his day flipping among four computer monitors lined up on his desk. Six hard drives sit at his feet. The data flow is enormous, with dozens of databases feeding separate computer networks that cannot interact with one another. *There is a long explanation for why these databases are still not connected, and it amounts to this: It's too hard, and some agency heads don't really want to give up the systems they have*"

5

Gartner

The Pervasive Silo Problem

- Hazards for medical patients due to a lack of data sharing between clinical organizations

    - According to *The Health Care Blog*, "We still have silos of information locked away in hospitals, offices, pharmacies, and labs. We still have redundant and unnecessary testing because our care is uncoordinated. We're still using a huge amount of paper in our healthcare facilities. Paper kills."

6

Gartner.

The Pervasive Silo Problem

- CFOs' ongoing and unfulfilled need for integrated financial information to help enterprises prosper in today's uncertain business environment

    - In the corporate world, CFOs' demand for financial information that is fully integrated instead of buried deep within silos is well documented but yet to be satisfied.

7

Gartner.

The Pervasive Silo Problem

- A lack of visibility into cross-site and cross-functional supply chains

  - The research report "The Product Supply Strategy: The Changed Face of the Traditional Manufacturing Strategy" states, "Most manufacturing companies struggle with the realities of complexity, lack of agility, redundant capacity, and poor performance in the product supply system because production operations are still site centric, fragmented, and structured in functional silos with poor cross-site and cross-functional visibility."

8

Gartner.

The Pervasive Silo Problem

- A profound inability to assess systemic risk, which contributed to the 2008 meltdown of the U.S. financial system

  - John Liechty, associate professor of Marketing and Statistics at Penn State University, co-founder of the Committee to Establish National Institute of Finance said the following in a Business Week interview last August 18: "Does anybody have the data in place to really deal with systemic risk"? Had this collection of data and analytics existed last fall, "regulators could have modeled the repercussions of a Lehman collapse using actual data…"
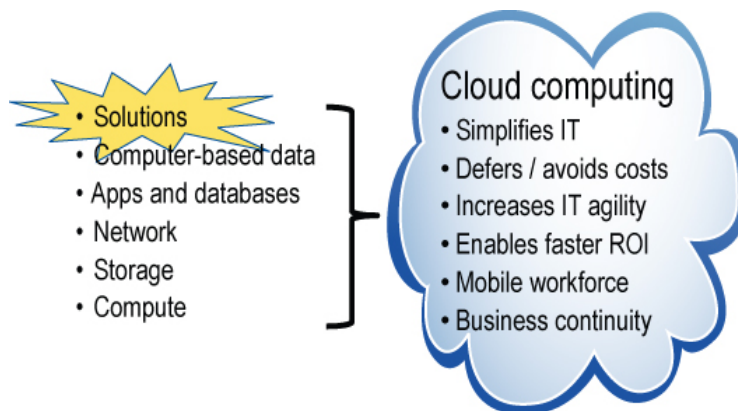
9

Gartner.

Agenda

- The Pervasive Silo Problem
- Silos, Cloud Computing, and Data Clouds
- Implementing a Data Cloud
- Data Cloud Infrastructure
- Recommendations
- Conclusion

10

Gartner.

Silos, Cloud Computing, and Data Clouds
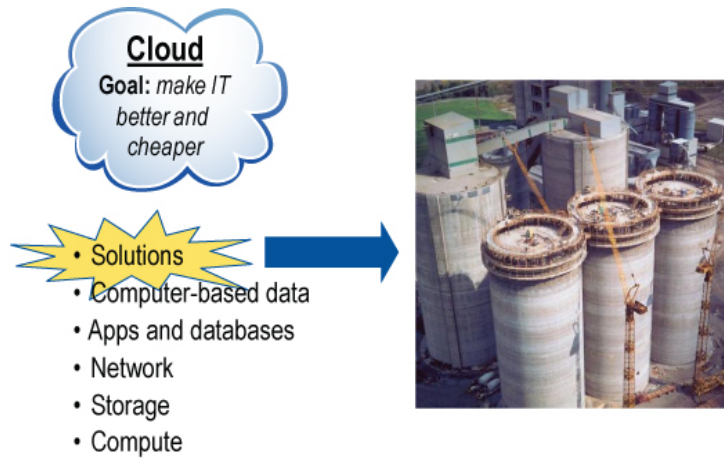
- The IT Solution Stack and Cloud Computing

11

Gartner.

