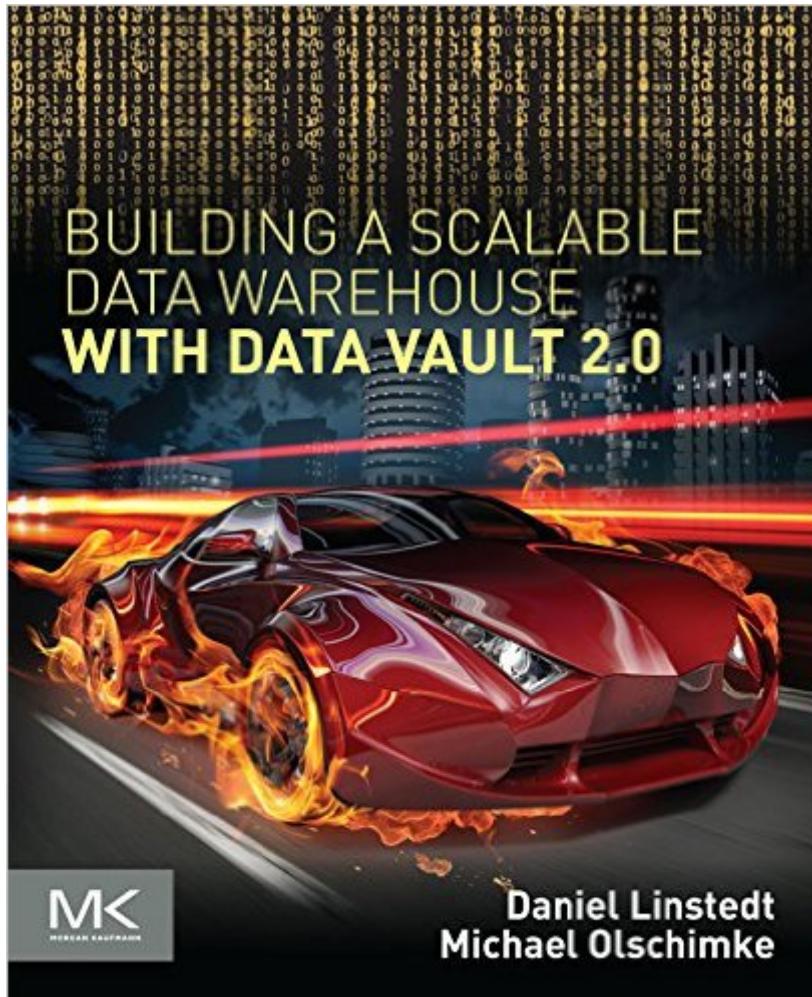


The book was found

# Building A Scalable Data Warehouse With Data Vault 2.0



## Synopsis

The Data Vault was invented by Dan Linstedt at the U.S. Department of Defense, and the standard has been successfully applied to data warehousing projects at organizations of different sizes, from small to large-size corporations. Due to its simplified design, which is adapted from nature, the Data Vault 2.0 standard helps prevent typical data warehousing failures. "Building a Scalable Data Warehouse" covers everything one needs to know to create a scalable data warehouse end to end, including a presentation of the Data Vault modeling technique, which provides the foundations to create a technical data warehouse layer. The book discusses how to build the data warehouse incrementally using the agile Data Vault 2.0 methodology. In addition, readers will learn how to create the input layer (the stage layer) and the presentation layer (data mart) of the Data Vault 2.0 architecture including implementation best practices. Drawing upon years of practical experience and using numerous examples and an easy to understand framework, Dan Linstedt and Michael Olschimke discuss: How to load each layer using SQL Server Integration Services (SSIS), including automation of the Data Vault loading processes. Important data warehouse technologies and practices. Data Quality Services (DQS) and Master Data Services (MDS) in the context of the Data Vault architecture. Provides a complete introduction to data warehousing, applications, and the business context so readers can get-up and running fast Explains theoretical concepts and provides hands-on instruction on how to build and implement a data warehouseDemystifies data vault modeling with beginning, intermediate, and advanced techniquesDiscusses the advantages of the data vault approach over other techniques, also including the latest updates to Data Vault 2.0 and multiple improvements to Data Vault 1.0