Silos, Cloud Computing, and Data Clouds
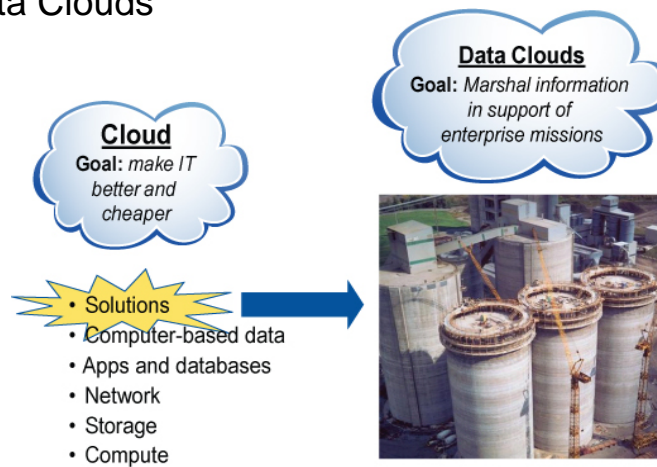
- Silos in Cloud Computing

Cloud
Goal: *make IT better and cheaper*

- Solutions
- Computer-based data
- Apps and databases
- Network
- Storage
- Compute

12

Gartner

Silos, Cloud Computing, and Data Clouds

- Data Clouds

Cloud
Goal: *make IT better and cheaper*

Data Clouds
Goal: *Marshal information in support of enterprise missions*

- Solutions
- Computer-based data
- Apps and databases
- Network
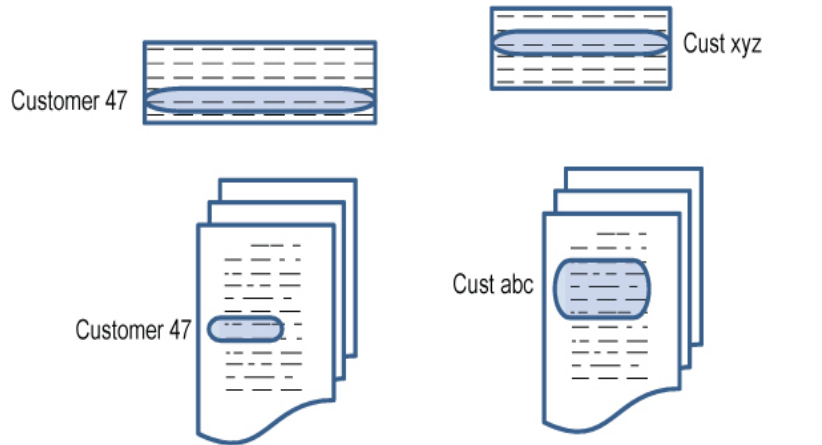- Storage
- Compute

13

Gartner

Silos, Cloud Computing, and Data Clouds

- The Problem of Identifiers
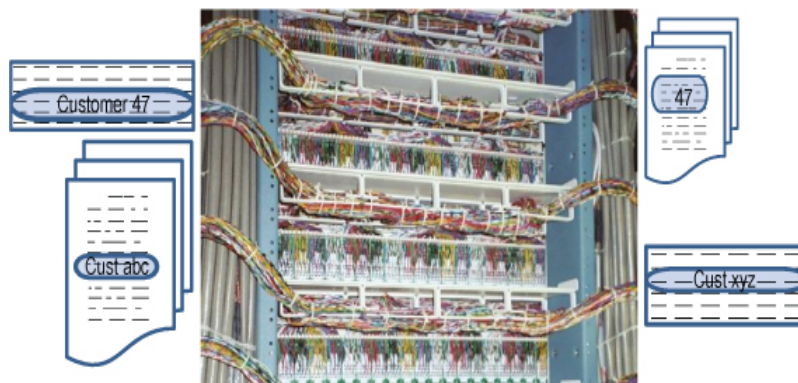

14

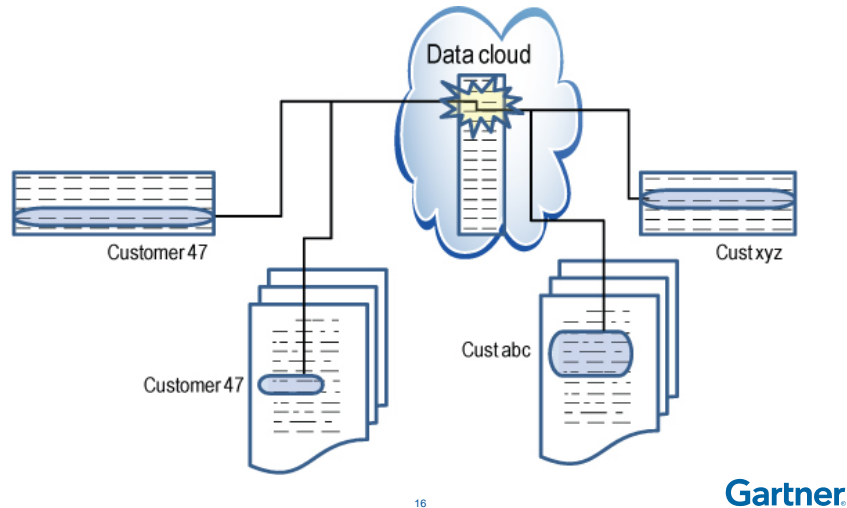# Silos, Cloud Computing, and Data Clouds

- Data Clouds as Patch Panels

15

Silos, Cloud Computing, and Data Clouds

- Connect Individual Data Items in Silos

Silos, Cloud Computing, and Data Clouds

- Defining Data Clouds

  - A data cloud is a means of reconciling, virtualizing, and integrating computer-based information that resides in silos, and making that integrated information appropriately available online so businesspeople can search, query, and analyze it in a freeform manner using web browsers, mobile devices, off-the-shelf software applications, and analysis tools. The data in the data cloud may or may not be duplicated from the siloed source systems into the cloud—the data can remain in place within its native silos. A data cloud often merely indexes and/or connects the data that resides in existing source systems, like a library card catalog does for books in a library or a patch panel does for telephones.
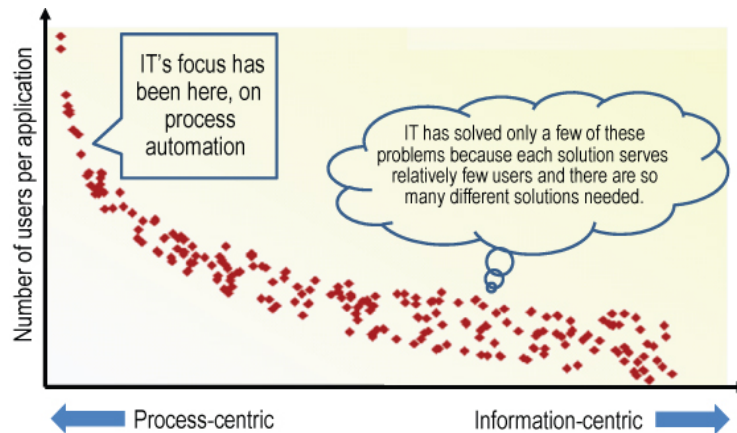
Agenda

- The Pervasive Silo Problem
- Silos, Cloud Computing, and Data Clouds
- **Implementing a Data Cloud**
- Data Cloud Infrastructure
- Recommendations
- Conclusion