## Book Information

File Size: 125451 KB

Print Length: 663 pages

Publisher: Morgan Kaufmann; 1 edition (September 15, 2015)

Publication Date: September 15, 2015

Sold by: Digital Services LLC

Language: English

ASIN: B015KKYFGO

Text-to-Speech: Enabled

X-Ray: Not Enabled

Word Wise: Not Enabled

Lending: Not Enabled

Enhanced Typesetting: Not Enabled

Best Sellers Rank: #209,953 Paid in Kindle Store (See Top 100 Paid in Kindle Store) #69

in Books > Computers & Technology > Databases & Big Data > Data Warehousing #128

in Books > Computers & Technology > Databases & Big Data > Data Modeling & Design #2760

in Kindle Store > Kindle eBooks > Computers & Technology

## Customer Reviews

This book fills a huge void in Data Vault 2.0 resources. It covers everything about data vault, including data modeling, ETL processing, error handling, metadata, data quality and more, all explained in depth with sufficient examples that can be immediately put to use. Every question that I could think of relating to data vault is covered in the book. It is an excellent reference manual to have on hand while doing data vault implementations. That said, I gave the book only four stars because there is always room for improvement. I was disappointed after reading Chapter 3, about the Data Vault 2.0 methodology. It appears that the authors threw in a bunch of various methodologies, frameworks, best practices and approaches without clearly explaining how all these parts fit together into a cohesive methodology. The authors appear out of their depth in the project management space. They incorrectly apply the term PMP (Project Management Professional, a certification from the Project Management Institute) as a best practice. What they probably meant is that the project management aspect of the Data Vault 2.0 methodology is taken from the PMBOK (Project Management Body of Knowledge), a project management standard issued by the Project Management Institute. I also don't see why the authors included Scrum in the mix. The explanation of how Scrum is applied in mini-waterfall sprints is confusing and contradicting. For example, it is not clear how the Scrum roles such as scrum master and product owner translate into the mini-waterfall sprints. They state that Scrum is used for team coordination/team organization and that the team is self-organized while at the same time they designate the project manager as the person who plans the tasks within a mini-waterfall sprint.

[Download to continue reading...](#)

Building a Scalable Data Warehouse with Data Vault 2.0 Data Architecture: A Primer for the Data Scientist: Big Data, Data Warehouse and Data Vault The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights Building Scalable Web Sites: Building, Scaling, and Optimizing the Next Generation of Web Applications The Data Warehouse Lifecycle Toolkit The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling Kimball's Data Warehouse

Toolkit Classics: 3 Volume Set Mastering Data Warehouse Aggregates: Solutions for Star Schema Performance Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business Leveraging the Power of Data Analytics, Data Science, ... (Hacking Freedom and Data Driven Book 2) Practical Node.js: Building Real-World Scalable Web Apps Web Development with Go: Building Scalable Web Apps and RESTful Services Microsoft Excel 2013 Building Data Models with PowerPivot: Building Data Models with PowerPivot (Business Skills) Big Data For Beginners: Understanding SMART Big Data, Data Mining & Data Analytics For improved Business Performance, Life Decisions & More! The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences Guerrilla Capacity Planning: A Tactical Approach to Planning for Highly Scalable Applications and Services Frontend Architecture for Design Systems: A Modern Blueprint for Scalable and Sustainable Websites uC/OS-III, The Real-Time Kernel, or a High Performance, Scalable, ROMable, Preemptive, Multitasking Kernel for Microprocessors, Microcontrollers & DSPs (Board NOT Included) The Art of Scalability: Scalable Web Architecture, Processes, and Organizations for the Modern Enterprise (2nd Edition) Programming Google App Engine with Python: Build and Run Scalable Python Apps on Google's Infrastructure Scalable Shared-Memory Multiprocessing

[Dmca](#)