Gartner

18

Implementing a Data Cloud
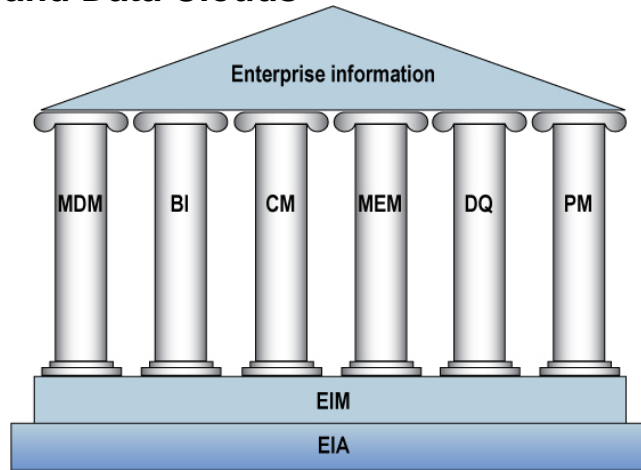
- The Long Tail of Enterprise IT
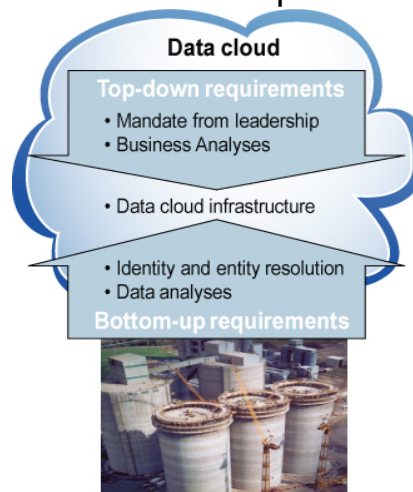
Implementing a Data Cloud

- **EIM and Data Clouds**

Implementing a Data Cloud

- Successful Data Clouds require Infrastructure

Agenda

- The Pervasive Silo Problem
- Silos, Cloud Computing, and Data Clouds
- Implementing a Data Cloud
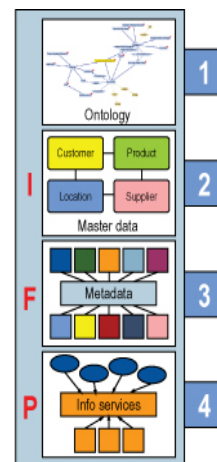- Data Cloud Infrastructure
- Recommendations
- Conclusion

22

Gartner.

Data Cloud Infrastructure

- Forming, Storming, Norming, Performing

1. **Forming:**
   Develop an ontology of conceptual entities and attributes

2. **Storming:**
   Designing the names for basic entities and attributes, such as master data

3. **Norming:**
   Design aggregate concepts (e.g., product profitability)

4. **Performing:**
   Design protocols for shared use of names/structures

Application neutral

Identifiers
Formats
and
Protocols

Source neutral

23

Gartner.

Data Cloud Infrastructure

- Gartner defines "cloud computing" as the set of disciplines, technologies, and business models used to deliver IT capabilities (software, platforms, hardware) as an on-demand, scalable, elastic service. Five essential characteristics of cloud computing:

  - It uses shared infrastructure.

  - It provides on-demand self-service.

  - It is elastic and scalable.

  - It is priced by consumption.

  - It is dynamic and virtualized.

24

Gartner

Data Cloud Infrastructure

- Four essential characteristics of data clouds:

  - Self-service for businesspeople.

  - Entity and identity resolution.

  - High-fidelity data.

  - Powerful analytics.

25

Gartner

Data Cloud Infrastructure

- Four categories or types of data clouds:
  - Think of Level "D" as a "Data warehouse-level data cloud"
  - Level "C" is a "Commercial data cloud"
  - Level "B" is a "Big data cloud"
  - Level "A" is an "All-purpose data cloud"

26

Gartner.

Agenda

- The Pervasive Silo Problem
- Silos, Cloud Computing, and Data Clouds
- Implementing a Data Cloud
- Data Cloud Infrastructure
- Recommendations
- Conclusion

27

Gartner.

Recommendations

- Data clouds are best built incrementally, in small, deliberate steps, each yielding tangible business benefits

  - Start Small

  - Start Right Away to Implement a Simple Data Cloud

  - Measure, Document, and Publish Your Progress and Results

  - Look for Opportunities to Move Beyond Simple Data Clouds

  - Build a Business Case and More-Capable Prototype

28

Gartner

Agenda

- The Pervasive Silo Problem

- Silos, Cloud Computing, and Data Clouds

- Implementing a Data Cloud

- Data Cloud Infrastructure

- Recommendations

- Conclusion

29

Gartner

Conclusion

- Marshaling an enterprise's computer-based information to support its mission is an audacious and powerful goal. Data clouds offer a way to achieve this goal. The technologies that enable a data cloud infrastructure are relatively new, but are readily available. Given the newness of this concept, a step-by-step approach is called for, and fortunately data clouds are readily implemented in stages. Each small step that IT takes towards data clouds will result in a giant leap in the effectiveness of modern enterprises.

30

Gartner